The University of Kansas
Department of Economics

High-Dimensional Challenges in Financial Economics and Macro-Finance - A Financial Econometric and Machine Learning Approach

Caio A. Vigo Pereira

June 2021

High-Dimensional Challenges in Financial Economics and Macro-Finance - A Financial Econometric and Machine Learning Approach

by

Caio A. Vigo Pereira
M.A., University of Kansas, 2018
M.Sc., Universidade de São Paulo, 2016
B.Sc., Universidade de São Paulo, 2011

Submitted to the graduate degree program in Economics and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. Tarun Sabarwal, Chair

Dr. John Keating, Member

Dr. John Zhu, Member

Dr. Yehuda Izhakian, External Member

Dr. Bozenna Pasik-Duncan, External Member

Date defended:

The Dissertation Defense Committee for Caio Vigo Pereira certifies that this is the approved version of the following dissertation defense :

High-Dimensional Challenges in Financial Economics and Macro-Finance - A Financial Econometric and Machine Learning Approach

Dr. Tarun Sabarwal, Dissertation Chair

Date approved: ____________________

Abstract

This work investigates how to properly deal with high-dimensional problems in applications in the fields of financial economics and macro-finance. The unifying theme for the essays here presented is the use of financial econometrics methods in high-dimension settings. The first essay is concerned with the time variation of the risk premia in U.S. Treasuries bonds. I take a different route from the prevailing literature, as I argue that the process of manually discovering and hand-picking a list of factors to understand the U.S. bond risk premia is leave out unseen relationships between the state variables in their derivation. I overcome many issues from previous works making use of deep neural networks to uncover relationships in the full set of information from the yield curve. I propose a novel approach for deriving a single state factor consistent with a dynamic term-structure model. My framework generates an econometric approach that fits with the existing theory of U.S. Treasury bond pricing. I show that this parsimonious number of state variables has predictive power for excess returns of bonds over a one-month holding period. In the second essay, I build efficient portfolios using high-dimensional information sets as conditioning information and conduct an extensive out-of-sample analysis comparing many different techniques in order to impose sparsity and dimensionality reduction when dealing with a large set of potential conditioning information. The main contribution of this essay is that it is possible to condense the large set of potential predictors to build meaningful efficient portfolios. Finally, in the third essay, I use penalized regression methods in which large datasets are necessary to capture information from one-minute tick data to identify short-lived signals in the intraday of FX markets.

Contents

List of Figures

List of Tables

Introduction

This dissertation is a collection of three essays, organized in three chapters, in the fields of macro-finance and financial economics. A unifying theme for these essays is the use of financial econometrics methods in high-dimension settings. In each essay, I seek to make contributions to the current literature, empirically and/or proposing new methodologies theoretically grounded in economic models. Interestingly, even though all three works are in the same field of econometrics methods applied to finance, each chapter deals with a different market and economic theory. Consequently, each work seeks to answer different research questions. In short, each paper deals with (i) U.S. Treasury bonds, (ii) stocks, and (iii) foreign exchange markets.

The first chapter is concerned with the time variation of the risk premia in U.S. Treasuries bonds. An alternative at the forefront of new research in financial economics is the use of machine learning techniques, which has been a promising and fruitful path to understand the behavior of financial markets. In the first essay, I take a different route from the prevailing literature, as I argue that the search for deriving, building, and estimating factors that represent state variables in macro-finance models may be limited. I show that the process of manually discovering and hand-picking this list of factors is leaving out unseen relationships between the state variables in their derivation. To do so, I make use of one of the most powerful approaches in machine learning, namely a deep neural network to uncover relationships in the full set of information from the yield curve. Considered a state-of-the-art technique, deep neural networks are among the most powerful approaches in machine learning. Their use is growing exponentially from areas ranging from autonomous robots and facial recognition to predictive analytics.

The literature in macro-finance argues that a dominant single factor structure for bond returns is possible, in such a way that risk premiums rise and fall together. In this essay I seek to provide a new interpretation for this long debated question as I provide a linear combination of forecasting variables that captures common movement in expected returns across Treasuries bonds. I propose a novel approach consistent with the spanning hypothesis to derive a single state variable factor from a deep neural network that provides a better approximation to the spanned space of all the information from the term-structure. I also introduce a way to obtain unspanned risks from the yield curve that is used to complete the state space. I show that this parsimonious number of state variables have predictive power for excess returns of bonds over 1-month holding period. Additionally, I provide an intuitive interpretation of derived factors and show what information from macroeconomic variables and sentiment-based measures they can capture.

Chapter two deals with the stock market. I combine the nascent, but fast-developing high-dimensional literature in finance to the portfolio theory problem. I build efficient portfolios using high-dimensional information sets as conditioning information and conduct an extensive out-of-sample analysis comparing the high-dimensional statistical methods of dealing with a large set of potential conditioning information. The literature has recognized a large number of variables with have predictive power. However, this search for all these instruments produced an enormous amount of variables that supposedly have or have had some predictive power.

The main argument of this chapter is to "bet in sparsity", in the sense that much of the information in the conditioning set can be summarized by few factors. I do this using many different techniques in order to impose sparsity and dimensionality reduction when finding the conditional mean, which is the most important driver in the formation of mean-variance efficient portfolios with conditioning information. I evaluate how penalized estimators, such as LASSO, Ridge and Elastic Net, as well as pure dimensionality reduction and latent factors approaches like Partial Least Squares (PLS) and Principal Components Regression (PCR), in addition to a generalization of the former (Three Pass Regression Filter) can produce different optimized portfolios.

In contrast to previous studies that made use of naive OLS and low-dimension information sets, I show that (i) accounting for large conditioning information sets, and (ii) the use of variable selection, shrinkage methods, and factors models, such as the principal component regression and the partial least squares provide better out-of-sample results as measured by Sharpe ratios.

Finally, chapter three deals with the FX market. I use high-frequency foreign exchange data to estimate short-lived signals. Given the complexity and interconnectedness in the current state of financial markets, movements in many other variables can be informative to a specific currency. In order to try to capture this possibly short-lived and unexpected signaling information, I make use of the Elastic-Net estimator to assist us in the process of selecting important sources of predictability, and shrinking and discarding uninformative information that simply adds noise to the process. I forecast one-minute-ahead rolling currency prices using other currency pairs, commodities, and stock market indices as predictors.

With such a large set of covariates, I impose sparsity to find unexpected signals in the intraday currency market, which allow us to deal dynamically with multicollinearity, while constraining the size of the estimates of uninformative variables. We investigate the existence of signal patterns for the five most liquid currencies (British Pound, Canadian Dollar, Euro, Japanese Yen, and Swiss-Franc).

Chapter 1

A Machine Learning Factor-Based Interpretation for the Bond Risk Premia in the U.S.

Abstract

In this paper, we study the time variation of the risk premia in U.S. Treasuries bonds. We propose a novel approach for deriving a single spanning state factor consistent with a dynamic term-structure with unspanned risks theoretically motivated model. Using deep neural networks to uncover relationships in the full set of information from the yield curve, we derive a single state variable factor that provides a better approximation to the spanned space of all the information from the term-structure. We also introduce a way to obtain unspanned risks from the yield curve that is used to complete our state space. We show that this parsimonious number of state variables have predictive power for excess returns of bonds over 1-month holding period. Additionally, we provide an intuitive interpretation of derived factors and show what information from macroeconomic variables and sentiment-based measures they can capture.

JEL classification: G12, E43, E44, E47.

Keywords: Bond Premia. Deep Learning. Machine Learning. Bond Returns. Yield Curve. Unspanned Risk Factors.

1.1 Introduction

In recent years, many studies had shed light on a critical assumption in macro-finance models, the expectations hypothesis. As more evidence is gathered, there is a growing consensus in the literature to refute it, implying that excess returns of Treasuries bonds in some extent should be forecastable. Equally important is the spanning hypothesis, that can be summarized in the idea that the yield curve incorporates all the information useful for forecasting interest rates, and consequently, bonds returns. However, to what extent the spanning hypothesis holds true is still open in the literature.

An important question that could assist to elucidate the whole bond premia problem is related with the factor structure of expected returns. Is there a factor representation? If so, what is its structure? In this article, we study the time variation of the risk premia in U.S. Treasuries bonds. We provide a new approach for the factor structure of the expected returns of bonds. Recently, Cochrane (2015) argued that it is possible that there is a dominant single factor structure for bond returns, in such a way that risk premiums rise and fall together. A central question, in his words, is: *what is the linear combination of forecasting variables that captures common movement in expected returns across assets*?

In Cochrane and Piazzesi (2005), the authors took this path. Ludvigson and Ng (2009) derived a single factor as well, however not consistent with the spanning hypothesis. Recent papers (see, e.g., Cieslak and Povala (2015); Lee (2018)) obtained other factors as well, some of them not necessarily aligned with the spanning hypothesis. Nonetheless, Bauer and Hamilton (2018) argued that evidence against the spanning hypothesis for several recent studies should be weaker when more robust tests are used.

In this paper, we take a different route. We argue that this search for deriving, building and estimating factors that represent state variables in macro-finance models may be limited. We claim that the process done by financial economists of manually discovering and hand picking this list of factors may be leaving out unseen relationships between the state variables in their derivation.

We propose a novel approach for deriving a single state factor consistent with a dynamic term-structure with unspanned risks theoretically motivated model. To do so, we make use of one of the most powerful approaches in machine learning, namely a deep neural network to uncover relationships in the full set of information from the yield curve. We derive a single state variable factor that should provide a better approximation to the spanned space of all the information from the term-structure.

In our methodology, we introduce a way to obtain unspanned risks from the yield curve that is used to complete our state space. This unspanned factor can fill the gap left by the spanning factor. The whole structure can be explained by a dynamic term-structure with unspanned risks, be macroeconomic, sentiment, or any other economy risk (since our methodology makes no differentiation or segregation among them) as an extension from the model proposed by Joslin et al. (2014). We show that a small numbers of state variables (in our framework only two), have predictive power for excess returns of bonds over 1-month holding period. Additionally, we provide an intuitive interpretation of derived factors, and show what information from macroeconomic variables and sentiment-based measures they can capture.

In our empirical analysis, we use the daily treasury parameters from Gürkaynak et al. (2007) to build the monthly excess returns. As we are interested in the short part of term structure, we deal with yields ranging up to five years ahead. Thus, we work with 60 observations at each month to better capture the information from the term structure. Our data spans from 1962 to 2017, and we use the data from 1962-1992 to initialize the process of obtaining our recursively updated latent factors in order to build a single spanning and unspanning factor for the period from 1993-2017.

The recursive process is done through a pre-designed architecture of a deep neural network that is fed solely with the high-dimensional set of monthly with different maturities. This architecture generates one latent factor for the two, three, four, and five years excess returns. Then, these recursively updated parameters are combined in a single spanning factor. The unspanned factor also has recursively updated parameters that are obtained through an orthogonalization process at each month, and then linearly combined into one single factor.

We evaluate our factors for the period of 1993-2017 and compare their predictability with the main factors in the literature, such as Cochrane and Piazzesi (2005), Ludvigson and Ng (2009), and Fama and Bliss (1987). We also perform an out-of-sample experiment to assess the statistical evidence in bond return predictability for the period 1997-2017.

We contribute to the literature in a number of different ways. First, to the best of our knowledge, we are the first to introduce a deep neural network-based structure to generate a single spanning factor, as well as an unspanned factor, with recursively updated latent factors. Thus, we are able to introduce nonlinearities when modeling the bond risk premia in our first step of the recursive process, while still making use of a linear combination of these latent factors in the second step, to provide a novel interpretation for the bond risk premia, as proposed by Cochrane's question. This is consistent with recent findings from some works in empirical finance (see, e.g., Gu et al. (2018); Bianchi et al. (2019)) that document the importance of allowing for nonlinearities. It is known that with neural networks we can introduce flexibility and complex nonlinear relationships from the inputs while approximating arbitrarily well a rich set of smooth functions.

Second, motivated by Bauer and Hamilton (2018) that document that the use of overlapping 12-month returns is prone to a number of problematic features, as it introduces substantial serial correlation in the predictive errors, we deviate from previous works that made use mainly of 12-month holding period, as we handle this issue with the use of non-overlapping returns. Furthermore, we are interested to obtain these factors to the short part of the term structure, and to do so we take into consideration the whole term structure at the higher frequency of 1-month holding period with maturities ranging up to 60 months ahead, allowing us to avoid to handpick only a subset of yields.

Third, we take a broader interpretation of the unspanning factor. Thus, our derived unspanned factor is more flexible, as it can be linked with other sources of risks, not limiting only to macroeconomics variables, but also with sentiment-based variables. And fourth, our approach avoids hand-picking the variables from the yield curve, as through our deep neural network we are able to recursively learn the best-approximating[1] function that condenses the yield curve into a single latent factor.

1.1.1 Related Literature

Our paper is related to several strands of the literature. This paper is related with the so known "spanning puzzle" which pertains a possible conflict between the theoretical spanning condition, in which the yield curve captures all the information for forecasting future yields and returns, and the use of unspanned macro information for these problems. Ludvigson and Ng (2009), Cooper and Priestley (2009), and Cieslak and Povala (2015) provide evidence that macroeconomic variables have predictive power for excess bond. On the other hand, Ghysels et al. (2018) show that the use of real time data substantially reduces the predictive power of macro variables for future bond. Bauer and Hamilton (2018) show that non spanning predictors proposed in the literature is weaker than expected, while Bauer and Rudebusch (2017) argue that the evidence from unspanned regressions cannot provide statistical basis for preferring either unspanned or spanned models.

Our work is also linked to a literature at the intersection of bond premia and sequential learning. Gargano et al. (2019) and Dubiel-Teleszynski et al. (2019) make use of a Bayesian learning approach in the context of bond risk premia. The former accounts for time-varying parameters, stochastic volatility, and parameter estimation error; while the latter implement the learning framework under a dynamic term structure model.

This paper is also related with the recent surge in the use of machine-learning and its growing impact in the field of economics, as reviewed in Mullainathan and Spiess (2017);

[1] For a given loss function.

Athey (2018); Athey and Imbens (2019); Varian (2014). This paper is also motivated by recent advances in the statistical properties of machine learning techniques, in especial the theoretical properties of inference using deep neural networks . Farrell et al. (2021), provide nonasymptotic bounds and convergence rates for nonparametric estimation using deep neural networks, and establishes valid inference on finite-dimensional parameters following first-step estimation using deep learning.

The use of recent machine-learning techniques in empirical finance has gotten special interest in the past years. Many papers have been interested in the dimensionality reduction, especially through penalized regression framework, such as LASSO, Ridge and Elastic-Net (see, e.g., Kozak et al. (2019); Freyberger et al. (2017)). Deep learning and some variations of deep neural networks, such as autoenconders, were used in some recent papers (Gu et al., 2018, 2020; Chen et al., 2019; Feng et al., 2018a,b; Heaton et al., 2017, 2016).

Specifically in macro-finance, Bianchi et al. (2019) and Huang and Shi (2019) make use of machine-learning techniques to model or evaluate bond return predictability. Huang and Shi (2019) use Supervised Adaptive Group LASSO to capture macroeconomic risks, and construct a single unspanned factor from a panel of 131 macro variables. Bianchi et al. (2019) seek to compare and evaluate several machine learning algorithms for the sole purpose of prediction of the US Treasury bonds excess returns. Their analysis ranges from penalized linear regressions, partial least squares, regression trees, random forests, and finally neural networks. The authors find evidence that non-linear methods can provide favorable out-of-sample prediction of bond excess returns. Our work builds on some of the insights from Bianchi et al. (2019) with regard of the use of deep neural networks to understanding bond premia. Nonetheless, our approach detours from their work in numerous ways, being the most important our goal to build a spanning and an unspanned factor from a deep neural network.

The structure of this paper is as follows. Next section introduces the general framework, contextualize the expectations and spanning hypothesis, and explain the deep-learning structure that we propose for bond premia. This section also provides an illustrative term-structure model. Section 1.3 explains our data, how we reconstruct the log yield of zero-coupons, and elucidate our empirical strategy. Section 1.4 presents the results. Finally, section 1.5 concludes. Additional results, tables and figures are presented in Appendix 1.6.1.

1.2 Framework

1.2.1 Notation

Following the standard notation in the literature, let $p_t^{(n)}$ denote the natural logarithm of the price for a bond with n-period maturity at time t, and y represent its yield, such that:

$$y_t^{(n)} \equiv -\frac{1}{n} p_t^{(n)} \tag{1.1}$$

The holding period returns of a n-period maturity bond from time t to $t+\Delta$ is given by:

$$r_{t+\Delta}^{(n)} \equiv p_{t+\Delta}^{(n-\Delta)} - p_t^{(n)} \tag{1.2}$$

If integers of Δ represent years, then:

$$\begin{aligned} r_{t+h/12}^{(n)} &\equiv p_{t+h/12}^{(n-h/12)} - p_t^{(n)} \\ &= n y_t^{(n)} - (n-h/12) y_{t+h/12}^{(n-h/12)} \end{aligned} \tag{1.3}$$

where h is the frequency of the returns, measured in months. Thus, we can define the excess returns as

$$\begin{aligned} rx_{t+h/12}^{(n)} &\equiv \underbrace{p_{t+h/12}^{(n-h/12)} - p_t^{(n)}}_{\text{holding period return } r_{t+h/12}^{(n)}} - (h/12) y_t^{(h/12)} \\ &= n y_t^{(n)} - (n-h/12) y_{t+h/12}^{(n-h/12)} - (h/12) y_t^{(h/12)} \end{aligned} \tag{1.4}$$

Finally, we can define the forward rates at time t for loans between time $t+n-h/12$ and $t+n$ as

$$\begin{aligned} f_t^{(n)} &\equiv p_t^{(n-h/12)} - p_t^{(n)} \\ &= n y_t^{(n)} - (n-h/12) y_t^{(n-h/12)} \end{aligned} \tag{1.5}$$

1.2.2 Expectation Hypothesis and the Spanning Hypothesis

In its most common form, the *expectation hypothesis* states that yields of long maturity bonds should be the average of the future expected yield of short maturity bonds. Hence, it is equivalent with the statement that excess returns should not be predictable. Setting

$h=1$ to express monthly frequency, the expectations hypothesis can be summarized as[2]

$$y_t^{(n)} \equiv \underbrace{\frac{1}{n}\mathbb{E}_t\left(y_t^{(1/12)} + y_{t+1/12}^{(1/12)} + \ldots + y_{t+n-1/12}^{(1/12)}\right)}_{\text{expectations component}} + \quad \text{yield risk premium} \quad . \tag{1.8}$$

In short, we can summarize the risk premium simply as the difference between a long rate and the expected average of future short rates. Knowing that we can express the *yield risk premium* as $\frac{1}{n}\mathbb{E}\left(rx_{t+1/12}^{(n)} + rx_{t+2/12}^{(n-1/12)} + rx_{t+3/12}^{(n-2/12)} + \ldots + +rx_{t+n-1/12}^{(2/12)}\right)$, then we can write

$$\begin{aligned} y_t^{(n)} \equiv\ & \frac{1}{n}\mathbb{E}_t\left(y_t^{(1/12)} + y_{t+1/12}^{(1/12)} + \ldots + y_{t+n-1/12}^{(1/12)}\right) + \\ & \frac{1}{n}\mathbb{E}_t\left(rx_{t+1/12}^{(n)} + rx_{t+2/12}^{(n-1/12)} + rx_{t+3/12}^{(n-2/12)} + \ldots + +rx_{t+n-1/12}^{(2/12)}\right) \quad . \end{aligned} \tag{1.9}$$

As in Duffee (2013), assuming that the agents' information set at time t can be summarized by a k-dimensional state vector $\mathbf{Z}_t$, from identity 1.9 we obtain

$$y_t^{(n)} = \frac{1}{n}\left(\sum_{j=0}^{12\cdot n/h-1} \mathbb{E}\left[y_{t+j\cdot h/12}^{(h/12)} | \mathbf{Z}_t\right]\right) \quad + \quad \frac{1}{n}\left(\sum_{j=0}^{12\cdot n/h-1} \left[rx_{t+h/12(j+1)}^{(n-j\cdot h/12)} | \mathbf{Z}_t\right]\right) \quad . \tag{1.10}$$

In equation (1.10), $\mathbf{Z}_t$ should contain all the information used by investors to forecast at time t the excess-returns for all future periods. If we stack all yields at time t in the vector $\boldsymbol{y}_t^{(n)}$, as

$$\boldsymbol{y}_t^{(n)} = f\left(\mathbf{Z}_t; N\right) \tag{1.11}$$

we can see that the yields must be a function only of the state vector $\boldsymbol{y}_t^{(n)}$ and the vector of

[2]An accounting identity makes the link between the yield of bond to the sum of one-periods (h-periods) with the its excess returns for a n-period maturity bond as:

$$y_t^{(n)} \equiv \frac{1}{n}\left(\sum_{j=0}^{12\cdot n/h-1} y_{t+j\cdot h/12}^{(h/12)}\right) \quad + \quad \frac{1}{n}\left(\sum_{j=0}^{12\cdot n/h-1} rx_{t+h/12(j+1)}^{(n-j\cdot h/12)}\right) \tag{1.6}$$

where j are multiple of h-periods of the defined frequency. For annual frequency, i.e., $h=12$ we have:

$$y_t^{(n)} \equiv \frac{1}{n}\left(\sum_{j=0}^{n-1} y_{t+j}^{(1)}\right) \quad + \quad \frac{1}{n}\left(\sum_{j=0}^{n-1} rx_{t+j+1}^{(n-j)}\right) \tag{1.7}$$

maturities N. The essential assumption is the existence of an inverse function $f(\cdot)^{-1}$ that allow us to write $\mathbf{Z}_t = f\left(\boldsymbol{y}_t^{(n)}; N\cdot\right)^{-1}$. This holds true, as long as exists a correspondence in such a way that each $\boldsymbol{z}_t \in \mathbf{Z}_t$ has its own effect on the yield curve $\boldsymbol{y}_t^{(n)}$. Thus, for a function $g(\cdot)$ we can write $\mathbb{E}_t\left(y_t^{(n)}\right) = g\left(\boldsymbol{y}_t^{(n)}; N\right)$[3].

As Duffee (2013) emphasizes, equation (1.9) determines that the expected returns depend on at most k state variables, and inverting equation (1.10) tells us that with the entire yield curve, we can disentangle shocks of the expected excess returns from shocks to expected future yields. What boils down to estimating the function $g(\cdot)$. The key takeaway is that the whole term-structure at time t contains all the information to predict $\mathbf{Z}_t$, and consequently the future yield curves.

However, the literature has gathered evidence against the expectations hypothesis. Influential studies from Fama and Bliss (1987), Campbell and Shiller (1991) and Cochrane and Piazzesi (2005) show some forecastability for excess returns. Among the most important approaches to test the predictability of the bonds' excess returns we have Fama and Bliss (1987), Cochrane and Piazzesi (2005), and Ludvigson and Ng (2009). Below we succinctly describe each one of them, as they will be used as our benchmarks.

Fama and Bliss (1987) builds forward rates spreads and use these variables as covariates. The forward rate spread between of a n-year maturity bond is defined as $fs_t^{(n,h)} \equiv f_t^{(n)} - y_t^{(h/12)}(h/12)$. The predictive regression in the Fama-Bliss approach is given by

$$rx_{t+h/12}^{(n)} = \beta_0 + \beta_1 fs_t^{(n,h)} + \epsilon_{t+h/12} \quad . \tag{1.12}$$

Cochrane and Piazzesi (2005) derive a single factor to use as predictor. The authors argue that their factor (CP_t^h), which has a peculiar tent-shape across maturities and is built from a linear combination of forward rates has a higher predictability of excess returns on one- to five-year maturity bonds. First, they estimate a vector $\boldsymbol{\gamma}$ by regressing the average of excess returns across maturities $n = 1, 2, 3, 4$ on all forward rates as

$$\begin{aligned} \tfrac{1}{4}\textstyle\sum_{n=2}^{5} rx_{t+h/12}^{(n)} &= \gamma_0 + \gamma_1 f_t^{(1)} + \gamma_2 f_t^{(2)} + \gamma_3 f_t^{(3)} + \gamma_4 f_t^{(4)} + \gamma_5 f_t^{(5)} + \bar{\epsilon}_{t+h/12} \\ \overline{rx}_{t+h/12} &= \underbrace{\boldsymbol{\gamma}^\top \mathbf{f}_t}_{CP_t^h} + \bar{\epsilon}_{t+h/12} \end{aligned} \tag{1.13}$$

where $\boldsymbol{f}$ and $\boldsymbol{\gamma}$ are 6×1 vectors given by $\boldsymbol{f} \equiv \begin{bmatrix} 1 & f_t^{(1)} & f_t^{(2)} & f_t^{(3)} & f_t^{(4)} & f_t^{(5)} \end{bmatrix}^\top$, and $\boldsymbol{\gamma} \equiv \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & \gamma_5 \end{bmatrix}^\top$. Denoting the estimated Cochrane-Piazzesi factor as $\widehat{CP}_t^h =$

[3]Which implies that $\mathbb{E}_t\left(rx_{t+h/12(j+1)}^{(n-j\cdot h/12)}\right) = g\left(\boldsymbol{y}_t^{(n)}; N\right)$ also holds.

$\widehat{\gamma}^\top \boldsymbol{f}_t$, the predictive regression in this approach is given by

$$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \widehat{CP}^h_t + \epsilon_{t+h/12} \quad . \tag{1.14}$$

Another important concept derived from the majority of macro-finance models is the *spanning hypothesis*. It says that all relevant information to forecast yields and excess returns can be found on the term-structure. Hence, under the *spanning hypothesis*, the yields curve fully spans all necessary information, and thus, no other variable or information already present in the term-structure should be necessary. As Bauer and Hamilton (2018) stress, the *spanning hypothesis* does not rules out the importance of macro variables (current or future). It only says that the yield curve completely reflects and spans this information.

Ludvigson and Ng (2009) show evidence against the spanning hypothesis. Using a large panel of macro variables, the authors build a single linear combination out of the first i estimated principal components $(\hat{g}_{i,t})$[4]. The authors start estimating a vector $\boldsymbol{\lambda}$ by regressing the average of excess returns across maturities $n = 1, 2, 3, 4$ on a subset of the first 8 principal components as

$$\begin{aligned} \tfrac{1}{4}\textstyle\sum_{n=2}^{5} rx^{(n)}_{t+h/12} &= \lambda_0 + \lambda_1 \hat{g}_{1,t} + \lambda_2 \hat{g}^3_{1,t} + \lambda_3 \hat{g}_{3,t} + \lambda_4 \hat{g}_{4,t} + \lambda_5 \hat{g}_{8,t} + \bar{\epsilon}_{t+h/12} \\ \overline{rx}_{t+h/12} &= \underbrace{\boldsymbol{\lambda}^\top \widehat{\boldsymbol{G}}_t}_{LN^h_t} + \bar{\epsilon}_{t+h/12} \end{aligned} \tag{1.16}$$

where $\widehat{\boldsymbol{G}}_t$ and $\boldsymbol{\lambda}$ are 5×1 vectors given by $\widehat{\boldsymbol{G}}_t \equiv \begin{bmatrix} \hat{g}_{1,t} & \hat{g}^3_{1,t} & \hat{g}_{3,t} & \hat{g}_{5,t} & \hat{g}_{8,t} \end{bmatrix}^\top$, and $\boldsymbol{\lambda} \equiv \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \end{bmatrix}^\top$. Denoting the estimated Ludvigson-Ng factor as $\widehat{LN}^h_t = \widehat{\boldsymbol{\lambda}}^\top \widehat{\boldsymbol{G}}_t$, the predictive regression in this approach is given by

$$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \widehat{LN}^h_t + \epsilon_{t+h/12} \quad . \tag{1.17}$$

1.2.3 A Deep-Learning Structure for Bond Premia

The three main approaches presented in the last section seek to provide an explanation for the bond premia. We can summarize these approaches with the following predictive regression

[4]The authors consider a $T \times M$ panel of macroeconomic variables and assume that each macro variable $\{z^{macro}_{j,t}\}$ has a factor structure taking the form

$$z^{macro}_{j,t} = \boldsymbol{\nu}^\top_t \boldsymbol{g}_t + e_{j,t} \tag{1.15}$$

where $\boldsymbol{g}_t$ is an $s \times 1$ vector of latent common factors obtained through principal components analysis, and $\boldsymbol{\nu}_t$ is an $s \times 1$ vector of latent factor loadings. The essential point here is that $s \ll M$.

$$rx_{t+h/12}^{(n)} = \boldsymbol{\beta}^{\top} \boldsymbol{Z}_t + \epsilon_{t+h/12} \tag{1.18}$$

where $\boldsymbol{Z}_t = \left\{ \boldsymbol{Z}_t^{\mathrm{y}}, \boldsymbol{Z}_t^{\mathrm{y}^{\complement}} \right\}$ is a set of state variables that could potentially forecast the excess returns, and thus, provide evidence against the expectations hypothesis. If they rely on the spanning hypothesis $\boldsymbol{Z}_t = \{\boldsymbol{Z}_t^{\mathrm{y}}\}$ and no macroeconomic variables are used to define the state space. Evidence against the spanning hypothesis is showed when $\boldsymbol{Z}_t^{\mathrm{y}^{\complement}} \neq \emptyset$.

In this paper we argue that this search for deriving, building and estimating factors that represent state variables in macro-finance models may be limited. We claim that the process done by financial economists of manually discovering and hand picking this list factors may be leaving unseen relationships between the state variables out in their derivation.

Hence, to assist in this process, we make use of one of the most powerful approaches in machine learning, namely a deep neural network. We aim to uncover relationships and derive a new single factor that could improve our understanding of the bond risk premia. We make use of deep feedforward network or multilayer perceptron (MLP)[5] and derive a single factor that has predictability in an our analysis.

Deep neural networks attempt to replicate the brain architecture in a computer, in a such a way that we must have many levels of processing information. As Murphy (2012) points out, it is believed that each level of learning features or representations at increasing levels of abstraction.

A deep feedforward network defines a mapping such as $rx_{t+h/12}^{(n)} = g\left(\boldsymbol{Z}_t, \boldsymbol{\theta}_t\right)$ to *learn* the parameter $\boldsymbol{\theta}_t$ that provides the best function approximation. In its most common structure, MLP can be represented in a direct acyclic graph with a chain of functions $g\left(\boldsymbol{Z}_t\right) = g^{(L)}\left(\ldots\left(g^{(2)}\left(g^{(1)}\left(\boldsymbol{Z}_t\right)\right)\right)\right)$. The name *network* comes from this chain and its interconnectedness architecture, and *feedforward* because the information flows in one direction from $\boldsymbol{Z}_t$ through these functions, to finally obtain an output $rx_{t+h/12}^{(n)}$. The number of these functions L defines the *depth* of the network, motivating the use of the name "deep learning" to refer to this structure. We say that $g^{(1)}$ is the first layer, while the last one $g^{(L)}(\cdot)$ is the output layer.

As Goodfellow et al. (2016) discuss, deep feedforward network can capture the information between any two inputs, which is a limitation that linear models such as logistic and linear regressions face. This is done as an extension from a linear model, in such a way that we apply a nonlinear function $\phi(\cdot)$ in $\boldsymbol{Z}_t$, transforming our independent variable. Thus, the model can be represented as $rx_{t+h/12}^{(n)} = g\left(\boldsymbol{Z}_t, \boldsymbol{\theta}_t, \boldsymbol{w}\right) = \phi\left(\boldsymbol{Z}_t, \boldsymbol{\theta}_t\right)^{\top} \boldsymbol{w}$.

In this approach ϕ defines a hidden layer, $\boldsymbol{\theta}_t$ the parameters used to learn ϕ, and $\boldsymbol{w}$ are

[5] They are also known as feedforward neural network

parameters mapping from $\phi(\boldsymbol{\theta}_t)$ to the output. An optimizing algorithm is responsible to find $\boldsymbol{\theta}_t$ that gives the best representation. The nonlinear function is called activation function, which is controlled by learned parameters. Hence, we define $g_l(\cdot) = g\left(\boldsymbol{W}_l^\top \boldsymbol{Z}_t + c\right)$, where $\boldsymbol{W}$ is a set of weights and c the biases.

One advantage of deep neural networks is based on the universal approximation theorem (Hornik et al., 1989; Cybenko, 1989) that states that feedforward network with a linear output layer and at least one hidden layer with any activation function can approximate any function[6] from one finite-dimensional space to another with any desired nonzero amount of error. In short, this theorem says that a simple neural network can represent a wide variety of functions. However it does not guarantee the training algorithm will be able to learn the function. One implication from the universal approximation theorem is that there exists a network large enough to achieve any degree of accuracy.

In our framework, the single factor capturing the information from the yield curve is built in the following way. First, at each t we use the cross-section of the information on the term-structure to feed MLPs to obtain as output a factor derived from a learning network. We denote this deep neural network factor as $\mathfrak{f}_{DNN}$.

Aligned with the results from Bauer and Hamilton (2018) who gathered evidence that rejections of the spanning hypothesis by some recent papers is significantly weaker when more robust methods are used to deal especially with overlapping data, we use as input in our networks only $\boldsymbol{Z}_t^{\mathrm{y}}$, which is formed by the full set of information from the yield curve. We argue that given the superiority of deep feedforward networks to uncover relationships between the information found in $\boldsymbol{Z}_t^{\mathrm{y}}$, especially its capacity to nonlinear and more complex associations in the data, there is a potential gain of extracting more information out of the yield curve.

Figure 1.1 shows the deep feed forward architecture to obtain the DNN factor $\mathfrak{f}_{DNN}$. The depth, the width, the activation function of the deep neural network, as well as the loss function used for training at each t are variations discussed in section 1.3.

Notice that there are 4 separate groups of networks. Each one of them seek to find the function that provides the approximation g, such that the mapping is given by $g^{(n)} : \boldsymbol{Z}_t^{\mathrm{y}} \mapsto rx_{t+h/12}^{(n)}$, where $n \in \{2, 3, 4, 5\}$, i.e., it is the mapping from the entire yield curve information to the excess returns in the next period $t + h/12$ for maturities ranging from 2 to 5 years.

Each group of network will deliver a factor associated with a maturity n at each t. After obtaining $\mathfrak{f}_{t,DNN}^{(n)}$, we estimate the single factor that summarizes the all the term-structure information to explain the excess returns. The idea is to describe the expected excess returns

[6] Precisely, any Borel measurable function, i.e., any continuous function on a closed and bounded subset of $\mathbb{R}^n$.

Figure 1.1: Deep Neural Network

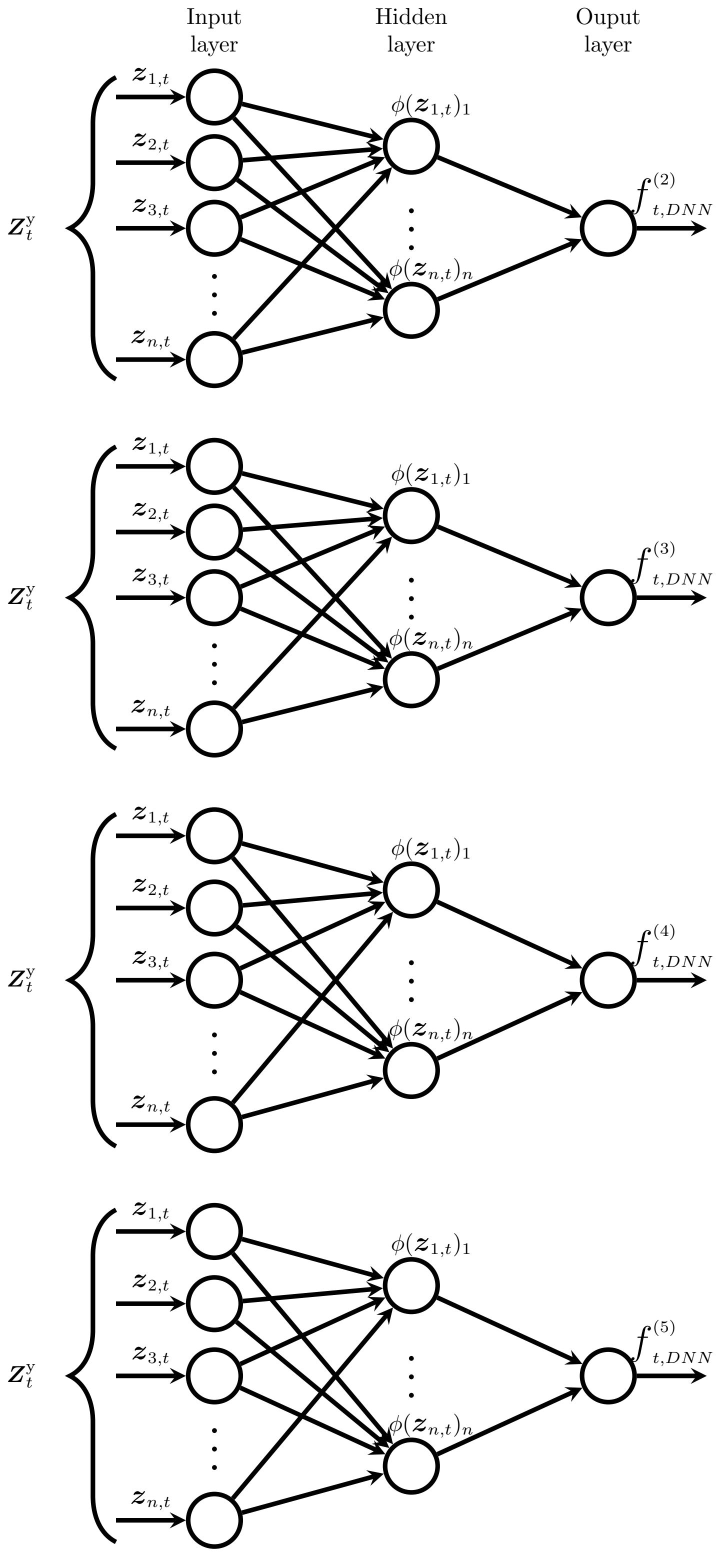

Figure 1.1 shows the general structure of the deep feed forward designed to obtain the DNN factor $\mathfrak{f}_{DNN}$. There are four groups of networks, each group for $n \in \{1, 2, 3, 4\}$. The inputs layer receives data from $\boldsymbol{Z}_t^{\mathrm{y}} = \{\boldsymbol{z}_{1,t}, \boldsymbol{z}_{2,t}, \ldots, \boldsymbol{z}_{n,t}\}$. Each group of network n outputs a factor (DNN factor), which we denote by $\mathfrak{f}^{(n),h}_{t,DNN}$.

of all maturities with a unique factor, as proposed initially by Cochrane and Piazzesi (2005), and extended by others (Ludvigson and Ng, 2009; Cieslak and Povala, 2015). First, we regress the average of the excess returns of maturities $2, 3, 4$ and 5 years on all four factors derived from our deep neural network, as below:

$$\begin{aligned} \tfrac{1}{4}\textstyle\sum_{n=2}^{5} rx_{t+h/12}^{(n)} &= \tau_0 + \tau_1 \mathfrak{f}_{t,DNN}^{(2),h} + \tau_2 \mathfrak{f}_{t,DNN}^{(3),h} + \tau_3 \mathfrak{f}_{t,DNN}^{(4),h} + \tau_4 \mathfrak{f}_{t,DNN}^{(5),h} + \bar{\epsilon}_{t+h/12} \\ &= \boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t^h + \bar{\epsilon}_{t+h/12} \end{aligned} \tag{1.19}$$

where $\widehat{\boldsymbol{\mathfrak{F}}}_t$ and $\boldsymbol{\tau}$ are 5×1 vectors given by $\widehat{\boldsymbol{\mathfrak{F}}}_t \equiv \begin{bmatrix} 1 & \mathfrak{f}_{t,DNN}^{(2),h} & \mathfrak{f}_{t,DNN}^{(3),h} & \mathfrak{f}_{t,DNN}^{(4),h} & \mathfrak{f}_{t,DNN}^{(5),h} \end{bmatrix}^\top$, and $\boldsymbol{\tau} \equiv \begin{bmatrix} \tau_0 & \tau_1 & \tau_2 & \tau_3 & \tau_4 \end{bmatrix}^\top$. The predictive regression in this approach is given by

$$rx_{t+h/12}^{(n)} = \beta_0 + \beta_1 \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t \right)_t^h + \epsilon_{t+h/12} \qquad n = 2, 3, 4, 5 \quad . \tag{1.20}$$

Equation 1.20 tells us that a single factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t \right)_t^h$ defines the state variable driving the excess returns. Thus, starting from the spanning hypothesis, we feed a MLP with the entire information from the yield curve to approximate a function, and then derive a single linear combination of factors to explain the time-varying expected returns across maturities.

From the deep neural network we also would like to estimate a factor that represent the information not spanned by the term-structure. To do so, we design a new approach in which at each t and each group $n \in \{2, 3, 4, 5\}$ of network for each maturity, we orthogonalize the excess returns by the deep neural network factor $\mathfrak{f}_{t,DNN}^{(n)}$, and denote it by $\boldsymbol{\xi}_t^{(n),h}$ as

$$\xi_{t+h/12}^{(n),h} = rx_{t+h/12}^{(n)} - \beta_0 - \beta_1 \mathfrak{f}_{t,DNN}^{(n),h} \quad . \tag{1.21}$$

From equation (1.21), the factor $\xi_{t+h/12}^{(n),h}$that lies in an orthogonal vector to the space spanned by $\mathfrak{f}_{t,DNN}^{(n)}$, can be seen as all the information not spanned by the term-structure captured by $\mathfrak{f}_{t,DNN}^{(n)}$ that affects the excess returns.

1.2.4 An Illustrative Term-Structure Model

In this section we make the link of our methodology with the main dynamic term-structure frameworks in the macro-finance literature. We follow Duffee (2013) and assume that interest rate dynamics are linear and homoskedastic with Gaussian shocks. The no-arbitrage assumption rely on the fundamental asset pricing equation:

$$P_t^{(n)} = \mathbb{E}_t \left(\mathcal{M}_{t+1} P_{t+1}^{(n-1)} \right) \tag{1.22}$$

where $P_t^{(n)}$ is the price of a bond and $\mathcal{M}_{t+h/12}$ is the stochastic discount factor (SDF).

The economic agents value nominal bonds using the following SDF:

$$\mathcal{M}_{t+h/12} = \exp^{-r_t \frac{1}{2}\Lambda_t^\top \Lambda_t - \Lambda_t^\top \epsilon_{t+h/12}} \tag{1.23}$$

where Λ_t are the market prices of the risks, i.e., the amount of compensation required by investors to face the unit normal shock $\epsilon_{t+h/12}$. The yield on a one-period bond $r_t \equiv y^{(1)}$ is a function of $\boldsymbol{Z_t}$, as

$$r_t = \rho_0 + \rho_1 \boldsymbol{Z_t} \quad . \tag{1.24}$$

As we defined $\boldsymbol{Z}_t = \left\{\boldsymbol{Z}_t^{\mathrm{y}}, \boldsymbol{Z}_t^{\mathrm{y}^\complement}\right\}$, we write the dynamics of $\boldsymbol{Z}_t$ that capture all the risks of the economy following a Gaussian VAR process given by:

$$\begin{aligned}
\begin{bmatrix} \boldsymbol{Z}_t^{\mathrm{y}} \\ \boldsymbol{Z}_t^{\mathrm{y}^\complement} \end{bmatrix} &= \boldsymbol{\mu} + \boldsymbol{\Phi} \begin{bmatrix} \boldsymbol{Z}_{t-1}^{\mathrm{y}} \\ \boldsymbol{Z}_{t-1}^{\mathrm{y}^\complement} \end{bmatrix} + \boldsymbol{\Sigma}\epsilon_t \\
\boldsymbol{Z}_t &= \boldsymbol{\mu} + \boldsymbol{\Phi}\boldsymbol{Z}_{t-1} + \boldsymbol{\Sigma}\epsilon_t \qquad \epsilon_t \sim N(0, \boldsymbol{I})
\end{aligned} \tag{1.25}$$

where $\boldsymbol{\mu}$ is a a $k \times 1$ vector, and $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ are $k \times k$ matrices, being k the number of state variables. In a similar fashion to Joslin et al. (2014), who developed an arbitrage-free dynamic term-structure model with unspanned macro risks, we can write:

$$\boldsymbol{Z}_t^{\mathrm{y}^\complement} = \gamma_0 + \gamma_1 \boldsymbol{Z}_t^{\mathrm{y}} + \boldsymbol{M}_{\boldsymbol{Z}_t^{\mathrm{y}}} \boldsymbol{Z}_t^{\mathrm{y}^\complement} \tag{1.26}$$

where $\boldsymbol{M}_{\boldsymbol{Z}_t^{\mathrm{y}}} \boldsymbol{Z}_t^{\mathrm{y}^\complement}$ is the annihilator matrix of the space spanned by $\boldsymbol{Z}_t^{\mathrm{y}}$, i.e., $\boldsymbol{M}_{\boldsymbol{Z}_t^{\mathrm{y}}} \boldsymbol{Z}_t^{\mathrm{y}^\complement} \equiv \boldsymbol{Z}_t^{\mathrm{y}^\complement} - \mathrm{Proj}\left[\boldsymbol{Z}_t^{\mathrm{y}^\complement} | \boldsymbol{Z}_t^{\mathrm{y}}\right]$. Previous models have assumed that the $\boldsymbol{Z}_t^{\mathrm{y}^\complement}$ was spanned by $\boldsymbol{Z}_t^{\mathrm{y}}$, thus imposing the restriction of $\boldsymbol{Z}_t^{\mathrm{y}^\complement} = \mathrm{Proj}\left[\boldsymbol{Z}_t^{\mathrm{y}^\complement} | \boldsymbol{Z}_t^{\mathrm{y}}\right]$ in equation (1.26).

Aligned with Joslin et al. (2014), our methodology is also based on (i) a small number of risk factors, and (ii) the unspanned components of $\boldsymbol{Z}_t^{\mathrm{y}^\complement}$ may contain predictive power for excess returns. However, we distinguish from Joslin et al. (2014) who provided the exact macroeconomic variables that are unspanned by the term-structure. Our unspanned factor, on the other hand, should be able to represent any other risk, be it macroeconomic or sentiment-based in the economy. In this sense, we say that our framework is more general. Additionally, to provide an intuitive interpretation, we analyze how $\boldsymbol{Z}_t^{\mathrm{y}^\complement}$ is correlated with macroeconomic variables and sentiment-based measures.

From the above illustrative term-structure model, we make a set of propositions that makes the link between our deep-learning framework and a dynamic term structure model.

Proposition 1. *The state vector* $\boldsymbol{Z}_t$ *that encompasses all risks in the economy can be partitioned as* $\boldsymbol{Z}_t = \left\{\boldsymbol{Z}_t^y, \boldsymbol{Z}_t^{y^\complement}\right\}$, *in such a way that* $\boldsymbol{Z}_t^y$ *contains information solely from the yield curve, and* $\boldsymbol{Z}_t^{y^\complement}$ *any other information not found in the term structure.*

Proposition 2. *Under the canonical arbitrage-free Gaussian term structure model as in Joslin et al. (2014),* $\boldsymbol{Z}_t^y$ *is given by the derived factor* $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ *from equation* (1.19), *and* $\boldsymbol{Z}_t^{y^\complement}$ *by a linear function* $f(\cdot)$ *of* $\boldsymbol{\xi}_{t+h/12}^h$.

Proposition 3. *As in the dynamic term structure model of Joslin et al. (2014),* $f(\boldsymbol{\xi}_{t+h/12}^h)$ *complete and fill the unspanned factor in the state space, in a such a way that* $\left[\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h, f(\boldsymbol{\xi}_{t+h/12}^h)\right]$ *and* $\boldsymbol{Z}_t$ *represent linear rotations of the same economy-wide risks underlying all tradable assets available to agents in the economy.*

1.3 Data & Strategy

1.3.1 Data

As emphasized by Bauer and Hamilton (2018), predictive regressions estimated using overlapping observations, approach commonly used by several previous studies, where monthly data is used and the annual excess bond return is the dependent variable, introduces serial correlation in the prediction errors, what results in inaccurate standard errors.

As done in Gargano et al. (2019), to overcome the issues generated by overlapping observations, we reconstruct the yield curve at the daily frequency, using the parameters estimated by Gürkaynak et al. (2007) and made available at the Federal Reserve Discussion Series website[7]. Thus, we reconstruct the log yield of a zero-coupon with n-period maturity at time t as

$$\begin{aligned} y_t^{(n)} = \beta_{0,t} \quad & +\beta_{1,t}\left(\tfrac{1-\exp(-n/\tau_1)}{n/\tau_1}\right) \\ & +\beta_{2,t}\left(\tfrac{1-\exp(-n/\tau_1)}{n/\tau_1} - \exp(-n/\tau_1)\right) \\ & +\beta_{3,t}\left(\tfrac{1-\exp(-n/\tau_2)}{n/\tau_2} - \exp(-n/\tau_2)\right) \end{aligned} \tag{1.27}$$

where the daily estimated parameters β_0, β_1, β_2, β_3, τ_1 and τ_2 are from Gürkaynak et al. (2007). The full period of analysis ranges from 1962:01 to 2017:12. We use these estimated

[7] https://www.federalreserve.gov/econres/feds/2006.htm

parameters from the last day of each month to construct a monthly derived zero-coupon bonds log yields with maturities up to 60 months from each t. Figure 1.2 plots the log yields for all maturities. Figure 1.3 shows the 1-month excess returns for maturities with $n = 2, 3, 4$ and 5 years. In Appendix 1.6.1, figure 1.14 plots for the same set of maturities the 12-month excess returns.

Figure 1.2: Derived zero-coupon bonds log yields for maturities (n) up to 60 months

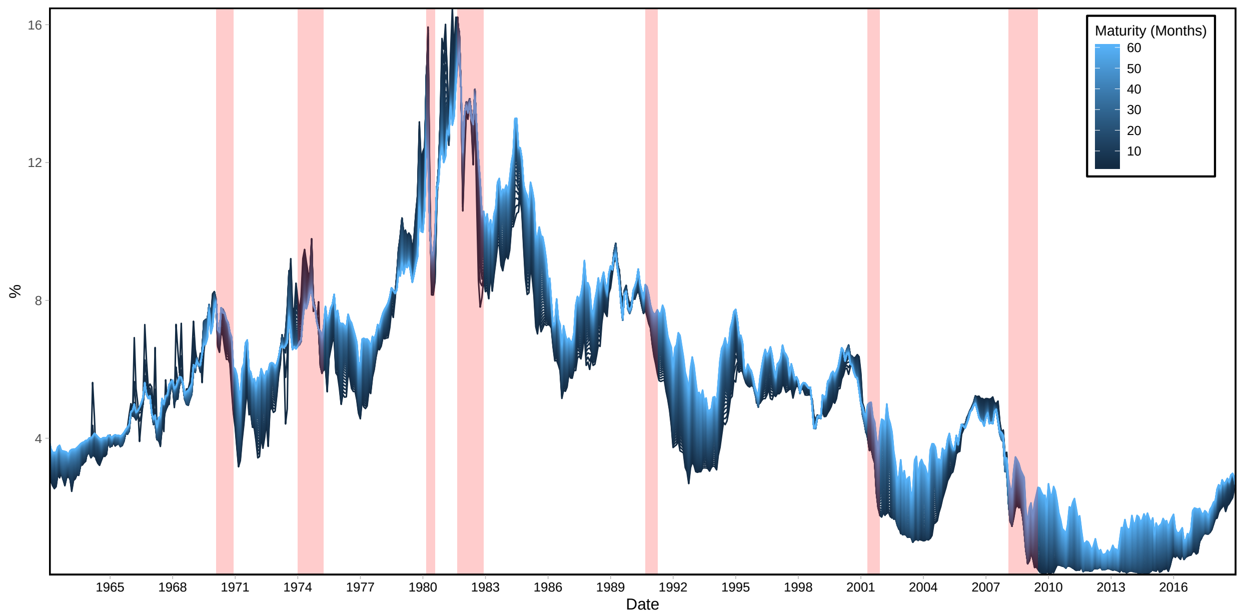

Figure 1.2 shows the log yields for all maturities we consider: $y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}$. At each month t, there are 60 yields represented by variation of color. The log yields of the zero-coupons bonds are reconstructed with equation (1.27), using the last day of each month estimated parameters from Gürkaynak et al. (2007) data. The entire sample ranges from 1962:01 to 2018:12.

Some papers have instead used the data from the Fama–Bliss Center for Research in Security Prices (CRSP) to build the series of log bond yields. Based on Fama and Bliss (1987), this approach constructs the yields sequentially from a set of estimated daily forward rates. As Gargano et al. (2019) point out, the differences between Fama and Bliss (1987) and Gürkaynak et al. (2007) are minimal. The correlation between both methods[8] when comparing yields and excess returns are both above 0.99 for all four maturities we use.

1.3.2 Empirical Strategy

In our first analysis we establish the period of evaluation from 1993:01 to 2017:12. We feed our deep neural network with three different sets of information from the term-structure:

[8]For a similar period: 1962:01 to 2015:12.

Figure 1.3: 1-Month Bonds Excess Returns

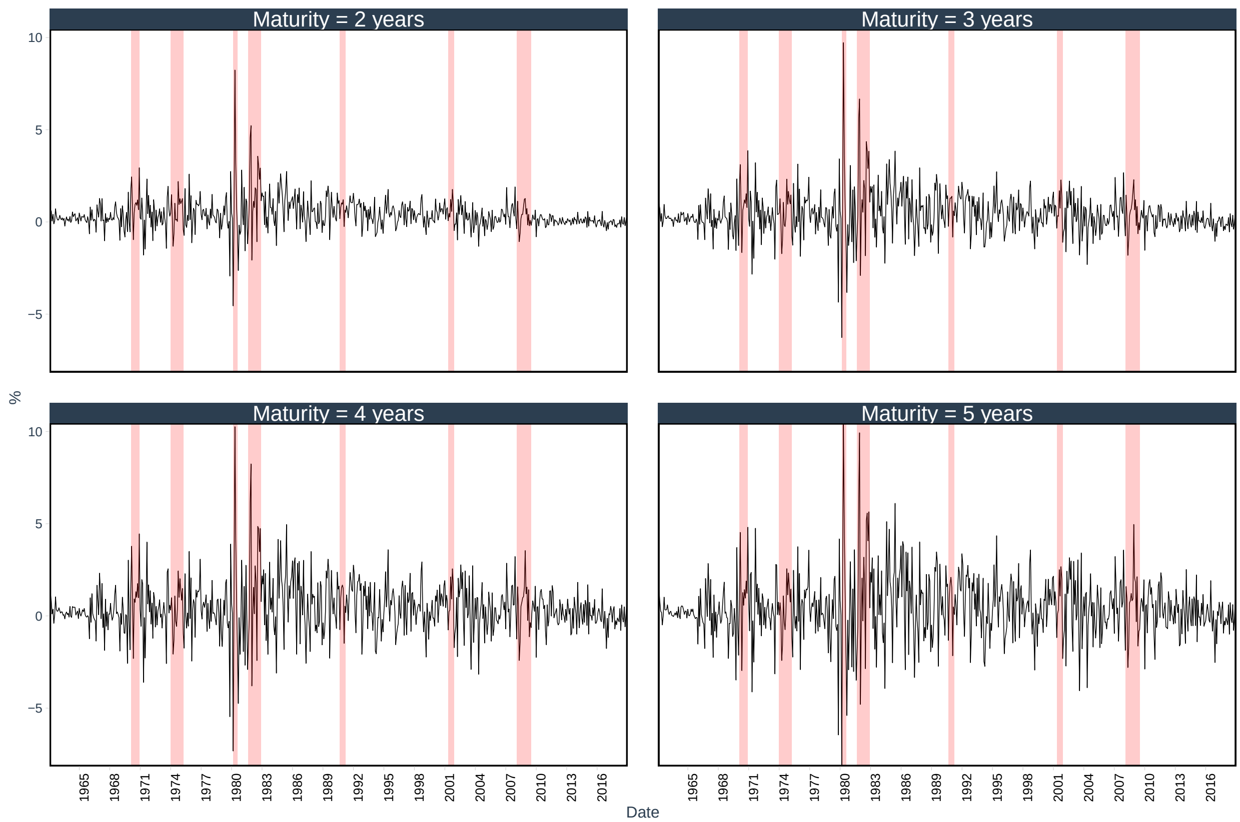

Figure 1.3 shows the 1-month excess returns for maturities with $n = 2, 3, 4$ and 5 years. The excess returns are calculeted as in equation (1.4), i.e., $rx^{(n)}_{t+1/12} = ny^{(n)}_t - (n + 1/12)y^{(n-h/12)}_{t+1/12} - y^n_t$. Each panel represents one of the four maturities. The y-axis shows values in percentage (%). NBER-classified recessions are shaded in light red.

(i) set of forward rates from 2 to 60 months from t, i.e., $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60/12)}\right\}$, (ii) set of zero-coupon yields with maturities ranging from 1 to 60 months from t, i.e., $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$, and finally (iii) a combination of the previous two groups, i.e., $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60/12)}, y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$.

Goodfellow et al. (2016) discuss that the design of the hidden units is an extremely active area of research. This leads to many potential options for the nonlinear function in the hidden layers. As the authors mention, the rectified linear activation function (ReLU) is the default and recommended for use with the majority of feedforward neural networks. In all our hidden layers, for all groups and architectures, we make use of this activation function defined as

$$ReLU(x) = \begin{cases} 0 & \text{, if } x \leq 0 \\ x & \text{, otherwise} \end{cases} . \tag{1.28}$$

As Goodfellow et al. (2016) mention, applying this function to the output of a linear transformation yields a nonlinear transformation. Notice that, since ReLU units are nearly linear, they have the advantage of also retaining many of the properties from linear models, such as (i) efficiency to optimize with gradient-based methods, and (ii) ability to preserve the properties that make linear models generalize well.

All our neural networks share the same architecture as show in figure 1.1. To make use of the flexibility that MLP allows us, we designed three variation for the whole network. Bianchi et al. (2019) also developed several designs in their study, and we use some of their intuitions to design our deep neural networks architectures. The first (**DNN 1**) and second model (**DNN 2**) are feedfoward neural networks with 2 hidden layers ($L = 2$), with 16 and 4 nodes respectively, and finally an output layer for each group of maturity $n \in \{1, 2, 3, 4\}$. What differentiates **DNN 1** from **DNN 2** is the regularization function, where we use a ℓ_1-norm for **DNN 1** and a ℓ_1- and ℓ_2-norm for **DNN 2**. On the other hand, **DNN 3** has 4 hidden layers ($L = 4$), with 64, 32, 16 and 4 nodes. For **DNN 3** we use a ℓ_1- and ℓ_2-norm regularization function.

The process of obtaining $\widehat{\boldsymbol{\mathfrak{F}}}_t^h$ from equation (1.19) at each t can be summarized in the following way. First, for each set of $\boldsymbol{Z}_t^{\mathrm{y}}$ in consideration, we feed each one of the three DNNs architectures with the entire past information of each variable in $\boldsymbol{Z}_t^{\mathrm{y}}$. We use the 10% most recent data in each $\boldsymbol{z}_t^{\mathrm{y}} \in \boldsymbol{Z}_t^{\mathrm{y}}$ for validation. After the set of weights are chosen, with the final set of weights and the final approximated function, we use $\boldsymbol{Z}_{t-1}^{\mathrm{y}}$ to predict $rx_t^{(n)}$ in each group of maturity $n \in \{2, 3, 4, 5\}$. Thus, we form the 4×1 vector of $\widehat{\boldsymbol{\mathfrak{F}}}_t^h$ as the factor at t generated by each DNN_i. As a final step we run univariate regressions to obtain $\boldsymbol{\xi}_t^h$, as shown in equation 1.21. Similarly, we build the 4×1 vector using the observation t residuals as $\widehat{\boldsymbol{\xi}}_t^h = \begin{bmatrix} \widehat{\xi}_t^{(2),h} & \widehat{\xi}_t^{(3),h} & \widehat{\xi}_t^{(4),h} & \widehat{\xi}_t^{(5),h} \end{bmatrix}$. The whole process is summarized in the pseudocode given in algorithm 1.

At the end of our period of analysis, we use the entire series of $\widehat{\boldsymbol{\mathfrak{F}}}_t^h$ to obtain our single factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ that spans the yield curve information as in equation (1.19). To complete the factor space of our dynamic term-structure model, we define the unspanned factor as a function of $\boldsymbol{\xi}_t^h$ to build a single factor as well for $\boldsymbol{Z}_t^{\mathrm{y}^{\complement}}$. We investigate two alternatives for $f(\boldsymbol{\xi}_t^h)$. In the first one, $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$ is the unique factor obtained as the projection of $\overline{rx}_{t+h/12}$ in $\widehat{\boldsymbol{\xi}}_t$. The second alternative is a similar projection, however for each maturity $n \in \{2, 3, 4, 5\}$ we regress $rx_{t+h/12}^{(n)}$ on $\widehat{\boldsymbol{\xi}}_t^{(-n),h} \equiv \widehat{\boldsymbol{\xi}}_t^h \setminus \widehat{\xi}_t^{(n),h}$, i.e., on the set $\boldsymbol{\xi}_t^h$ excluding its own $\widehat{\xi}_t^{(n)}$. We denote the factors generated by this second approach as $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$.

Consistent with our adapted dynamic term-structure model, the orthogonal vector from $\text{Proj}\left[f(\boldsymbol{\xi}_t^{(n)}) | \boldsymbol{Z}_t^{\mathrm{y}}\right]$ has predictive power for excess returns. Thus, we use the projection error $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^h$ for alternative 1 and $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$ for alternative 2 in our predictive analysis in the following section.

The intuition that motivates our construction of $\boldsymbol{Z}_t^{\mathrm{y}^{\complement}}$ lies in the fact that at each t, $\widehat{\xi}_t^{(n)}$ is orthogonal to $\mathfrak{f}_{t,DNN}^{(n),h}$, allowing the interpretation that, for each maturity group n in our DNN, anything not captured by the neural network process of approximating $g(\cdot)$ from the yield curve information $\boldsymbol{Z}_t^{\mathrm{y}}$, are unspanned and should be in an orthogonal space. Hence, the unspanned information in $\widehat{\boldsymbol{\xi}}_t^h$ could be capturing macroeconomic information or sentiment measures not spanned by the term-structure information that affects the bonds' excess returns. Alternative 1 builds an unique factor for $\boldsymbol{Z}_t^{\mathrm{y}^{\complement}}$ in a such a way that a single linear combination of orthogonal variables is the state variable that completes the state space for time-varying expected returns on all maturities. On the other hand, alternative 2 tells us that a single linear combination of three orthogonal variables from the remaining maturities complete the state space for time-varying expected returns for maturity n.

Algorithm 1: Recursively generated factors with updated parameters

Initialization:

Start with a set of information from the term structure collected in $\boldsymbol{Z}^{\mathrm{y}}$. Partitionate your sample $\{t_0, \ldots, t_{split}, \tau, \tau+1, \ldots, T\}$ between the data to be used to initialize the process $\{t_0, \ldots, t_{split}\}$, and to obtain the recursively generated factors $\{\tau, \tau+1, \ldots, T\}$;

Define the deep neural network architecture to be used (number of hidden layers L, and number of nodes in each layer $l \in L$);

For a pre-defined deep neural network architecture DNN_i, set the activation function $\phi(\cdot)$ in each node at layer $l \in L$ as the ReLu defined in equation (1.28);

for $n \in \{2,3,4,5\}$ **do**

  for $t \in \{\tau, \tau+1, \ldots, T\}$ **do**

    Feed DNN_i with lagged data $\boldsymbol{Z}^{\mathrm{y}}_{t-1} = \left\{\boldsymbol{z}^{\mathrm{y}}_{t_0}, \boldsymbol{z}^{\mathrm{y}}_{t_0+1}, \ldots, \boldsymbol{z}^{\mathrm{y}}_{t-1}\right\}$ to learn/aproximate with output $rx_t^{(n)}$, and use the last 10% of the data for validation;

    Obtain the learned parameters;

$$\widehat{\mathfrak{f}}^{(n),h}_{t,DNN} \leftarrow g\left(\boldsymbol{Z}^{\mathrm{y}}_{t-1}, \boldsymbol{\theta}_{t-1}\right)$$

    Obtain the t-th element that lies in the orthogonal vector from the space generated by the $\mathfrak{f}^{(n),h}_{t-1,DNN}$ through:

$$\widehat{\xi}^{(n),h}_t \leftarrow rx_t^{(n)} - \widehat{\beta}_0 - \widehat{\beta}_1 \mathfrak{f}^{(n),h}_{t-1,DNN_i}$$

  end

end

Result:

$$\widehat{\boldsymbol{\mathfrak{F}}}_{t,DNN_i} \equiv \begin{bmatrix} \widehat{\mathfrak{f}}^{(2),h}_{t,DNN_i} \\ \widehat{\mathfrak{f}}^{(3),h}_{t,DNN_i} \\ \widehat{\mathfrak{f}}^{(4),h}_{t,DNN_i} \\ \widehat{\mathfrak{f}}^{(5),h}_{t,DNN_i} \end{bmatrix} = \begin{bmatrix} \widehat{\mathfrak{f}}^{(2),h}_{\tau,DNN_i} & \widehat{\mathfrak{f}}^{(3),h}_{\tau,DNN_i} & \widehat{\mathfrak{f}}^{(4),h}_{\tau,DNN_i} & \widehat{\mathfrak{f}}^{(5),h}_{\tau,DNN_i} \\ \widehat{\mathfrak{f}}^{(2),h}_{\tau+1,DNN_i} & \widehat{\mathfrak{f}}^{(3),h}_{\tau+1,DNN_i} & \widehat{\mathfrak{f}}^{(4),h}_{\tau+1,DNN_i} & \widehat{\mathfrak{f}}^{(5),h}_{\tau+1,DNN_i} \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\mathfrak{f}}^{(2),h}_{T,DNN_i} & \widehat{\mathfrak{f}}^{(3),h}_{T,DNN_i} & \widehat{\mathfrak{f}}^{(4),h}_{T,DNN_i} & \widehat{\mathfrak{f}}^{(5),h}_{T,DNN_i} \end{bmatrix}$$

And,

$$\widehat{\boldsymbol{\xi}}^h_t \equiv \begin{bmatrix} \widehat{\xi}^{(2),h}_{\tau,DNN_i} & \widehat{\xi}^{(3),h}_{\tau,DNN_i} & \widehat{\xi}^{(4),h}_{\tau,DNN_i} & \widehat{\xi}^{(5),h}_{\tau,DNN_i} \\ \widehat{\xi}^{(2),h}_{\tau+1,DNN_i} & \widehat{\xi}^{(3),h}_{\tau+1,DNN_i} & \widehat{\xi}^{(4),h}_{\tau+1,DNN_i} & \widehat{\xi}^{(5),h}_{\tau+1,DNN_i} \\ \vdots & \vdots & vdots & \vdots \\ \widehat{\xi}^{(2),h}_{T,DNN_i} & \widehat{\xi}^{(3),h}_{T,DNN_i} & \widehat{\xi}^{(4),h}_{T,DNN_i} & \widehat{\xi}^{(5),h}_{T,DNN_i} \end{bmatrix}$$

Notes: For our analysis the derived factors are calculated for the period τ=1993:01 to T=2017:12, while the data ranging from t_0 =1962:01 to t_{split} =1992:12 is used as a burn-in data to initiate the recursive process of obtaining the the derived factors $\widehat{\mathfrak{f}}^{(n),h}_{t,DNN}$ and $\widehat{\xi}^{(n),h}_t$.

1.4 Empirical Results

For the period 1993:01 to 2017:12, we generated $\mathfrak{f}_{t,DNN}^{(n),h}$ in a recursive way. Figure 1.4 shows the derived DNN factor for each scenario under consideration. Each column uses a different set of information from the term-structure to derive the factor $\mathfrak{f}_{t,DNN}^{(n),h}$. Column (1) shows the the derived DNN factors when we feed the MLP with all the forward rates, i.e., $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60/12)}\right\}$, column (2) when the set of yields is used, i.e., $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ and column (3) when both previous sets are combined, i.e., $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60)}, y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$. Each row represents one of the four groups of maturities. Finally, different colors represent the three variations of DNN as explained in section 1.3.2. A quick inspection in figure 1.4 shows how the different structures of neural networks result in different factors. Clearly, **DNN 3** distinguishes from the other two. We also see that **DNN 1** and **DNN 2** have an evident mean reverting tendency.

In order to better investigate how the factors $\mathfrak{f}_{t,DNN}^{(n),h}$ behave, we plot in figure 1.5 only for the **DNN 2** factors generated by the set of yields in terms of maturity for the period of analysis (1993:01 - 2017:12). Some patterns become evident when we inspect this figure. First, on average the set $\left\{\mathfrak{f}_{t,DNN}^{(2),h}, \mathfrak{f}_{t,DNN}^{(3),h}, \mathfrak{f}_{t,DNN}^{(4),h}, \mathfrak{f}_{t,DNN}^{(5),h}\right\}$ throughout the period of analysis, we notice that it behaves as an increasing function of the maturity (n). In the first months we see that the DNN factors behave quite erratically, what could be interpreted as the neural network changing the weights in its functions more intensively to try to improve the learning process. Another clear pattern inferred from figure 1.5 is that the curve generated at each t apparently moves in synchrony across maturities. This is more evident when we take in consideration the two recessions (2001:04 - 2001:11 and 2008:01 - 2009:06) in the period of analysis. We see that the curve of generated factors move down for all maturities following a recession and for some time after it the values of $\mathfrak{f}_{t,DNN}^{(n),h}$ are low. As the recession fades, the curve $\mathfrak{f}_{t,DNN}^{(n),h}$ slowly start to move up as well.

Following our methodology, we use equation (1.19) to obtain our single factor $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ as a linear combination of the derived factors $\mathfrak{f}_{t,DNN}^{(n),h}$. Figure 1.6 plots $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ for each DNN architecture and the three different sets of $\boldsymbol{Z}_t^{\mathrm{y}}$, our unique factor from the state space in the dynamic term-structure model that captures all the information from the yield curve. Notice that, based on which DNN structure we use, the factor $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ behaves quite differently. In the first years of analysis, the single factor seems to be correlated. However, consistent with our comments from figure 1.4, as the training process of the neural network advances, the three DNNs produce distincts $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, being more evident the contrast of the factor

Figure 1.4: DNN Factor $\mathfrak{f}^{(n),h}_{t,DNN}$ by MLP Architecture and Choice of $\boldsymbol{Z}^{\mathrm{y}}_t$

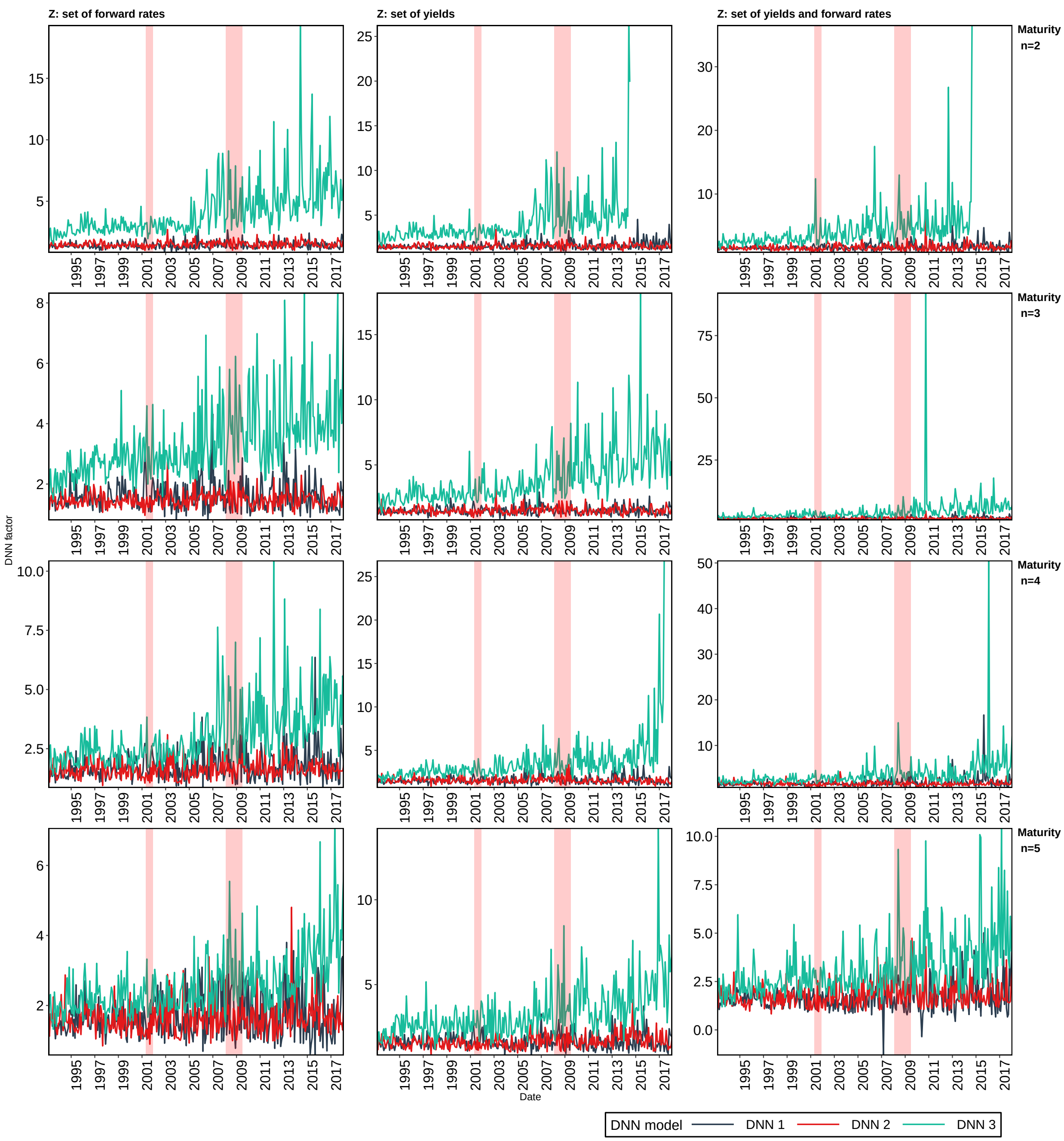

Figure 1.4 shows the derived $\mathfrak{f}^{(n),h}_{t,DNN}$ for each scenario under consideration. Each column uses a different set of information from the term-structure to derive the factor $\mathfrak{f}^{(n),h}_{t,DNN}$. Column (1) shows the the derived DNN factors for $\boldsymbol{Z}^{\mathrm{y}}_t = \left\{ f^{(2/12)}_t, f^{(3/12)}_t, \ldots, f^{(60)}_t \right\}$, column (2) for $\boldsymbol{Z}^{\mathrm{y}}_t = \left\{ y^{(1/12)}_t, y^{(2/12)}_t, \ldots, y^{(60/12)}_t \right\}$ and column (3) for $\boldsymbol{Z}^{\mathrm{y}}_t = \left\{ f^{(2/12)}_t, f^{(3/12)}_t, \ldots, f^{(60/12)}_t, y^{(1)}_t, y^{(2/12)}_t, \ldots, y^{(60/12)}_t \right\}$. Each row represents one of the four groups of maturities. Finally, different colors represent the three variations of DNN considered, as explained in section 1.3.2. The derived factors are calculated for the period 1993:01 to 2017:12, where we use the data from 1962:01 to 1992:12 as a burn-in data to initiate the recursive process of obtaining the the derived factors $\mathfrak{f}^{(n),h}_{t,DNN}$.

Figure 1.5: Derived Factors $\mathfrak{f}_{t,DNN}^{(n),h}$ for **DNN 2** Generated Using the Set of Yields

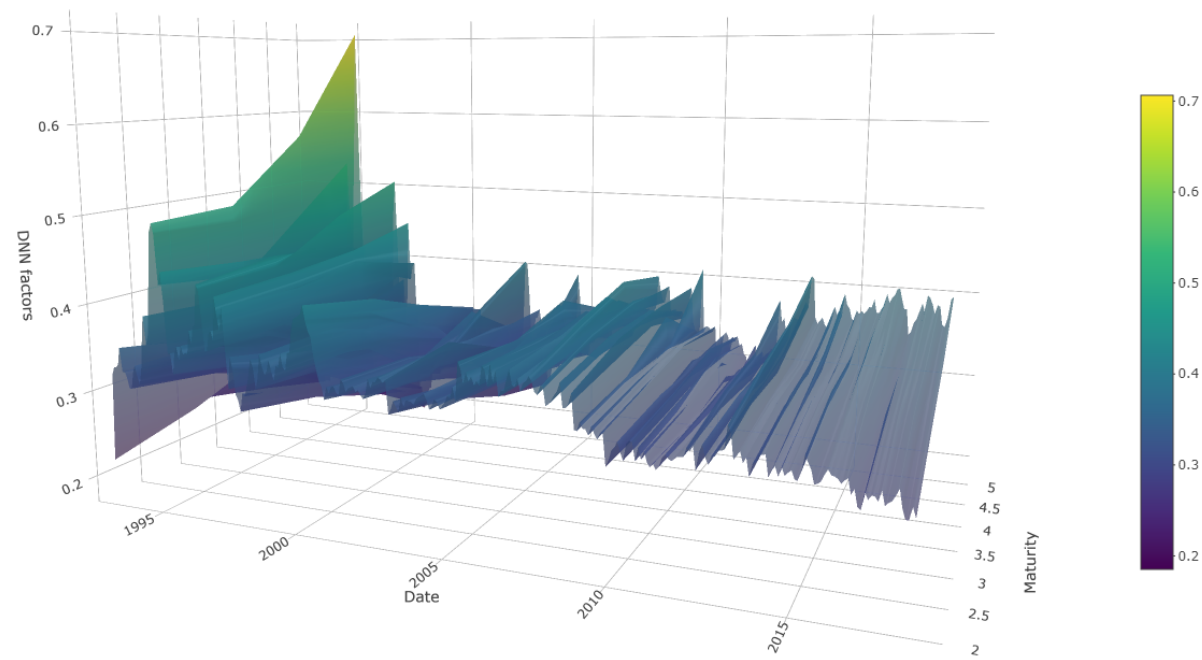

Figure 1.5 shows a 3D representation of $\mathfrak{f}_{t,DNN}^{(n),h}$ generated by the MLP architecture **DNN 2** using the set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{ y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)} \right\}$ in terms of maturity for the period of analysis. **DNN 2** is a feedfoward neural networks architecture with 2 hidden layers ($L = 2$), with 16 and 4 nodes respectively, and an output layer for each group of maturity $n \in \{1, 2, 3, 4\}$. Period of analysis ranges from 1993:01 to 2017:12.

Figure 1.6: Single Factor $\left(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}_t\right)_t^h$ Series by DNN Architecture and Choice of $\boldsymbol{Z}_t^{\mathrm{y}}$

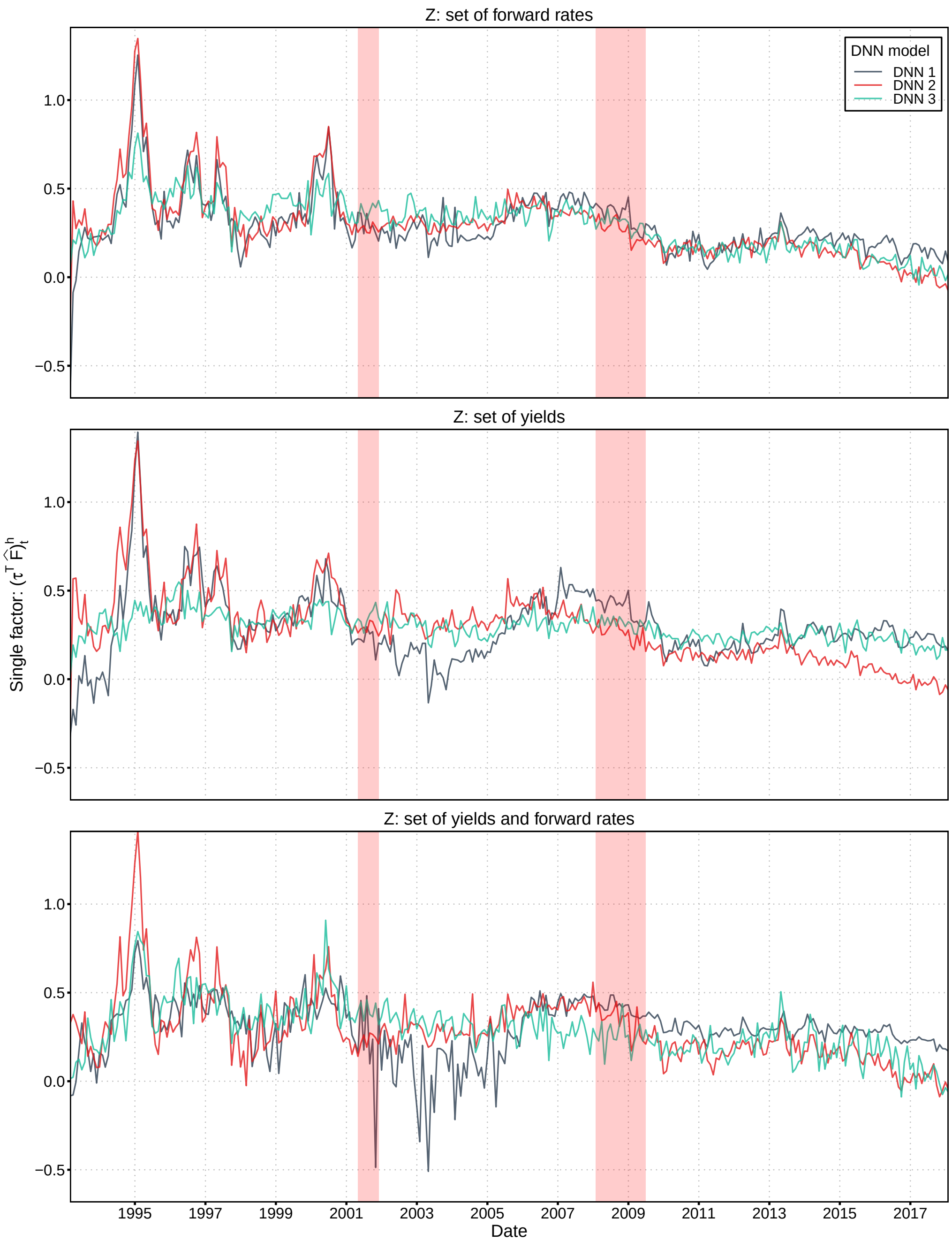

Figure 1.6 shows $\left(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}_t\right)_t^h$ for each DNN architecture and the three different sets of $\boldsymbol{Z}_t^{\mathrm{y}}$. The first panel plots $\left(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}_t\right)_t^h$ when $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60)}\right\}$ is used to obtain $\left(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}_t\right)_t^h$ from the DNN derived factors $\mathfrak{f}_{t,DNN}^{(n),h}$. The panel in the center plots the single factor when $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ is used. Finally, the third panel plots $\left(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}_t\right)_t^h$ when $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60/12)}, y_t^{(1)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ is used. Different colors represent the three variations of DNN considered, as explained in section 1.3.2. The derived factors are calculated for the period 1993:01 to 2017:12, where we use the data from 1962:01 to 1992:12 as a burn-in data to initiate the recursive process.

produced by **DNN 3**, since the its structure is the most different one.

1.4.1 Predictive Regressions

In table 1.1 we have the predictive regressions for the period from 1993:01 to 2017:12 using our derived state variables: $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ and $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$ (alternative 1) or $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$ (alternative 2). We split the regressions in 4 panels, one for each maturity. We evaluate three different predictive regression models. The first one is shown in equation (1.20), where we only use $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ as the state variable. In our model, to complete the state space, we use the orthogonal vector from the projection of $f(\boldsymbol{\xi}_{t+h/12}^{(n)})$ on $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$. The two alternatives for the single factor that captures the unspanned information from the yield curve are the following two regression models. In each panel, we show the these three regressions depending on wich DNN architecture was used to build the single state factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$.

From table 1.1 we see that for 1-month holding period, with no overlapping returns to affect the robustness of our tests, our state variable $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, when used as the only predictor, is always statistically significant for **DNN 1** and **DNN 2**. For **DNN 3**, the single factor loses statistically significance when the maturity increases. More importantly, the adjusted R^2 ranges for maturity of 2 years, for maturity of 2 years, for maturity of 2 years, and for maturity of 5 years. When we add the second state variable that captures the unspanned factors, we keep seeing statistically significance for the same cases, and the adjusted R^2 raises quite substantially, either for alternative 1 $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$, or alternative 2 $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$.

As we discussed above, for each DNN architecture and each set of Z used, we obtained a state factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ for the time varying of the expected returns across all maturities. Out of the 9 different specifications for this single factor, we will focus only on the one formed using the $\mathfrak{f}_{t,DNN}^{(n),h}$ from the **DNN 2** fed with the entire set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$. We do so motivated by two reasons. First, because as shown in Gu et al. (2018), higher complexity with a much "deeper" network it is not necessary associated with better out-of-sample results. And second, because this pair of choices result in smaller MSE in our period of analysis.

Table 1.1: Predictive Regressions Using $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$, $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$ and $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$ as State Variables

Panel A: $rx_{t+h/12}^{(2)}$

	DNN 1			**DNN 2**			**DNN 3**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}})_t^h$	0.810***	0.810***	0.810***	0.811***	0.811***	0.811***	1.419***	1.419***	1.419***
	(0.160)	(0.149)	(0.147)	(0.131)	(0.119)	(0.119)	(0.414)	(0.377)	(0.356)
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-2),h}$		0.760***			0.779***			0.875***	
		(0.204)			(0.180)			(0.211)	
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{h}$			0.591***			0.525***			0.679***
			(0.139)			(0.126)			(0.138)
Constant	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010	−0.189*	−0.189*	−0.189**
	(0.054)	(0.050)	(0.049)	(0.039)	(0.035)	(0.035)	(0.110)	(0.101)	(0.094)
Observations	300	300	300	300	300	300	300	300	300
Adjusted R^2	0.100	0.148	0.159	0.119	0.178	0.175	0.046	0.105	0.124

Panel B: $rx_{t+h/12}^{(3)}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}})_t^h$	0.959***	0.959***	0.959***	0.943***	0.943***	0.943***	1.175*	1.175**	1.175**
	(0.248)	(0.234)	(0.233)	(0.199)	(0.188)	(0.184)	(0.630)	(0.566)	(0.559)
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-3),h}$		0.799***			0.789***			0.984***	
		(0.234)			(0.219)			(0.236)	
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{h}$			0.765***			0.757***			0.929***
			(0.225)			(0.205)			(0.224)
Constant	−0.008	−0.008	−0.008	−0.003	−0.003	−0.003	−0.072	−0.072	−0.072
	(0.087)	(0.082)	(0.082)	(0.063)	(0.060)	(0.059)	(0.169)	(0.153)	(0.150)
Observations	300	300	300	300	300	300	300	300	300
Adjusted R^2	0.055	0.092	0.093	0.063	0.100	0.109	0.010	0.067	0.067

(Continued)

Table 1.1: Predictive Regressions Using $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^h$ and $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$ as State Variables *(Continued)*

Panel C:	$rx_{t+h/12}^{(4)}$								
$(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}})_t^h$	1.073***	1.073***	1.073***	1.065***	1.065***	1.065***	0.864	0.864	0.864
	(0.334)	(0.320)	(0.317)	(0.264)	(0.253)	(0.248)	(0.835)	(0.759)	(0.755)
$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-4),h}$		0.795***			0.807***			1.038***	
		(0.291)			(0.288)			(0.289)	
$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{h}$			0.902***			0.945***			1.144***
			(0.312)			(0.284)			(0.313)
Constant	0.002	0.002	0.002	0.004	0.004	0.004	0.063	0.063	0.063
	(0.120)	(0.116)	(0.115)	(0.088)	(0.086)	(0.085)	(0.228)	(0.209)	(0.207)
Observations	300	300	300	300	300	300	300	300	300
Adjusted R^2	0.036	0.060	0.063	0.042	0.069	0.080	0.001	0.046	0.046

Panel D:	$rx_{t+h/12}^{(5)}$								
$(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}})_t^h$	1.158***	1.158***	1.158***	1.181***	1.181***	1.181***	0.542	0.542	0.542
	(0.415)	(0.395)	(0.398)	(0.325)	(0.312)	(0.309)	(1.025)	(0.949)	(0.939)
$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-5),h}$		0.854**			0.848***			1.069***	
		(0.336)			(0.318)			(0.339)	
$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{h}$			1.000**			1.081***			1.322***
			(0.398)			(0.363)			(0.404)
Constant	0.017	0.017	0.017	0.010	0.010	0.010	0.198	0.198	0.198
	(0.152)	(0.146)	(0.147)	(0.114)	(0.111)	(0.111)	(0.284)	(0.267)	(0.263)
Observations	300	300	300	300	300	300	300	300	300
Adjusted R^2	0.025	0.049	0.046	0.032	0.060	0.062	-0.002	0.033	0.036

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.1 reports the predictive regressions using $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^h$ and $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$ as state variables for 1-month holding period ($h = 1$). Panel A reports the predictive regressions for maturity $n = 2$ years. Panel B reports the predictive regressions for maturity $n = 3$ years. Panel C reports the predictive regressions for maturity $n = 4$ years. Panel D reports the predictive regressions for maturity $n = 5$ years. The state factor $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ reported in this table used only the set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ to feed the MLP. We use Newey-West robust standard errors. Sample ranges from 1993 : 01 to 2017 : 12.

Table 1.2 presents the correlation between our state variables, i.e., $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$ and $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$, as well as with the Cochrane-Piazzesi and Ludvingson-Ng factors calculated as explained in section 1.2.2. By definition the correlation between our factor that summarizes the information from the term-structure and the alternatives for the one(s) that complete the state space is 0, which we can see in table 1.2. Now the correlation between our factors for the unspanned information from the yield curve are always high, ranging from .84 to .99. We see that the correlation between $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$ and the factor for low maturities, especially $n = 2$, is high (.99). For the remaining one, we notice that the correlation decays.

Table 1.2: Correlation Matrix

	$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})_t^h$	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^h$	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-2),h}$	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-3),h}$	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-4),h}$	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-5),h}$	$\widehat{CP}_t^h$	$\widehat{LN}_t^h$
$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})_t^h$	1	0	0	0	0	0	0.556	-0.059
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^h$	0	1	0.995	0.912	0.904	0.919	0.129	0.171
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-2),h}$	0	0.995	1	0.938	0.900	0.888	0.135	0.174
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-3),h}$	0	0.912	0.938	1	0.947	0.849	0.170	0.203
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-4),h}$	0	0.904	0.900	0.947	1	0.959	0.173	0.204
$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-5),h}$	0	0.919	0.888	0.849	0.959	1	0.146	0.178
$\widehat{CP}_t^h$	0.556	0.129	0.135	0.170	0.173	0.146	1	-0.007
$\widehat{LN}_t^h$	-0.059	0.171	0.174	0.203	0.204	0.178	-0.007	1

Table 1.2 reports the correlation between our single factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, with the factors that complete the state space in our dynamic term-structure model. The first alternative is $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$, which is the unique factor obtained as the projection of $\overline{rx}_{t+h/12}$ in $\widehat{\boldsymbol{\xi}}_{t+h/12}$. The second alternative is a similar projection, however for each maturity $n \in \{2,3,4,5\}$ we regress $rx_{t+h/12}^{(n)}$ on $\widehat{\boldsymbol{\xi}}_{t+h/12}^{(-n),h} \equiv \widehat{\boldsymbol{\xi}}_{t+h/12}^{h} \setminus \widehat{\xi}_{t+h/12}^{(n),h}$. We use orthogonal vector from the projection of each one of them on $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ to complete our state space. The table also reports the correlation with the Cochrane-Piazzesi and Ludvingson-Ng factors calculated as explained in section 1.2.2. The period of analysis ranges from 1993 : 01 to 2017 : 12.

1.4.2 Comparison with Other Factors from the Literature

In this section, we are interested in evaluating how our derived and theoretically motivated factors compare with the other factors and frameworks that were proposed in the literature to explain the time-varying expected excess returns. Figure 1.7 shows in two separated panels our single factor that spans the information from the term-structure, $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, as well as the factor with the spanned risks from alternative 1, $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$, along with the Cochrane-Piazzesi and Ludvingson-Ng factors. Aligned with the correlation in table 1.2, we see that our factor has some positive correlation (.56) with the Cochrane-Piazzesi factor. However, this correlation is not strong enough to claim that both are capturing the same

Figure 1.7: Time Series of our Derived Factors $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t$ and $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)^h_t$, along with $\widehat{CP}^h_t$ and $\widehat{LN}^h_t$

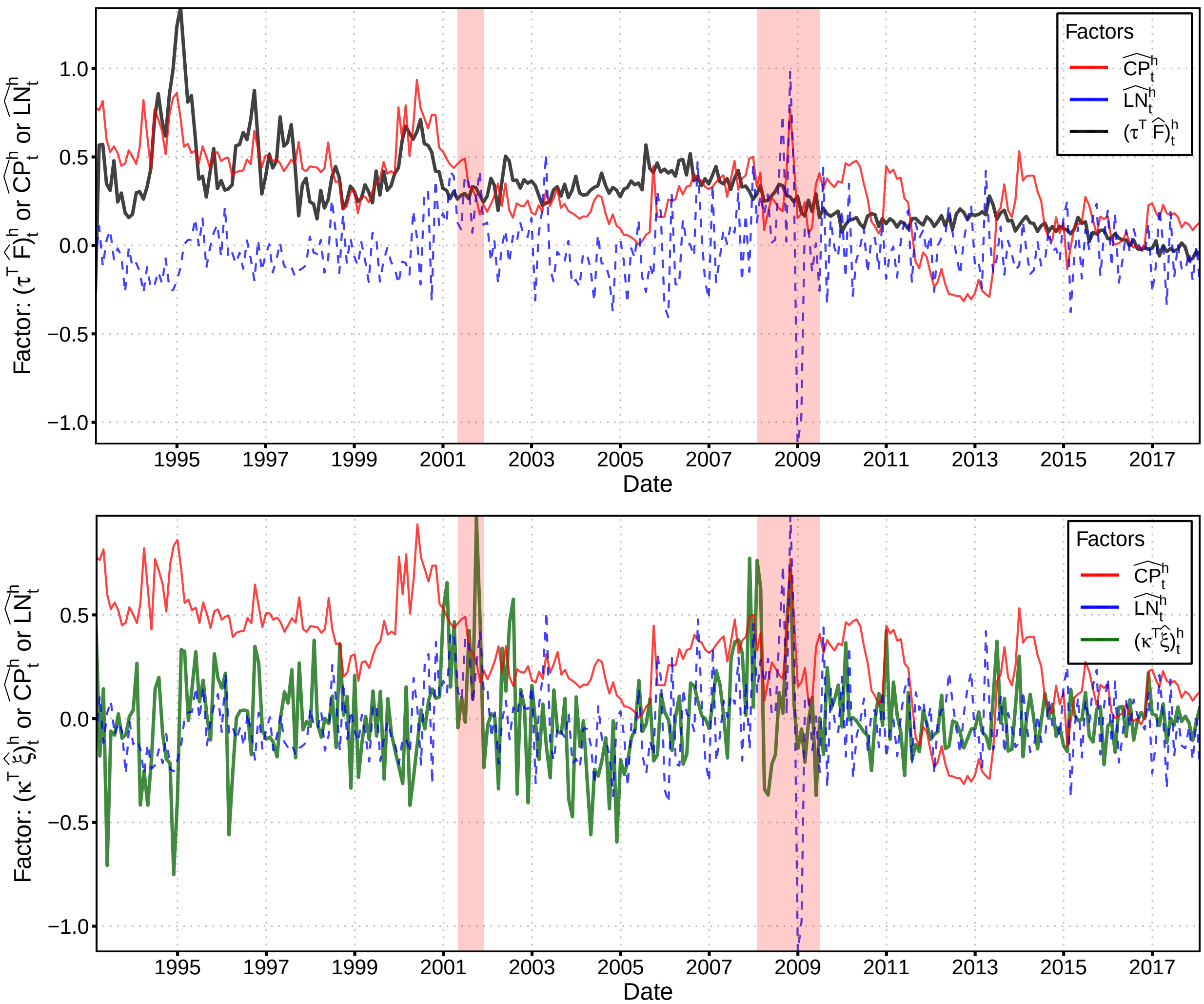

Figure 1.7 plots in two separated panels our single factor that spans the information from the term-structure, as well as the factor with the spanned risks (alternative 1). The graph in the top plots $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t$ along with the Cochrane-Piazzesi and Ludvingson-Ng factors. The bottom graph plots $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)^h_t$ along with the same factors. The period of analysis ranges from 1993 : 01 to 2017 : 12.

Table 1.3: Predictive Regressions with $\left(\boldsymbol{\tau}^\top\widehat{\mathfrak{F}}_t\right)_t^h$ and $\left(\boldsymbol{\kappa}^\top\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$, along with the Cochrane-Piazzesi and Ludvingson-Ng factors, and Fama-Bliss Regressions with Forward Spreads

Panel A: $rx^{(2)}_{t+h/12}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$(\boldsymbol{\tau}^\top\widehat{\mathfrak{F}})_t^h$	0.847***	0.842***	0.853***	0.824***	0.525***	0.582***	0.582***	0.614***
	(0.124)	(0.115)	(0.128)	(0.117)	(0.154)	(0.140)	(0.145)	(0.135)
$\boldsymbol{M}_{\boldsymbol{\tau}^\top\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top\widehat{\boldsymbol{\xi}})_t^{(-2),h}$		0.658***		0.745***		0.704***		0.558***
		(0.172)		(0.182)		(0.182)		(0.185)
$\widehat{LN}_t^h$	0.617***	0.529***					0.559***	0.518***
	(0.127)	(0.120)					(0.110)	(0.110)
$fs_t^{(n,h)}$			−0.746	−0.225			−0.570	−0.172
			(0.476)	(0.438)			(0.437)	(0.429)
$\widehat{CP}_t^h$					0.454***	0.364***	0.465***	0.375***
					(0.126)	(0.112)	(0.112)	(0.109)
Constant	−0.013	−0.012	0.031	0.002	−0.060	−0.050	−0.031	−0.044
	(0.037)	(0.034)	(0.051)	(0.047)	(0.039)	(0.036)	(0.045)	(0.043)
Observations	300	300	300	300	300	300	300	300
Adjusted R^2	0.183	0.223	0.128	0.177	0.150	0.197	0.215	0.240

Panel B: $rx^{(3)}_{t+h/12}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$(\boldsymbol{\tau}^\top\widehat{\mathfrak{F}})_t^h$	0.996***	0.989***	0.940***	0.947***	0.559**	0.648***	0.626***	0.719***
	(0.190)	(0.184)	(0.199)	(0.188)	(0.245)	(0.234)	(0.238)	(0.237)
$\boldsymbol{M}_{\boldsymbol{\tau}^\top\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top\widehat{\boldsymbol{\xi}})_t^{(-3),h}$		0.620***		0.852***		0.692***		0.585**
		(0.209)		(0.228)		(0.226)		(0.237)
$\widehat{LN}_t^h$	0.921***	0.800***					0.900***	0.823***
	(0.209)	(0.201)					(0.194)	(0.191)
$fs_t^{(n,h)}$			−0.215	0.410			−0.053	0.394
			(0.554)	(0.532)			(0.525)	(0.542)
$\widehat{CP}_t^h$					0.608***	0.467**	0.583***	0.437**
					(0.205)	(0.195)	(0.188)	(0.198)
Constant	−0.007	−0.006	0.021	−0.049	−0.070	−0.054	−0.064	−0.098
	(0.060)	(0.059)	(0.091)	(0.087)	(0.063)	(0.061)	(0.082)	(0.082)
Observations	300	300	300	300	300	300	300	300
Adjusted R^2	0.120	0.141	0.060	0.099	0.084	0.111	0.136	0.151

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.3 reports the predictive regressions using $\left(\boldsymbol{\tau}^\top\widehat{\mathfrak{F}}_t\right)_t^h$ and $\left(\boldsymbol{\kappa}^\top\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$, along with $\widehat{CP}_t^h$, $\widehat{LN}_t^h$ and the Fama-Bliss regressions with forward spreads for 1-month holding period ($h = 1$). Panel A reports the predictive regressions for maturity $n = 2$ years. Panel B reports the predictive regressions for maturity $n = 3$ years. The state factor $\left(\boldsymbol{\tau}^\top\widehat{\mathfrak{F}}_t\right)_t^h$ reported in this table used only the set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ to feed the MLP. We use Newey-West robust standard errors. Sample ranges from 1993 : 01 to 2017 : 12.

Table 1.4: **(Continued)** Predictive Regressions with $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ and $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$, along with the Cochrane-Piazzesi and Ludvingson-Ng factors, and Fama-Bliss Forward Spreads

Panel C: $rx_{t+h/12}^{(4)}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}})_t^h$	1.135***	1.127***	1.082***	1.108***	0.547	0.651**	0.685**	0.790**
	(0.254)	(0.247)	(0.270)	(0.257)	(0.335)	(0.323)	(0.329)	(0.329)
$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-4),h}$		0.609**		0.872***		0.688**		0.555**
		(0.262)		(0.289)		(0.291)		(0.274)
$\widehat{LN}_t^h$	1.218***	1.079***					1.222***	1.118***
	(0.307)	(0.287)					(0.285)	(0.273)
$fs_t^{(n,h)}$			0.260	0.665			0.386	0.655
			(0.622)	(0.595)			(0.593)	(0.587)
$\widehat{CP}_t^h$					0.822***	0.657**	0.755***	0.606**
					(0.290)	(0.276)	(0.265)	(0.272)
Constant	−0.0003	0.0002	−0.038	−0.103	−0.085	−0.068	−0.144	−0.171
	(0.085)	(0.084)	(0.130)	(0.124)	(0.089)	(0.087)	(0.121)	(0.118)
Observations	300	300	300	300	300	300	300	300
Adjusted R^2	0.095	0.108	0.039	0.070	0.063	0.081	0.112	0.122

Panel D: $rx_{t+h/12}^{(5)}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}})_t^h$	1.268***	1.258***	1.247***	1.263***	0.511	0.626	0.736*	0.834**
	(0.315)	(0.305)	(0.334)	(0.318)	(0.422)	(0.401)	(0.409)	(0.400)
$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-5),h}$		0.673**		0.872***		0.738**		0.590**
		(0.281)		(0.312)		(0.315)		(0.279)
$\widehat{LN}_t^h$	1.501***	1.337***					1.518***	1.386***
	(0.421)	(0.381)					(0.387)	(0.360)
$fs_t^{(n,h)}$			0.633	0.789			0.739	0.848
			(0.698)	(0.656)			(0.658)	(0.632)
$\widehat{CP}_t^h$					1.064***	0.882**	0.967***	0.818**
					(0.380)	(0.352)	(0.343)	(0.337)
Constant	0.005	0.005	−0.116	−0.147	−0.106	−0.086	−0.248	−0.253*
	(0.111)	(0.109)	(0.166)	(0.158)	(0.117)	(0.115)	(0.158)	(0.152)
Observations	300	300	300	300	300	300	300	300
Adjusted R^2	0.082	0.098	0.031	0.062	0.054	0.074	0.103	0.114

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.4 reports the predictive regressions using $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ and $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$, along with $\widehat{CP}_t^h$, $\widehat{LN}_t^h$ and the Fama-Bliss regressions with forward spreads for 1-month holding period ($h = 1$). Panel C reports the predictive regressions for maturity $n = 4$ years. Panel D reports the predictive regressions for maturity $n = 5$ years. The state factor $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ reported in this table used only the set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ to feed the MLP. We use Newey-West robust standard errors. Sample ranges from 1993 : 01 to 2017 : 12.

information. We must say that this should be an expected result, given that both factors capture information from the term-structure.

On the other hand, $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ seems to be uncorrelated ($-.06$) with the Ludvingson-Ng factor. Now, the time series in figure 1.7 with the correlation shows us an interesting result. Consistent with our framework, the unspanned risks from the term-structure should be captured by our orthogonal factor $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$, or $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$. Given that Ludvingson-Ng factor is solely based in macroeconomic variables information, we see that the correlation of $\widehat{LN}_t^h$ with our unspanned risks factors ranges from .17 to .20. This could be understood as the risk factors not spanned by the yield-curve, that are captured by our orthogonal state variable and Ludvingson-Ng approach.

Next, we run predictive regressions using our factors with the main factors proposed in the literature. Tables 1.3 and 1.4 reports the predictive regressions using $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$ and $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$, along with $\widehat{CP}_t^h$, $\widehat{LN}_t^h$ and the Fama-Bliss regressions with forward spreads for 1-month holding period ($h = 1$). For each maturity (in each one of the four panels), there are 8 different regressions specifications. In pairs, we run predictive regressions first only with our state variable that spans the term-structure along with a proposed factor from the literature. Then, we add our state variable of the unspanned risks.

The results, consistent with table 1.1, shows that our state variables are still significant when adding either CP LN or forward spreads especially for maturities $n = 2$ and $n = 3$. Interestingly, the forward spreads loose statistical significance. In column (8) we see that our factors remain significant if we still add all factors, including $\widehat{CP}_t^h$, $\widehat{LN}_t^h$ and the forward spreads. As already mentioned, nonetheless, for higher maturities our factors loose some statistical significance.

1.4.3 Economic Interpretation

Some natural questions may arise at this stage. What are the economic interpretation of these factors derived from a deep neural network? How are they linked with the macroeconomic variables? What macroeconomic and possibly other risk measures do they capture? In order to answer these questions, we make use of the the FRED-MD dataset (McCracken and Ng, 2016), which is a large macroeconomic database and monthly updated by the FRED[9] that shares the predictive content of a widespread dataset known in the literature as Stock-Watson (Stock and Watson (1996)). It is a balanced panel consisting of 128 macroeconomic and financial variables. The variables are split in 8 groups: (1) output and income, (2) labor market, (3) housing, (4) consumption, orders, and inventories, (5) money and credit, (6)

[9] https://research.stlouisfed.org/econ/mccracken/fred-databases/

Figure 1.8: Marginal R^2 of the factors $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$ and $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^h$

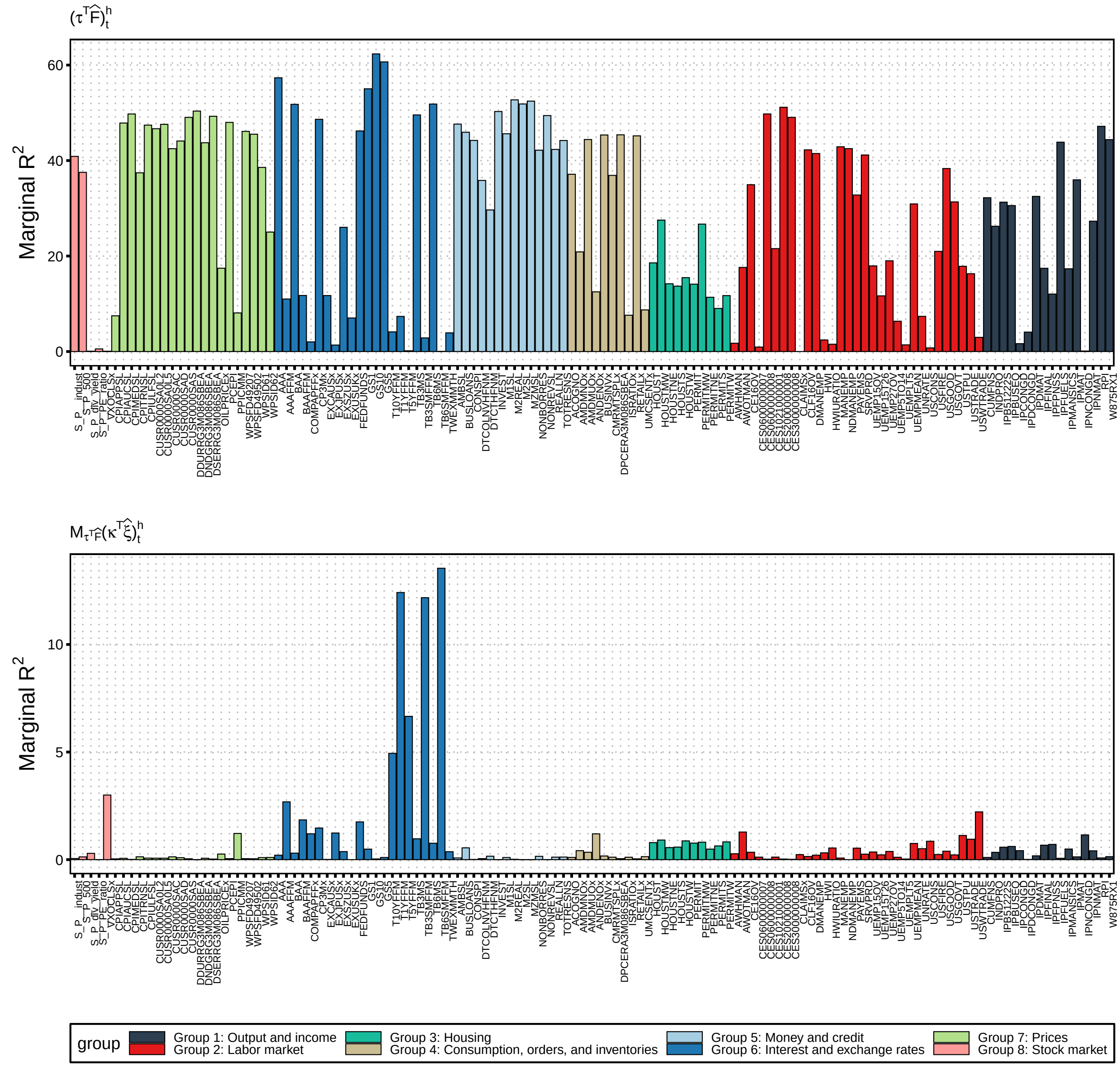

Figure 1.8 reports the marginal R^2 of the factor $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$ in the top panel, and $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^h$ in the bottom panel. The marginal R^2 is obtained with the regression of each one the 128 variables from the FRED-MD on $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$ or $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^h$. Sample ranges from 1993 : 01 to 2017 : 12.

Figure 1.9: Marginal R^2 of the factors $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_{t+h/12}$

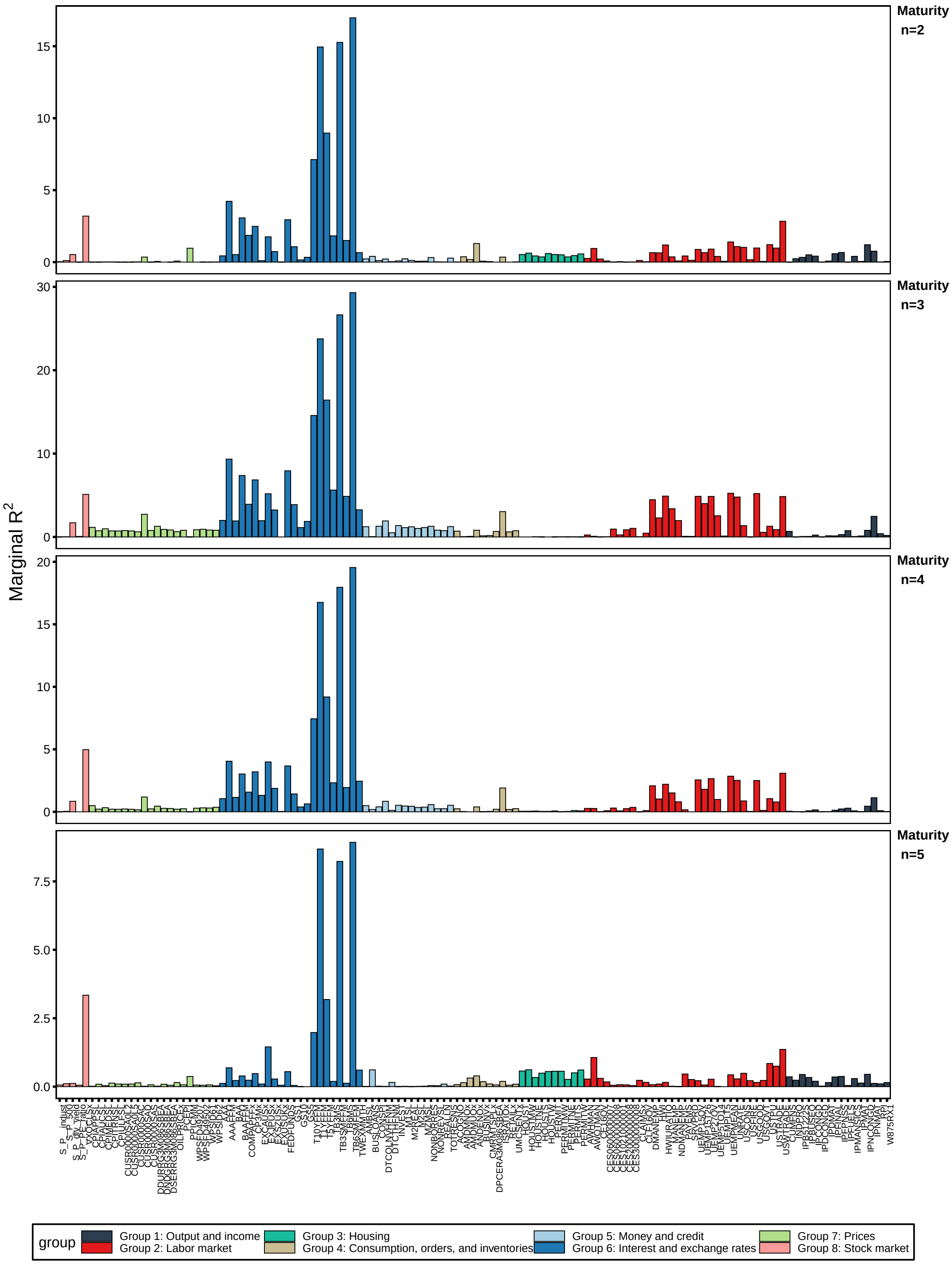

Figure 1.9 reports the marginal R^2 of the factors $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_{t+h/12}$ by maturity. The marginal R^2 is obtained with the regression of each one the 128 variables from the FRED-MD on $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_{t+h/12}$. Sample ranges from 1993 : 01 to 2017 : 12.

Figure 1.10: Marginal R^2 Using Sentiment-Based Measures

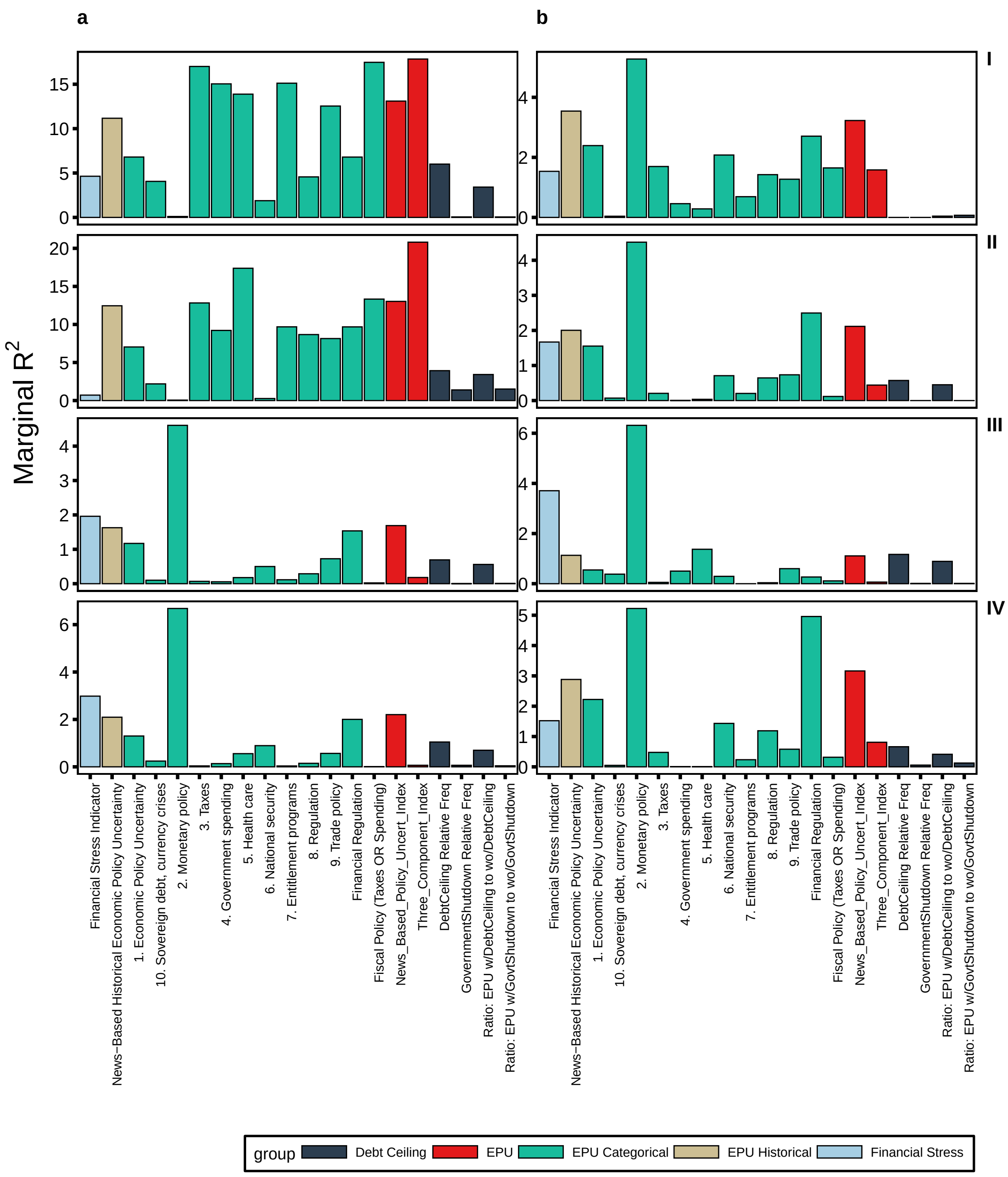

Figure 1.10 reports the marginal R^2 obtained from sentiment-based measures. It is obtained with the regression of each one these indexes on our state variables. For comparison, we also report for the Cochrane-Piazzesi and Ludvingson-Ng factors. Row (I), panel (a) shows the marginal R^2 for $\widehat{CP}_t^h$, and panel (b) plots for $\widehat{LN}_t^h$. Row (II) panel (a) plots for our spanning factor $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$, and panel (b) for the unspanned factor $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^h$. Rows (III) and (IV) plots for the other derived unspanned state variables: in panel (III-a) we have $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-2),h}$, panel (III-b) plots $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-3),h}$, panel (IV-a) shows the marginal R^2 for $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-4),h}$, and panel (IV-b) for $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-5),h}$. Sample ranges from 1993 : 01 to 2017 : 12.

interest and exchange rates, (7) prices, and (8) stock market. In Appendix 1.6.1, table 1.11 list all the variables, codes and their groups.

In a similar fashion to Ludvigson and Ng (2009), we find the marginal R^2 of our factors $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_{t+h/12}^{(-n),h}$ and $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_{t+h/12}^{h}$. The marginal R^2 simply is the goodness-of-fit of the regression of each one the 128 variables from the FRED-MD on our state variables. Figure 1.8 reports the marginal R^2 as bar charts using colors to split the 8 groups. A quick inspection in this figure reveals that $\left(\boldsymbol{\tau}^{\top}\widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ has a high R^2 with many macroeconomic variables. However, this is not evenly distributed within and across groups. We can see that especially the groups (7) *prices* and (5) *money and credit* have R^2 above or around .40 for most of their variables. Even though the groups (6) *interest and exchange rates* and (2) *labor market* have some variables with high R^2, there are many others within the group that do not. Thus, apparently, the state variable spanning the yield curve loads more in monetary variables movements, what should be expected. Nonetheless, it also captures a wide range of macroeconomic variables.

An interesting pattern stands out when we calculate the marginal R^2 with our unspanned factor from the term-structure. The graph in the bottom of figure 1.8 shows the marginal R^2 for $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^h$, while figure 1.9 shows for each $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$ in different panels. It is clear that our state variable of the unspanned risks capture important information left out by the spanning factor. We see that especially some variables from the group (6) *interest and exchange rates* stand out. This pattern is consistent either for $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^h$, or the four factors $\left(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}}\right)_t^{(-n),h}$. Among these variables that have a high R^2 with our unspanned factor, there are relevant variables, such as *6-Month Treasury C Minus FEDFUNDS* (TB6SMFFM), *1-Year Treasury C Minus FEDFUNDS* (T1YFFM) and *3-Month Treasury C Minus FEDFUNDS* (TB3SMFFM).

Next, we evaluate if our factors capture any sentiment information. To do so, we make use of several indexes recently proposed in the literature that seek to estimate the state of the sentiment in the economy. The first one is the economic policy uncertainty measure (EPU) from Baker et al. (2016). The EPU is an index that proxies for movements in policy-related economic uncertainty for U.S., being based on newspaper coverage frequency. The authors also calculated a categorical EPU, which is derived using results from the Access World News database of over 2,000 US newspapers, in such a way that each one of the sub-indexes requires the economic uncertainty term, as well as a set of categorical policy terms[10].

[10] As an example, the category *Monetary policy* has the following terms: Monetary policy - federal reserve, the fed, money supply, open market operations, quantitative easing, monetary policy, fed funds rate, overnight lending rate, Bernanke, Volcker, Greenspan, central bank, interest rates, fed chairman, fed chair, lender of last resort, discount window, European Central Bank, ECB, Bank of England, Bank of Japan,

In the sense of the EPU, we also use the financial stress indicator (FSI) for the U.S from Püttmann (2018). The essence of the FSI is being an indicator of negative financial sentiment. It is based on the reporting in five major US newspapers[11]. Püttmann (2018) shows that the FSI is a robust indicator, such that an increase in negative financial sentiment is followed by a fall in output, higher unemployment, lower stock market returns and rising corporate bond spreads.

Figure 1.10 plots in each panel the marginal R^2 obtained using these sentiment-based measures, where we use colors to split between each index. Row (I), panel (a) shows the marginal R^2 for $\widehat{CP}_t^h$, and panel (b) plots for $\widehat{LN}_t^h$ for comparison. Row (II) panel (a) plots for our spanning factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, and panel (b) for the unspanned factor $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^h$. Rows (III) and (IV) plots for the other derived unspanned state variables: in panel (III-a) we have $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-2),h}$, panel (III-b) plots $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-3),h}$, panel (IV-a) shows the marginal R^2 for $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-4),h}$, and panel (IV-b) for $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)_t^{(-5),h}$.

It is clear 1.10 from figure three facts: (i) our spanning factor and the Cochrane-Piazzesi factor have similar marginal R^2, (ii) our unspanned state factors and the Ludvingson-Ng also have similar marginal R^2, and most important (iii) our unspanned factors has their highest R^2 with the categorical EPU related to monetary policy. Therefore, there is some evidence that the spanned factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$, or even the Cochrane-Piazzesi factor, cannot capture some economy sentiment associated with possible changes in the monetary policy.

1.4.4 Out-of-Sample Forecasting Performance

In this section we are interested in to know how the predictive regressions using our DNN derived state variables behave in an out-of-sample (OoS) analysis. Following Campbell and Thompson (2007); Gargano et al. (2019), we compute the out-of-sample R^2 for all possible predictive regression models from tables 1.1 and 1.3. Additionally, we also consider univariate predictive regressions using only $\widehat{LN}_t^h$, or $fs_t^{(n,h)}$, or $\widehat{CP}_t^h$. We set the out-of-sample period to range from 1997 : 01 to 2017 : 12, where the data from 1993 : 01 to 1996 : 12 is used to initiate the analysis. To avoid any look-ahead bias, at each $\tau \in \tau_{OoS}$, where τ_{OoS} is the OoS subsample, we use all the previous information up to $\tau - 1$ to obtain the point forecast of $rx^{(n)}$ for the month τ. The out-of-sample R^2 is computed as

BOJ, Bank of China, Bundesbank, Bank of France, Bank of Italy

[11]Boston Globe, Chicago Tribune, Los Angeles Times, Wall Street Journal and Washington Post.

$$R^{2(n)}_{OoS,i} = 1 - \frac{\sum_{\tau \in \tau_{OoS}} \left(rx^{(n)}_{t+h/12|t} - \widehat{rx}^{(n)}_{t+h/12|t} \right)^2}{\sum_{\tau \in \tau_{OoS}} \left(rx^{(n)}_{t+h/12|t} - \overline{rx}^{(n)}_{t+h/12|t} \right)^2} \tag{1.29}$$

where $\widehat{rx}^{(n)}_{t+h/12|t}$ is the estimate of the conditional mean of the excess returns for the bond with maturity (n), and $\overline{rx}^{(n)}_{t+h/12|t}$ is the estimate of the conditional mean assuming that the excess returns are constant (as under the expectation hypothesis), implying that the βs from all predictive regressions are assumed to be zero for the same bond with maturity (n). Notice that evidence of time-varying return predictability is obtained when the out-of-sample R^2 is positive.

Table 1.5 summarizes the R^2_{OoS} of our predictive regressions. The first three rows present the same regressions models from table 1.1, the last three rows report the univariate predictive regression using the other factors from the literature: the Cochrane-Piazzesi and Ludvingson-Ng factors, and Fama-Bliss regressions with forward spreads. Finally, the remaining rows are the same regressions models from table 1.3.

It is clear that for $n = 2$ and $n = 3$, we see evidence of time-varying return predictability. Also, we can see indication that the parsimonious regressions using either $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t$ or our unspanned factors, provide comparable better R^2_{OoS}, especially for low maturities. Notice that our factors provide higher R^2_{OoS} when compared to univariate predictive regressions using other factors from the literature. For longer maturities, especially $n = 5$, no regression model provided evidence of time-varying return predictability. However, the higher R^2_{OoS} are still those obtained using the DNN factors.

1.4.5 Relation with PCs

A natural question is how these factors relate with the first principal components from the term-structure. In table 1.6 we present the correlation between the first five principal components and our state variables, the spanning factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t$ and the unspannned factor $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t$. We also computed the correlations with the other representation of the unspanned factors $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_t$, with the Cochrane-Piazzesi and Ludvingson-Ng factors, and for the sake of completeness with the latent DNN factors $\mathfrak{f}^{(n),h}_{t,DNN}$ derived from the selected neural network.

Table 1.5: Out-of-Sample R^2

Regression	**Maturity** $n=2$	**Maturity** $n=3$	**Maturity** $n=4$	**Maturity** $n=5$
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t)^h_t + \epsilon_{t+h/12}$	0.17	0.03	-0.02	-0.04
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_t + \epsilon_{t+h/12}$	0.21	0.05	-0.01	-0.02
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{h}_t + \epsilon_{t+h/12}$	0.22	0.05	-0.01	-0.03
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \widehat{LN}^h_t + \epsilon_{t+h/12}$	0.21	0.04	-0.03	-0.05
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_t + \beta_3 \widehat{LN}^h_t + \epsilon_{t+h/12}$	0.23	0.04	-0.02	-0.05
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 fs^{(n,h)}_t + \epsilon_{t+h/12}$	0.26	0.08	0.02	-0.00
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_t + \beta_3 fs^{(n,h)}_t + \epsilon_{t+h/12}$	0.27	0.08	0.02	-0.00
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \widehat{CP}^h_t + \epsilon_{t+h/12}$	0.20	0.01	-0.06	-0.09
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_t + \beta_3 \widehat{CP}^h_t + \epsilon_{t+h/12}$	0.22	0.01	-0.06	-0.08
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \widehat{LN}^h_t + \beta_3 fs^{(n,h)}_t + \beta_4 \widehat{CP}^h_t + \epsilon_{t+h/12}$	0.19	-0.03	-0.10	-0.13
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t + \beta_2 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_t + \beta_3 \widehat{LN}^h_t + \beta_4 fs^{(n,h)}_t + \beta_5 \widehat{CP}^h_t + \epsilon_{t+h/12}$	0.19	-0.04	-0.11	-0.13
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \widehat{LN}^h_t + \epsilon_{t+h/12}$	0.12	-0.02	-0.06	-0.07
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 fs^{(n,h)}_t + \epsilon_{t+h/12}$	0.18	0.05	0.00	-0.01
$rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \widehat{CP}^h_t + \epsilon_{t+h/12}$	0.15	-0.02	-0.08	-0.10

Table 1.5 reports the OoS R^2 of our predictive regressions. The first three rows present the same regressions models from table 1.1, the last three rows report the univariate predictive regression using the other factors from the literature: $\widehat{LN}^h_t$, $fs^{(n,h)}_t$, and $\widehat{CP}^h_t$. Finally, the remaining rows are the same regressions models from table 1.3. The out-of-sample period ranges from 1997 : 01 to 2017 : 12, where the data from 1993 : 01 to 1996 : 12 is used to initiate the analysis.

Table 1.6: Correlations with Principal Components of the Term-Structure

		PC1 Level	**PC2** Slope	**PC3** Curvature	**PC4**	**PC5**
Spanning Factor	$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})_t^h$	-0.7549	0.1390	-0.0498	0.1685	-0.0731
Unspanning Factors	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^h$	-0.0212	-0.1854	0.4276	-0.0313	0.0907
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-2),h}$	-0.0532	-0.2055	0.4353	-0.0635	0.0953
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-3),h}$	-0.1426	-0.2531	0.4837	-0.1288	0.0909
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-4),h}$	-0.0709	-0.1939	0.4791	-0.0424	0.0744
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^{(-5),h}$	0.0204	-0.1285	0.4284	0.0397	0.0671
Other Factors	$\widehat{CP}_t^h$	-0.7170	0.2799	0.2859	0.3689	0.0413
	$\widehat{LN}_t^h$	0.0492	-0.1539	0.2745	-0.0328	-0.0240
Derived $\mathfrak{f}_{t,DNN}^{(n),h}$	$\mathfrak{f}_{t,DNN}^{(2),h}$	-0.7378	-0.0243	-0.0424	0.1556	-0.0399
	$\mathfrak{f}_{t,DNN}^{(3),h}$	-0.4654	-0.1406	-0.1674	0.1332	0.0350
	$\mathfrak{f}_{t,DNN}^{(4),h}$	-0.1252	-0.0483	-0.3031	0.1284	0.0389
	$\mathfrak{f}_{t,DNN}^{(5),h}$	0.1284	0.0498	-0.3747	0.0690	0.1224

Table 1.6 reports the correlation of the first five principal components, PC1 (level), PC2 (slope), PC3 (curvature), PC4, and PC5 computed from the the monthly set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$. The correlations are calculated against the spanning factor, the unspanned factor(s), the Cochrane-Piazzesi and Ludvingson-Ng factors, and for the sake of completeness with the latent DNN factors. The period of analysis ranges from 1993:01 to 2017:12.

The first five principal components is correlated out of the monthly set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ from 1993 to 2017. We see that the first principal component (level) is negatively correlated with the spanning factor (-0.7549) and almost zero correlation with the unspanned factor(s). The PC1 (level) has a similar correlation with the Cochrane-Piazzesi factor (-0.717), and therefore an irrelevant correlation with the Ludvingson-Ng factor (0.0492) as well. The following two principal components have a small correlation with the spanned factor (0.139 and -0.0498, respectively). However, we see a somewhat interesting positive correlation (0.428) between the curvature component and the unspanned factor. The takeaway from table 1.6 is that the spanning factor captures negatively the slope of the yield curve, and the unspanning factor has some positive linear association with the curvature of the term-structure.

Additionally, we project our spanning and unspanning factors in an increasing set of first principal components (up to PC5). The R-squared of these regressions are reported in table 1.7. The regressions are computed in such a way that the first column uses PC1 as the independent variable, the second column PC1 and PC2, and so on. We see that our single spanning factor is associated with a larger set of principal components. In column (III), where the R-squared of the regressions using level, slope and curvature components

Table 1.7: R-Squared of Projections - Principal Components of the Term-Structure

		(I)	(II)	(III)	(IV)	(V)
		PC1	(I) + PC2	(II) + PC3	(III) + PC4	(IV) + PC5
Spanning Factor	$(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}})_t^h$	0.5699	0.5892	0.5917	0.6201	0.6255
Unspanning Factors	$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^h$	0.0005	0.0348	0.2177	0.2187	0.2269
	$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-2),h}$	0.0028	0.0451	0.2346	0.2386	0.2477
	$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-3),h}$	0.0203	0.0844	0.3184	0.3350	0.3432
	$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-4),h}$	0.0050	0.0426	0.2722	0.2740	0.2795
	$\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^{(-5),h}$	0.0004	0.0169	0.2005	0.2021	0.2066
Other Factors	$\widehat{CP}_t^h$	0.5140	0.5924	0.6741	0.8102	0.8119
	$\widehat{LN}_t^h$	0.0024	0.0261	0.1014	0.1025	0.1031
Derived $\mathfrak{f}_{t,DNN}^{(n),h}$	$\mathfrak{f}_{t,DNN}^{(2),h}$	0.5443	0.5449	0.5467	0.5709	0.5725
	$\mathfrak{f}_{t,DNN}^{(3),h}$	0.2166	0.2363	0.2643	0.2821	0.2833
	$\mathfrak{f}_{t,DNN}^{(4),h}$	0.0157	0.0180	0.1099	0.1264	0.1279
	$\mathfrak{f}_{t,DNN}^{(5),h}$	0.0165	0.0190	0.1593	0.1641	0.1791

Table 1.7 reports the R-squared of the projection of our spanning and unspanning factor(s) in an increasing set of the first five principal components: PC1 (level), PC2 (slope), PC3 (curvature), PC4, and PC5. The principal components are computed from the the monthly set of yields $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$. We also compute the regressions for the Cochrane-Piazzesi and Ludvingson-Ng factors, and for the sake of completeness with the latent DNN factors. In column **(I)**, we regress the factors on PC1, in column **(II)**, we regress the factors on PC1 and PC2, in column **(III)**, we regress the factor on PC1, PC2, and PC3, in column **(IV)**, we regress the factors on PC1, PC2, PC3, and PC4, and finally in column **(V)**, we regress the factors on PC1, PC2, PC3, PC4, and PC5. The period of analysis ranges from 1993:01 to 2017:12.

are used, we see an R-squared of 0.592 for the spanning factor, 0.218 for the unspanning factor $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_t^h$. For the first three principal component, we see a similar pattern with the Cochrane-Piazzesi factor, with a little higher values; and it is expected given the nature of the derivation of the Ludvingson-Ng factor, the R-squared remain low for all cases.

1.4.6 A Robustness Check

Recent studies in the forecasting literature raised the issue that defining the sample split may be data-mined (Hansen and Timmermann, 2012; Kelly and Pruitt, 2013; Rossi and Inoue, 2012). As a robustness check, we seek to know if the results reported of the statistical significance of our state factors could be a sample-specific fact. To demonstrate the robustness of our estimates to alternative sample splits, we re-run the same regressions from table 1.1 restricting the series up to the last month of each year from 1994 up to the last year of analysis, 2017.

Figure 1.11 reports the coefficients estimates of $\left(\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}_t\right)_t^h$, $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_{t+h/12}^{(-n),h}$ and $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})_{t+h/12}^{h}$. In a recursive approach we seek to show how the estimates of the pa-

Figure 1.11

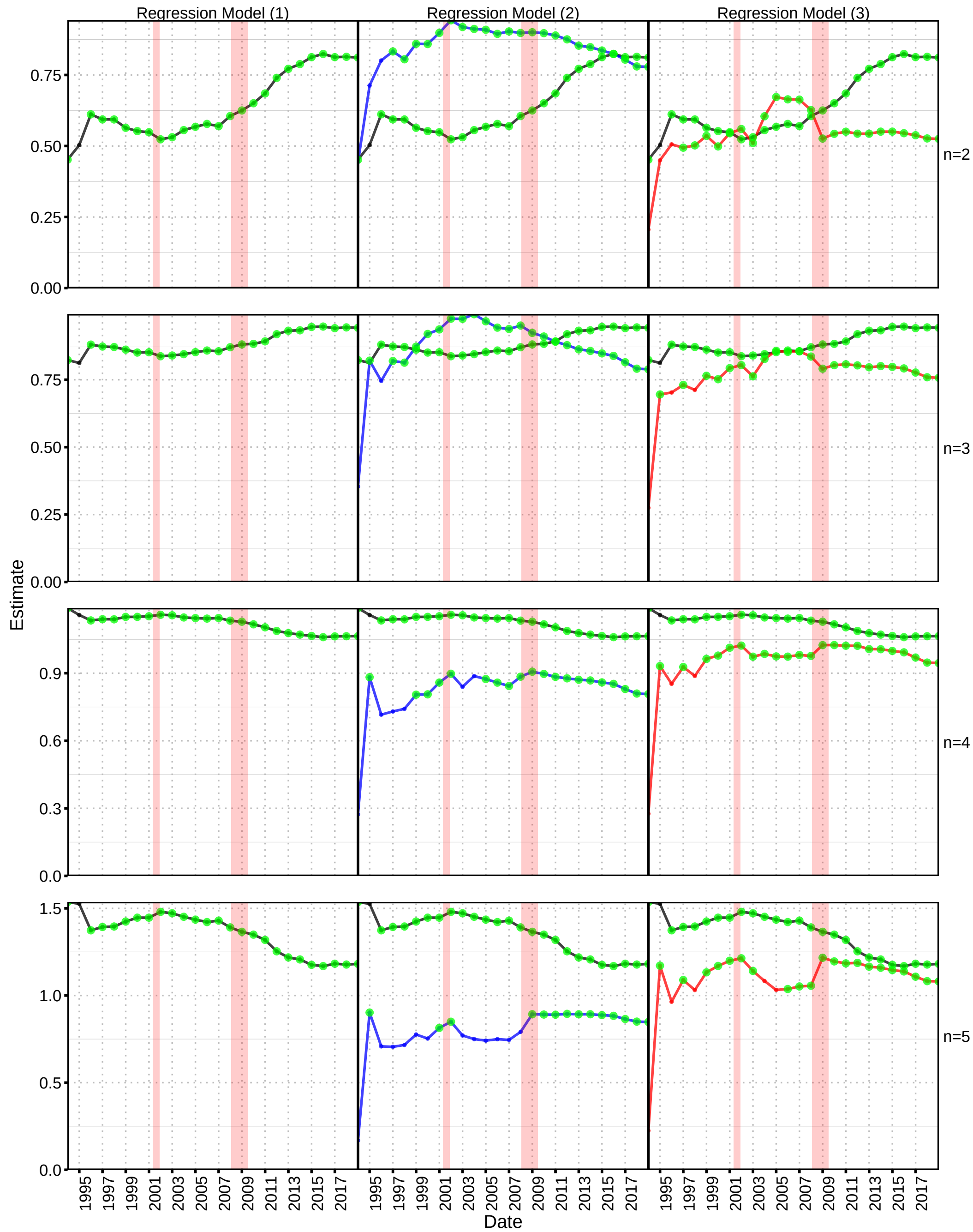

Figure 1.11 reports the estimates of each one of the three regressions models reported similiar to those summarized in table 1.1, where the x-axis defines the end of the sample. All samples start in 1993 : 01. Regression Model (1) is given by $rx^{(n)}_{t+h/12} = \beta_0 + \beta_1(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t)^h_t + \epsilon_{t+h/12}$. Regression Model (2) is given by $rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_{t+h/12} + \epsilon_{t+h/12}$. Regression Model (3) is given by $rx^{(n)}_{t+h/12} = \beta_0 + \beta_1 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{h}_{t+h/12} + \epsilon_{t+h/12}$. The black line represents the estimates of $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t$. The blue line represents the estimates of $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{(-n),h}_{t+h/12}$. Finally, the red represents the estimates of $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^{h}_{t+h/12}$. The figure is split in four panels, each panel representing one maturity. Statistically significant coefficients are presented as green points.

Figure 1.12: Regression Coefficients of $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$ Over Time as a Function of Maturity (n)

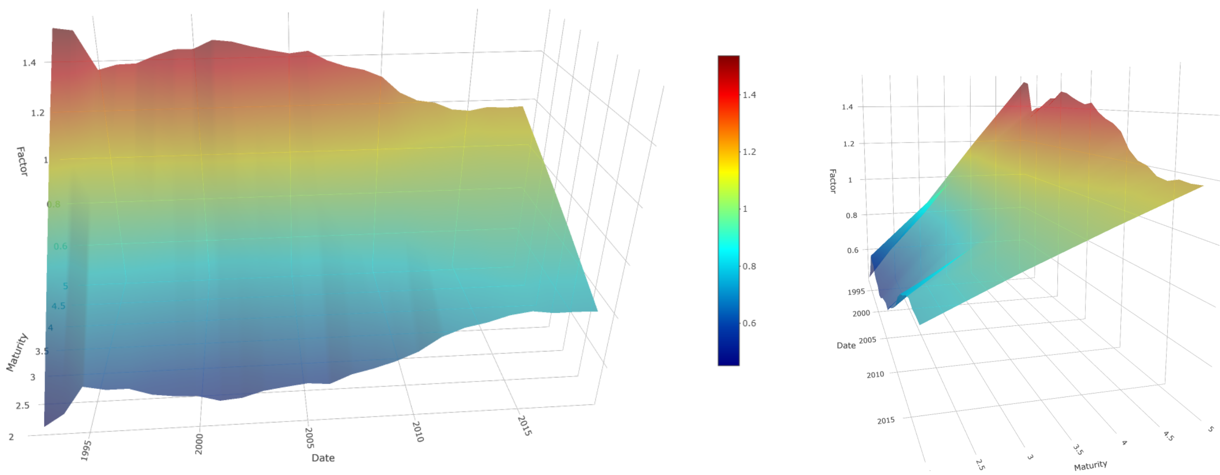

Figure 1.12 plots the behavior of our spanning factor $\left(\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}_t\right)_t^h$ as a function of maturity (n) over the period of analysis (1993-2017).

Figure 1.13: Regression Coefficients of $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-n),h}$ Over Time as a Function of Maturity (n)

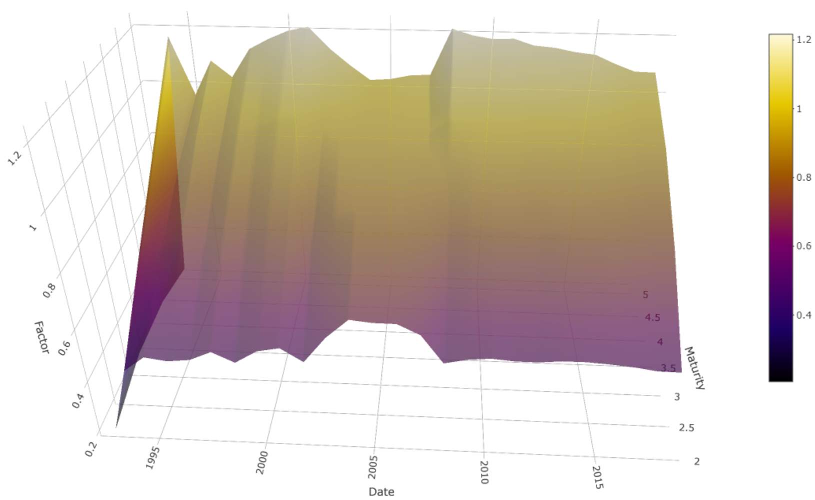

Figure 1.13 plots the behavior of our unspanned factor $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-n),h}$ as a function of maturity (n) over the period of analysis (1993-2017).

rameters varies across an expanding sample size and the statistical significance as well. The figure has four panels, each panel representing one maturity. Statistically significant coefficients are presented as green points. Clearly we see that, despite the initial variation in the estimates for the first years, what is expected given the limited sample size, the (i) estimates do not behave erratically with abrupt variations, and (ii) the vast majority of the estimates for each year from 1994 to 2017 is statistically significant.

In figure 1.12 we plot the estimates obtained in these regressions ranging from 1994 to 2017 across all maturities. The figure shows a clear pattern for the estimates of $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ is increasing in the maturity (n). Notice that the estimates of this increasing line shifted during both recessions in the period of analysis. Another pattern that can be inferred from this figure is that over time the difference between longer maturities and shorter shrunk over the period 1993 to 2017, what can be seen as the curve of $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h$ becoming more flat over time.

Similarly, figure 1.12 plots the regression coefficients of $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-n),h}$ as a function of Maturity (n), and figure 1.15 in Appendix 1.6.1 plots the regression coefficients of $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{h}$. The pattern mentioned above maintains for the unspanned factor $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{(-n),h}$. However, for $\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}}(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_{t+h/12}^{h}$ the curve as a function of maturities are much flatter when compared to the other state variables, and analogously to the Cochrane-Piazzesi factor, it has a more clear tent shape. This format becomes more evident during the recessions, when mid levels of maturity have the highest value for this factor, while low and high level of maturities are smaller.

1.4.7 GMM/GEL Estimation

As an additional robustness check, in a similar fashion to Lee (2018), we can use of GMM to estimate jointly the parameters of our state variables and obtain better standard errors estimates for the inference of our parameters. This is especially important as we have some generated regressors in our analysis, and that the dependent variables $(rx_{t+h}^{(n)})$ have clear cross-sectional correlations among them. The states variables are obtained as:

$$\overline{rx}_t = \kappa_0 + \kappa_1 \xi_t^{(2)} + \kappa_2 \xi_t^{(3)} + \kappa_3 \xi_t^{(4)} + \kappa_4 \xi_t^{(5)} + \overline{u}_t \tag{1.30}$$

$$(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^h = \delta_0 + \delta_1 \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)_t^h + e_t \tag{1.31}$$

where $\overline{rx}_t = \sum_{n=2}^5 rx_t^{(n)}/4$ and $(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})_t^h = \widehat{\kappa}_{1,t}\xi_t^{(2)} + \widehat{\kappa}_{2,t}\xi_t^{(3)} + \widehat{\kappa}_{3,t}\xi_t^{(4)} + \widehat{\kappa}_{4,t}\xi_t^{(5)}$, which is obtained in equation (1.30). Then, in the second stage for the risk premium forecasts we run:

$$rx^{(n)}_{t+h/12} = \beta^{(n)}_0 + \beta^{(n)}_1 \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t + \beta^{(n)}_2 \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t + \epsilon^{(n)}_{t+h/12} \tag{1.32}$$

for each of the four maturities $n \in \{2, 3, 4, 5\}$. We can define the vector of moments of our GMM as below:

$$g_T(\boldsymbol{\theta}) = \begin{bmatrix} \overline{u}_t \otimes \begin{pmatrix} 1 & \xi^{(2)}_t & \xi^{(3)}_t & \xi^{(4)}_t & \xi^{(5)}_t \end{pmatrix} \\ e_t \otimes \begin{pmatrix} 1 & \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t \end{pmatrix} \\ \epsilon^{(2)}_{t+1} \otimes \begin{pmatrix} 1 & \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t & \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t \end{pmatrix} \\ \epsilon^{(3)}_{t+1} \otimes \begin{pmatrix} 1 & \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t & \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t \end{pmatrix} \\ \epsilon^{(4)}_{t+1} \otimes \begin{pmatrix} 1 & \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t & \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t \end{pmatrix} \\ \epsilon^{(5)}_{t+1} \otimes \begin{pmatrix} 1 & \left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t & \boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t \end{pmatrix} \end{bmatrix} \tag{1.33}$$

where $\boldsymbol{\theta}$ is the 19×1 vector of all the parameters as in
$\boldsymbol{\theta} = \begin{bmatrix} \kappa_0 & \kappa_1 & \kappa_2 & \kappa_3 & \kappa_4 & \delta_0 & \delta_1 & \beta^{(2)}_0 & \beta^{(2)}_1 & \beta^{(2)}_2 \dots & \beta^{(5)}_0 & \beta^{(5)}_1 & \beta^{(5)}_2 \end{bmatrix}^\top$.

Table 1.8: GMM and GEL estimations of the spanning factor $\left(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}_t\right)^h_t$ and the unspanned factor $\left(\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}}\right)^{(h}_t$

	Variable	Estimate	se(GMM)	se(GEL)	se(OLS)	GMM t-stat	GEL t-stat	OLS t-stat
$rx^{(2)}_{t+h/12}$	$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t$	0.811	0.130	0.141	0.110	6.259***	5.741***	7.355***
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t$	0.525	0.080	0.087	0.124	6.579***	6.002***	4.228***
$rx^{(3)}_{t+h/12}$	$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t$	0.943	0.196	0.219	0.173	4.802***	4.307***	5.447***
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t$	0.757	0.075	0.080	0.202	10.089***	9.445***	3.743***
$rx^{(4)}_{t+h/12}$	$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t$	1.065	0.260	0.294	0.236	4.097***	3.624***	4.519***
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t$	0.945	0.119	0.124	0.282	7.972***	7.62***	3.355***
$rx^{(5)}_{t+h/12}$	$(\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}})^h_t$	1.181	0.320	0.364	0.296	3.695***	3.248**	3.99***
	$\boldsymbol{M}_{\boldsymbol{\tau}^\top \widehat{\boldsymbol{\mathfrak{F}}}} (\boldsymbol{\kappa}^\top \widehat{\boldsymbol{\xi}})^h_t$	1.081	0.194	0.216	0.359	5.558***	5.006***	3.011**

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.8 reports the the standard errors and t-statistics from the Generalized Empirical Likelihood (GEL), Generalized Empirical Likelihood (GEL), and Ordinary Least Squares (OLS). The vector of moments for the GMM and GEL is given in equation (1.33). The standard errors for GMM and OLS use HAC variance-covariance matrix, as in Newey and West (1987).

Table 1.8 presents the standard error estimates using OLS and GMM for the risk premium forecasts parameters as in equation 1.32, for the four maturities considered. The takeaway from this table is that even though that the standard errors from the GMM are a little bit larger when compared to the ones generated by OLS, the inference over the interested parameters keeps still being strongly significant. We control the auto-correlation with a HAC variance-covariance matrix for both methods, as in Newey and West (1987).

To corroborate the robustness of our generated regressors, we also make of use of the Generalized Empirical Likelihood (GEL). As Newey and Smith (2004) and Anatolyev (2005) had shown, GEL has a significant advantage over the GMM estimation, since the bias of the latter does not increase with the number of moment conditions, what does not necessarily holds for GMM. Thus, the efficiency improves when the number of conditions goes up. We can use the same vector of moments as in equation (1.33) to estimate the risk premium forecasts parameters. Table 1.8 also reports the standard errors and the t-statistics for the same regressions. We see that the standard errors slightly larger than GMM and OLS ones, but still strongly statistically significant. In short, these results endorse the robustness of our factors in the forecasting the risk premium of the Treasury bonds.

1.5 Conclusion

In this paper we proposed a novel approach for deriving a single state factor consistent with a dynamic term-structure with unspanned risks. We make use of deep neural networks to uncover relationships in the full set of information from the yield curve. This allows us through an approximation to to derive a single state variable factor that spans the space of all the information from the term-structure. We also introduced a way to obtain unspanned risks from the yield curve that is used to complete our state space.

We show that this parsimonious number of state variables have predictive power for excess returns of bonds over 1-month holding period. Additionally, we provide an intuitive interpretation of derived factors, and show what information from macroeconomic variables and sentiment-based measures they can capture.

1.6 Appendix

1.6.1 Data

Figure 1.14: 12-Month Bonds Excess Returns

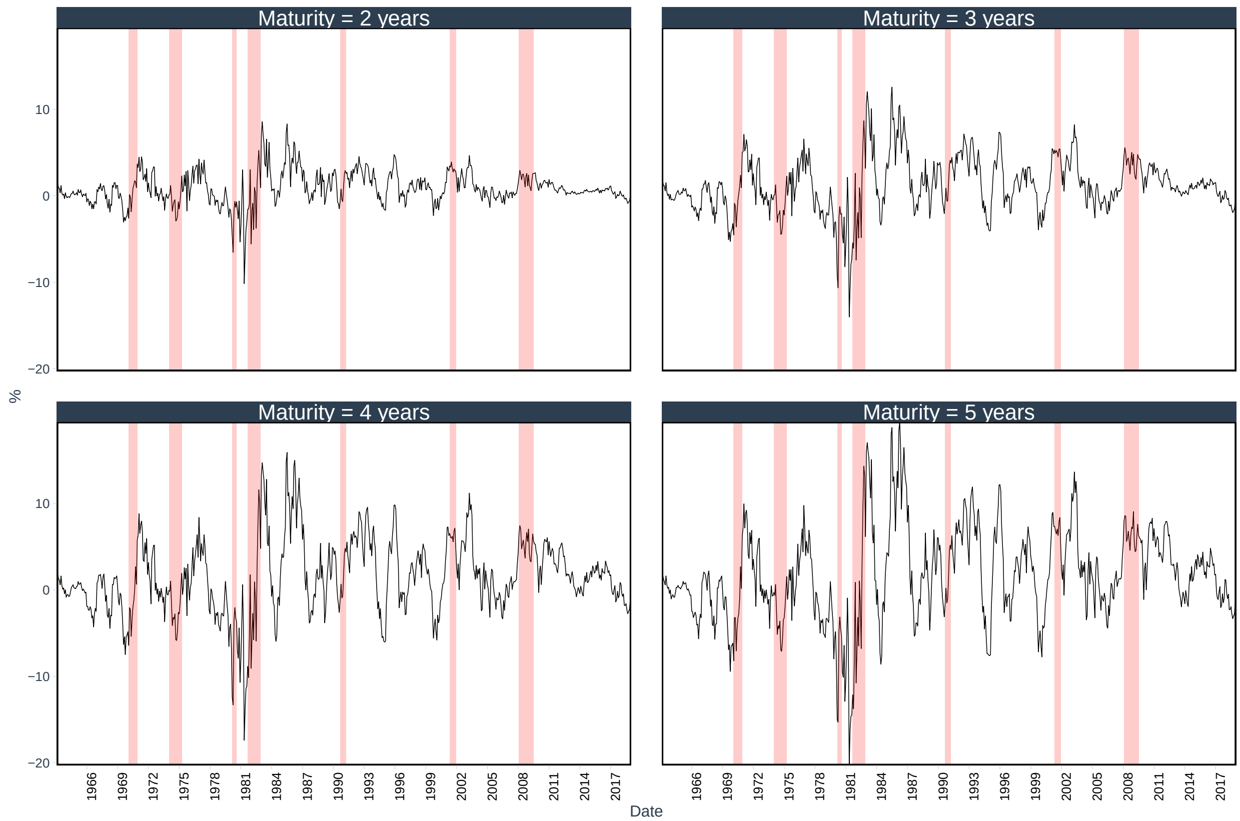

Figure 1.14 shows the 12-month excess returns for maturities with $n = 2, 3, 4$ and 5 years. The excess returns are calculeted as in equation (1.4), i.e., $rx_{t+1}^{(n)} = ny_t^{(n)} - (n+1)y_{t+1}^{(n-1)} - y_t^n$. Each panel represents one of the four maturities. The y-axis shows values in percentage (%). NBER-classified recessions are shaded in light red.

Table 1.9: Descriptive Statistics - DNN Factor $\mathfrak{f}_{t,DNN}^{(n),h}$ by MLP Architecture and Choice of $\boldsymbol{Z}_t^{\mathrm{y}}$

Panel A: $\boldsymbol{Z}_t$: set of forward rates

	$\mathfrak{f}_{t,DNN}^{(2),h}$			$\mathfrak{f}_{t,DNN}^{(3),h}$			$\mathfrak{f}_{t,DNN}^{(4),h}$			$\mathfrak{f}_{t,DNN}^{(5),h}$		
Model	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$
DNN 1	0.3300	0.0368	0.8099	0.3361	0.0458	0.8372	0.3377	0.0639	0.7711	0.2874	0.1114	0.7993
DNN 2	0.3360	0.0364	0.8964	0.3369	0.0342	0.8780	0.3463	0.0563	0.8708	0.3214	0.0772	0.8462
DNN 3	0.4274	0.0225	0.8092	0.4270	0.0237	0.7397	0.4199	0.0341	0.8338	0.4054	0.0413	0.8563

Panel B: $\boldsymbol{Z}_t$: set of yields

	$\mathfrak{f}_{t,DNN}^{(2),h}$			$\mathfrak{f}_{t,DNN}^{(3),h}$			$\mathfrak{f}_{t,DNN}^{(4),h}$			$\mathfrak{f}_{t,DNN}^{(5),h}$		
Model	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$
DNN 1	0.3441	0.0373	0.8106	0.3359	0.0393	0.8324	0.3308	0.0534	0.7812	0.2781	0.1122	0.9090
DNN 2	0.3370	0.0360	0.8821	0.3379	0.0327	0.8508	0.3351	0.0432	0.8066	0.3288	0.0628	0.8350
DNN 3	0.4363	0.0204	0.7833	0.4335	0.0226	0.7264	0.4298	0.0305	0.8150	0.4226	0.0403	0.8411

Panel C: $\boldsymbol{Z}_t$: set of yields and forward rates

	$\mathfrak{f}_{t,DNN}^{(2),h}$			$\mathfrak{f}_{t,DNN}^{(3),h}$			$\mathfrak{f}_{t,DNN}^{(4),h}$			$\mathfrak{f}_{t,DNN}^{(5),h}$		
Model	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$
DNN 1	0.3421	0.0512	0.6709	0.3313	0.0690	0.6546	0.3088	0.0969	0.7587	0.2642	0.1110	0.6448
DNN 2	0.3418	0.0356	0.8732	0.3421	0.0340	0.7917	0.3434	0.0449	0.8122	0.3429	0.0647	0.8116
DNN 3	0.4375	0.0202	0.7627	0.4321	0.0232	0.7543	0.4246	0.0310	0.8092	0.4190	0.0381	0.8319

Table 1.9 reports the mean, standard deviation and the first autocorrelation ($\hat{\rho}_1$) of the derived $\mathfrak{f}_{t,DNN}^{(n),h}$ for each scenario under consideration. Each panel considers a different set of information from the term-structure to derive the factor $\mathfrak{f}_{t,DNN}^{(n),h}$. Panel A shows the the derived DNN factors for $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60)}\right\}$, Panel B for $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ and Panel C for $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60/12)}, y_t^{(1)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$. The descriptive statistics is computed for for each group of maturity $n \in \{1, 2, 3, 4\}$. Period of analysis ranges from 1993:01 to 2017:12.

Table 1.10: Descriptive Statistics - $\xi_t^{(n),h}$ by MLP Architecture and Choice of $\boldsymbol{Z}_t^{\mathrm{y}}$

Panel A: $\boldsymbol{Z}_t$: set of forward rates

	$\xi_t^{(2),h}$			$\xi_t^{(3),h}$			$\xi_t^{(4),h}$			$\xi_t^{(5),h}$		
Model	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$
DNN 1	-0.0311	0.4799	0.2417	0.0187	0.7713	0.1718	0.0610	1.0581	0.1324	0.1558	1.3409	0.1178
DNN 2	-0.0265	0.4797	0.2427	0.0145	0.7716	0.1645	0.0550	1.0651	0.1323	0.1240	1.3424	0.1099
DNN 3	-0.0394	0.4796	0.2428	-0.0015	0.7729	0.1636	0.0402	1.0630	0.1261	0.0898	1.3419	0.1036

Panel B: $\boldsymbol{Z}_t$: set of yields

	$\xi_t^{(2),h}$			$\xi_t^{(3),h}$			$\xi_t^{(4),h}$			$\xi_t^{(5),h}$		
Model	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$
DNN 1	-0.0344	0.4810	0.2462	0.0144	0.7705	0.1625	0.0701	1.0591	0.1263	0.1804	1.3347	0.1059
DNN 2	-0.0254	0.4793	0.2422	0.0159	0.7721	0.1657	0.0646	1.0600	0.1279	0.1135	1.3410	0.1077
DNN 3	-0.0685	0.4952	0.2903	-0.0032	0.7724	0.1640	0.0366	1.0636	0.1276	0.0874	1.3429	0.1062

Panel C: $\boldsymbol{Z}_t$: set of yields and forward rates

	$\xi_t^{(2),h}$			$\xi_t^{(3),h}$			$\xi_t^{(4),h}$			$\xi_t^{(5),h}$		
Model	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$	**Mean**	**sd**	$\hat{\rho}_1$
DNN 1	-0.0240	0.4760	0.2492	0.0336	0.7683	0.1747	0.0999	1.0560	0.1423	0.1932	1.3370	0.1004
DNN 2	-0.0269	0.4795	0.2389	0.0146	0.7711	0.1650	0.0590	1.0612	0.1296	0.1051	1.3428	0.1084
DNN 3	-0.0622	0.4956	0.2867	0.0012	0.7728	0.1603	0.0450	1.0620	0.1221	0.0872	1.3419	0.1030

Table 1.9 reports the mean, standard deviation and the first autocorrelation ($\hat{\rho}_1$) of the $\xi_t^{(n),h}$ for each scenario under consideration. Each panel considers a different set of information from the term-structure to derive the factor $\mathfrak{f}_{t,DNN}^{(n),h}$. Panel A shows the the derived DNN factors for $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60)}\right\}$, Panel B for $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{y_t^{(1/12)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$ and Panel C for $\boldsymbol{Z}_t^{\mathrm{y}} = \left\{f_t^{(2/12)}, f_t^{(3/12)}, \ldots, f_t^{(60/12)}, y_t^{(1)}, y_t^{(2/12)}, \ldots, y_t^{(60/12)}\right\}$. The descriptive statistics is computed for for each group of maturity $n \in \{1, 2, 3, 4\}$. Period of analysis ranges from 1993:01 to 2017:12.

Figure 1.15: Regression Coefficients of $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})^{h}_{t+h/12}$ Over Time as a Function of Maturity (n)

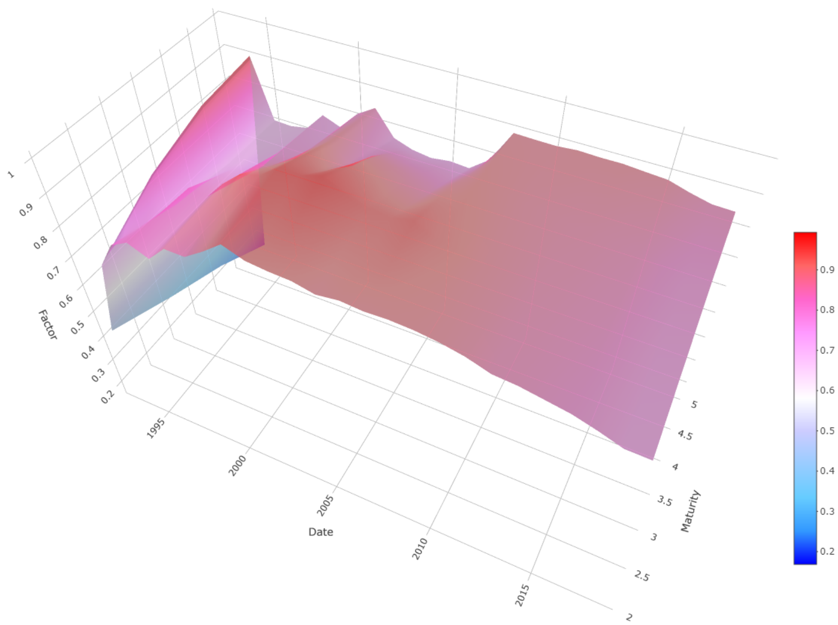

Figure 1.15 plots the behavior of our unspanned factor $\boldsymbol{M}_{\boldsymbol{\tau}^{\top}\widehat{\mathfrak{F}}}(\boldsymbol{\kappa}^{\top}\widehat{\boldsymbol{\xi}})^{h}_{t+h/12}$ as a function of maturity (n) over the period of analysis (1993-2017).

Table 1.11: FRED-MD

Group	FRED Code	Description
Output and income	RPI	Real Personal Income
	W875RX1	Real personal income ex transfer receipts
	INDPRO	IP Index
	IPFPNSS	IP: Final Products and Nonindustrial Supplies
	IPFINAL	IP: Final Products (Market Group)
	IPCONGD	IP: Consumer Goods
	IPDCONGD	IP: Durable Consumer Goods
	IPNCONGD	IP: Nondurable Consumer Goods
	IPBUSEQ	IP: Business Equipment
	IPMAT	IP: Materials
	IPDMAT	IP: Durable Materials
	IPNMAT	IP: Nondurable Materials
	IPMANSICS	IP: Manufacturing (SIC)
	IPB51222s	IP: Residential Utilities
	IPFUELS	IP: Fuels
	CUMFNS	Capacity Utilization: Manufacturing
Labor market	HWI	Help-Wanted Index for United States
	HWIURATIO	Ratio of Help Wanted/No. Unemployed
	CLF16OV	Civilian Labor Force
	CE16OV	Civilian Employment
	UNRATE	Civilian Unemployment Rate
	UEMPMEAN	Average Duration of Unemployment (Weeks)
	UEMPLT5	Civilians Unemployed - Less Than 5 Weeks
	UEMP5TO14	Civilians Unemployed for 43599 Weeks
	UEMP15OV	Civilians Unemployed - 15 Weeks & Over
	UEMP15T26	Civilians Unemployed for 15-26 Weeks
	UEMP27OV	Civilians Unemployed for 27 Weeks and Over
	CLAIMSx	Initial Claims
	PAYEMS	All Employees: Total Nonfarm
	USGOOD	All Employees: Goods-Producing Industries
	CES1021000001	All Employees: Mining and Logging: Mining
	USCONS	All Employees: Construction
	MANEMP	All Employees: Manufacturing
	DMANEMP	All Employees: Durable goods
	NDMANEMP	All Employees: Nondurable goods
	SRVPRD	All Employees: Service-Providing Industries
	USTPU	All Employees: Trade, Transportation & Utilities
	USWTRADE	All Employees: Wholesale Trade
	USTRADE	All Employees: Retail Trade
	USFIRE	All Employees: Financial Activities
	USGOVT	All Employees: Government
	CES0600000007	Avg Weekly Hours : Goods-Producing
	AWOTMAN	Avg Weekly Overtime Hours : Manufacturing
	AWHMAN	Avg Weekly Hours : Manufacturing
	CES0600000008	Avg Hourly Earnings : Goods-Producing
	CES2000000008	Avg Hourly Earnings : Construction
	CES3000000008	Avg Hourly Earnings : Manufacturing
Housing	HOUST	Housing Starts: Total New Privately Owned
	HOUSTNE	Housing Starts, Northeast
	HOUSTMW	Housing Starts, Midwest
	HOUSTS	Housing Starts, South
	HOUSTW	Housing Starts, West
	PERMIT	New Private Housing Permits (SAAR)
	PERMITNE	New Private Housing Permits, Northeast (SAAR)
	PERMITMW	New Private Housing Permits, Midwest (SAAR)
	PERMITS	New Private Housing Permits, South (SAAR)
	PERMITW	New Private Housing Permits, West (SAAR)
Consumption, orders, and inventories	DPCERA3M086SBEA	Real Personal Consumption Expenditures
	CMRMTSPLx	Real Manu. and Trade Industries Sales
	RETAILx	Retail and Food Services Sales
	ACOGNO	New Orders for Consumer Goods
	AMDMNOx	New Orders for Durable Goods
	ANDENOx	New Orders for Nondefense Capital Goods
	AMDMUOx	Unfilled Orders for Durable Goods
	BUSINVx	Total Business Inventories
	ISRATIOx	Total Business: Inventories to Sales Ratio
	UMCSENTx	Consumer Sentiment Index
Money and credit	M1SL	M1 Money Stock
	M2SL	M2 Money Stock
	M2REAL	Real M2 Money Stock
	AMBSL	St. Louis Adjusted Monetary Base
	TOTRESNS	Total Reserves of Depository Institutions
	NONBORRES	Reserves Of Depository Institutions
	BUSLOANS	Commercial and Industrial Loans
	REALLN	Real Estate Loans at All Commercial Banks
	NONREVSL	Total Nonrevolving Credit
	CONSPI	Nonrevolving Consumer Credit to Personal Income
	MZMSL	MZM Money Stock
	DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding
	DTCTHFNM	Total Consumer Loans and Leases Outstanding
	INVEST	Securities in Bank Credit at All Commercial Banks
Interest and exchange rates	FEDFUNDS	Effective Federal Funds Rate
	CP3Mx	3-Month AA Financial Commercial Paper Rate
	TB3MS	3-Month Treasury Bill
	TB6MS	6-Month Treasury Bill
	GS1	1-Year Treasury Rate
	GS5	5-Year Treasury Rate
	GS10	10-Year Treasury Rate
	AAA	Moody's Seasoned Aaa Corporate Bond
	BAA	Moody's Seasoned Baa Corporate Bond
	COMPAPFFx	3-Month Commercial Paper Minus
	TB3SMFFM	3-Month Treasury C Minus FEDFUNDS
	TB6SMFFM	6-Month Treasury C Minus FEDFUNDS
	T1YFFM	1-Year Treasury C Minus FEDFUNDS
	T5YFFM	5-Year Treasury C Minus FEDFUNDS
	T10YFFM	10-Year Treasury C Minus FEDFUNDS
	AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS
	BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS
	TWEXMMTH	Trade Weighted U.S. Dollar Index: Major Currencies
	EXSZUSx	Switzerland / U.S. Foreign Exchange Rate
	EXJPUSx	Japan / U.S. Foreign Exchange Rate
	EXUSUKx	U.S. / U.K. Foreign Exchange Rate
	EXCAUSx	Canada / U.S. Foreign Exchange Rate
Prices	WPSFD49207	PPI: Finished Goods
	WPSFD49502	PPI: Finished Consumer Goods
	WPSID61	PPI: Intermediate Materials
	WPSID62	PPI: Crude Materials
	OILPRICEx	Crude Oil, Spliced WTI and Cushing
	PPICMM	PPI: Metals and Metal Products:
	CPIAUCSL	CPI : All Items
	CPIAPPSL	CPI : Apparel
	CPITRNSL	CPI : Transportation
	CPIMEDSL	CPI : Medical Care
	CUSR0000SAC	CPI : Commodities
	CUSR0000SAD	CPI : Durables
	CUSR0000SAS	CPI : Services
	CPIULFSL	CPI : All Items Less Food
	CUSR0000SA0L2	CPI : All Items Less Shelter
	CUSR0000SA0L5	CPI : All Items Less Medical Care
	PCEPI	Personal Cons. Expend.: Chain Index
	DDURRG3M086SBEA	Personal Cons. Exp: Durable goods
	DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods
	DSERRG3M086SBEA	Personal Cons. Exp: Services
Stock market	S&P 500	S&P's Common Stock Price Index: Composite
	S&P div yield	S&P's Composite Common Stock: Dividend Yield
	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio
	VXOCLSx	VXO

Table 1.11 lists the 128 macroeconomic and financial variables from the FRED-MD dataset. The table reports the group, FRED code and a description of each variable. The variables are split in one of the 8 groups: (1) output and income, (2) labor market, (3) housing, (4) consumption, orders, and inventories, (5) money and credit, (6) interest and exchange rates, (7) prices, and (8) stock market.

Table 1.12: Descriptive Statistics - FRED-MD

				Full Sample		1962-1992		1993-2017	
Fred Code	**tcode**	**Group**	**Description**	**Mean**	**sd**	**Mean**	**sd**	**Mean**	**sd**
RPI	5	Output and income	Real Personal Income	0.0030	0.0060	0.0030	0.0040	0.0020	0.0070
W875RX1	5	Output and income	Real personal income ex transfer receipts	0.0020	0.0060	0.0030	0.0040	0.0020	0.0070
INDPRO	5	Output and income	IP Index	0.0020	0.0070	0.0030	0.0080	0.0020	0.0060
IPFPNSS	5	Output and income	IP: Final Products and Nonindustrial Supplies	0.0020	0.0070	0.0030	0.0080	0.0010	0.0060
IPFINAL	5	Output and income	IP: Final Products (Market Group)	0.0020	0.0080	0.0030	0.0080	0.0010	0.0070
IPCONGD	5	Output and income	IP: Consumer Goods	0.0020	0.0090	0.0020	0.0090	0.0010	0.0070
IPDCONGD	5	Output and income	IP: Durable Consumer Goods	0.0020	0.0210	0.0030	0.0220	0.0020	0.0190
IPNCONGD	5	Output and income	IP: Nondurable Consumer Goods	0.0010	0.0070	0.0020	0.0080	0.0000	0.0070
IPBUSEQ	5	Output and income	IP: Business Equipment	0.0040	0.0120	0.0040	0.0120	0.0030	0.0120
IPMAT	5	Output and income	IP: Materials	0.0020	0.0090	0.0020	0.0100	0.0020	0.0080
IPDMAT	5	Output and income	IP: Durable Materials	0.0030	0.0130	0.0030	0.0150	0.0030	0.0110
IPNMAT	5	Output and income	IP: Nondurable Materials	0.0020	0.0110	0.0030	0.0100	0.0000	0.0110
IPMANSICS	5	Output and income	IP: Manufacturing (SIC)	0.0020	0.0080	0.0030	0.0090	0.0020	0.0070
IPB51222s	5	Output and income	IP: Residential Utilities	0.0020	0.0350	0.0030	0.0280	0.0010	0.0420
IPFUELS	5	Output and income	IP: Fuels	0.0010	0.0190	0.0010	0.0200	0.0010	0.0180
CUMFNS	2	Output and income	Capacity Utilization: Manufacturing	-0.0060	0.6230	-0.0050	0.7110	-0.0080	0.5000
HWI	2	Labor market	Help-Wanted Index for United States	8.2700	169.6090	3.1960	106.1450	14.3210	222.8650
HWIURATIO	2	Labor market	Ratio of Help Wanted/No. Unemployed	0.0010	0.0330	0.0000	0.0320	0.0030	0.0340
CLF16OV	5	Labor market	Civilian Labor Force	0.0010	0.0030	0.0020	0.0030	0.0010	0.0020
CE16OV	5	Labor market	Civilian Employment	0.0010	0.0030	0.0020	0.0030	0.0010	0.0020
UNRATE	2	Labor market	Civilian Unemployment Rate	-0.0030	0.1730	0.0040	0.1880	-0.0110	0.1550
UEMPMEAN	2	Labor market	Average Duration of Unemployment (Weeks)	0.0090	0.6000	0.0090	0.4740	0.0090	0.7230
UEMPLT5	5	Labor market	Civilians Unemployed - Less Than 5 Weeks	0.0000	0.0550	0.0020	0.0480	-0.0010	0.0610
UEMP5TO14	5	Labor market	Civilians Unemployed for 43599 Weeks	0.0010	0.0540	0.0020	0.0560	-0.0010	0.0520
UEMP15OV	5	Labor market	Civilians Unemployed - 15 Weeks & Over	0.0010	0.0510	0.0030	0.0550	-0.0020	0.0440
UEMP15T26	5	Labor market	Civilians Unemployed for 15-26 Weeks	0.0010	0.0750	0.0030	0.0780	-0.0020	0.0710
UEMP27OV	5	Labor market	Civilians Unemployed for 27 Weeks and Over	0.0010	0.0700	0.0030	0.0780	-0.0010	0.0580
CLAIMSx	5	Labor market	Initial Claims	0.0000	0.0480	0.0000	0.0530	-0.0010	0.0400
PAYEMS	5	Labor market	All Employees: Total nonfarm	0.0010	0.0020	0.0020	0.0020	0.0010	0.0020
USGOOD	5	Labor market	All Employees: Goods-Producing Industries	0.0000	0.0040	0.0000	0.0050	0.0000	0.0040
CES1021000001	5	Labor market	All Employees: Mining and Logging: Mining	0.0000	0.0180	0.0000	0.0230	0.0010	0.0090
USCONS	5	Labor market	All Employees: Construction	0.0010	0.0090	0.0010	0.0100	0.0020	0.0060
MANEMP	5	Labor market	All Employees: Manufacturing	0.0000	0.0040	0.0000	0.0050	-0.0010	0.0030
DMANEMP	5	Labor market	All Employees: Durable goods	0.0000	0.0060	0.0000	0.0070	-0.0010	0.0040
NDMANEMP	5	Labor market	All Employees: Nondurable goods	0.0000	0.0030	0.0000	0.0030	-0.0010	0.0020
SRVPRD	5	Labor market	All Employees: Service-Providing Industries	0.0020	0.0020	0.0020	0.0020	0.0010	0.0010
USTPU	5	Labor market	All Employees: Trade, Transportation & Utilities	0.0010	0.0020	0.0020	0.0020	0.0010	0.0020
USWTRADE	5	Labor market	All Employees: Wholesale Trade	0.0010	0.0020	0.0020	0.0020	0.0010	0.0020
USTRADE	5	Labor market	All Employees: Retail Trade	0.0020	0.0030	0.0020	0.0030	0.0010	0.0020
USFIRE	5	Labor market	All Employees: Financial Activities	0.0020	0.0020	0.0020	0.0020	0.0010	0.0020
USGOVT	5	Labor market	All Employees: Government	0.0010	0.0030	0.0020	0.0030	0.0010	0.0020
CES0600000007	1	Labor market	Avg Weekly Hours : Goods-Producing	40.2940	0.6340	39.9960	0.4900	40.6490	0.6020
AWOTMAN	2	Labor market	Avg Weekly Overtime Hours : Manufacturing	0.0020	0.1350	0.0030	0.1510	0.0010	0.1130
AWHMAN	1	Labor market	Avg Weekly Hours : Manufacturing	40.7920	0.7240	40.4110	0.5750	41.2460	0.6130
CES0600000008	6	Labor market	Avg Hourly Earnings : Goods-Producing	0.0000	0.0040	0.0000	0.0050	0.0000	0.0030
CES2000000008	6	Labor market	Avg Hourly Earnings : Construction	0.0000	0.0080	0.0000	0.0100	0.0000	0.0050
CES3000000008	6	Labor market	Avg Hourly Earnings : Manufacturing	0.0000	0.0050	0.0000	0.0050	0.0000	0.0030
HOUST	4	Housing	Housing Starts: Total New Privately Owned	7.2230	0.3190	7.3070	0.2360	7.1240	0.3730
HOUSTNE	4	Housing	Housing Starts, Northeast	5.0590	0.4130	5.2750	0.3400	4.8020	0.3390
HOUSTMW	4	Housing	Housing Starts, Midwest	5.5580	0.4240	5.6950	0.3130	5.3960	0.4800
HOUSTS	4	Housing	Housing Starts, South	6.4150	0.3030	6.4390	0.2600	6.3860	0.3450
HOUSTW	4	Housing	Housing Starts, West	5.7800	0.3870	5.8540	0.3210	5.6910	0.4370
PERMIT	4	Housing	New Private Housing Permits (SAAR)	7.1750	0.3120	7.2020	0.2570	7.1440	0.3650
PERMITNE	4	Housing	New Private Housing Permits, Northeast (SAAR)	5.0780	0.3900	5.2540	0.3460	4.8690	0.3340
PERMITMW	4	Housing	New Private Housing Permits, Midwest (SAAR)	5.5070	0.3880	5.5810	0.3200	5.4190	0.4410
PERMITS	4	Housing	New Private Housing Permits, South (SAAR)	6.3070	0.3370	6.2350	0.3210	6.3930	0.3360
PERMITW	4	Housing	New Private Housing Permits, West (SAAR)	5.7960	0.3860	5.8590	0.3280	5.7200	0.4340
DPCERA3M086SBEA	5	Consumption, orders, and inventories	Real personal consumption expenditures	0.0030	0.0050	0.0030	0.0060	0.0020	0.0030
CMRMTSPLx	5	Consumption, orders, and inventories	Real Manu. and Trade Industries Sales	0.0020	0.0100	0.0030	0.0120	0.0020	0.0080
RETAILx	5	Consumption, orders, and inventories	Retail and Food Services Sales	0.0050	0.0120	0.0060	0.0130	0.0030	0.0100
ACOGNO	5	Consumption, orders, and inventories	New Orders for Consumer Goods	0.0010	0.0130	0.0000	0.0040	0.0030	0.0180
AMDMNOx	5	Consumption, orders, and inventories	New Orders for Durable Goods	0.0040	0.0380	0.0050	0.0340	0.0020	0.0410
ANDENOx	5	Consumption, orders, and inventories	New Orders for Nondefense Capital Goods	0.0030	0.0780	0.0050	0.0750	0.0020	0.0810
AMDMUOx	5	Consumption, orders, and inventories	Unfilled Orders for Durable Goods	0.0050	0.0100	0.0060	0.0110	0.0030	0.0100
BUSINVx	5	Consumption, orders, and inventories	Total Business Inventories	0.0040	0.0060	0.0060	0.0060	0.0030	0.0050
ISRATIOx	2	Consumption, orders, and inventories	Total Business: Inventories to Sales Ratio	0.0000	0.0170	0.0000	0.0200	0.0000	0.0140
UMCSENTx	2	Consumption, orders, and inventories	Consumer Sentiment Index	0.0210	3.2860	0.0200	2.7990	0.0230	3.7910

(Continued)

Table 1.12: FRED-MD *(Continued)*

Fred Code	tcode	Group	Description	Full Sample Mean	Full Sample sd	1962-1992 Mean	1962-1992 sd	1993-2017 Mean	1993-2017 sd
M1SL	6	Money and credit	M1 Money Stock	0.000	0.009	0.000	0.005	0.000	0.012
M2SL	6	Money and credit	M2 Money Stock	0.000	0.003	0.000	0.002	0.000	0.004
M2REAL	5	Money and credit	Real M2 Money Stock	0.002	0.005	0.002	0.005	0.003	0.005
AMBSL	6	Money and credit	St. Louis Adjusted Monetary Base	0.000	0.018	0.000	0.005	0.000	0.025
TOTRESNS	6	Money and credit	Total Reserves of Depository Institutions	0.000	0.064	0.000	0.032	0.000	0.089
NONBORRES	7	Money and credit	Reserves Of Depository Institutions	0.000	1.089	0.000	0.035	0.000	1.613
BUSLOANS	6	Money and credit	Commercial and Industrial Loans	0.000	0.006	0.000	0.006	0.000	0.007
REALLN	6	Money and credit	Real Estate Loans at All Commercial Banks	0.000	0.005	0.000	0.003	0.000	0.007
NONREVSL	6	Money and credit	Total Nonrevolving Credit	0.000	0.008	0.000	0.008	0.000	0.008
CONSPI	2	Money and credit	Nonrevolving consumer credit to Personal Income	0.000	0.001	0.000	0.001	0.000	0.001
MZMSL	6	Money and credit	MZM Money Stock	0.000	0.006	0.000	0.006	0.000	0.005
DTCOLNVHFNM	6	Money and credit	Consumer Motor Vehicle Loans Outstanding	0.000	0.025	0.000	0.021	0.000	0.029
DTCTHFNM	6	Money and credit	Total Consumer Loans and Leases Outstanding	0.000	0.021	0.000	0.013	0.000	0.027
INVEST	6	Money and credit	Securities in Bank Credit at All Commercial Banks	0.000	0.011	0.000	0.010	0.000	0.012
FEDFUNDS	2	Interest and exchange rates	Effective Federal Funds Rate	0.000	0.521	0.002	0.691	-0.002	0.162
CP3Mx	2	Interest and exchange rates	3-Month AA Financial Commercial Paper Rate	-0.001	0.508	0.001	0.667	-0.003	0.192
TB3MS	2	Interest and exchange rates	3-Month Treasury Bill	0.000	0.431	0.002	0.562	-0.003	0.173
TB6MS	2	Interest and exchange rates	6-Month Treasury Bill	-0.001	0.405	0.001	0.527	-0.003	0.170
GS1	2	Interest and exchange rates	1-Year Treasury Rate	-0.001	0.423	0.001	0.548	-0.003	0.188
GS5	2	Interest and exchange rates	5-Year Treasury Rate	-0.002	0.326	0.006	0.389	-0.011	0.230
GS10	2	Interest and exchange rates	10-Year Treasury Rate	-0.002	0.282	0.007	0.326	-0.013	0.218
AAA	2	Interest and exchange rates	Moody's Seasoned Aaa Corporate Bond	-0.001	0.224	0.010	0.257	-0.013	0.177
BAA	2	Interest and exchange rates	Moody's Seasoned Baa Corporate Bond	0.000	0.216	0.010	0.228	-0.012	0.201
COMPAPFFx	1	Interest and exchange rates	3-Month Commercial Paper Minus	0.084	0.419	0.031	0.518	0.148	0.241
TB3SMFFM	1	Interest and exchange rates	3-Month Treasury C Minus	-0.482	0.711	-0.729	0.860	-0.189	0.263
TB6SMFFM	1	Interest and exchange rates	6-Month Treasury C Minus	-0.349	0.765	-0.573	0.941	-0.081	0.311
T1YFFM	1	Interest and exchange rates	1-Year Treasury C Minus	0.026	0.770	-0.074	0.964	0.145	0.409
T5YFFM	1	Interest and exchange rates	5-Year Treasury C Minus	0.723	1.371	0.447	1.611	1.051	0.915
T10YFFM	1	Interest and exchange rates	10-Year Treasury C Minus	1.061	1.658	0.595	1.834	1.617	1.205
AAAFFM	1	Interest and exchange rates	Moody's Aaa Corporate Bond Minus	2.080	1.982	1.245	1.993	3.077	1.433
BAAFFM	1	Interest and exchange rates	Moody's Baa Corporate Bond Minus	3.100	2.084	2.323	2.096	4.026	1.646
TWEXMMTH	5	Interest and exchange rates	Trade Weighted U.S. Dollar Index: Major Currencies	0.000	0.015	0.000	0.014	0.000	0.016
EXSZUSx	5	Interest and exchange rates	Switzerland / U.S. Foreign Exchange Rate	-0.002	0.026	-0.003	0.026	-0.001	0.025
EXJPUSx	5	Interest and exchange rates	Japan / U.S. Foreign Exchange Rate	-0.002	0.024	-0.003	0.023	0.000	0.026
EXUSUKx	5	Interest and exchange rates	U.S. / U.K. Foreign Exchange Rate	-0.001	0.022	-0.002	0.023	-0.001	0.021
EXCAUSx	5	Interest and exchange rates	Canada / U.S. Foreign Exchange Rate	0.000	0.014	0.001	0.009	0.000	0.018
WPSFD49207	6	Prices	PPI: Finished Goods	0.000	0.007	0.000	0.006	0.000	0.008
WPSFD49502	6	Prices	PPI: Finished Consumer Goods	0.000	0.009	0.000	0.007	0.000	0.010
WPSID61	6	Prices	PPI: Intermediate Materials	0.000	0.007	0.000	0.006	0.000	0.009
WPSID62	6	Prices	PPI: Crude Materials	0.000	0.041	0.000	0.030	0.000	0.052
OILPRICEx	6	Prices	Crude Oil, spliced WTI and Cushing	0.000	0.095	0.000	0.090	0.000	0.100
PPICMM	6	Prices	PPI: Metals and metal products:	0.000	0.033	0.000	0.025	0.000	0.040
CPIAUCSL	6	Prices	CPI : All Items	0.000	0.003	0.000	0.003	0.000	0.003
CPIAPPSL	6	Prices	CPI : Apparel	0.000	0.005	0.000	0.005	0.000	0.006
CPITRNSL	6	Prices	CPI : Transportation	0.000	0.011	0.000	0.006	0.000	0.015
CPIMEDSL	6	Prices	CPI : Medical Care	0.000	0.003	0.000	0.003	0.000	0.002
CUSR0000SAC	6	Prices	CPI : Commodities	0.000	0.005	0.000	0.004	0.000	0.007
CUSR0000SAD	6	Prices	CPI : Durables	0.000	0.003	0.000	0.003	0.000	0.002
CUSR0000SAS	6	Prices	CPI : Services	0.000	0.003	0.000	0.003	0.000	0.001
CPIULFSL	6	Prices	CPI : All Items Less Food	0.000	0.003	0.000	0.003	0.000	0.003
CUSR0000SA0L2	6	Prices	CPI : All items less shelter	0.000	0.004	0.000	0.003	0.000	0.004
CUSR0000SA0L5	6	Prices	CPI : All items less medical care	0.000	0.003	0.000	0.003	0.000	0.003
PCEPI	6	Prices	Personal Cons. Expend.: Chain Index	0.000	0.002	0.000	0.002	0.000	0.002
DDURRG3M086SBEA	6	Prices	Personal Cons. Exp: Durable goods	0.000	0.003	0.000	0.003	0.000	0.003
DNDGRG3M086SBEA	6	Prices	Personal Cons. Exp: Nondurable goods	0.000	0.006	0.000	0.004	0.000	0.008
DSERRG3M086SBEA	6	Prices	Personal Cons. Exp: Services	0.000	0.002	0.000	0.002	0.000	0.002
S&P 500	5	Stock market	S&P's Common Stock Price Index: Composite	0.005	0.036	0.005	0.036	0.006	0.036
S&P: indust	5	Stock market	S&P's Common Stock Price Index: Industrials	0.006	0.036	0.005	0.037	0.006	0.035
S&P div yield	2	Stock market	S&P's Composite Common Stock: Dividend Yield	-0.001	0.121	0.000	0.148	-0.002	0.077
S&P PE ratio	5	Stock market	S&P's Composite Common Stock: Price-Earnings Ratio	0.000	0.046	0.000	0.039	-0.001	0.053
VXOCLSx	1	Stock market	VXO	18.929	7.390	18.299	5.981	19.679	8.730

As in McCracken and Ng (2016) we transform the variables following the code presented in column 'tcode'. The transformation for a series x are: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$, (5) $\Delta log(x_t)$, (6) $\Delta^2 log(x_t)$, and (7) $\Delta(x_t/x_{t-1} - 1)$. The column 'gsi' and 'gsi:description' present the comparable series in Global Insight.

Table 1.13: Descriptive Statistics - Sentiment-Based Measures

	Full Sample		**1985-1992**		**1993-2018**	
Description	**Mean**	**sd**	**Mean**	**sd**	**Mean**	**sd**
Three Component Index	107.76	32.68	114.78	20.65	105.19	35.78
News-Based Policy Uncert Index	108.06	39.55	106.42	28.04	108.66	43.03
News-Based Historical Economic Policy Uncertainty	146.45	41.47	145.15	30.45	146.93	44.88
1. Economic Policy Uncertainty	101.56	41.96	120.50	39.45	94.63	40.77
2. Monetary policy	97.11	59.30	119.85	64.74	88.78	55.00
Fiscal Policy (Taxes OR Spending)	106.62	65.85	122.12	59.68	100.94	67.19
3. Taxes	106.08	65.32	113.14	55.13	103.49	68.58
4. Government spending	110.91	101.51	152.86	101.14	95.54	97.40
5. Health care	115.76	91.37	68.91	57.97	132.93	95.36
6. National security	96.20	81.23	127.60	85.12	84.69	76.77
7. Entitlement programs	112.41	85.69	79.55	67.82	124.45	88.47
8. Regulation	105.40	55.31	106.27	47.06	105.08	58.13
Financial Regulation	105.80	119.82	123.88	109.99	99.17	122.76
9. Trade policy	93.43	107.95	107.88	70.54	88.14	118.42
10. Sovereign debt, currency crises	116.59	202.91	101.42	164.86	122.14	215.20
DebtCeiling Relative Frequency	0.00	0.00	0.00	0.00	0.00	0.00
GovernmentShutdown Relative Frequency	0.00	0.00	0.00	0.00	0.00	0.00
Ratio: EPU w/DebtCeiling to wo/DebtCeiling	1.00	0.01	1.00	0.00	1.00	0.01
Ratio: EPU w/GovtShutdown to wo/GovtShutdown	1.00	0.01	1.00	0.00	1.00	0.02
Financial Stress Indicator	101.23	0.73	101.05	0.37	101.29	0.81

Table 1.9 reports the mean and standard deviation of the sentiment-based variables used in section 1.4.3 for the full sample (1895:01-2017:12) and the period of analysis (1993:01-2017:12). The first one is the economic policy uncertainty measure (EPU) from Baker et al. (2016). The sentiment-based variables are: *Three Component Index, News-Based Policy Uncertainty Index, News-Based Historical Economic Policy Uncertainty.* The categorical EPU considers a range of sub-indexes based solely in news data from *Access World News* of over 2,000 US newspapers. The categories are: *Economic Policy Uncertainty, Monetary policy, Fiscal Policy (Taxes or Spending), Taxes, Government Spending, Health Care, National Security, Entitlement Programs, Regulation, Financial Regulation, Trade Policy, and Sovereign Debt, Currency Crises.* Finally, is the Financial Stress Indicator (FSI) for the U.S from Püttmann (2018).

Chapter 2

Portfolio Efficiency with High-Dimensional Data as Conditioning Information

Abstract

In this paper, we build efficient portfolios using different frameworks proposed in the literature and drawing upon several datasets that contain an increasing number of predictors as conditioning information. We carry an extensive empirical study to investigate approaches that impose sparsity and dimensionality reduction, as well as possible latent factors driving the returns of the risky assets. In contrast to previous studies that made use of naive OLS and low-dimension information sets, we find that (i) accounting for large conditioning information sets, and (ii) the use of variable selection, shrinkage methods and factor models, such as the principal component regression and the partial least squares, provides better out-of-sample results as measured by Sharpe ratios, implied Sharpe ratios, and higher certainty equivalent returns (CER).

JEL classification: G11, G17, C32, C38.

Keywords: Dimensionality reduction. Shrinkage. Efficient Portfolios. Principal Components Regression (PCR). Partial Least Squares (PLS). Three-Pass Regression Filter (3PRF). Ridge Regression, LASSO.

2.1 Introduction

The use of return predictability to form efficient portfolios is a cornerstone in the asset pricing literature. Theoretically, this has been addressed in the mean-variance efficiency framework where many ways of bringing the conditioning information during the optimization process was proposed. Over the last decades, the literature has acknowledged several variables for their predictive power and thus potential to serve as conditioning information. However, the search for all these instruments produced an enormous amount of variables that supposedly has or had some predictability power. This applies to firm characteristics, whole market financial variables, and macroeconomic datasets as well.

The question we propose to investigate is how to properly deal with an extensive set of potential candidates while constructing mean-variance efficient portfolios with respect to conditioning information? Cochrane (2011), Goyal (2012), and Harvey et al. (2016) emphasize the profusion of variables that seem to explain the cross-sectional pattern of returns, and therefore, the clear need to reduce this dimensionality. Thus, we are interested in determining whether there is a better approach for generating useful portfolios by shrinking this high-dimensional set while estimating the optimal weights of the efficient portfolios.

In this paper, we apply the high-dimensional literature in finance to the above portfolio theory problem. We incorporate conditioning information in the efficient portfolios framework as it has been proposed in the literature, while drawing upon the most recent statistical techniques for high-dimensional predictive estimations. Specifically, we build efficient portfolios using high-dimensional information sets as conditioning information, and conduct an extensive out-of-sample analysis that compares high-dimensional statistical methods with a naive OLS approach for dealing with a large set of potential conditioning information.

We exploit the wealth set of data as predictive signals and condense this information when estimating the conditional mean using high-dimensional data. We conduct this analysis by leveraging several techniques that impose sparsity and dimensionality reduction when finding the conditional mean, which is the most important driver in the formation of mean-variance efficient portfolios with conditioning information. In short, we make a “bet in sparsity”, in the sense that much of the information in the conditioning set can be summarized by few factors. We evaluate how penalized estimators, such as LASSO, Ridge and Elastic Net, as well as pure dimensionality reduction and latent factors approaches such as Partial Least Squares (PLS) and Principal Components Regression (PCR), in addition to a generalization of the former (Three Pass Regression Filter) can produce different optimized portfolios.

The main contribution of this paper is that it is possible to condense the large set of potential predictors to build meaningful efficient portfolios. We show that conditioning in a large set of instruments yields better outcomes when compared to a low-dimensional set, and that relying on OLS to obtain the conditional mean of these efficient portfolios is not enough. In fact, the use of OLS can be damaging to results, even in cases when not conditioning in a large set of instruments. We document that a few methods, PCR and PLS especially, provide good out-of-sample results from an investor standpoint. Given that we considered a large out-of-sample (1996-2017) for the US market, which comprises two recessions, these results highlight the importance of (i) the use of high-dimensional sets of instruments when building mean-variance efficient portfolios with respect to conditioning information, and (ii) the use of a proper technique to reduce the dimension of the signals when finding the conditional mean.

In this paper we document evidence of the superior outcomes when the information in the conditioning set is summarized by few factors. First, independently of the size of the conditioning information set, we show that both methods are able to convey economically meaningful results as there are clear improvements in the out-of-sample Sharpe ratios, and especially higher out-of-sample monthly implied Sharpe ratios, which is when an investor can exploit signals from the conditioning information set (Campbell and Thompson, 2007). Second, we document improvements in economic gains obtained primarily when these two techniques are used to shrink the high-dimensional conditioning sets. By evaluating the certainty equivalent return performance, we see that PLS and PCR consistently rank above the other methods, and in some cases, Ridge also performs well. Third, we find that PLS and PCR can enhance the formation of optimized portfolios, generating large and significant alphas when evaluated in factor models, such as Fama-French 3 and 5. These results are robust to different approaches of building efficient portfolios with conditioning information, as we evaluated three different methods. Finally, we also assess which one, among the large set of signals, has larger impacts in the excess returns conditional means. We find that a small set of variables capturing the dynamics of the treasury bonds and the labor markets are the ones that contribute the most to the estimated conditional means.

Since the purpose of using conditioning information is to provide signals about the state of the economy or the condition of the risky assets, exploring a diverse, large set of possibilities for the predictive instruments, as well as flexible functional forms in which these signals are used in the optimization process can enhance the construction of efficient portfolios. Another motivation for working with high-dimensional information sets can be related to the Hansen-Richard critique (see Cochrane, 2009). Though there might be unobservable conditioning information used by economic agents, we could posit that, when faced with a large set of potential candidates, enlarging the set would increase the chance of capturing more information from a wide variety of information sets of heterogeneous economic agents.

Moreover, with the use of recently proposed measures of economic uncertainty (see Baker et al., 2016; Püttmann, 2018; Jurado et al., 2015), we could partially capture some of what used to be non-observable information.

We can model the data generating process of the excess returns as a conditional mean plus a noise term (see e.g., Ferson and Siegel (2001)). The fundamental idea is that this conditional mean, which is a function of a set of lagged variables, can be approximated by a broad class of functions. Previous studies that analyzed empirically efficient portfolios for a specific market and asset class, predominantly made use of OLS to estimate this function when deriving the efficient portfolios weights from the mean-variance optimization problem (Ferson and Siegel, 2009; Abhyankar et al., 2012; Fletcher and Basu, 2016; Peñaranda, 2016). This was mainly motivated by the fact they used only a few or a handful of instruments as a conditioning information set[1].

However, in doing so, we face several issues. First, as Welch and Goyal (2007) shows in a comprehensive study, most of the variables taken as having predictability power and commonly used in academic research and by market practitioners perform badly out-of-sample, and in some cases even within the in-sample analysis[2]. Second, even in the case that a variable has some predictability power, this can only exist for a specific period (perhaps only in the past). Welch and Goyal (2007) sheds light on the problem of short-lived predictors, in contraposition to steady long-lived predictors. Some of the variables could have some predictive power that existed in a specific period. In a broad analysis McLean and Pontiff (2016) show that after an academic paper is published, the average predictor's long-short return shrinks post-publication[3].

Third, modern financial markets are fast and complex, and with access to so many sources, the process of data collection and processing reached the level commonly known as big data. It is known that in such cases the ordinary least squares (OLS) estimates will have a poor behavior in forecasting, or not even exist for wide data for instance. An appropriate way

[1] For example, Peñaranda (2016) used only three instruments (the U.S. dividend–price ratio, the default spread, and the term spread), Abhyankar et al. (2012) used seven (the market index, the dividend yield, the 1-month T-bill, the term spread, the convexity of the term structure, the credit yield spread, and inflation), Ferson and Siegel (2009) used five (the 1-month T-bill, the dividend yield, the default spread, and two different term spreads), and Fletcher and Basu (2016) used nine (the dividend yield, the 1-month Treasury Bill, the term spread, the default spread, inflation, the output gap, two excess market returns (local and U.S.), the quarterly log growth index).

[2] The authors selected variables from previously published articles and show that most of them behave poorly with unstable predictive performance.

[3] The authors evaluated the out-of-sample post-publication return predictability of 97 signals from 79 different academic studies. Portfolio returns dropped 26% out-of-sample and 58%, yielding an estimated effect of 32% (58%–26%) lower return from publication-informed trading. They also show that the publication has an impact on the predictor portfolio returns. They argue that post-published correlation with already-published predictor portfolio returns increases.

to deal with these types of data is via some sort of shrinkage-type regularization. Sparsity and dimensionality reduction are the goals of these problems. There are compelling arguments that support that we cannot rely on a small set of conditioning information. In order to obtain better approximations to the conditional mean, the set of instruments must be enriched to resemble the exponential growth of available information. Finally, the act of finding the conditional mean to explain the variation of the risky assets is a risk premia problem. By far, it is known that the problem of explaining the risk premia has been one of the most debated and scrutinized topics for a long period in the literature. Clearly, approximating the conditional mean is a hard task because market efficiency *forces* return variation to be dominated by unforecastable news that obscures risk premia Gu et al. (2018). We can say that not only the amount, but also the complexity of information driving the conditional mean of the risky assets could be enhanced. Modeling the conditional mean with few observable factors entering linearly, with no interactions, penalization, or dimension reduction techniques, could also create an incomplete structure for explaining the stochastic discount factor.

Hence, in order to exploit the wealth of data and try to condense them when estimating the conditional mean using high-dimensional data, we seek to impose sparsity and dimensionality reduction when finding the conditional mean that enters in all functions to build efficient portfolios. Importantly, it is the fact that we can consider the function that seeks to approximate the conditional mean driving the DGP of the excess returns to have a rich and flexible format. To obtain better approximations, we assume that this aproximating function is asset dependent, in the sense that the factors driving excess returns of each risky asset are built using high-dimensional data for each different asset.

2.1.1 Related Literature

Our paper is connected to several strands of the literature. First, this paper is related to the theoretical implications derived by Hansen and Richard (1987) Hansen and Richard (1987), and the different ways of defining the minimum variance efficient portfolios: the straightforward conditionally mean-variance (CMV), and the unconditionally mean-variance with conditioning information (UMV) portfolios. Ferson and Siegel (2001) presented closed-form solutions to the UMV portfolios, that is a more applicable way of working with conditioning information with roots on information asymmetry. More recently, Chiang (2015) proposed a modern portfolio management version of the UMV, known as mean-variance active tracking error (MVATE). In this paper, we assess all three approaches for building efficient portfolios with conditioning information.

Second, this paper also relates to the dimension reduction literature in finance, which is a growing topic with several research methodologies for tackling the issue of high-dimensionality in applied work. One of the most straightforward and commonly used technique is the principal components regression (PCR). Kelly et al. (2019) proposed a dimension reduction method based on the principal components analysis (PCA)[4] to explain the cross-section of average returns. Gu et al. (2018) made use of PCR, among other machine learning techniques, for measuring asset risk premia. Another classical dimension reduction technique is the partial least squares (PLS). PLS is a simple regression-based procedure designed to parsimoniously forecast a single time series using a large panel of predictors. Light et al. (2017) use PLS to estimate the expected returns on individual stocks from a large set of firm characteristics. Haddad et al. (2020) show that some dimensionality-reduction methods, such as principal components, can help to construct portfolios that time many anomalies simultaneously. Kelly and Pruitt (2015) proposed recently a generalization of PLS, known as the Three-Pass Regression Filter (3PRF). The 3PRF is a constrained least squares estimator and reduces to partial least squares as a special case. Kelly and Pruitt (2013) uses the 3PRF to forecast the conditional expectation of market returns and cash flow.

Included in the body of dimensionality reduction literature are penalized regression methods that are applied to asset pricing. Kozak et al. (2019) adapt Ridge and LASSO estimators to estimate the stochastic discount factor to summarize and impose sparsity of an extensive list of predictors in a high-dimension setting. Freyberger et al. (2017) use a variation of LASSO to select characteristics from a large number of characteristics and estimate a nonlinear function for expected returns. Chinco et al. (2019) use LASSO to make intraday return forecasts for a large number of stocks. Other papers utilized penalized estimators for macro variables and bond risk premia as well (see Bianchi et al. (2019); Huang and Shi (2011); Bai and Ng (2006)).

Third, this paper is also motivated by the financial literature concerned with the abundance of predictors, as outlined by Cochrane (2011), and Goyal (2012). Several recent studies attempted to compile these possible predictors. Green et al. (2013) assessed 330 signals (stock-level) reported in the past years. Feng et al. (2017) studied a list of 150 risk factors, what they called a "zoo" of factors proposed in the literature in the past 30 years. Harvey et al. (2016) analyzed the statistical significance of more than 300 signals published over the last forty years. Finally, this work is also linked to a literature at the intersection of high-dimensional predictors and portfolio construction. DeMiguel et al. (2020) examine how transaction costs change the number of characteristics that are jointly significant for an investor's optimal portfolio and, hence, how they change the dimension of the cross-section of

[4] Instrumented Principal Components Analysis (IPCA).

stock returns. DeMiguel et al. (2021) use Elastic Net, and other machine learning methods, to exploit the predictive ability of a large set of mutual fund characteristics.

The structure of this paper is as follows. Next section introduces the general framework and presents the three approaches mentioned above, the optimization problems and the weights' solution of building mean-variance efficient portfolios. Section 2.3 summarizes the techniques for imposing sparsity in a high-dimensional setting, presenting the penalized estimators, the PLS, PCR, and the 3PRF. Section 2.4 explains our research design and the datasets used. Section 2.5 presents the results, and how the performance is evaluated. Finally, section 2.6 concludes the analysis.

2.2 General Framework

2.2.1 Optimal Portfolios with Conditioning Information

Consider an investment set with N risky assets, indexed by $i = 1, \dots, N$ and the existence of a risk-free asset. Let $R_{f,t}$ be the return of the risk-free asset at time t, and denote $\mathbf{R}_t$[5] as the N-dimensional vector with the gross returns of the N risky assets at time t. The vector of excess of returns $\mathbf{r}_t$ at time t is given by $\mathbf{r}_t = \mathbf{R}_t - \mathbf{1}R_f$, where $\mathbf{1}$ is an N-dimensional vector of ones. Let $\mathbf{x}_t(\mathbf{Z}_t)$ denote the vector of portfolio weights on the N risky assets at time t, which is a function of the investor's conditioning information $\mathbf{Z}_t$ at time t. At each period of time t, using the conditioning information $\mathbf{Z}_t$ the investor sets the weights of the vector $\mathbf{x}_t$, investing the remainder of the funds in the risk-free asset. Thus, the weight on the risk-free asset is given by $1 - \mathbf{x}_t(\mathbf{Z}_t)^\top \mathbf{1}$, being the final payoff of the portfolio at $t+1$ is given by $R_{p,t+1} = \mathbf{x}_t(\mathbf{Z}_t)^\top \mathbf{r}_{t+1} + R_{f,t+1}$[6]. Below, we briefly present some of the approaches that were proposed in the literature to build efficient portfolios with conditioning information, summarizing three frameworks.

[5] A note on the notation throughout our analysis. Unless mentioned otherwise, we use lowercase boldface characters to represent vectors (except for $\mathbf{R}_t$), uppercase boldface characters for matrices, and regular characters for scalars.

[6]

$$
\begin{aligned}
R_{p,t+1} &= \mathbf{R}_{t+1}^\top \mathbf{x}_t(\mathbf{Z}_t) + \left(1 - \mathbf{x}_t(\mathbf{Z}_t)^\top \mathbf{1}\right) R_{f,t+1} && (2.1)\\
&= \left(\mathbf{R}_{t+1} - \mathbf{1}R_{f,t+1}\right)^\top \mathbf{x}_t(\mathbf{Z}_t) + R_{f,t+1} && (2.2)\\
&= \mathbf{x}_t(\mathbf{Z}_t)^\top \mathbf{r}_{t+1} + R_{f,t+1} && (2.3)
\end{aligned}
$$

Conditionally Efficient Portfolios

As in Hansen and Richard (1987), conditionally efficient returns are the solution of the following problem:

$$\begin{aligned} \min_{x(Z)} \quad & \mathbb{E}\left(r_{p,t+1}^2 | \mathbf{Z}_t\right) \\ \text{s.t.} \quad & \mathbb{E}\left(r_{p,t+1} | \mathbf{Z}_t\right) = \alpha_{p,t+1} \end{aligned} \tag{2.4}$$

where $r_{p,t+1}$ is the portfolio excess returns at time $t+1$, and $\alpha_{p,t+1}$ is the target conditional expected excess return. The solution of the problem in equation (2.4) providing the weights allocated in each risky asset is given by:

$$\mathbf{x}_t(\mathbf{Z}_{t-1}) = \left(\frac{\mathbb{E}\left(r_{p,t+1} | \mathbf{Z}_t\right)}{\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)} \right) \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right) \tag{2.5}$$

where $\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)$ is the n-dimensional vector of conditional expected excess returns at time t, and $\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)$ is the $N \times N$ conditional second moment matrix at time t.

Unconditionally Efficient Portfolios with Respect to the Information

The unconditionally mean-variance (UMV) efficient portfolio strategy takes another path. Ferson and Siegel (2001) study the properties of unconditional minimum-variance portfolios in the presence of conditioning information. Different from the CMV approach above, here the objective is to maximize the unconditional mean relative to the unconditional variance, being the strategies a function of the available information[7]. UMV is related to information asymmetry, a typical situation in which a portfolio manager using conditioning information builds a conditionally efficient portfolio. However, for an uninformed investor, the portfolio does not appear efficient. The solution of the UMV problem providing the weights allocated in each risky asset is given by:

$$\mathbf{x}_t(\mathbf{Z}_{t-1}) = \left(\frac{\mathbb{E}\left(r_{p,t+1}\right)}{\mathbb{E}\left(\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)\right)} \right) \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right) \tag{2.6}$$

where $\mathbb{E}\left(r_{p,t+1}\right)$ is the unconditional expected return target, and $\mathbb{E}\left(\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)\right)$ is the unconditional mean of $\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)$.

[7] Hansen and Richard (1987) show that unconditionally efficient returns are a subset of conditionally efficient returns.

2.2.2 Unconditionally Tracking Efficient Portfolios

Using tracking error strategies, in which an active portfolio manager uses conditioning information to optimize unconditional performance measures relative to a benchmark, Chiang (2015) proposes the *unconditional tracking efficiency* to exploit conditioning information in the mean-variance framework. This structure should be seen as a common practice between market practitioners, since a portfolio manager conducting an optimization uses more information than his clients, who possibly do not have access to the entire set being used as conditioning information, which drives managers to seek higher performance from the clients' perspective. Under the unconditional mean-variance tracking error approach (UMVTE), the solution to the problem[8] is determined by the following function for the portfolios weights[9]:

$$\mathbf{x}_t(\mathbf{Z}_{t-1}) \quad = \lambda_1 \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right) + \left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\gamma}_t\left(\mathbf{Z}_{t-1}\right) \quad , \tag{2.8}$$

where $\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)$ is the conditional mean, $\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)$ is the conditional second moment of the excess returns, $\boldsymbol{\gamma}_t\left(\mathbf{Z}_{t-1}\right) = \mathbb{E}\left(\mathbf{r}_t\left(r_{b,t} - \mu_{b,t}\right)|\mathbf{Z}_{t-1}\right)$, and λ_1 a scalar [10]. Note that the CMV, UMV and UMVTE are not the only frameworks proposed in the literature to build efficient portfolios taking into account conditioning information. Brandt and Santa-Clara (2006) Peñaranda (2016) also proposed different approaches to build mean-variance efficient portfolios with conditioning information.

2.3 Estimation

In general, the OLS regression to obtain the K parameters of $\boldsymbol{\theta}_i$ will provide nonzero estimates. When dealing with high-dimensional sets with a large number of predictors,

[8]In this framework, a portfolio manager uses conditioning information $\mathbf{Z}_t$, which is not available to his clients, to build a vector $\mathbf{x}_t(\mathbf{Z}_{t-1})$ of unrestricted weights of the N available risky assets. Under the unconditional mean-variance tracking error approach (UMVTE), this active manager uses conditioning instruments to form portfolios that optimize unconditional performance measure. Thus, for a benchmark b, with returns given by $R_{b,t}$, the UMVTE problem is to minimize the unconditional error variance, $\mathbb{V}\text{ar}\left(R_{p,t+1} - R_{b,t+1}\right)$, for a given unconditional expected return target $\alpha_{p,t+1} = \mathbb{E}(R_{p,t+1} - R_{b,t+1})$,

$$\begin{aligned} \min_{\mathbf{x}(\mathbf{Z})} \quad & \mathbb{V}\text{ar}\left(R_{p,t+1} - R_{b,t+1}\right) \\ \text{s.t.} \quad & \mathbb{E}\left(R_{p,t+1} - R_{b,t+1}\right) = \alpha_{p,t+1} \end{aligned} \tag{2.7}$$

Chiang (2015) solves this minimization problem using calculus of variations.

[9]To the case without constraint on portfolio risk.

[10]Precisely,

$$\lambda_1 = \frac{\left(\alpha_{p,t+1} + \mu_{b,t}\right) - \eta_2}{\eta_1} \tag{2.9}$$

where $\eta_1 = \mathbb{E}\left(\boldsymbol{\mu}_t^\top\left(\mathbf{Z}_{t-1}\right)\left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)\right)$, and $\eta_2 = \mathbb{E}\left(\boldsymbol{\mu}_t^\top\left(\boldsymbol{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1} \boldsymbol{\gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)$. For a full proof of the solution of the UMVTE problem, see Chiang (2015).

usually we face two issues using OLS. First, whenever we have $K > T$, OLS cannot find a unique solution. Second, the infinite set of solutions has a tendency to overfit the data instead of extracting informative signals for the conditional mean. This fact, known as "curse of dimensionality" is well known in the literature. Gu et al. (2018) mention that this issue is more exacerbated by the low state of the signal-noise ratio in return predictions. This causes the OLS estimations to have a poor behavior in forecasting. Below we summarize some of the approaches proposed in the current state of the literature to deal with high-dimensional set of predictors, imposing sparsity through shrinkage and selection type techniques, and latent variables estimators as well.

2.3.1 Penalized Regression

To deal with the inconsistency and the inefficiency of the OLS in high-dimensional data, shrinking the number of estimated parameters is essential to avoid overfitting the data. Thus, some sort of shrinkage or variable selection might be necessary to deal with this type of data. This can be done by penalizing the objective function being optimized. In general, penalized estimators share the following common structure:

$$\mathcal{L}(\boldsymbol{\theta};\cdot) = \underbrace{\mathcal{L}(\boldsymbol{\theta})}_{\text{Loss Function}} + \underbrace{\phi(\boldsymbol{\theta};\cdot)}_{\text{Penalty}} \quad . \tag{2.10}$$

An estimator to impose some sort of sparsity in the set of regressors is the Least Absolute Selection and Shrinkage Operator (LASSO), proposed in the seminal paper by Tibshirani (1996). This estimator regularizes the estimation process constraining the size of the estimates. Following Hastie et al. (2015); Friedman et al. (2001), the three main estimators from the penalized regression class can be adapted for our predictive portfolio setting as:

$$\hat{\boldsymbol{\theta}}_i := \arg\min_{\theta_{0,i},\boldsymbol{\theta}_i} \frac{1}{T}\sum_{t=1}^{T}\left(r_{i,t} - \theta_{0,i} - \sum_{k=1}^{K} Z_{k,t-1}\theta_{k,i}\right)^2 + \underbrace{\phi(\boldsymbol{\theta};\cdot)}_{\text{Penalty}} \begin{cases} \text{LASSO: } \lambda\sum_{k=1}^{K}|\theta_{k,i}| \\ \\ \text{Ridge: } \lambda\sum_{k=1}^{K}\theta_{k,i}^2 \\ \\ \text{ENet: } \lambda\sum_{k=1}^{K}\left(\alpha\,|\theta_{k,i}| + (1-\alpha)\,\theta_{k,i}^2\right) \end{cases}$$

where $\alpha \in [0,1]$. The LASSO estimator is an ℓ_1 regularized regression, being the parameter λ what controls the amount of shrinkage, i.e., the larger the lambda, the greater is the shrinkage. The penalty imposed in equation (2.11) makes the solutions nonlinear in $r_{i,t}$. This is associated with the fact that there is no closed-form solution. The lasso problem is a convex program, and computing its solution is a quadratic programming problem. On the other hand, the Ridge estimator sets an ℓ_2 norm penalty, i.e., $\phi(\boldsymbol{\theta};\cdot)$ as $\|\theta\|_2^2 = \sum_{k=1}^{K} \theta_{k,i}^2$. An important characteristic of the Ridge approach is the one-to-one correspondence between the parameters λ and t. This fact makes Ridge estimation to attenuate multicollinearity, when it is present in the data. Since, when K is large, the large number of regressors may result in high correlation among some of them. With standard "naive" OLS, multicollinearity causes poorly determined coefficients with large variances (and large variance inflator factor - VIF), what Ridge can mitigate.

Finally, the Elastic Net (ENet) estimator (Zou and Hastie, 2005) seeks to combine the penalties of both estimators above (LASSO and Ridge) via a convex combination. Setting $\alpha = 1$ this estimator reduces to the LASSO regression with an ℓ_1-norm, with $\alpha = 0$ it reduces to the ridge regression with an ℓ_2-penalty. Among the positive features of the ENet, when adding both penalties, this estimator automatically controls for strong within-group correlations. The ENet is also a strictly convex problem, thus, providing a unique solution independently of correlations or duplications in the $Z_{k,t-1}$. However, in equation (2.11) we have an additional tuning parameter α that has to be determined or defined *ad hoc*.

2.3.2 Principal Components Regression (PCR)

Whenever we are dealing with a high dimensional set of explanatory variables, high collinearities may be an issue to generate good conditional means estimates. One technique to deal with multicollinearity is the Principal Components Regression (PCR). PCR is a dimension reduction approach based on the Principal Components Analysis (PCA). As a regression technique, PCR regresses the dependent variable on the principal components generated by the PCA.

In summary, PCR first performs PCA on the matrix of predictors (here, the $\mathbf{Z}_{t-1}$ of lagged instruments) and obtains the principal components associated with it. Out of the M components, we can select a subset $m \in M$ and regress the dependent variable using the first m principal components as regressors[11]. Since PCR is essentially a shrinkage estimator that seeks to capture components with high variance[12], at the cost of those with low variance, this approach performs dimension reduction by selecting m components. Further, it also has a shrinkage effect by removing those components with low variance to form the final model for obtaining the conditional mean.

[11] A final step is done to scale back to the original regressors using the PCA loadings.

[12] Those with higher eigenvalues in $\mathbf{Z}^\top\mathbf{Z}$.

Using the singular value decomposition (SVD), we can factorize $\mathbf{Z}$[13] as $\mathbf{Z} = \mathbf{U}\mathbf{D}\mathbf{V}^{\top}$, where $\mathbf{U}_{T\times K}$ and $\mathbf{V}_{K\times K}$ are orthogonal matrices[14], and $\mathbf{D}_{K\times K} = \operatorname{diag}[\delta_1, \delta_2, \ldots, \delta_K]$, being each δ_m the singular values of $\mathbf{Z}$. We can obtain the principal values of $\mathbf{Z}^{\top}\mathbf{Z}$ as $\mathbf{\Lambda}_{K\times K} = \operatorname{diag}[\delta_1^2, \delta_2^2, \ldots, \delta_K^2]$ through a spectral decomposition given by $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\top}$. Finally, we can multiply the matrix $\mathbf{Z}$ by the first m principal components $\mathbf{V}^{(m)} = \left[\boldsymbol{\omega}^{(1)}, \ldots, \boldsymbol{\omega}^{(m)}\right]$, where $m \in \{1, 2, \ldots, K\}$, to obtain our derived covariates. Each m principal component $\boldsymbol{\omega}^{(m)}$ can be seen as m linear combinations of $\mathbf{Z}_{t-1}$. Thus, using the PCR approach, the conditional mean at time t of the excess returns of risky asset i can be approximated by running the following regression

$$\mathbf{r}_i = \underbrace{\left(\mathbf{Z}_{t-1}\mathbf{V}_{t-1}^{(m)}\right)}_{\text{Derived Covariates}} \boldsymbol{\theta}_i^{(m)} + \boldsymbol{\epsilon}_i \quad , \tag{2.11}$$

so that $\widehat{\boldsymbol{\mu}}_i(\mathbf{Z}_{t-1}) = \left(\mathbf{Z}_{t-1}\mathbf{V}_{t-1}^{(m)}\right)\hat{\boldsymbol{\theta}}_i^{(m)}$, where $\hat{\boldsymbol{\theta}}_i^{(m)} = \left(\mathbf{V}_{t-1}^{(m)\top}\mathbf{Z}_{t-1}^{\top}\mathbf{Z}_{t-1}\mathbf{V}_{t-1}^{(m)}\right)\mathbf{V}_{t-1}^{(m)\top}\mathbf{Z}_{t-1}^{\top}r_{i,t}$. Equation (2.11) makes clear that PCR reduces dimensions by using m principal components that capture the largest common variation in the conditioning information set, and weights through $\boldsymbol{\omega}^{(m)}$ the original covariates. The m number of PCA components used to form the derived covariates in the final model is a parameter that needs to be chosen adaptively. In section 2.4.1 we discuss how this is done.

2.3.3 Partial Least Squares (PLS)

When PCR finds the components with the largest common variation in the conditioning information set, it takes into consideration in its objective function only the information from $\mathbf{Z}_{t-1}$. As Gu et al. (2018) argues, doing so it fails to incorporate the information from the dependent variable (excess returns) when performing dimension reduction, which can be suboptimal to generate good conditional means approximations. The Partial Least Squares (PLS), on the other hand, uses the common components from $\mathbf{Z}_{t-1}$ by conditioning on the joint distribution of $r_{i,t}$ and $\mathbf{Z}_{t-1}$. For this reason, PLS is considered a latent approach that models the covariance between the spaces generated by both matrices ($r_{i,t}$ and $\mathbf{Z}_{t-1}$).

As mentioned by Friedman et al. (2001) this method seeks to obtain the directions that provide the highest variance and highest correlation[15] between $r_{i,t}$ and $\mathbf{Z}_{t-1}$ Stone and Brooks

[13]We dropped the time subscripts $t-1$ for simplicity.

[14]These matrices are known as left and right singular vectors of $\mathbf{Z}$ respectively.

[15]A clear contrast to PCR that only seeks directions that maximize only the variance between $r_{i,t}$ and $\mathbf{Z}_{t-1}$

(1990); Frank and Friedman (1993). Following Friedman et al. (2001), the idea of PLS is to weight each vector $\mathbf{z}_{t-1,k} \in \mathbf{Z}_{t-1}$ by its partial least squares direction $\hat{\varphi}_k = \langle\ \mathbf{z}_{t-1,k}, r_{i,t}\rangle$. The derived input $\mathfrak{z}_k = \sum_k \widehat{\varphi}_k \mathbf{z}_{t-1,k}$ is then used to obtain the the estimated coefficient $\widehat{\theta}_k$ regressing $r_{i,t}$ on the derived input $\mathfrak{z}_k$. This is done from $k = 1, 2, \ldots, K$, where we orthogonalize the original predictors with respect to the previous component. This process repeats until $m \leq K$ desired directions to reduce dimensions have been found. Clearly, if we build $m = K$ directions, we return to the standard least squares estimates.

The idea of the partial least squares direction $\hat{\varphi}_k$ is to obtain the weight of the strength (or the partial sensitivity) of the univariate effect each variable in the conditioning set $\mathbf{Z}_{t-1}$ on $r_{i,t}$. As with the PCR, the m number of directions is a parameter that needs to be chosen appropriately.

2.3.4 Three-Pass Regression Filter (3PRF)

Kelly and Pruitt (2015) proposed the three-pass regression filter (3PRF), a generalization of the PLS estimator, in the sense that the latter is a special case of the former. In this setting, assuming that the data can be represented by an approximate factor model to reduce the dimension of the predictive information, the target variable is the time series of the excess returns of each risky asset i, having the following general form:

$$r^i_{t+h} = \beta^i_0 + \boldsymbol{\beta}^{i^\top} \boldsymbol{F}^i_t + \eta^i_{t+h} \quad , \tag{2.12}$$

where β^i_0 is a scalar, $\boldsymbol{F}^i_t$ is a K_F-dimensional vector of factors, $\boldsymbol{\beta}^i$ is a K_F-dimensional vector of parameters, and η^i is the error term.

For each risky asset i in $i = 1, 2, \ldots, N$ there are three steps summarized in table 2.1. For a each risky asset i, the first pass of the estimator runs L separate time series regressions on each predictor. The estimated coefficients from this regression describe the sensitivity of conditioning information (predictors) to the latent factor driving the forecast target. In the second pass, we use the estimated first-pass coefficients in T separate cross section regressions, where our conditioning information variable l are the dependent variable, and the first-stage loadings $\hat{\phi}^i_l$ are the independent variables. Therefore, the first-stage coefficient estimates map the cross-sectional distribution of predictors to the latent factors. Second-stage cross section regressions use this map to back out estimates of the factors at each point in time (Kelly and Pruitt, 2015). The third stage is the final forecasting step. Running a single time series regression of the target variable r^i_{t+1} on the second-pass estimated factors

$\hat{\boldsymbol{F}}_t^i$, it provides the forecast, which is the fitted value $r_{t+1}^i = \hat{\beta}_0^i + \hat{\boldsymbol{F}}_t^{i\top} \hat{\boldsymbol{\beta}}^i$[16]. Kelly and Pruitt (2015) prove that, under a set of assumptions[17] the 3PRF is a consistent estimator as L and T become large.

Table 2.1: The Three Pass Regression Filter (3PRF)

Pass	Description
1	run time series regression of $\boldsymbol{z}_l^i$ on $\mathbf{r}^i$ for $j = 1, 2, \ldots, L$ predictors
2	run cross section regression of $\boldsymbol{z}_l^i$ on $\hat{\boldsymbol{\phi}}_l^i$ for $t = 1, 2, \ldots, T$
3	run time series regression of r_{t+1}^i on predictive factors $\hat{\mathbf{F}}_t^i$

.

2.4 Empirical Strategy

Our focus is from July 1963 to December 2017 for U.S. data. The out-of-sample cut off starts in Jan-1996 and we use recursive windows to update the estimations with the new available information. Using different estimators to impose dimensionality reduction, we compare the performance of each approach to build CMV, UMV and MVATE efficient portfolios, as presented on section 2.2.1.

At each t we form the weights of our portfolios using equation (2.5) for the CMV, equation (2.6) for the UMV, and equation (2.8) for the UMVTE portfolio. For each one of them, we use the estimators presented in section 2.3 to obtain the conditional expected excess returns $\boldsymbol{\mu}_t$. To exploit the wealth of information available for predictors, we make use of high dimensional datasets as our conditioning information. Hence, we assess how PCR, PLS, 3PRF, LASSO, RIDGE, and the ENet behave, comparing them with our benchmark estimator, the OLS.

Thus, for all three methodologies we replace $\boldsymbol{\mu}_t(\mathbf{Z}_{t-1})$ by its sample counterpart $\widehat{\boldsymbol{\mu}}_t(\mathbf{Z}_{t-1})$ on $\mathbf{x}_t(\mathbf{Z}_t)$, being $\widehat{\boldsymbol{\mu}}_t(\mathbf{Z}_{t-1})$ the fitted values of the estimation of the excess returns of all N risky assets in the investment set on the L lagged instruments. Here we have an important difference from previous studies. Ferson and Siegel (2009), Fletcher and Basu (2016), Abhyankar et al. (2012), Chiang (2015) among others have used standard OLS to obtain the fitted values, and a small set of instruments selected in advance as conditioning information.

[16] The authors also show that the 3PRF has a one-step closed form:

$$\hat{\boldsymbol{r}}^i = \mathbf{1}_T \bar{\boldsymbol{r}}^i + \boldsymbol{J}_T \boldsymbol{Z} \boldsymbol{J}_L \boldsymbol{Z}^\top \boldsymbol{J}_T \boldsymbol{r}^i \left(\boldsymbol{r}^{i\top} \boldsymbol{J}_T \boldsymbol{Z} \boldsymbol{J}_L \boldsymbol{Z}^\top \boldsymbol{J}_T \boldsymbol{Z} \boldsymbol{J}_L \boldsymbol{Z}^\top \boldsymbol{J}_T \boldsymbol{r} \right)^{-1} \boldsymbol{r}^{i\top} \boldsymbol{J}_T \boldsymbol{Z} \boldsymbol{J}_L \boldsymbol{Z}^\top \boldsymbol{J}_T \boldsymbol{r}^i \tag{2.13}$$

where $\bar{\boldsymbol{r}}^i = \mathbf{1}_T^\top \boldsymbol{r}^i / T$, for $\mathbf{1}_T$ a T-dimensional vector of ones, and $\boldsymbol{J}_T \equiv \boldsymbol{I}_T - \frac{1}{T} \mathbf{1}_T \mathbf{1}_T^\top$, for $\boldsymbol{I}_T$ a T-dimensional identity matrix, being $\boldsymbol{J}_L$ analogous.

[17] For further details, please refer to the original paper.

Given that returns of risky assets is given by the conditional mean plus a noise term, we can estimate the conditional second moment matrix as which is given by:

$$\begin{aligned}\mathbf{\Gamma}_t\left(\mathbf{Z}_{t-1}\right) &= \mathbb{E}\left(\mathbf{r}_t\mathbf{r}_t^\top|\mathbf{Z}_{t-1}\right)\\ &= \mathbf{\Sigma}_t\left(\mathbf{Z}_{t-1}\right)+\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top\end{aligned} \tag{2.14}$$

where $\mathbf{r}_t$ is an $N\times 1$ vector of the N risky assets, and $\mathbf{\Sigma}_t\left(\mathbf{Z}_{t-1}\right)$ is the nonsigular conditional covariance matrix of the noise. Plugging the estimates of the conditional expected excess returns on $\mathbf{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)$, we can estimate the second moment matrix, assuming that the conditional covariance matrix $\mathbf{\Sigma}_t\left(\mathbf{Z}_{t-1}\right)$ is constant[18]. Notice, that $\mathbf{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)$ is time-varying due to the time-varying characteristic of $\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)$. We estimate the conditional covariance matrix $\mathbf{\Sigma}_t\left(\mathbf{Z}_{t-1}\right)$ as the estimate of the residual covariance matrix obtained with the regressions for $\hat{\boldsymbol{\mu}}_t\left(\mathbf{Z}_{t-1}\right)$.

Following Fletcher and Basu (2016), we also calculate $\mathbb{E}\left(\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top\left(\mathbf{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1}\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)\right)$ on equation (2.6) as the average value of $\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)^\top\left(\mathbf{\Gamma}_t\left(\mathbf{Z}_{t-1}\right)\right)^{-1}\boldsymbol{\mu}_t\left(\mathbf{Z}_{t-1}\right)$ during the estimation window. Following Kirby and Ostdiek (2012) and Fletcher and Basu (2016) we set a small monthly target return. Thus, in equation (2.5) for the CMV, we set the conditional expected excess returns $\mathbb{E}\left(r_{p,t+1}|\mathbf{Z}_t\right)$ equal to 0.5%. Similarly, in equation (2.6) for the UMV framework, we also set the unconditional expected excess returns $\mathbb{E}\left(r_{p,t+1}\right)$ equal to 0.5%. For the UMVTE, we set α_p to be $\mathbb{E}\left(R_{p,t+1}-R_{b,t+1}\right)=0.5\%$, where we consider the tracking portfolio to be the risk-free asset of the economy. In the Internet Appendix, we assess how the robustness of our results for different values of α, i.e., $\alpha=1\%$, $\alpha=1.5\%$, and $\alpha=2\%$. There are no significant changes.

2.4.1 Validation

All the techniques described in section 2.3 have hyperparameters that need to be chosen to define the final model. This process of determining the best hyperparameters is commonly known as *tuning*. As previously mentioned, penalized regressions do not have a closed form solution, mainly because of the penalty λ imposed in the objective function. For ENet, λ and α need to be chosen, while for LASSO and RIDGE, only λ. For PCR, the number of principal components m in $\mathbf{V}^{(m)}=\left[\boldsymbol{\omega}^{(1)},\ldots,\boldsymbol{\omega}^{(m)}\right]$ needs to be defined to determine the final regression model. For PLS, the m number of principal directions should also be chosen[19].

To tune the hyperparameters above, we recursively define the training and testing periods.

[18]Previous studies such as Ferson and Siegel (2009) and Fletcher and Basu (2016) have found that time-varying conditional covariance matrix leads to marginal changes.

[19]For the 3PRF, given its complexity, we pre-determined a small number of three latent factors to impose sparsity.

The prevailing approach in this literature is to use the tuning subsample to adaptively determine the hyperparameters. At each $\tau \in \tau_{OOS}$, where τ_{OOS} is the OOS subsample, we use all the previous information up to $\tau - 1$ to build the training group. Within this group, the data ranges from $t = 1, 2, \ldots, \tau - 1$. For example, for the first estimation window, 389 observations[20] are used as the training sample to obtain the next t conditional mean forecast.

As Bergmeir et al. (2018) point out the usage for time-series applications, we decided to make use of K-fold cross-validation in the training subsamples. For a list of possible different hyperparameters for each estimator (i.e., whether λ, α, or the number of components m), the mean square error (MSE) is computed in this training sample. The hyperparameter(s) value(s) that produce(s) the lowest MSE is(are) chosen. Thus, in the first estimation window, we validate the hyperparameters splitting the 389 training observations into 10 subsamples (folds) and selecting the chosen hyperparameter(s) value(s) that yield(s) the smallest MSE among them. For the next t, we enlarge our training window by one month, and repeat the K-fold cross-validation in the training subsamples. This recursive process of refitting is done monthly until the last observation in our full sample (2017:12). This recursion scheme allows us to incorporate the most recent data from the set of conditioning information $\mathbf{Z}$ when enlarging the training samples by one t (i.e., one month), while dynamically selecting the best hyperparameter(s) for a given objective function. Finally, the observation τ is used for testing using the best-tunned model estimates for all techniques. Therefore, at each recursion, the conditional mean forecast with tuned hyperparameters obtained from the training samples is tested in the next t, in our out-of-sample data $\tau_{OOS} = \tau \ldots T$ (1996:01-2017:12).

2.4.2 Datasets

As it is standard in this literature, we use size- and value- sorted and grouped portfolios from Ken French's website to represent our universe of risky assets of U.S. stocks. In order to obtain a large cross-sectional dispersion in average returns and a higher degree of granularity, the chosen number N of available portfolios are 25 and 100[21]. In the Internet Appendix, we

[20] From 1963:07 - 1995:12 there are 390 observations, as the monthly series of the high-dimensional conditioning information sets $\mathbf{Z}_{t-1}$ is lagged one period, we drop one t, yielding 389 observations for the first estimation window.

[21] Having a set of $N = 100$ and $N = 25$ portfolios, we expect a better representation of the entire tradable assets grouped by importance in size and value. It has been standard practice in empirical applications in this literature the use of these sorted portfolios. For example, DeMiguel et al. (2009) use a similar set of datasets for evaluating their framework of constructing portfolios in the presence of estimation error. Kelly and Pruitt (2013) use 6, 25, 100 portfolios similarly sorted in their out-of-sample exercise. Abhyankar et al. (2012) use 5-industry portfolios and the 6 sorted on size and value in their empirical study of optimal use of conditioning information to construct dynamically managed unconditional efficient portfolios. Several other studies made use of similar set of sorted portfolios ranging from $N = 6$ to $N = 100$ (Kirby and Ostdiek, 2012;

extend the analysis for other portfolios with a smaller number of assets. Precisely, we assess *5 Industry Portfolios* and *6 portfolios formed on size and boot-to-market*. We use monthly returns for these portfolios, and the risk-free rate used to compute the excess returns is also extracted from Ken French's website. Summary statistics for all the portfolios is presented in the Internet Appendix[22].

With high-dimensional datasets and a vast number of suggested predictors, we use several instruments and split them into groups. The first group is formed by the predictive candidates used in Welch and Goyal (2007) and made available on Amit Goyal's website. We consider a set of 10 variables: book-to-market (b/m), default return spread (dfr), default yield spread (dfy), inflation ($infl$), long-term yield (lty), long-term rate of returns (ltr), net equity expansion ($ntis$), stock variance ($svar$), term spread (tms) and treasury bill rate (tbl). Many subsequent studies made use of this list of variables in empirical works after Welch and Goyal (2007) comprehensive study of the performance of these variables been suggested by the academic literature (Kelly and Pruitt, 2013; Pettenuzzo and Ravazzolo, 2016; Peñaranda and Sentana, 2016; Faria and Verona, 2017; Cooper and Maio, 2019).

The second group makes use of the FRED-MD dataset as presented in McCracken and Ng (2016). The FRED-MD is a large macroeconomic database and monthly updated by the FRED[23] that shares the predictive content of the Stock-Watson dataset (Stock and Watson (1996)). It is a balanced panel consisting of 128 macroeconomic and financial variables. The variables are split into 8 groups: (1) output and income, (2) labor market, (3) housing, (4) consumption, orders, and inventories, (5) money and credit, (6) interest and exchange rates, (7) prices, and (8) stock market. In order to mitigate the consequences of the presence of cointegration, we apply the suggested transformations to obtain stationary time series. The transformations can be grouped in seven categories: (i) no transformation; (ii) Δx_t; (iii) $\Delta^2 x_t$; (iv) $log(x_t)$, (v) $\Delta log(x_t)$, (vi) $\Delta^2 log(x_t)$, and (vii) $\Delta(x_t/x_{t-1} - 1)$[24].

The third group is based on the economic policy uncertainty measure (EPU) from Baker et al. (2016) and related indexes. The EPU is an index that proxies for movements in policy-related economic uncertainty for U.S., being based on newspaper coverage frequency. In the sense of the EPU, we also use the financial stress indicator (FSI) for the U.S from Püttmann (2018). The essence of the FSI is being an indicator of negative financial sentiment. It is based on the reporting in five major US newspapers[25]. Püttmann (2018) shows that the FSI

Goto and Xu, 2015; Peñaranda, 2016; Coqueret, 2015; McLean and Pontiff, 2016; Dai and Wang, 2019).

[22]Table 2.9.2 in the Internet Appendix.

[23]`https://research.stlouisfed.org/econ/mccracken/fred-databases/`

[24]Table 2.9.1 in the Internet Appendix 2.9 presents the entire list of variables, groups, corresponding transformations, and sample means and standard deviations for the full sample, IS and OOS.

[25]Boston Globe, Chicago Tribune, Los Angeles Times, Wall Street Journal, and Washington Post.

is a robust indicator, such that an increase in negative financial sentiment is followed by a fall in output, higher unemployment, lower stock market returns, and rising corporate bond spreads.

2.5 Empirical Results

2.5.1 Out-of-Sample Analysis - Sharpe Ratios

Performance Evaluation

To answer whether we can exploit signals provided by a large set of conditioning information using penalized estimators to impose sparsity or latent factors models with small numbers of factors, we first evaluate how the estimators behave with different sets of conditioning information to provide mean-standard deviation ratios of the returns in the out-of-sample data.

Table 2.2 shows the Sharpe ratios of the seven estimators for all three mean-variance efficient portfolio frameworks. For the 25 Portfolios formed on size and BTM, we see that only the OLS failed to deliver a significant SR for the MVATE with large sets of conditioning information. Overall we note that all the other estimators provided similar SR, while the PCR and the PLS alternate to deliver a higher ratio to the remaining estimators. Using Goyal's instruments we that OLS, 3PRF, LASSO, RIDGE, and ENet produce similar SR. As expected, OLS cannot handle large set of covariates, what makes the SR drop whenever we increase the set of conditioning information.

For the 100 Portfolios formed on size and BTM, a similar pattern stands out. OLS cannot extract efficiently information from a large set of instruments, making its SR to be considerably lower and not statistically significant (in most cases) compared to the rest of the estimators. Again, we notice that the PCR and the PLS deliver higher SR in most cases. Differently from the case with a smaller number of risky assets, we see that all estimators, except PLS and PCR, failed to extract signals and deliver statistically significant SRs when using a small set of instruments (Goyal). An important point to notice is that, in general, as we move to a high-dimensional setting, increasing the number of instruments, the dimensionality reduction approaches deliver a higher SR.

Statistical Inference for the Difference of Sharpe Ratios

The results from the previous section have shown that some of the methods of estimating the conditional mean driving the formation of efficient portfolios delivered a clear improvement when compared among them in our out-of-sample experiment in terms of Sharpe ratios. Given our universe of mean-variance efficient portfolios with respect to high-dimensional conditioning information sets, we assess if the monthly updated optimal weights that generated the risk-return tradeoff are different from any of the results from the other methods and sets of conditioning information. To answer if the Sharpe ratio generated by an estimator is statistically different from another SR generated by any other estimator and conditioning set, we need to test the difference between them. We follow Ledoit and Wolf (2008) to perform these tests. In short, consider the Sharpe ratios of the optimal portfolios generated by two different estimators, say a and b. The difference between SR_a and SR_b is given by

$$\Delta = SR^{(a)} - SR^{(b)} = \frac{\mu^{(a)}}{\sigma^{(a)}} - \frac{\mu^{(b)}}{\sigma^{(b)}} \tag{2.15}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the OOS unconditional mean and variance of the excess returns from the optimal portfolio generated by an estimator. Ledoit and Wolf (2008) suggest the construction of a studentized time series bootstrap confidence interval for the difference of the SRs. This method has been shown to be robust when returns have tails heavier than the normal distribution or are of time series nature. The bootstrap data is generated using the circular block bootstrap of Politis and Romano (1992). The two-sided distribution function of the studentized statistic can be obtained via bootstrap as follows:

$$f\left(\frac{|\widehat{\Delta} - \Delta|}{se\left(\widehat{\Delta}\right)}\right) \approx f\left(\frac{|\widehat{\Delta}^{\text{boot}} - \Delta|}{se\left(\widehat{\Delta}^{\text{boot}}\right)}\right) \tag{2.16}$$

where $f(\cdot)$ is the distribution of a random variable, Δ is populational difference between SR_a and SR_b, $\widehat{\Delta}$ is the sample counterpart of this difference obtained in the data in the estimation window, and $\widehat{\Delta}^{\text{boot}}$ is the estimated difference computed from bootstrap. The standard errors are denoted by $se(\cdot)$. Out of the distribution obtained from the bootstrap in the equation (2.16), we can find the confidence interval for Δ and the p-values of the test[26].

Table 2.3 reports for the dataset with 100 portfolios formed on Size/BTM for the null hypothesis given by no difference in Sharpe rations ($H_0 : \Delta = 0$). Table 2.4 presents the test of difference of SRs for the dataset with 25 portfolios formed on Size/BTM[27]. In order to see that the Sharpe ratios generated using not only two different estimators, but also different set of conditioning information, these tables also present the tests of difference between these cases too. We notice in table 2.3 that in the case with 100 risky assets, PLS and PCR

[26] For further details in this procedure, see Ledoit and Wolf (2008).

[27] In the Internet Appendix, tables 2.8.2 and table 2.8.3 report the same tests for the dataset with 6 portfolios formed on Size/BTM and the dataset with 5 industry portfolios respectively.

indeed generate different SRs for all other 5 estimators when using the Goyal variables as conditioning information. The p-values of the difference of Sharpe ratios are always low, providing strong evidence for this difference. This fact is valid for all three mean-variance frameworks. Combined with the results from table 2.2, we can claim that the SRs of 0.245, 0.241 and 0.279 (CMV, UMV, MVATE respectively) for the PLS, and the SRs of 0.303, 0.315 and 0.312 (CMV, UMV, MVATE respectively) produced by the PCR are indeed higher than the remaining ones.

Table 2.2: Out-of-Sample Sharpe ratios delivered by each estimator and set of conditioning information

	25 Portfolios Formed on Size and Book-to-Market						100 Portfolios Formed on Size and Book-to-Market					
	CMV		UMV		MVATE		CMV		UMV		MVATE	
Estimator	*SR*	*p*-val	*SR*	*p*-val	*SR*	*p*-val	*SR*	*p*-val	*SR*	*p*-val	*SR*	*p*-val
Panel A: Goyal												
OLS	0.307	0.000	0.293	0.000	0.270	0.000	0.076	0.221	0.072	0.245	0.003	0.962
3PRF	0.327	0.000	0.330	0.000	0.307	0.000	0.117	0.060	0.133	0.032	0.059	0.341
PLS	0.307	0.000	0.322	0.000	0.300	0.000	0.245	0.000	0.241	0.000	0.279	0.000
PCR	0.383	0.000	0.386	0.000	0.380	0.000	0.303	0.000	0.315	0.000	0.312	0.000
LASSO	0.276	0.000	0.279	0.000	0.278	0.000	0.087	0.158	0.100	0.107	0.084	0.173
RIDGE	0.313	0.000	0.311	0.000	0.299	0.000	0.060	0.333	0.083	0.178	0.044	0.472
ENET	0.315	0.000	0.316	0.000	0.259	0.000	0.097	0.117	0.103	0.096	0.043	0.483
Panel B: FRED-MD												
OLS	0.161	0.010	0.166	0.008	0.006	0.924	0.128	0.040	0.128	0.040	0.047	0.446
3PRF	0.282	0.000	0.262	0.000	0.261	0.000	0.210	0.001	0.197	0.002	0.135	0.029
PLS	0.296	0.000	0.299	0.000	0.285	0.000	0.350	0.000	0.346	0.000	0.230	0.000
PCR	0.371	0.000	0.361	0.000	0.374	0.000	0.203	0.001	0.209	0.001	0.198	0.002
LASSO	0.215	0.001	0.209	0.001	0.212	0.001	0.312	0.000	0.305	0.000	0.240	0.000
RIDGE	0.274	0.000	0.277	0.000	0.275	0.000	0.274	0.000	0.265	0.000	0.188	0.003
ENET	0.280	0.000	0.267	0.000	0.286	0.000	0.307	0.000	0.301	0.000	0.129	0.037
Panel C: All Instruments												
OLS	0.201	0.003	0.204	0.003	-0.054	0.419	0.118	0.078	0.118	0.078	-0.026	0.694
3PRF	0.290	0.000	0.278	0.000	0.254	0.000	0.240	0.000	0.228	0.000	0.145	0.020
PLS	0.331	0.000	0.303	0.000	0.314	0.000	0.309	0.000	0.315	0.000	0.263	0.000
PCR	0.307	0.000	0.293	0.000	0.310	0.000	0.269	0.000	0.272	0.000	0.271	0.000
LASSO	0.231	0.000	0.230	0.000	0.232	0.000	0.323	0.000	0.315	0.000	0.253	0.000
RIDGE	0.305	0.000	0.306	0.000	0.299	0.000	0.243	0.000	0.247	0.000	0.117	0.079
ENET	0.302	0.000	0.280	0.000	0.302	0.000	0.359	0.000	0.352	0.000	0.179	0.008

Table 2.2 summarises the OOS (Jan-1996 to Dec-2017) Sharpe ratios (SR) by estimator and optimal portfolio framework (CMV, UMV and MVATE) for both portfolios. Panel A reports the Sharpe ratios generated when the variables from Goyal's website are used as **Z**. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. Panel B presents the Sharpe ratios obtained using the FRED-MD variables. The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, panel C shows the Sharpe rations when all variables are used as conditioning information. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). The *p*-val is the p-value from the two-sided test of the SR ($H_0 : SR = 0$).

Table 2.3: Test for the difference of the Sharpe ratios - OOS (Jan-1996 - Dec-2017) - 100 Portfolios

Panel A: CMV

Goyal

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.092	0.001*	0.000*	0.594	0.428	0.358
	3PRF		0.001*	0.000*	0.105	0.002*	0.322
	PLS			0.111	0.001*	0.000*	0.002*
	PCR				0.000*	0.000*	0.001*
	LASSO					0.087	0.407
	RIDGE						0.022

FRED-MD

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.023	0.000*	0.042	0.004*	0.059	0.002*
	3PRF		0.000*	0.092	0.030	0.165	0.021
	PLS			0.204	0.596	0.819	0.579
	PCR				0.928	0.800	0.959
	LASSO					0.113	0.013
	RIDGE						0.003*
FRED-MD	OLS	0.166	0.001*	0.425	0.022	0.138	0.022
	3PRF		0.003*	0.930	0.102	0.469	0.112
	PLS			0.024	0.587	0.406	0.510
	PCR				0.317	0.561	0.291
	LASSO					0.564	0.901
	RIDGE						0.578

All Instr.

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.004*	0.000*	0.005*	0.001*	0.038	0.000*
	3PRF		0.003*	0.010	0.010	0.183	0.005*
	PLS			0.872	0.654	0.806	0.451
	PCR				0.847	0.371	0.877
	LASSO					0.116	0.002*
	RIDGE						0.000*
FRED-MD	OLS	0.070	0.014	0.128	0.002*	0.109	0.002*
	3PRF		0.122	0.515	0.046	0.705	0.023
	PLS			0.031	0.241	0.043	0.582
	PCR				0.309	0.818	0.194
	LASSO					0.034	0.670
	RIDGE						0.347
All Instr.	OLS	0.037	0.010	0.107	0.003*	0.132	0.002*
	3PRF		0.343	0.850	0.200	0.847	0.101
	PLS			0.591	0.818	0.348	0.443
	PCR				0.585	0.815	0.390
	LASSO					0.087	0.187
	RIDGE						0.011

Panel B: UMV

Goyal

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.013	0.001*	0.000*	0.224	0.590	0.190
	3PRF		0.006*	0.001*	0.087	0.013	0.138
	PLS			0.082	0.001*	0.000*	0.004*
	PCR				0.001*	0.000*	0.002*
	LASSO					0.349	0.788
	RIDGE						0.184

FRED-MD

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.029	0.000*	0.023	0.005*	0.063	0.002*
	3PRF		0.000*	0.137	0.069	0.255	0.048
	PLS			0.311	0.614	0.856	0.594
	PCR				0.917	0.632	0.869
	LASSO					0.169	0.029
	RIDGE						0.016
FRED-MD	OLS	0.230	0.001*	0.380	0.027	0.161	0.025
	3PRF		0.001*	0.885	0.082	0.436	0.078
	PLS			0.037	0.559	0.358	0.477
	PCR				0.373	0.639	0.334
	LASSO					0.519	0.915
	RIDGE						0.530

All Instr.

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.005*	0.001*	0.002*	0.001*	0.030	0.000*
	3PRF		0.028	0.022	0.035	0.254	0.017
	PLS			0.978	0.690	0.860	0.476
	PCR				0.660	0.284	0.961
	LASSO					0.165	0.005*
	RIDGE						0.002*
FRED-MD	OLS	0.097	0.013	0.108	0.004*	0.094	0.003*
	3PRF		0.074	0.356	0.035	0.489	0.016
	PLS			0.023	0.178	0.040	0.493
	PCR				0.375	0.820	0.234
	LASSO					0.053	0.680
	RIDGE						0.337
All Instr.	OLS	0.053	0.012	0.090	0.004*	0.116	0.003*
	3PRF		0.236	0.669	0.176	0.941	0.085
	PLS			0.608	0.992	0.280	0.546
	PCR				0.660	0.810	0.433
	LASSO					0.133	0.174
	RIDGE						0.017

Panel C: MVATE

Goyal

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.030	0.000*	0.000*	0.082	0.238	0.174
	3PRF		0.000*	0.000*	0.493	0.601	0.503
	PLS			0.413	0.000*	0.000*	0.000*
	PCR				0.001*	0.000*	0.000*
	LASSO					0.102	0.197
	RIDGE						0.963

FRED-MD

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.019	0.001*	0.020	0.000*	0.011	0.077
	3PRF		0.002*	0.053	0.009*	0.105	0.349
	PLS			0.101	0.668	0.321	0.081
	PCR				0.385	0.150	0.062
	LASSO					0.273	0.577
	RIDGE						0.275
FRED-MD	OLS	0.081	0.137	0.282	0.008*	0.031	0.358
	3PRF		0.260	0.547	0.028	0.295	0.928
	PLS			0.537	0.901	0.662	0.275
	PCR				0.678	0.926	0.528
	LASSO					0.311	0.059
	RIDGE						0.374

All Instr.

		3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.009*	0.001*	0.002*	0.000*	0.102	0.018
	3PRF		0.016	0.005*	0.006*	0.449	0.137
	PLS			0.371	0.519	0.094	0.243
	PCR				0.320	0.040	0.138
	LASSO					0.786	0.413
	RIDGE						0.176
FRED-MD	OLS	0.079	0.003*	0.130	0.005*	0.357	0.127
	3PRF		0.047	0.221	0.024	0.631	0.583
	PLS			0.986	0.853	0.185	0.421
	PCR				0.737	0.480	0.800
	LASSO					0.012	0.283
	RIDGE						0.796
All Instr.	OLS	0.017	0.000*	0.041	0.001*	0.028	0.025
	3PRF		0.078	0.276	0.052	0.521	0.688
	PLS			0.951	0.856	0.020	0.274
	PCR				0.872	0.224	0.464
	LASSO					0.010	0.143
	RIDGE						0.321

p-val	$[0 \leq p \leq 0.05]$	$(0.05 < p < 0.1)$	$(0.1 \leq p \leq 1]$
Gradient color bounds			

Test for the differences of the Sharpe ratios of the OOS (Jan-1996 - Dec-2017) returns of the efficient portfolios formed from the dataset with 100 Size/BTM portfolios using 7 different estimators (OLS, 3PRF, PLS, PCR, LASSO, RIDGE and ENet) and three different set of conditioning information(Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios. Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

Table 2.4: Test for the difference of the Sharpe ratios - OOS (Jan-1996 - Dec-2017) - 25 Portfolios

Panel A: CMV

		Goyal						FRED-MD						All Instr.					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.487	0.998	0.153	0.592	0.862	0.836	0.715	0.871	0.280	0.116	0.567	0.650	0.783	0.655	0.397	0.148	0.328	0.819
	3PRF		0.677	0.234	0.354	0.732	0.790		0.638	0.449	0.082	0.353	0.447		0.585	0.333	0.098	0.281	0.525
	PLS			0.066	0.673	0.916	0.898			0.149	0.151	0.518	0.643			0.273	0.180	0.312	0.875
	PCR				0.160	0.257	0.314				0.008*	0.034	0.053				0.009*	0.023	0.113
	LASSO					0.553	0.311					0.975	0.962					0.629	0.692
	RIDGE						0.949						0.608						0.790
FRED-MD	OLS							0.077	0.112	0.007*	0.454	0.099	0.111	0.053	0.116	0.176	0.436	0.123	0.047
	3PRF								0.833	0.092	0.220	0.797	0.971		0.582	0.837	0.340	0.783	0.693
	PLS									0.163	0.234	0.719	0.807			0.803	0.310	0.799	0.951
	PCR										0.008*	0.033	0.074				0.012	0.103	0.252
	LASSO											0.213	0.107					0.094	0.022
	RIDGE												0.900						0.406
All Instr.	OLS													0.139	0.139	0.190	0.352	0.127	0.079
	3PRF														0.697	0.976	0.265	0.998	0.879
	PLS															0.724	0.302	0.678	0.991
	PCR																0.415	0.975	0.670
	LASSO																	0.377	0.104
	RIDGE																		0.604

Panel B: UMV

		Goyal						FRED-MD						All Instr.					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.190	0.510	0.074	0.783	0.611	0.539	0.652	0.931	0.267	0.165	0.785	0.665	0.802	0.572	0.496	0.233	0.535	0.714
	3PRF		0.847	0.211	0.301	0.612	0.718		0.639	0.608	0.071	0.382	0.330		0.364	0.327	0.096	0.311	0.307
	PLS			0.144	0.532	0.819	0.927			0.433	0.074	0.413	0.344			0.211	0.085	0.231	0.415
	PCR				0.116	0.199	0.250				0.006*	0.054	0.025				0.007*	0.033	0.039
	LASSO					0.581	0.312					0.975	0.870					0.775	0.967
	RIDGE						0.873						0.516						0.573
FRED-MD	OLS							0.189	0.093	0.012	0.564	0.109	0.181	0.106	0.223	0.240	0.507	0.106	0.120
	3PRF								0.568	0.074	0.383	0.662	0.939		0.604	0.677	0.588	0.345	0.745
	PLS									0.189	0.192	0.681	0.616			0.485	0.254	0.674	0.640
	PCR										0.013	0.049	0.078				0.018	0.119	0.167
	LASSO											0.183	0.190					0.057	0.063
	RIDGE												0.841						0.766
All Instr.	OLS													0.223	0.272	0.285	0.518	0.135	0.195
	3PRF														0.888	0.990	0.373	0.737	0.997
	PLS															0.892	0.454	0.964	0.994
	PCR																0.465	0.806	0.875
	LASSO																	0.217	0.247
	RIDGE																		0.948

Panel C: MVATE

		Goyal						FRED-MD						All Instr.					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.262	0.585	0.070	0.890	0.499	0.863	0.903	0.831	0.114	0.376	0.935	0.796	0.803	0.935	0.885	0.495	0.736	0.747
	3PRF		0.891	0.150	0.599	0.850	0.455		0.764	0.289	0.157	0.590	0.740		0.721	0.673	0.232	0.504	0.792
	PLS			0.056	0.755	0.985	0.600			0.092	0.172	0.620	0.810			0.403	0.230	0.338	0.916
	PCR				0.175	0.198	0.143				0.009*	0.041	0.072				0.012	0.028	0.126
	LASSO					0.728	0.753					0.969	0.915					0.628	0.723
	RIDGE						0.484						0.843						0.959
FRED-MD	OLS							0.000*	0.001*	0.000*	0.016	0.000*	0.001*	0.001*	0.001*	0.001*	0.007*	0.001*	0.000*
	3PRF								0.720	0.057	0.417	0.736	0.683		0.521	0.511	0.625	0.499	0.472
	PLS									0.112	0.297	0.862	0.994			0.949	0.399	0.834	0.955
	PCR										0.007*	0.035	0.091				0.011	0.081	0.212
	LASSO											0.205	0.092					0.112	0.026
	RIDGE												0.828						0.420
All Instr.	OLS													0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
	3PRF														0.535	0.541	0.706	0.498	0.552
	PLS															0.950	0.436	0.811	0.874
	PCR																0.377	0.846	0.804
	LASSO																	0.434	0.131
	RIDGE																		0.610

p-val	$[0 \leq p \leq 0.05]$	$(0.05 < p < 0.1)$	$(0.1 \leq p \leq 1]$
Gradient color bounds			

Test for the differences of the Sharpe ratios of the OOS (Jan-1996 - Dec-2017) returns of the efficient portfolios formed from the dataset with 25 Size/BTM portfolios using 7 different estimators (OLS, 3PRF, PLS, PCR, LASSO, RIDGE and ENet) and three different set of conditioning information(Goyal, FRED-MD and “All Instruments”). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios. Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, “All Instruments” is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

Figure 2.1: Distribution of Optimized Portfolio Weights - 25 Portfolios Formed on Size and Book-to-Market

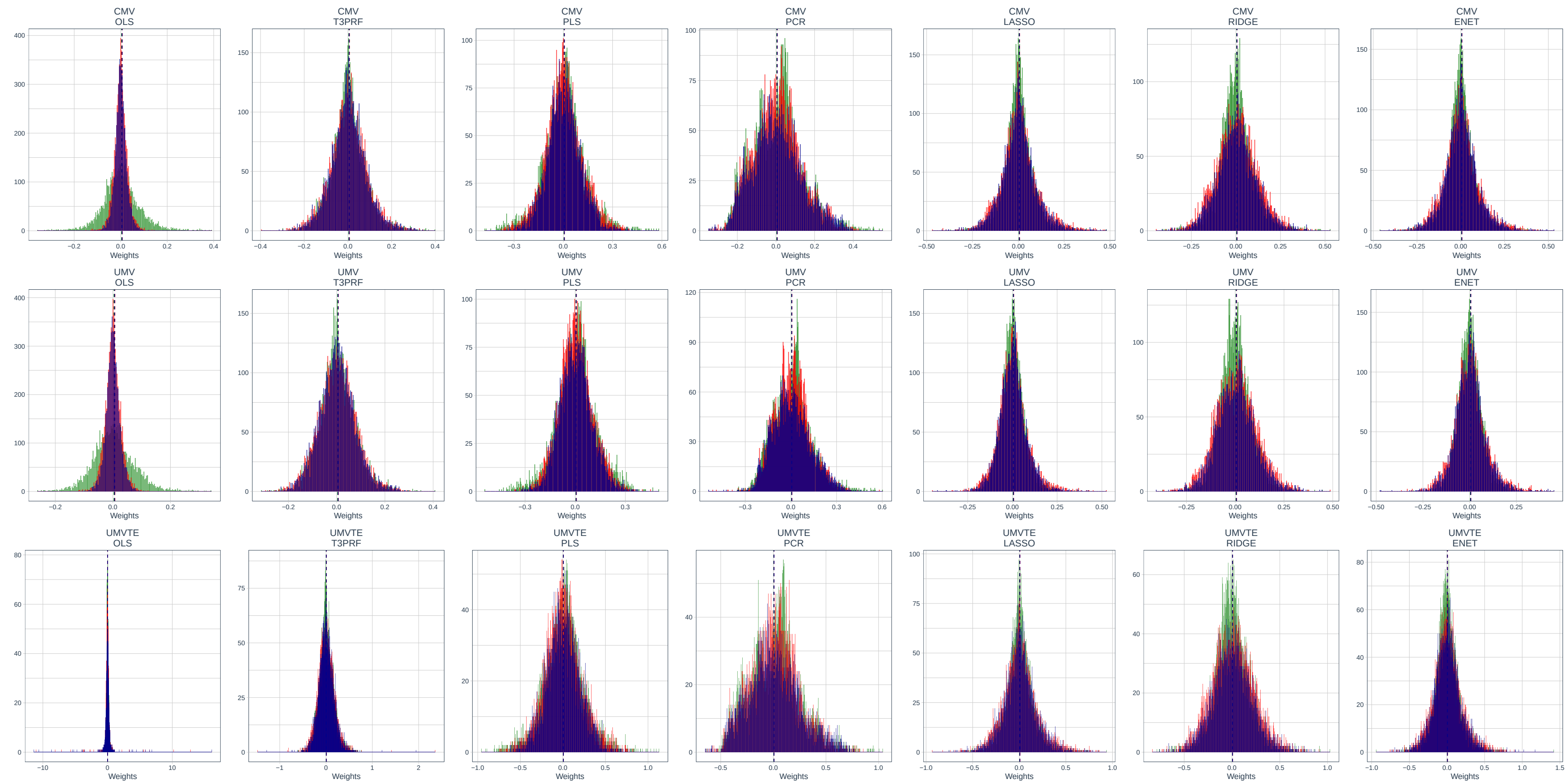

Figure 2.1 plots the distribution of the optimized portfolios weights generated by each estimator (columns) and mean-variance optimal framework (rows) for the 25 portfolios formed on Size/BTM. The first row reports the CMV strategies, the second row reports the UMV, and the third one plots the UMVTE. The set of conditioning information used in **Z** are plotted in different colors: (i) Goyal's in green, (ii) FRED-MD in red, and (iii) "All Instruments" in blue. Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. FRED-MD is a large dataset containing 128 macroeconomic and financial variables. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

Figure 2.2: Distribution of Optimized Portfolio Weights - 100 Portfolios Formed on Size and Book-to-Market

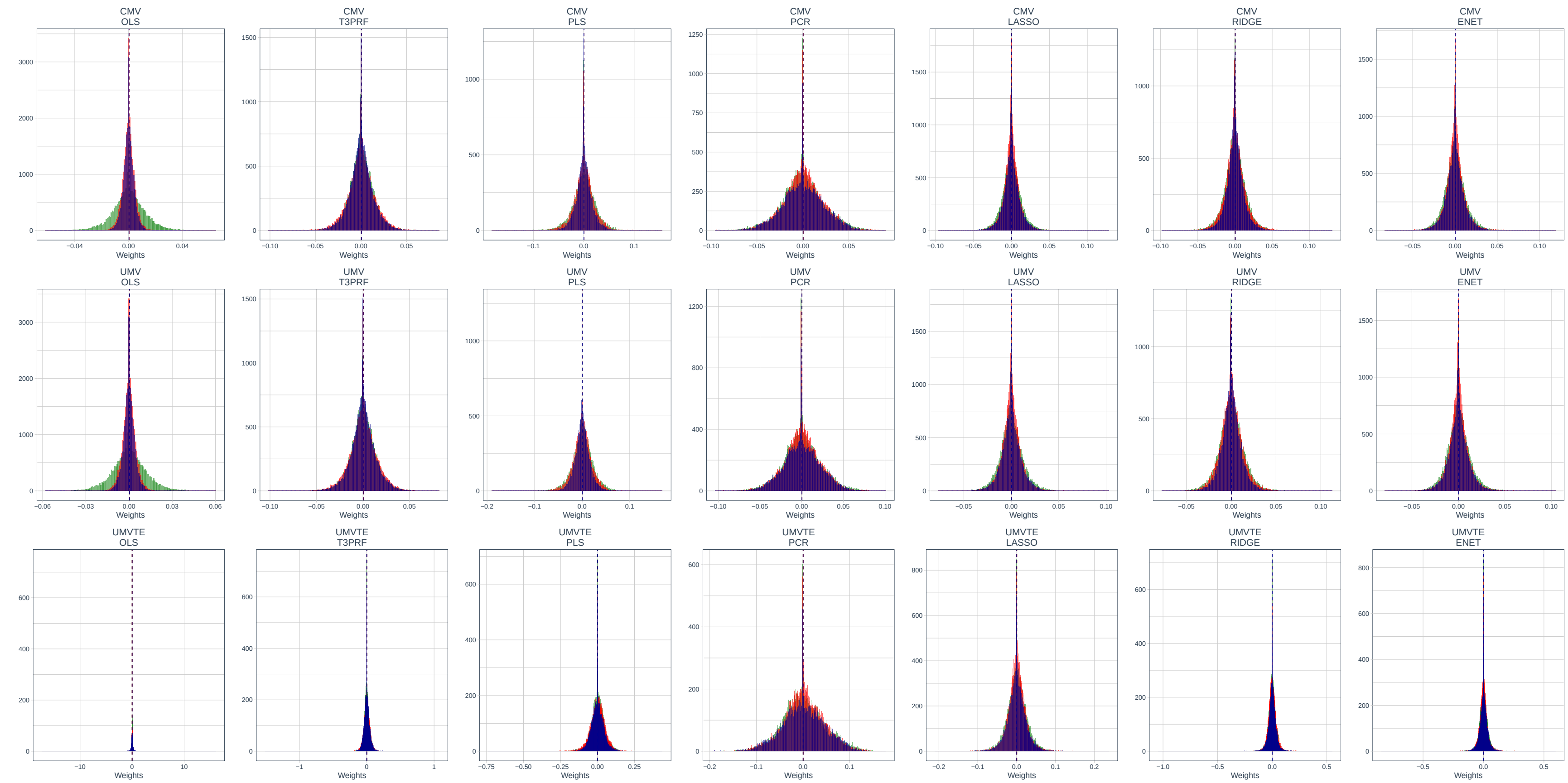

Figure 2.2 plots the distribution of the optimized portfolios weights generated by each estimator (columns) and mean-variance optimal framework (rows) for the 100 portfolios formed on Size/BTM. The first row reports the CMV strategies, the second row reports the UMV, and the third one plots the UMVTE. The set of conditioning information used in $\mathbf{Z}$ are plotted in different colors: (i) Goyal's in green, (ii) FRED-MD in red, and (iii) "All Instruments" in blue. Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. FRED-MD is a large dataset containing 128 macroeconomic and financial variables. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

2.5.2 Characteristics and the Distribution of Optimized Portfolio Weights

In order to understand why a specific estimator can generate mean-variance optimized portfolios with higher return-volatility ratios, we first analyze how the weights of the optimized portfolio behave for each estimator and set of conditioning information. For each month t in the out-of-sample period, each estimator produced a set of weights depending on the matrix $\mathbf{Z}$ used. We pooled the N weights for each t and plotted the distribution of these weights in separate figures for each dataset employed. Figure 2.1 plots the weights for the dataset with 25 portfolios formed on Size/BTM. Figure 2.2 shows the distribution of the weights $\mathbf{x}_t(\mathbf{Z}_t)$ for the dataset with 100 portfolios formed on Size/BTM[28]. Overall we see that across all three efficient strategies to build portfolios, PCR generates weights with higher variance than compared to the other estimators. Another point that can be inferred from these figures is that, in general, most of the estimators generate symmetric weights distributions. However, PCR does not, having in most cases a right-skewed distribution. We also notice that OLS is the estimator that does not respond much to predictive information, a fact that can be seen from its highly concentrated distribution around zero.

2.5.3 Variable Contribution

The results from the previous sections raise the natural question of what variables are important to each estimator. To answer this question we evaluate how each lagged variable contributed to produce the estimated conditional means in the OOS. After we obtain the estimates of the conditional mean $\widehat{\mu}_i\left(\mathbf{Z}_{t-1}\right) = \widehat{\mathbb{E}}\left(r_{i,t}|\mathbf{Z}_{t-1}\right)$ for each risky asset $i = 1, \ldots, N$ at each t, we evaluate how each lagged variable in our sets of conditioning information contribute to generate $\widehat{\boldsymbol{\mu}}_t$ to be used to form the weights in $\mathbf{x}_t(\mathbf{Z}_t)$.

We take a straightforward approach to assess this contribution. At each month t, we compute the absolute values of the estimated coefficients for each asset i and lagged variable k in the conditioning set. For the estimators that we standardized the instruments before the regression, we multiply each $\widehat{\theta}_{i,l}$ by its own standard deviation computed in the estimation window. Grouping by estimators and $\mathbf{Z}$ used, we pool all the estimates of the N assets and calculate the average of the absolute values. In order to make the comparison clear, we normalize these means to sum one. Doing so, we can rank the most influential covariates driving the conditional means on a percentage scale.

Figure 2.3 reports for the dataset with 100 portfolios formed on size and BTM the 10

[28]In the Internet Appendix, we present the weights for the dataset with 6 portfolios formed on Size/BTM and the dataset with 5 industry portfolios.

Figure 2.3: Variable Contribution by Estimator and Set of Conditioning Information - 100 Portfolios Formed on Size and Book-to-Market

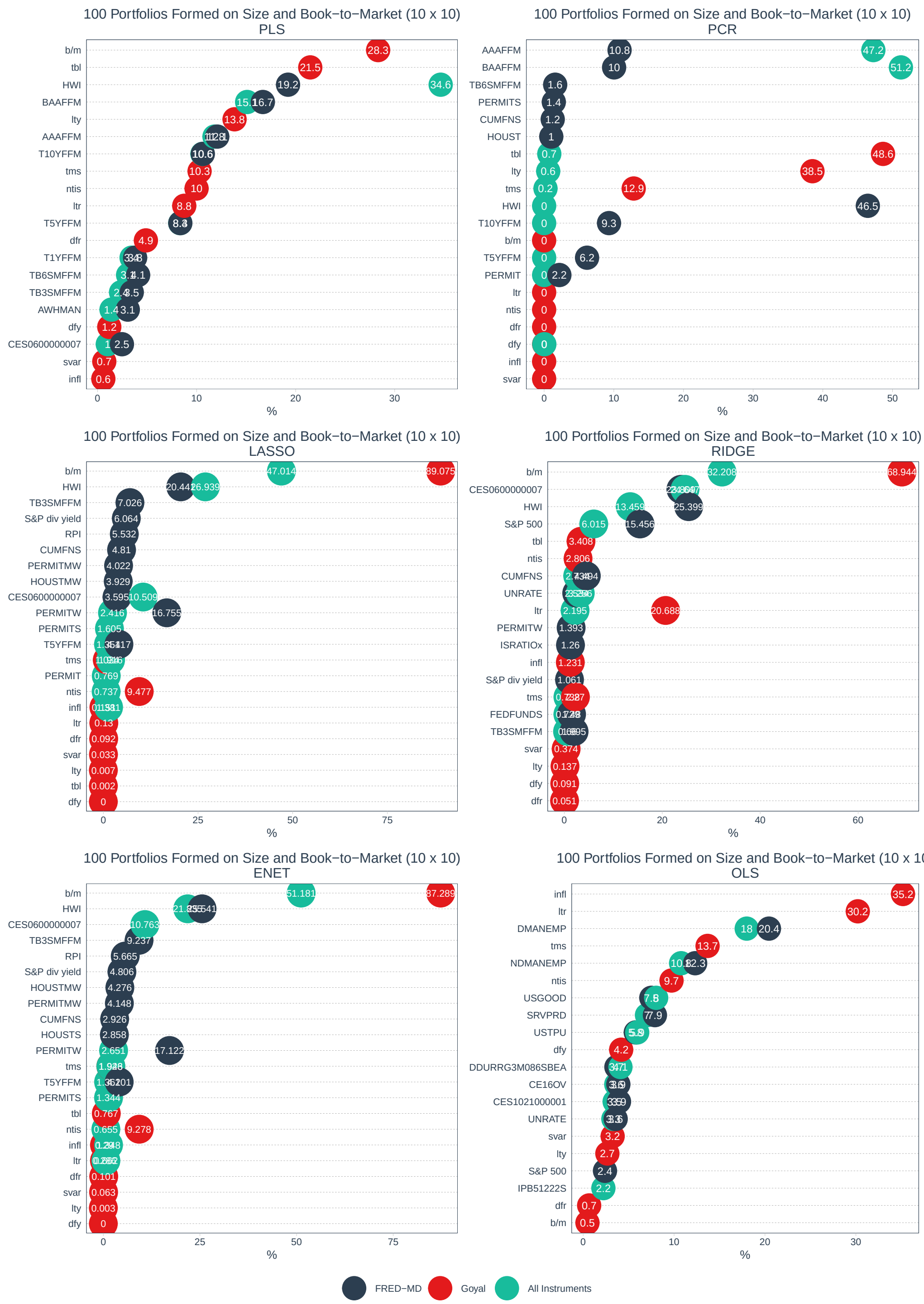

Figure 2.3 reports the 10 most influential variables by estimator (panels) and set of conditioning information used (colors) for the for the 100 portfolios formed on Size/BTM. We take a direct approach to obtain the contribution for each variable. At each month t, we compute the absolute values of the estimated coefficients for each asset i and lagged variable k in the conditioning set. We destandardize the variables whenever necessary. Grouping by estimators and $\mathbf{Z}$ used, we pool all the estimates of the N assets and calculate the average of the absolute values. In order to make the comparison clear, we normalize these means to sum one.

Figure 2.4: Variable Contribution by Estimator and Set of Conditioning Information - 25 Portfolios Formed on Size and Book-to-Market

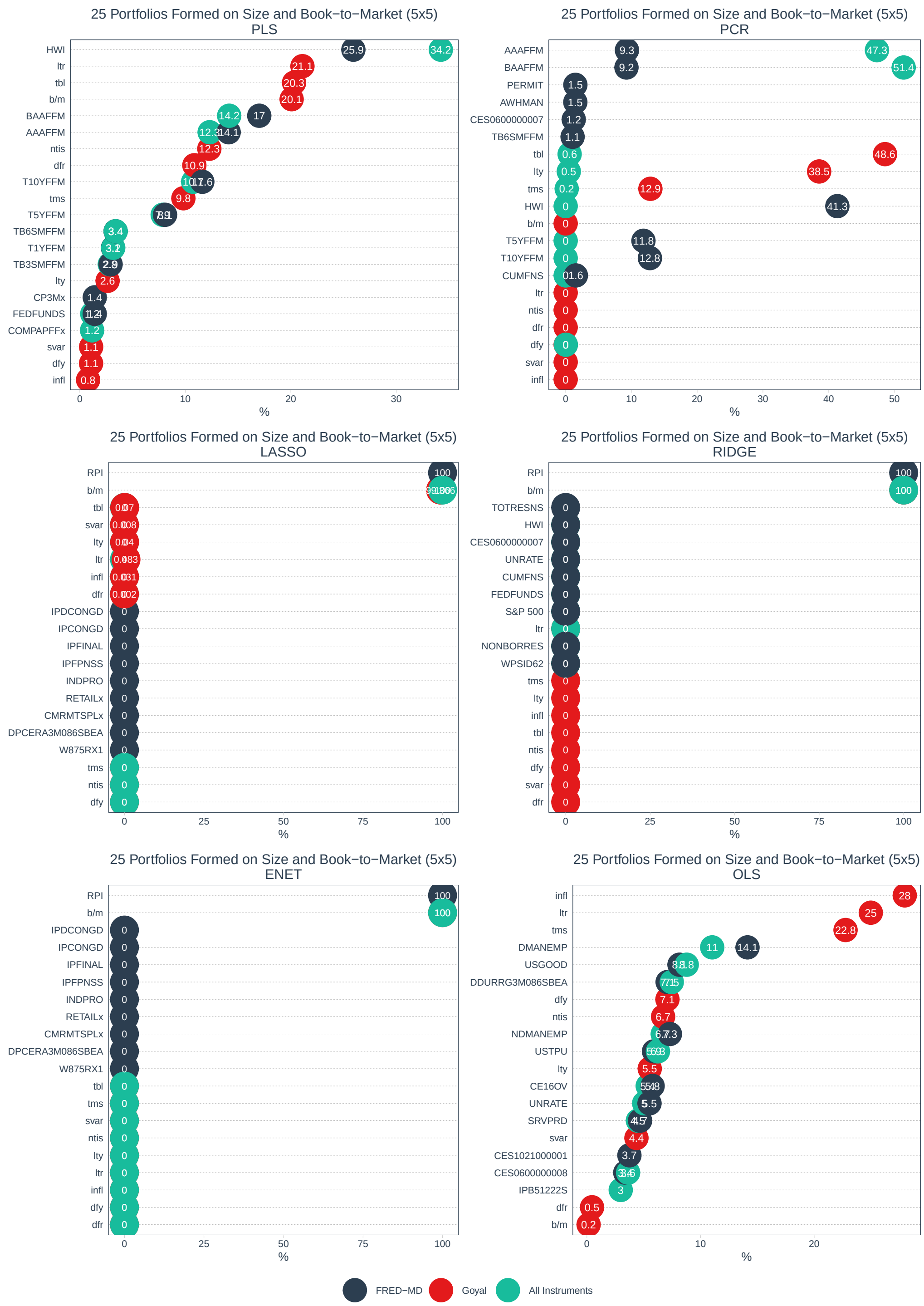

Figure 2.4 reports the 10 most influential variables by estimator (panels) and set of conditioning information used (colors) for the for the 25 portfolios formed on Size/BTM. We take a direct approach to obtain the contribution for each variable. At each month t, we compute the absolute values of the estimated coefficients for each asset i and lagged variable k in the conditioning set. We destandardize the variables whenever necessary. Grouping by estimators and $\mathbf{Z}$ used, we pool all the estimates of the N assets and calculate the average of the absolute values. In order to make the comparison clear, we normalize these means to sum one.

most influential lagged variables for all the estimators[29]. Interestingly, the PLS and PCR have some similarities regarding variable contribution, however this is not consistent across all three sets of conditioning information. Using FRED-MD, PLS and PCR ranked similarly the variables. The five most influential are: *HWI* (Help-Wanted Index for US), *AAAFFM* (Moody's Aaa Corporate Bond minus FEDFUNDS Aaa-FF spread), *BAAFFM* (Moody's Baa Corporate Bond minus FEDFUNDS Baa-FF spread), *T10YFFM* (10-Year Treasury C Minus FEDFUNDS 10 yr-FF spread), *T5YMFFM* (5-Year Treasury C minus FEDFUNDS 5 yr-FF spread) for both estimators. On the other hand, using all conditioning information, there is a clear difference in weights. The majority of the influence for PCR is split by two variables (*AAAFFM* and *BAAFFM*), while for PLS these two variables respond less than 30%, with a labor market variable (*HWI*) also having a large impact (35%), and small contributions from interest and exchange rates variables (*T1YFFM*, *TB6SMFFM*, *TB3SMFFM*). Out of the ten variables from Goyal's dataset, only three contributes to build efficient portfolios when using PCR (*tbl*, *lty* and *tms*).

For LASSO, Ridge, and Enet we see a similar pattern for all three estimators depending on which set of conditioning information was used. Finally, for OLS we see a completely different pattern compared to the previous estimators. We notice that labor market variables such as *DMANEMP* (All Employees: Durable goods), *NDMANEMP* (All Employees: Nondurable goods) and *USGOOD* (All Employees: Goods-Producing Industries) have a large impact when using large sets of predictors. These differences may provide us information on why some estimators produced better OOS mean-variance ratios.

Figure 2.4 illustrates the same analysis for the dataset with 25 portfolios formed on size and BTM[30]. The pattern for PCR and PLS is remarkably similar as seen for 100 portfolios formed on size and BTM. The same can be said regarding the OLS been driven mostly by labor market variables with high-dimensional conditioning information sets. However, for the penalized estimators we see a different story. We see that LASSO, Ridge, and ENet did a large shrinkage and selection in the set of lagged variables. For each one, the conditioning information set practically one variable was responsible for the full contribution: an output and income variable, *RPI* (Real Personal Income) using FRED-MD, and *b/m* (book-to-market) using Goyal and a combination of all lagged instruments.

Figure 2.5: Out-of-Sample R^2 - 100 and 25 Portfolios Formed on Size and Book-to-Market

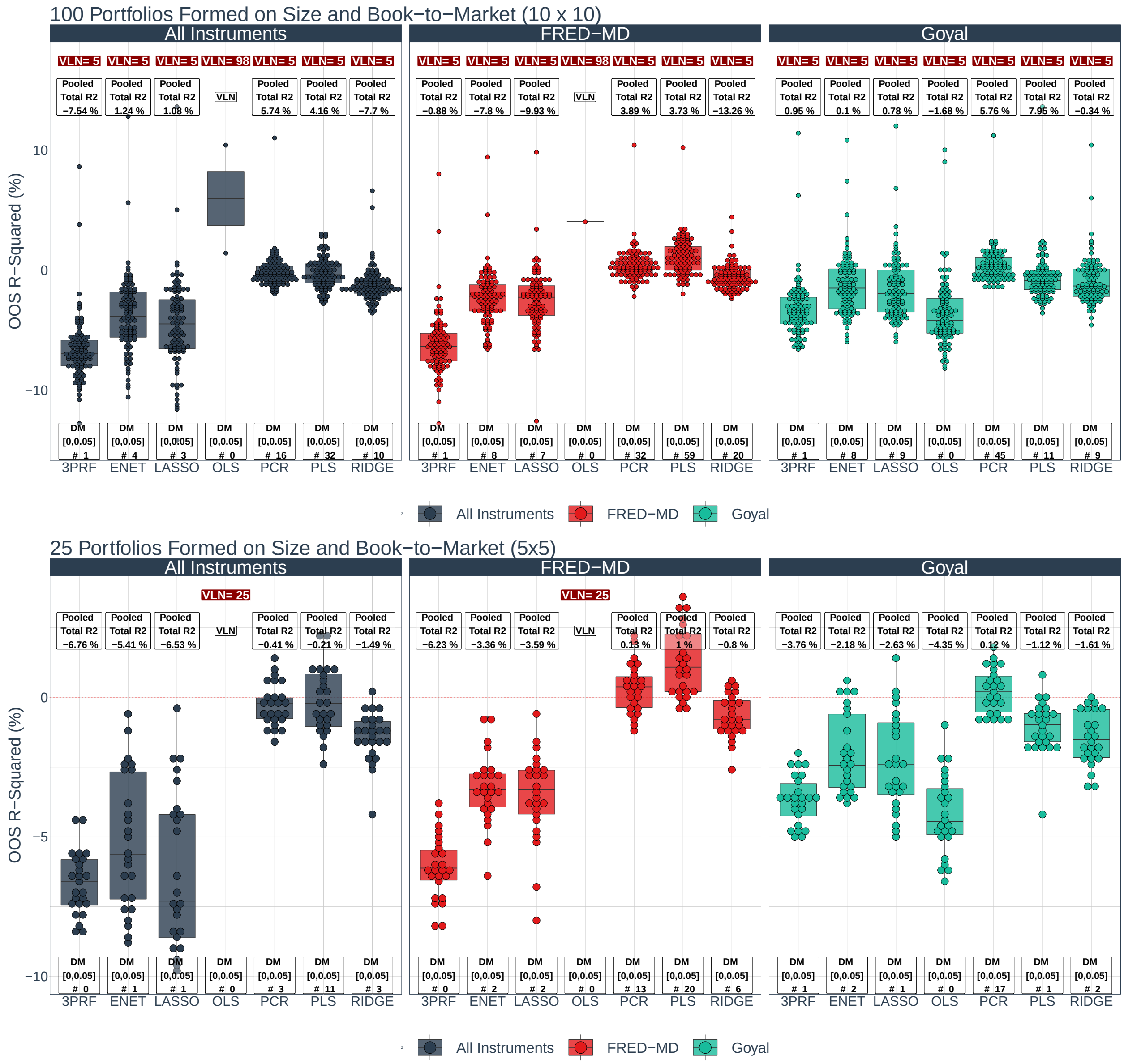

Figure 2.5 reports the OOS R^2 for the datasets with 100 and 25 portfolios formed on Size/BTM by estimator and set of conditioning information used, along with overlapping boxplots. The pooled OOS R^2 across all N risky assets in each dataset is also reported in a white box on the top of each plot. In order to make the plots readable and comparable, we filtered the OOS R^2 larger than -0.5 in absolute values and present the amount of assets that generated these very large numbers (VLN). The amount of VLN per estimator is reported in a red box on top of each plot. We also report at the bottom of each plot the number of significant cases (p-values ≤ 0.05) in terms of a forecast accuracy test. We used the Harvey et al. (1997) modified version of the Diebold and Mariano (1995) t-test among all assets for each estimator and set of conditioning information. There is no significant change when the original Diebold and Mariano (1995) test is used. We assess a one-sided test, in which the alternative hypothesis can be interpreted as our zero forecast benchmark is less accurate than each one of the seven estimators' forecasts.

2.5.4 Out-of-Sample Analysis - The Behavior of the Conditional Mean Estimates

To form the weights of the optimized portfolios, many factors could play a role to drive the formation of mean-variance efficient portfolios. Being the conditional mean one of these components, how the estimates behave? In order to answer this question, one way to evaluate the quality of the estimates is to assess the predictive performance of each estimator when using each set of lagged variables. Following Gu et al. (2018); Kelly et al. (2019) we calculate the out-of-sample R^2 for all 7 estimators and three different sets of conditioning information for each $\tau \in \tau_{OOS}$ as

$$R_i^2 = 1 - \frac{\sum_{t \in \tau_{OOS}} \left(r_{i,t+1} - \hat{r}_{i,t+1}\right)^2}{\sum_{t \in \tau_{OOS}} \left(r_{i,t+1}^2\right)} \quad , \tag{2.17}$$

where τ_{OOS} represents the OOS periods not used for testing or tuning. Usually, the OOS R^2 is computed demeaning the denominator. As Gu et al. (2018) argues, doing so can be a mistaken analysis, since the historical averages usually underperform a naive forecast of zero. Thus, a way to overcome this issue is to compare against a benchmark of zero. Given the large number of results[31], we present them in different figures.

Figure 2.5 reports the OOS R^2 for the 100 and 25 portfolios formed on Size/BTM[32]. It is clear how PLS and PCR generate higher OOS R^2 compared to most of the estimators. This result is consistent across different sets of conditioning information and sets of risky assets. The figure also reports the pooled[33] OOS R^2, in which we pooled all risky assets and calculated the generated total OOS pooled R^2. In general, we see that again PLS and PCR can generate higher and in most cases positive R^2. Recall that these are monthly returns, and therefore positive R^2, even small ones, has a meaningful economic impact.

We can use Diebold and Mariano (1995) t-test to evaluate the robustness of these results. In the same figure, we report the number of significant cases among all assets for each estimator and set of conditioning information, at the bottom of each plot. To determine significance, we consider the Harvey et al. (1997) modification of the Diebold and Mariano

[29] Except for the 3PRF which is a latent factor model

[30] In the Internet Appendix, we present similar analysis for the dataset with 6 portfolios formed on Size/BTM and the dataset with 5 industry portfolios.

[31] One for each asset $i = 1, 2, \ldots, N$, estimator (7), and set of lagged variables (3).

[32] In the Internet Appendix, we present similar analysis for the dataset with 6 portfolios formed on Size/BTM and the dataset with 5 industry portfolios respectively.

[33]

$$R_{i,t}^{2^{(\text{pooled})}} = 1 - \frac{\sum_{i=1}^{N} \sum_{t \in \tau_{OOS}} \left(r_{i,t+1} - \hat{r}_{i,t+1}\right)^2}{\sum_{i=1}^{N} \sum_{t \in \tau_{OOS}} \left(r_{i,t+1}^2\right)} \tag{2.18}$$

(1995)[34] for a one-sided test, in which the alternative hypothesis can be interpreted as our zero forecast benchmark[35] is less accurate than each one of the seven estimators' forecasts[36]. The threshold considered in these plots is for p-values ≤ 0.05. In the Internet Appendix we show the complete results for all four portfolios in the tables 2.8.4 and 2.8.5.

There are important takeaways from this exercise. First, we need to remember that our goal is to form efficient portfolios through exploiting conditional information to generate conditional mean estimates. While we would prefer strong individual asset forecasts, the formation, and thus, the right combination of these assets to form the efficient portfolio, is what dictates the final outcome. Despite this fact, we see that for some cases, specifically for PLS and PCR we are still able to obtain better results in terms of OOS R^2, and in terms of statistically significant improvements in accuracy when compared to our benchmark. Finally, it is not surprising that we obtained some negative OOS R^2. In fact, this is consistent with Welch and Goyal (2007) that document in a comprehensive study poor out-of-sample predictability. However, in our case, we can go one step further as we see clear improvements in some cases when there is a better use of the conditional information, in the sense of sparsity and dimension reduction.

2.5.5 Economic Value

It is known that the R^2 are generally extremely low or negative when dealing with forecasting excess returns. To assess if the OOS R^2 generated by different approaches deliver economically meaningful results, we follow Campbell and Thompson (2007); Gu et al. (2018) and calculate the implied SR (SR*). The implied SR seeks to measure the improvement produced in the SR when an investor can exploit signals from the conditioning information set $\mathbf{Z}$. Under mild assumptions[37], the implied SR is given by

$$SR^* = \sqrt{\frac{SR^2 + R^2}{1 - R^2}} \quad . \tag{2.19}$$

The interpretation is straightforward. In the case the conditioning information produces valuable signal, the OOS R^2 is large relative to SR^2, so the investor can exploit the information in $\mathbf{Z}$ to obtain a large proportional increase in portfolio return.

[34] We also test the significance with the original Diebold and Mariano (1995) t-test, and there is no significant change.

[35] As previously mentioned, doing is a way to overcome issue of using historical averages as they usually underperform a naive forecast of zero (Gu et al., 2018).

[36] We thank an anonymous referee for this suggestion.

[37] Assuming an investor with a mean-variance preferences, single-period horizon, and γ risk aversion coefficient.

Table 2.5 reports the implied Sharpe ratios for an investor with mean-variances preferences exploiting predictive information from the conditioning set. It is clear that PLS and PCR deliver higher monthly implied Sharpe ratios, and thus conveying economically meaningful results. Notice that when the OOS R^2 is negative, larger absolute R^2 values than the Sharpe ratio will cause the impossibility of the implied Sharpe ratio to be calculated.

Table 2.5: Implied Sharpe Ratios

	25 Portfolios Formed on Size and Book-to-Market			100 Portfolios Formed on Size and Book-to-Market		
	CMV	UMV	MVATE	CMV	UMV	MVATE
Panel A: Goyal						
OLS	0.220	0.202	0.167	-	-	-
3PRF	0.258	0.262	0.234	0.153	0.166	0.115
PLS	0.286	0.302	0.279	0.389	0.387	0.413
PCR	0.384	0.388	0.382	0.398	0.408	0.406
LASSO	0.221	0.225	0.223	0.125	0.134	0.122
RIDGE	0.284	0.281	0.268	0.014	0.060	-
ENET	0.275	0.277	0.211	0.102	0.108	0.053
Panel B: FRED-MD						
OLS	-	-	-	-	-	-
3PRF	0.128	0.076	0.074	0.187	0.173	0.097
PLS	0.314	0.317	0.304	0.407	0.404	0.306
PCR	0.373	0.363	0.376	0.289	0.293	0.285
LASSO	0.099	0.086	0.094	-	-	-
RIDGE	0.258	0.261	0.259	-	-	-
ENET	0.208	0.191	0.216	0.123	0.109	-
Panel C: All Instruments						
OLS	-	-	-	-	-	-
3PRF	0.125	0.096	-	-	-	-
PLS	0.327	0.300	0.311	0.378	0.383	0.340
PCR	0.299	0.285	0.303	0.371	0.374	0.372
LASSO	-	-	-	0.341	0.333	0.275
RIDGE	0.278	0.279	0.271	-	-	-
ENET	0.187	0.152	0.187	0.378	0.371	0.212

Table 2.5 reports the implied SR^* as given in equation (2.19) for all seven estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge and ENet), three different sets of conditioning information (Goyal's, FRED-MD, and "All Instruments", wich is the combination of the previous two with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI)) and three different mean-variance approaches (CMV, UMV, MVATE) to build efficient portfolios. Notice that when the OOS R^2 is negative, larger absolute R^2 values than the Sharpe ratio will cause the impossibility of the implied Sharpe ratio to be calculated. These cases are omitted in the table.

Economic Gains

We use two measures to capture possible economic gains produced for optimal portfolios generated using high-dimensional data imposing sparsity in the different estimators. First, we consider the certainty equivalent return (CER) performance of each optimal portfolio, which can be calculated as

$$\text{CER} = \hat{\mu}_p - \frac{\gamma}{2}\hat{\sigma}_p \tag{2.20}$$

where $\hat{\mu}_p$ and $\hat{\sigma}_p$ are the OOS unconditional mean and variance of the excess returns from the optimal portfolio generated by an estimator, γ is the risk aversion coefficient. We follow Brandt (2010); Goto and Xu (2015); Fletcher and Basu (2016) and set γ equal to 5. Higher CER values for a given optimal portfolio indicate that this portfolio has a better risk-return characteristic.

Table 2.6 reports the monthly CER as a percentage for all cases. Inside brackets, we also report the ranking among all 7 estimators within strategy and set of conditioning information employed. We see that PLS and PCR generate a higher CER for an investor with risk aversion coefficient $\gamma = 5$ when N available risky assets is large (either 25 or 100)[38]. For 25 portfolios formed on Size/BTM we see that Ridge also generates high monthly CER. Overall, another fact that we can infer from the table is that the CER ranking changes just marginally across the three approaches (CMV, UMV, MVATE).

Another way to infer possible economic gains of using a better approach to form mean-variance efficient portfolios is to measure the maximum fee an investor would pay to switch from one approach to another. This management fee, motivated by Fleming et al. (2001), assumes a risk-averse investor with preferences given by a quadratic von Neumann-Morgenstern utility function[39]. This fee can be obtained by solving the following problem:

$$\mathbb{E}\left(r_{p,t}^{(a)} - \mathcal{F}\right) - \frac{\gamma}{2\left(1+\gamma\right)}\mathbb{E}\left(\left(r_{p,t}^{(a)} - \mathcal{F}\right)^2\right) = \mathbb{E}\left(r_{p,t}^{(b)}\right) - \frac{\gamma}{2\left(1+\gamma\right)}\mathbb{E}\left(r_{p,t}^{2^{(b)}}\right) \tag{2.22}$$

where $r_{p,t}^{(a)}$ is the OOS optimal portfolio return generated by a strategy making use of esti-

[38] Table 2.8.7 in the Internet Appendix presents the CER for 6 portfolios formed on Size/BTM and the dataset with 5 industry portfolios.

[39] The expected utility function from the risk-averse investor is given by

$$U = W_0\left(\mathbb{E}\left(r_{p,t}\right) - \frac{\gamma}{2\left(1+\gamma\right)}\mathbb{E}\left(r_{p,t}\right)^2\right) \tag{2.21}$$

where W_0 represents the invertor's initial wealth, γ the relative risk aversion, which is assumed to remain constant, and $r_{p,t}$ the portfolio return.

Table 2.6: Certainty Equivalent Excess Returns (CER) (Monthly %)

	25 Portfolios Formed on Size and Book-to-Market						100 Portfolios Formed on Size and Book-to-Market					
	CMV		UMV		MVATE		CMV		UMV		MVATE	
Panel A: Goyal												
OLS	0.19	[5]	0.18	[5]	0.30	[5]	0.05	[6]	0.04	[7]	-0.13	[7]
3PRF	0.22	[4]	0.22	[3]	0.37	[4]	0.10	[3]	0.11	[3]	-0.01	[4]
PLS	0.29	[2]	0.32	[2]	0.49	[2]	0.22	[2]	0.22	[2]	0.40	[2]
PCR	0.38	[1]	0.38	[1]	0.67	[1]	0.26	[1]	0.25	[1]	0.46	[1]
LASSO	0.16	[7]	0.14	[7]	0.29	[7]	0.06	[5]	0.07	[5]	0.07	[3]
RIDGE	0.24	[3]	0.21	[4]	0.40	[3]	0.04	[7]	0.06	[6]	-0.05	[6]
ENET	0.18	[6]	0.16	[6]	0.29	[6]	0.07	[4]	0.07	[4]	-0.02	[5]
Panel B: FRED-MD												
OLS	0.03	[7]	0.03	[7]	-0.35	[7]	0.02	[7]	0.02	[7]	-0.86	[7]
3PRF	0.19	[5]	0.17	[5]	0.34	[5]	0.14	[3]	0.13	[3]	0.18	[5]
PLS	0.26	[3]	0.24	[3]	0.45	[3]	0.23	[1]	0.23	[1]	0.35	[1]
PCR	0.39	[1]	0.36	[1]	0.70	[1]	0.19	[2]	0.19	[2]	0.31	[2]
LASSO	0.16	[6]	0.14	[6]	0.28	[6]	0.10	[6]	0.09	[6]	0.18	[4]
RIDGE	0.26	[2]	0.25	[2]	0.46	[2]	0.13	[4]	0.12	[4]	0.23	[3]
ENET	0.21	[4]	0.17	[4]	0.38	[4]	0.12	[5]	0.11	[5]	0.12	[6]
Panel C: All Instruments												
OLS	0.03	[7]	0.03	[7]	-0.29	[7]	0.02	[7]	0.02	[7]	-2.04	[7]
3PRF	0.19	[5]	0.17	[4]	0.31	[5]	0.16	[3]	0.15	[3]	0.19	[4]
PLS	0.27	[3]	0.25	[3]	0.47	[3]	0.17	[2]	0.16	[2]	0.31	[2]
PCR	0.35	[1]	0.32	[1]	0.61	[1]	0.24	[1]	0.23	[1]	0.44	[1]
LASSO	0.17	[6]	0.14	[6]	0.30	[6]	0.13	[5]	0.13	[5]	0.23	[3]
RIDGE	0.29	[2]	0.27	[2]	0.49	[2]	0.12	[6]	0.12	[6]	0.12	[6]
ENET	0.22	[4]	0.17	[5]	0.40	[4]	0.14	[4]	0.14	[4]	0.18	[5]

Table 2.6 summarises the CER (monthly %) by estimator and optimal portfolio framework (CMV, UMV and MVATE) for both portfolios. Panel A reports the Sharpe ratios generated when the variables from Goyal's website are used as **Z**. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. Panel B presents the Sharpe ratios obtained using the FRED-MD variables. The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, panel C shows the Sharpe rations when all variables are used as conditioning information. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). Inside brackets, we also report the ranking among all 7 estimators within strategy and set of conditioning information employed

Figure 2.6: Management Fee

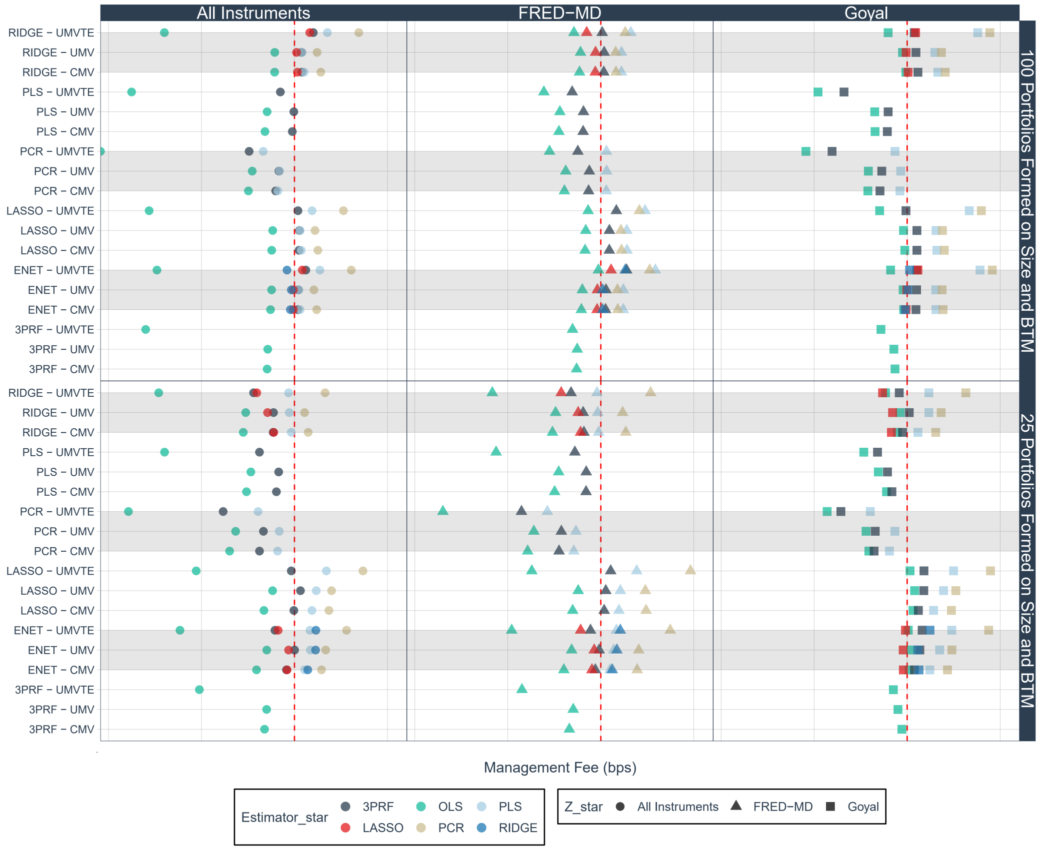

Figure 2.6 presents the management fee in bps, computed as the solution for $\mathcal{F}$ in equation (2.22), from an investor switching from an optimal portfolio formed by the estimator and mean-variance framework given in the left axis to another portfolio plotted in different colors and shape. The comparison is done in pairs of optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and sets of conditioning information.

mator (a), while $r_{p,t}^{(b)}$ is the benchmark OOS optimal portfolio return generated by estimator (b). Solving for $\mathcal{F}$ we can find the management fee that an investor would be willing to pay to have access to a better formation of an optimal portfolio.

Figure 2.6 reports the management fee, computed as the solution for $\mathcal{F}$ in equation (2.22). The comparison is made in pairs, assessing the management fee generated switching from (a) to (b). We can interpret the results in this figure, as the management fee of switching from a strategy given in the vertical axis, which shows the combination of the estimator employed and the mean-variance approach used, to another one plotted in different colors for all seven techniques and set of conditioning information. Essentially, we see that a positive management fee is generated when switching to either PLS or PCR. There are few cases in which switching from a penalized estimator to PLS or PCR produces a negative management fee.

2.5.6 Out-of-Sample Analysis - Portfolio Efficiency

Can these dimensionality reduction techniques generate significant alphas in a standard factor model? We answer this question in this section. We evaluate how the returns of the optimal portfolios produced by all seven estimators behave when evaluated in a standard factor model to explain the risk premia. We compare all possible combinations in the Fama-French three- and five-factor models, as well as with a Fama-French five-factor with momentum. The p-values of the generated alphas are plotted in figure 2.7, while the R^2 of the regressions are presented in figure 2.8. It is clear that in most cases, for N large PLS and PCR can generate statistically significant alphas. Table 2.7 present more details from the regression, reporting the alphas, t-statistics, p-values and the R^2 of the regressions using only the FRED-MD as the set of lagged instruments. In the Internet Appendix, we present the regressions for the remaining sets of conditioning information, as well as additional results.

Figure 2.7: p-values

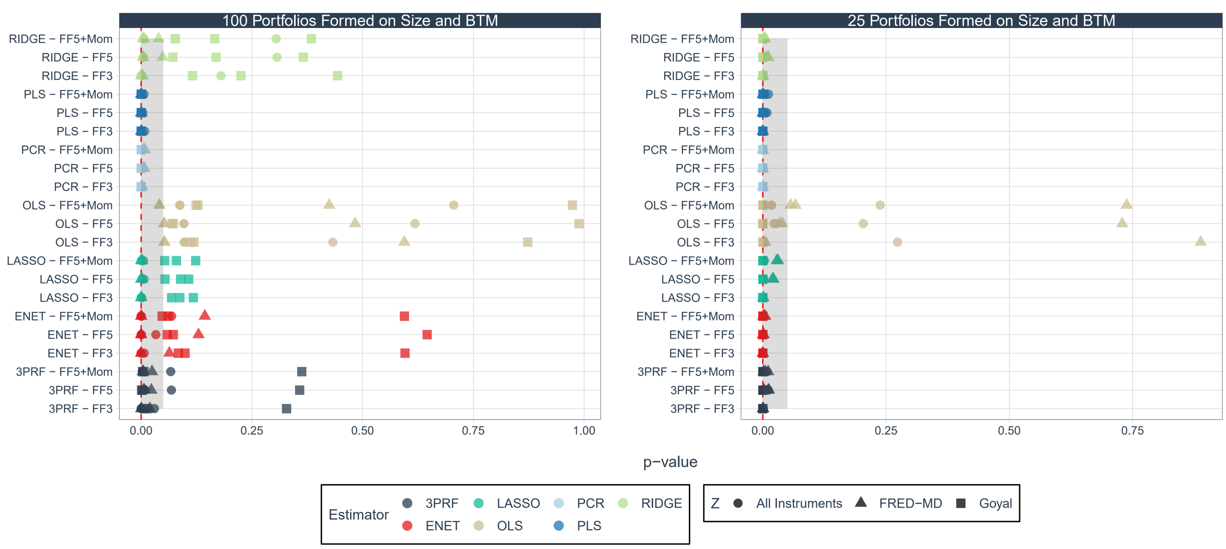

Figure 2.7 reports the p-values of alphas of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and set of conditioning information used on the Fama-French 3, 5, and 5 + momentum factor models. The p-values are calculated from Newey–West t-statistics computed with one lag.

Figure 2.8: R-Squared

Figure 2.8 reports the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and set of conditioning information used on the Fama-French 3, 5, and 5 + momentum factor models.

Table 2.7: Alphas (Monthly %)

		25 Portfolios Formed on Size and Book-to-Market									100 Portfolios Formed on Size and Book-to-Market								
		CMV			UMV			MVATE			CMV			UMV			MVATE		
		FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom
Z: FRED-MD																			
OLS	α (%)	0.027	0.021	0.019	0.027	0.021	0.019	-0.037	-0.082	-0.077	0.017	0.016	0.016	0.017	0.016	0.016	0.188	0.219	0.248
	t-statistics	[2.83]	[2.07]	[1.84]	[2.9]	[2.13]	[1.92]	[-0.14]	[-0.35]	[-0.34]	[1.95]	[1.95]	[2.05]	[1.94]	[1.95]	[2.05]	[0.53]	[0.7]	[0.8]
	p-val	0.005**	0.04*	0.066	0.004**	0.034*	0.056	0.888	0.729	0.738	0.052	0.052	0.041*	0.053	0.053	0.042*	0.594	0.483	0.425
	R^2	0.01	0.03	0.07	0.01	0.03	0.07	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.06	0.06	0.06
3PRF	α (%)	0.172	0.140	0.133	0.156	0.122	0.115	0.321	0.248	0.238	0.142	0.130	0.129	0.139	0.127	0.126	0.264	0.223	0.233
	t-statistics	[3.81]	[3.04]	[3.2]	[3.54]	[2.57]	[2.84]	[3.4]	[2.52]	[2.56]	[3.65]	[2.8]	[2.94]	[3.49]	[2.69]	[2.83]	[2.35]	[2.29]	[2.26]
	p-val	0**	0.003**	0.002**	0**	0.011*	0.005**	0.001**	0.012*	0.011*	0**	0.006**	0.004**	0.001**	0.008**	0.005**	0.02*	0.023*	0.025*
	R^2	0.12	0.15	0.18	0.09	0.13	0.16	0.14	0.17	0.19	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.06	0.06
PLS	α (%)	0.246	0.215	0.208	0.224	0.188	0.181	0.479	0.404	0.394	0.240	0.245	0.236	0.235	0.238	0.230	0.461	0.507	0.501
	t-statistics	[4.61]	[4.08]	[4.03]	[4.28]	[4.12]	[4.05]	[4.08]	[3.77]	[3.64]	[4.58]	[4.06]	[4.4]	[4.4]	[4.14]	[4.38]	[3.27]	[3.41]	[3.33]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.001**	0.001**	0.001**
	R^2	0.14	0.16	0.17	0.15	0.17	0.19	0.10	0.13	0.14	0.05	0.06	0.10	0.05	0.06	0.09	0.05	0.06	0.06
PCR	α (%)	0.379	0.321	0.300	0.346	0.291	0.271	0.742	0.622	0.582	0.217	0.205	0.195	0.211	0.196	0.187	0.427	0.397	0.379
	t-statistics	[6.19]	[5.47]	[5.32]	[5.99]	[5.02]	[4.99]	[6.09]	[5.53]	[5.19]	[3.1]	[2.84]	[2.79]	[3.23]	[2.9]	[2.77]	[3.07]	[2.69]	[2.64]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.002**	0.005**	0.006**	0.001**	0.004**	0.006**	0.002**	0.008**	0.009**
	R^2	0.25	0.28	0.37	0.28	0.31	0.40	0.25	0.28	0.37	0.00	0.01	0.03	0.00	0.01	0.04	0.00	0.01	0.03
LASSO	α (%)	0.157	0.114	0.103	0.138	0.101	0.092	0.296	0.213	0.193	0.099	0.084	0.081	0.096	0.080	0.077	0.190	0.155	0.149
	t-statistics	[3.35]	[2.35]	[2.21]	[3.39]	[2.31]	[2.19]	[3.34]	[2.32]	[2.17]	[4.87]	[4.2]	[3.99]	[4.75]	[4]	[3.8]	[3.55]	[3.11]	[3.01]
	p-val	0.001**	0.019*	0.028*	0.001**	0.022*	0.029*	0.001**	0.021*	0.031*	0**	0**	0**	0**	0**	0**	0**	0.002**	0.003**
	R^2	0.10	0.13	0.18	0.09	0.13	0.17	0.10	0.14	0.18	0.04	0.08	0.10	0.05	0.08	0.11	0.03	0.06	0.07
RIDGE	α (%)	0.247	0.171	0.150	0.238	0.169	0.149	0.470	0.325	0.290	0.127	0.106	0.096	0.122	0.101	0.091	0.263	0.207	0.189
	t-statistics	[4.19]	[2.56]	[3.04]	[3.53]	[2.56]	[2.86]	[4.5]	[2.75]	[3.01]	[4.32]	[2.92]	[2.9]	[4.21]	[2.83]	[2.74]	[2.91]	[1.98]	[2.07]
	p-val	0**	0.011*	0.003**	0**	0.011*	0.005**	0**	0.006**	0.003**	0**	0.004**	0.004**	0**	0.005**	0.007**	0.004**	0.049*	0.039*
	R^2	0.18	0.25	0.34	0.17	0.23	0.34	0.19	0.25	0.34	0.09	0.12	0.23	0.09	0.12	0.24	0.08	0.10	0.15
ENET	α (%)	0.195	0.159	0.149	0.163	0.128	0.121	0.395	0.315	0.302	0.117	0.102	0.098	0.112	0.097	0.093	0.151	0.117	0.106
	t-statistics	[4.3]	[3.47]	[3.34]	[4.32]	[3.16]	[2.84]	[4.21]	[3.51]	[3.03]	[4.63]	[4.11]	[3.89]	[4.58]	[3.93]	[3.8]	[1.86]	[1.52]	[1.47]
	p-val	0**	0.001**	0.001**	0**	0.002**	0.005**	0**	0.001**	0.003**	0**	0**	0**	0**	0**	0**	0.064	0.130	0.144
	R^2	0.13	0.16	0.20	0.13	0.17	0.19	0.09	0.12	0.14	0.05	0.07	0.11	0.05	0.07	0.11	0.00	0.02	0.05

Table 2.7 reports the alphas, t-statistics, p-values and the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) using the FRED-MD as conditioning information on the Fama-French 3, 5, and 5 + momentum factor models. FRED-MD is a large dataset containing 128 macroeconomic and financial variables.

2.5.7 Financial Metrics

As a final analysis, we also assess how sparsity and dimension reduction can affect common financial metrics for the returns generated by the efficient portfolios. We check how the turnover is affected by different estimators. Additionally, we assess another common risk metric, which is the maximum drawdown. Related to it, we evaluate the pain ratio[40]. As in Kirby and Ostdiek (2012), we define the turnover as

$$\text{Turnover}_p = \frac{1}{T} \sum_{t \in \tau_{OOS}} \sum_{i}^{N} |x_{p,i,t} - x_{p,i,t-1}| + \left| \sum_{i}^{N} (x_{p,i,t} - x_{p,i,t-1}) \right| \tag{2.24}$$

where $x_{p,i,t}$ is the optimal weight of asset i for a portfolio p at t in the OOS, and $x_{p,i,t-1}$ is the optimal weight of asset i for a portfolio p at $t-1$. Following Gu et al. (2018), we define the maximum drawdown as

$$\text{MaxDD}_p = \max_{0 \leq Y_{p,t_1} \leq Y_{p,t_2} \leq T} (Y_{p,t_1} - Y_{p,t_2}) \tag{2.25}$$

where $Y_{p,t}$ is the cumulative log return for a portfolio p from the first OOS period to t. Given the large amount of results, we report only for the CMV approach[41].An inspection in table 2.8 show that overall all estimator generate high turnover, especially when the available number of risky assets N is low. Notice that PLS, in general, produces a lower rate of turnover compared to the penalized regressors for most of the cases. This behavior, as similarly noted by Gu et al. (2018), should be understood in light of the large role of price trend predictors selected by these dimension reduction techniques. The maximum drawdown for PLS and PCR is in line with the other estimators. It is clear that, with more risky assets, the diversification generates a lower drawdown. We see that the pain ratio also shows that, in general, for large N, PCR and PLS perform well compared to most estimators.

2.6 Conclusion

The literature has assembled multiple variables with the potential to influence the conditioning information sets of investors. In this paper, we investigated how one can account for extensive sets of potential signals while constructing mean-variance efficient portfolios with

[40]The Pain Ratio can be calculated as

$$\text{Pain Ratio}_p = \frac{r_{p,t}}{\sum_{t \in \tau_{OOS}} \frac{|D_t|}{T}} \tag{2.23}$$

where D_τ is the drawdown since previous peak.

[41]For the other two approaches, UMV and MVATE, the results are similar.

Table 2.8: Turnover and Financial Metrics (Monthly %)

	25 Portfolios Formed on Size and Book-to-Market			100 Portfolios Formed on Size and Book-to-Market		
	Turnover (%)	Max DD (%)	Pain Ratio	Turnover (%)	Max DD (%)	Pain Ratio
Panel A: Goyal						
OLS	92.35	3.50	0.323	61.26	15.20	0.072
3PRF	89.61	3.14	0.375	66.88	15.98	0.239
PLS	119.98	6.11	0.324	82.92	10.98	0.361
PCR	131.74	5.56	0.467	93.62	8.61	0.373
LASSO	126.59	4.36	0.277	75.93	14.22	0.124
RIDGE	120.74	3.44	0.368	73.03	22.05	0.053
ENET	129.82	2.39	0.418	75.89	14.26	0.132
Panel B: FRED-MD						
OLS	72.13	1.15	0.098	45.99	1.37	0.045
3PRF	162.75	4.66	0.197	121.68	4.22	0.289
PLS	121.16	7.82	0.308	87.83	4.57	0.528
PCR	126.98	4.98	0.543	100.89	13.13	0.154
LASSO	180.06	7.04	0.138	81.94	1.21	0.465
RIDGE	180.59	7.16	0.182	95.36	3.43	0.386
ENET	178.36	4.74	0.261	86.89	1.81	0.453
Panel C: All Instruments						
OLS	61.63	1.27	0.106	40.35	1.56	0.044
3PRF	144.62	4.51	0.229	100.64	4.21	0.336
PLS	120.02	9.78	0.225	82.15	3.46	0.354
PCR	154.23	6.06	0.335	123.91	9.84	0.273
LASSO	166.84	4.94	0.171	91.41	1.59	0.428
RIDGE	178.70	5.89	0.249	86.11	2.99	0.228
ENET	175.27	3.45	0.378	93.85	2.02	0.410

Table 2.8 reports several standard financial metrics to evaluate portfolios by estimator for the CMV framework for both portfolios. The Turnover is computed following equation (2.24). MaxDD is maximum drawdown as presented in equation (2.25). Pain Ratio is the standard metric, as shown in equation (2.23). Panel A reports the financial metrics generated when the variables from Goyal's website are used as $\mathbf{Z}$. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. Panel B presents the financial metrics obtained using the FRED-MD variables. Finally, panel C shows the financial metrics when all variables are used as conditioning information. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

respect to conditioning information. We followed the guidance of Cochrane (2011) and Goyal (2012) and sought to reduce the dimension of informative signals to generate our conditional means and build efficient portfolios.

We exploited the wealth set of data as predictive signals, and condensed this information while estimating the conditional mean using high-dimensional data. Since the conditional mean is the most important driver of the formation of mean variance efficient portfolios with regard to conditioning information, we utilized an extensive out-of-sample analysis to assess different techniques for imposing sparsity and dimensionality reduction when locating the conditional mean. We evaluated how penalized estimators, such as LASSO, Ridge and Elastic Net, as well as pure dimensionality reduction and latent factors approaches, such as Partial Least Squares (PLS) and Principal Components Regression (PCR), in addition to a generalization of the former (Three Pass Regression Filter) can produce different optimized portfolios.

The main takeaway of this paper is the evidence gathered showing to be possible to condense the large set of potential predictors to build meaningful mean-variance efficient portfolios with respect to the conditioning information. Standard approaches relying on a few lagged variables and in a naive OLS estimation can be problematic and inefficient, as shown in this work. The use of the proper technique to reduce the dimensionality of the conditional mean that enters the solution of the efficient portfolios is a key point. Few methods, especially PCR and PLS, were able to provide better out-of-sample results from an investor standpoint. We document on average, higher Sharpe ratios OOS, implied Sharpe ratios (SR^*), and higher certainty equivalent returns (CER). In short, from an investor's point of view is possible to obtain favorable outcomes properly exploiting a high-dimensional set of conditioning information.

Appendix

2.7 Appendix - Statistical Inference for the Difference of Sharpe Ratios

To test for the difference of the Sharpe ratios of two different optimal portfolios, we follow Ledoit and Wolf (2008) approach to perform these tests. The main advantage of this method, is its robustness to (i) when returns have tails heavier than the normal distribution, and (ii) serial correlation of the actual returns or of the squared returns (volatility clustering). Following Ledoit and Wolf (2008) the difference between SR_a and SR_b, for two optimal portfolios produced by two different estimators, a and b, is given by

$$\Delta = SR^{(a)} - SR^{(b)} = \frac{\mu^{(a)}}{\sigma^{(a)}} - \frac{\mu^{(b)}}{\sigma^{(b)}} \tag{2.26}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the OOS unconditional mean and variance of the excess returns from the optimal portfolio generated by an estimator. Thus, the null hypothesis can be written as: $H_0 : \Delta = 0$. Denote the vector of uncentered population and sample moments of the excess returns generated by two different estimators by

$$\begin{aligned} \zeta &= \left[\mu^{(a)}, \mu^{(b)}, \mathbb{E}\left(r^{2(a)}\right), \mathbb{E}\left(r^{2(b)}\right)\right]^\top \\ \widehat{\zeta} &= \left[\widehat{\mu^{(a)}}, \widehat{\mu^{(b)}}, \widehat{\mathbb{E}\left(r^{2(a)}\right)}, \widehat{\mathbb{E}\left(r^{2(b)}\right)}\right]^\top \end{aligned} \tag{2.27}$$

Under the assumption that the differences of these moments converges in distribution to a Normal distribution with mean 0 and variance Ψ, using the delta method we have

$$\sqrt{T}\left(\widehat{\Delta} - \Delta\right) \xrightarrow{d} N\left(0, \nabla^\top f(\zeta)\Psi\nabla f(\zeta)\right) \tag{2.28}$$

where $f(\zeta) = \dfrac{\mu^{(a)}}{\sqrt{\mathbb{E}\left(r^{2(a)}\right) - \mu^{2(a)}}} - \dfrac{\mu^{(b)}}{\sqrt{\mathbb{E}\left(r^{2(b)}\right) - \mu^{2(b)}}}$ and the standard error of $\widehat{\Delta}$ is given by

$$se\left(\widehat{\Delta}\right) = \sqrt{\frac{\nabla^\top f(\widehat{\zeta})\widehat{\Psi}\nabla f(\widehat{\zeta})}{T}} \tag{2.29}$$

Since neither the returns, nor the squared returns of financial assets are generally an i.i.d. process, we need to use a robust estimator. One evident way to obtain $\widehat{\Psi}$ is using HAC kernel estimation. Another way is to use bootstrap inference. Ledoit and Wolf (2008) suggest to construct a studentized time series bootstrap confidence interval for the difference of the SRs. This method has been shown to be robust when returns have tails heavier than the normal distribution or are of time series nature. The bootstrap data is generated using the circular block bootstrap of Politis and Romano (1992). The two-sided distribution function of the studentized statistic can be obtained via bootstrap as follows:

$$f\left(\frac{|\widehat{\Delta}-\Delta|}{se\left(\widehat{\Delta}\right)}\right) \approx f\left(\frac{|\widehat{\Delta}^{\text{boot}}-\Delta|}{se\left(\widehat{\Delta}^{\text{boot}}\right)}\right) \tag{2.30}$$

where $f(\cdot)$ is the distribution of a random variable, Δ is populational difference between SR_a and SR_b, $\widehat{\Delta}$ is the sample counterpart of this difference obtained in the data in the estimation window, and $\widehat{\Delta}^{\text{boot}}$ is the estimated difference computed from bootstrap. The standard errors are denoted by $se(\cdot)$. Out of the distribution obtained from the bootstrap in the equation (2.30), we can find the confidence interval for Δ and the p-values of the test[42].

[42]For further details in this procedure, see Ledoit and Wolf (2008).

2.8 Appendix - Additional Results

2.8.1 Out-of-Sample Analysis - Sharpe Ratios

Performance Evaluation

Table 2.8.1: Out-of-Sample Sharpe ratios delivered by each estimator and set of conditioning information

	5 Industry Portfolios						6 Portfolios Formed on Size and Book-to-Market					
	CMV		UMV		MVATE		CMV		UMV		MVATE	
Estimator	SR	p-val	SR	p-val	SR	p-val	SR	p-val	SR	p-val	SR	p-val
Panel A: Goyal												
OLS	0.133	0.032	0.100	0.107	0.135	0.030	0.207	0.001	0.218	0.001	0.203	0.001
3PRF	0.125	0.044	0.083	0.181	0.125	0.043	0.222	0.000	0.258	0.000	0.223	0.000
PLS	0.121	0.051	0.109	0.078	0.119	0.054	0.247	0.000	0.246	0.000	0.247	0.000
PCR	0.089	0.149	0.076	0.217	0.085	0.168	0.298	0.000	0.299	0.000	0.300	0.000
LASSO	0.209	0.001	0.141	0.024	0.209	0.001	0.168	0.007	0.187	0.003	0.166	0.008
RIDGE	0.131	0.035	0.112	0.072	0.131	0.034	0.173	0.006	0.224	0.000	0.174	0.005
ENET	0.163	0.009	0.130	0.037	0.168	0.007	0.231	0.000	0.260	0.000	0.236	0.000
Panel B: FRED-MD												
OLS	0.005	0.934	0.009	0.881	-0.046	0.455	0.154	0.013	0.158	0.011	0.037	0.550
3PRF	0.072	0.246	0.009	0.883	0.073	0.239	0.186	0.003	0.197	0.002	0.188	0.003
PLS	0.102	0.099	0.046	0.455	0.101	0.102	0.210	0.001	0.198	0.002	0.201	0.001
PCR	0.115	0.064	0.019	0.764	0.112	0.070	0.272	0.000	0.255	0.000	0.275	0.000
LASSO	0.156	0.012	0.056	0.363	0.152	0.014	0.226	0.000	0.201	0.001	0.226	0.000
RIDGE	0.116	0.061	0.028	0.655	0.116	0.062	0.242	0.000	0.230	0.000	0.246	0.000
ENET	0.127	0.042	0.032	0.601	0.125	0.044	0.261	0.000	0.261	0.000	0.261	0.000
Panel C: All Instruments												
OLS	-0.035	0.594	-0.002	0.979	-0.078	0.240	0.160	0.017	0.168	0.012	-0.002	0.973
3PRF	0.067	0.281	0.029	0.640	0.068	0.271	0.193	0.002	0.213	0.001	0.191	0.002
PLS	0.002	0.976	0.037	0.581	-0.001	0.991	0.264	0.000	0.263	0.000	0.265	0.000
PCR	0.133	0.047	0.040	0.552	0.129	0.054	0.238	0.000	0.237	0.001	0.240	0.000
LASSO	0.122	0.067	0.052	0.435	0.118	0.078	0.271	0.000	0.228	0.000	0.274	0.000
RIDGE	0.113	0.091	0.035	0.602	0.112	0.094	0.248	0.000	0.229	0.001	0.247	0.000
ENET	0.069	0.301	0.047	0.483	0.069	0.300	0.255	0.000	0.234	0.000	0.266	0.000

Table 2.8.1 summarises the OOS (Jan-1996 to Dec-2017) Sharpe ratios (SR) by estimator and optimal portfolio framework (CMV, UMV and MVATE) for the 5 industry portfolios and 6 portfolios formed on Size/BTM. Panel A reports the Sharpe ratios generated when the variables from Goyal's website are used as **Z**. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl.* Panel B presents the Sharpe ratios obtained using the FRED-MD variables. The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, panel C shows the Sharpe rations when all variables are used as conditioning information. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). The p-val is the p-value from the two-sided test of the SR.

Statistical Inference for the Difference of Sharpe Ratios

Table 2.8.2: Test for the difference of the Sharpe ratios - OOS (Jan-1996 - Dec-2017) - 6 Portfolios Formed on Size and Book-to-Market

Panel A: CMV

		Goyal						FRED-MD						All Instr.					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.734	0.505	0.116	0.316	0.492	0.593	0.695	0.974	0.241	0.747	0.507	0.351	0.780	0.332	0.642	0.251	0.457	0.485
	3PRF		0.621	0.077	0.298	0.339	0.860		0.825	0.234	0.938	0.655	0.431		0.621	0.990	0.309	0.843	0.443
	PLS			0.163	0.128	0.159	0.776			0.462	0.656	0.898	0.776			0.860	0.480	0.973	0.728
	PCR				0.012	0.012	0.221				0.079	0.071	0.344				0.714	0.034	0.535
	LASSO					0.905	0.182					0.120	0.084					0.045	0.175
	RIDGE						0.131						0.129						0.117
FRED-MD	OLS							0.655	0.525	0.156	0.334	0.252	0.175	0.591	0.244	0.395	0.126	0.286	0.170
	3PRF								0.659	0.038	0.437	0.052	0.120		0.035	0.091	0.056	0.011	0.151
	PLS									0.069	0.741	0.426	0.268			0.488	0.194	0.350	0.315
	PCR										0.233	0.281	0.757				0.785	0.300	0.956
	LASSO											0.676	0.228					0.685	0.209
	RIDGE												0.621						0.525
All Instr.	OLS													0.790	0.200	0.371	0.072	0.241	0.118
	3PRF														0.050	0.147	0.094	0.026	0.247
	PLS															0.586	0.477	0.641	0.709
	PCR																0.161	0.769	0.366
	LASSO																	0.248	0.666
	RIDGE																		0.459

Panel B: UMV

		Goyal						FRED-MD						All Instr.					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.184	0.601	0.122	0.345	0.865	0.236	0.665	0.762	0.578	0.766	0.816	0.445	0.901	0.356	0.806	0.809	0.845	0.799
	3PRF		0.815	0.281	0.080	0.310	0.961		0.351	0.952	0.284	0.554	0.947		0.952	0.637	0.616	0.425	0.732
	PLS			0.218	0.308	0.646	0.804			0.855	0.376	0.686	0.777			0.745	0.788	0.565	0.930
	PCR				0.019	0.082	0.478				0.064	0.087	0.420				0.201	0.055	0.273
	LASSO					0.259	0.071					0.371	0.198					0.371	0.418
	RIDGE						0.294						0.505						0.770
FRED-MD	OLS							0.588	0.646	0.278	0.556	0.343	0.185	0.440	0.218	0.419	0.360	0.390	0.311
	3PRF								0.980	0.356	0.930	0.269	0.254		0.139	0.393	0.418	0.152	0.412
	PLS									0.184	0.957	0.494	0.233			0.384	0.534	0.531	0.469
	PCR										0.363	0.595	0.900				0.767	0.498	0.886
	LASSO											0.497	0.060					0.401	0.165
	RIDGE												0.472						0.724
All Instr.	OLS													0.608	0.226	0.448	0.372	0.413	0.364
	3PRF														0.258	0.621	0.605	0.462	0.631
	PLS															0.589	0.715	0.444	0.765
	PCR																0.944	0.867	0.901
	LASSO																	0.794	0.775
	RIDGE																		0.762

Panel C: MVATE

		Goyal						FRED-MD						All Instr.					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.646	0.482	0.104	0.362	0.577	0.449	0.787	0.969	0.207	0.697	0.431	0.316	0.810	0.316	0.595	0.224	0.442	0.360
	3PRF		0.655	0.082	0.276	0.350	0.803		0.688	0.235	0.959	0.633	0.443		0.632	0.992	0.316	0.865	0.354
	PLS			0.156	0.121	0.169	0.840			0.437	0.655	0.962	0.782			0.877	0.463	0.998	0.646
	PCR				0.009*	0.013	0.235				0.074	0.070	0.303				0.720	0.029	0.621
	LASSO					0.853	0.125					0.091	0.073					0.040	0.111
	RIDGE						0.092						0.127						0.081
FRED-MD	OLS							0.045	0.046	0.002*	0.009*	0.006*	0.002*	0.041	0.009*	0.012	0.002*	0.009*	0.002*
	3PRF								0.808	0.034	0.459	0.037	0.124		0.030	0.090	0.055	0.011	0.117
	PLS									0.028	0.607	0.268	0.187			0.367	0.130	0.278	0.189
	PCR										0.213	0.297	0.710				0.787	0.263	0.955
	LASSO											0.631	0.236					0.698	0.143
	RIDGE												0.690						0.447
All Instr.	OLS													0.028	0.002*	0.003*	0.000*	0.002*	0.000*
	3PRF														0.038	0.129	0.080	0.022	0.177
	PLS															0.608	0.479	0.615	0.623
	PCR																0.167	0.823	0.328
	LASSO																	0.248	0.831
	RIDGE																		0.384

p-val	$[0 \leq p \leq 0.05]$	$(0.05 < p < 0.1)$	$(0.1 \leq p \leq 1]$
Gradient color bounds			

Test for the differences of the Sharpe ratios of the OOS (Jan-1996 - Dec-2017) returns of the efficient portfolios formed from the dataset with 6 Size/BTM portfolios using 7 different estimators (OLS, 3PRF, PLS, PCR, LASSO, RIDGE and ENet) and three different set of conditioning information (Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios. Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

Table 2.8.3: Test for the difference of the Sharpe ratios - OOS (Jan-1996 - Dec-2017) - 5 Portfolios Formed on Size and Book-to-Market

Panel A: CMV

		Goyal						**FRED-MD**						**All Instr.**					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.888	0.841	0.539	0.105	0.961	0.528	0.406	0.644	0.769	0.713	0.807	0.916	0.359	0.037	0.740	0.608	0.522	0.219
	3PRF		0.937	0.588	0.056	0.915	0.451		0.733	0.872	0.642	0.899	0.986		0.088	0.949	0.817	0.732	0.360
	PLS			0.649	0.133	0.874	0.493			0.931	0.604	0.947	0.937			0.991	0.899	0.805	0.430
	PCR				0.098	0.584	0.332				0.381	0.713	0.605				0.746	0.831	0.710
	LASSO					0.035	0.223					0.166	0.203					0.060	0.016
	RIDGE						0.442						0.943						0.194
FRED-MD	OLS							0.285	0.224	0.118	0.017	0.089	0.086	0.302	0.874	0.076	0.074	0.073	0.312
	3PRF								0.701	0.524	0.163	0.305	0.371		0.400	0.428	0.550	0.500	0.923
	PLS									0.851	0.458	0.856	0.722			0.833	0.936	0.975	0.574
	PCR										0.515	0.982	0.850				0.991	0.892	0.480
	LASSO											0.440	0.409					0.431	0.114
	RIDGE												0.844						0.390
All Instr.	OLS													0.114	0.665	0.042	0.034	0.052	0.177
	3PRF														0.395	0.388	0.536	0.433	0.937
	PLS															0.127	0.142	0.227	0.403
	PCR																0.892	0.742	0.413
	LASSO																	0.875	0.207
	RIDGE																		0.457

Panel B: UMV

		Goyal						**FRED-MD**						**All Instr.**					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.666	0.828	0.733	0.346	0.730	0.439	0.193	0.548	0.221	0.446	0.271	0.287	0.281	0.282	0.331	0.366	0.231	0.351
	3PRF		0.603	0.918	0.237	0.532	0.328		0.703	0.355	0.715	0.420	0.502		0.478	0.518	0.604	0.432	0.558
	PLS			0.609	0.572	0.954	0.687			0.175	0.340	0.183	0.191			0.243	0.269	0.148	0.210
	PCR				0.400	0.612	0.466				0.816	0.514	0.606				0.668	0.499	0.608
	LASSO					0.357	0.782					0.083	0.129					0.063	0.116
	RIDGE						0.590						0.244						0.223
FRED-MD	OLS							0.997	0.696	0.892	0.440	0.751	0.723	0.709	0.585	0.509	0.299	0.454	0.388
	3PRF								0.699	0.881	0.398	0.633	0.682		0.697	0.555	0.454	0.489	0.495
	PLS									0.744	0.903	0.845	0.863			0.843	0.941	0.806	0.899
	PCR										0.583	0.886	0.845				0.688	0.842	0.738
	LASSO											0.569	0.365					0.651	0.823
	RIDGE												0.929						0.724
All Instr.	OLS													0.640	0.635	0.582	0.378	0.554	0.468
	3PRF														0.907	0.833	0.687	0.862	0.750
	PLS															0.969	0.852	0.981	0.905
	PCR																0.871	0.940	0.924
	LASSO																	0.742	0.869
	RIDGE																		0.820

Panel C: MVATE

		Goyal						**FRED-MD**						**All Instr.**					
		3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET	3PRF	PLS	PCR	LASSO	RIDGE	ENET
Goyal	OLS	0.871	0.804	0.495	0.122	0.938	0.493	0.409	0.611	0.724	0.778	0.789	0.883	0.365	0.034	0.704	0.567	0.517	0.225
	3PRF		0.912	0.551	0.060	0.915	0.405		0.719	0.842	0.687	0.893	0.997		0.083	0.911	0.769	0.725	0.369
	PLS			0.628	0.128	0.853	0.433			0.922	0.626	0.959	0.935			0.990	0.873	0.816	0.451
	PCR				0.093	0.553	0.285				0.383	0.682	0.584				0.747	0.798	0.760
	LASSO					0.039	0.285					0.172	0.204					0.066	0.019
	RIDGE						0.386						0.925						0.207
FRED-MD	OLS							0.065	0.059	0.039	0.003*	0.016	0.015	0.070	0.505	0.026	0.026	0.020	0.101
	3PRF								0.718	0.560	0.191	0.324	0.396		0.375	0.477	0.606	0.533	0.909
	PLS									0.869	0.474	0.848	0.723			0.855	0.964	0.981	0.595
	PCR										0.531	0.955	0.838				0.989	0.912	0.510
	LASSO											0.483	0.448					0.465	0.136
	RIDGE												0.859						0.403
All Instr.	OLS													0.035	0.364	0.017	0.013	0.016	0.060
	3PRF														0.369	0.440	0.594	0.469	0.923
	PLS															0.132	0.148	0.222	0.387
	PCR																0.888	0.780	0.449
	LASSO																	0.922	0.256
	RIDGE																		0.476

p-val	$[0 \leq p \leq 0.05]$	$(0.05 < p < 0.1)$	$(0.1 \leq p \leq 1]$
Gradient color bounds			

Test for the differences of the Sharpe ratios of the OOS (jan-1996 - Dec-2017) returns of the efficient portfolios formed from the dataset with 5 industry portfolios using 7 different estimators (OLS, 3PRF, PLS, PCR, LASSO, RIDGE and ENet) and three different set of conditioning information (Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios. Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

2.8.2 Characteristics and the Distribution of Optimized Portfolio Weights

Figure 2.8.1: Distribution of Optimized Portfolio Weights - 6 Portfolios Formed on Size and Book-to-Market

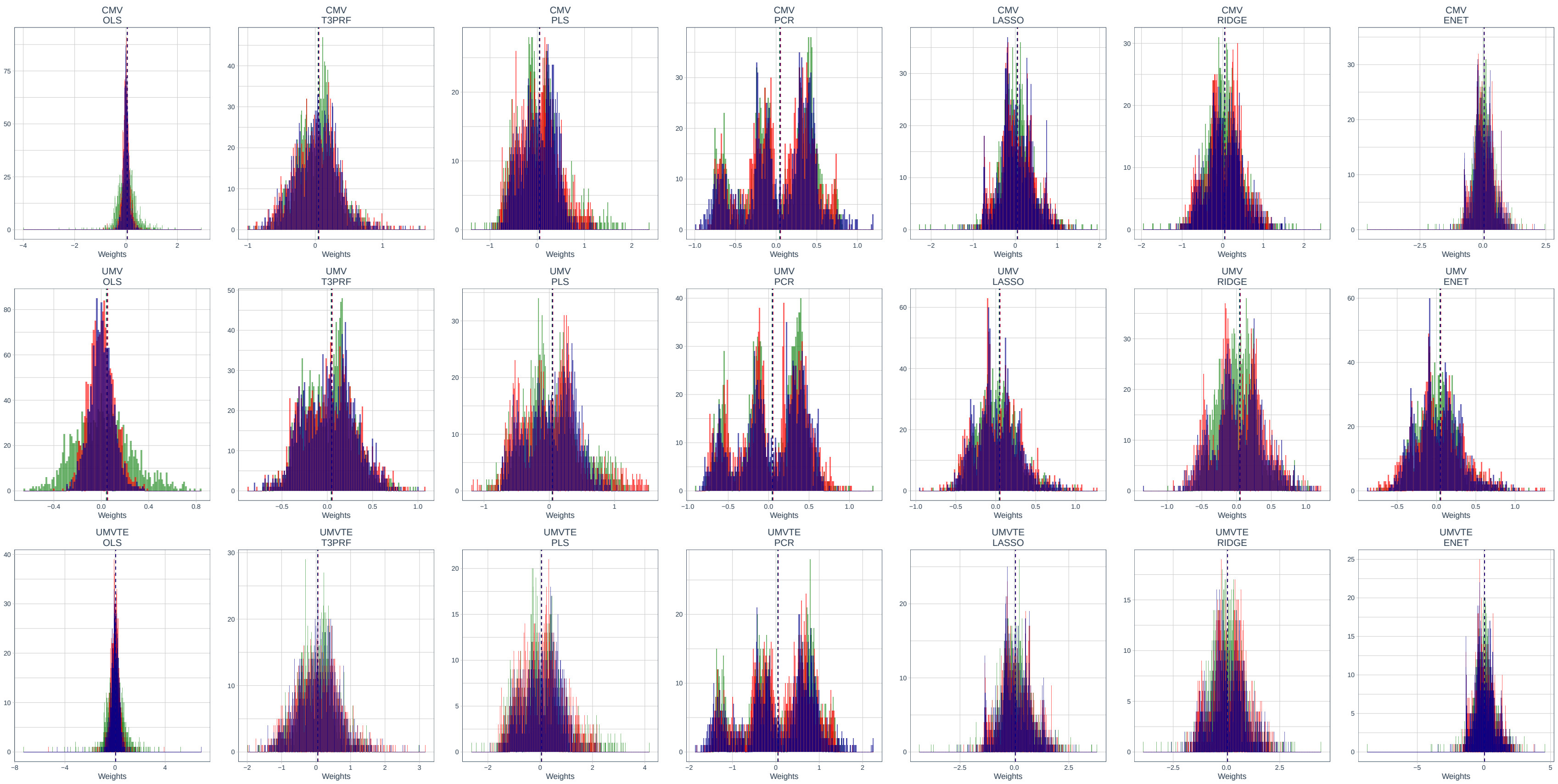

Figure 2.8.1 plots the distribution of the optimized portfolios weights generated by each estimator (columns) and mean-variance optimal framework (rows) for the 6 portfolios formed on Size/BTM. The first row reports the CMV strategies,, the second row reports the UMV, and the third one plots the UMVTE. The set of conditioning information used in **Z** are plotted in different colors: (i) Goyal's in green, (ii) FRED-MD in red, and (iii) "All Instruments" in blue. Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. FRED-MD is a large dataset containing 128 macroeconomic and financial variables. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

Figure 2.8.2: Distribution of Optimized Portfolio Weights - 5 Industry Portfolios

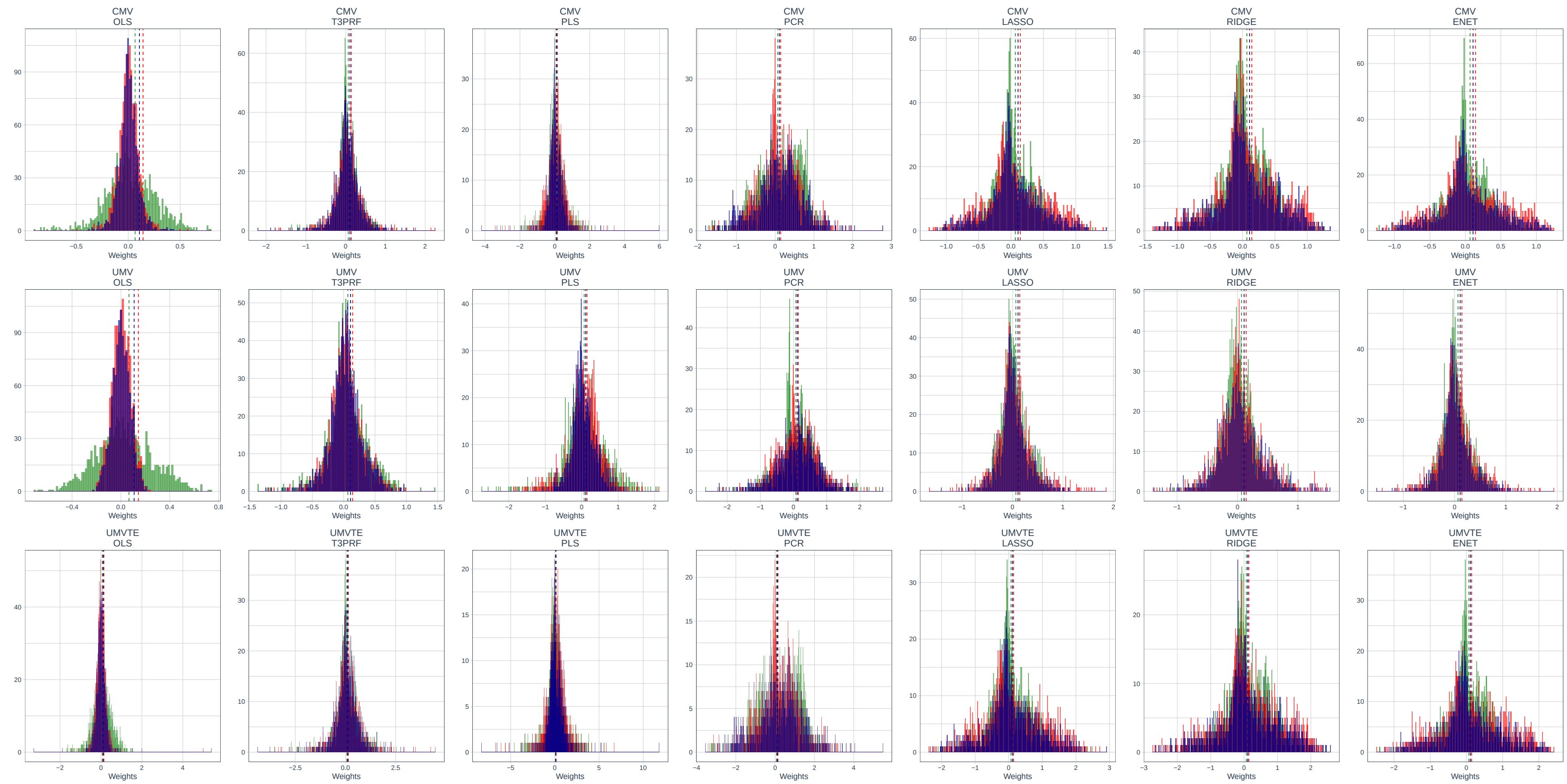

Figure 2.8.2 plots the distribution of the optimized portfolios weights generated by each estimator (columns) and mean-variance optimal framework (rows) for the 5 industry portfolios. The first row reports the CMV strategies,, the second row reports the UMV, and the third one plots the UMVTE. The set of conditioning information used in **Z** are plotted in different colors: (i) Goyal's in green, (ii) FRED-MD in red, and (iii) "All Instruments" in blue. Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. FRED-MD is a large dataset containing 128 macroeconomic and financial variables. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

2.8.3 Variable Contribution

Figure 2.8.3: Variable Contribution by Estimator and Set of Conditioning Information - 6 Portfolios Formed on Size and Book-to-Market

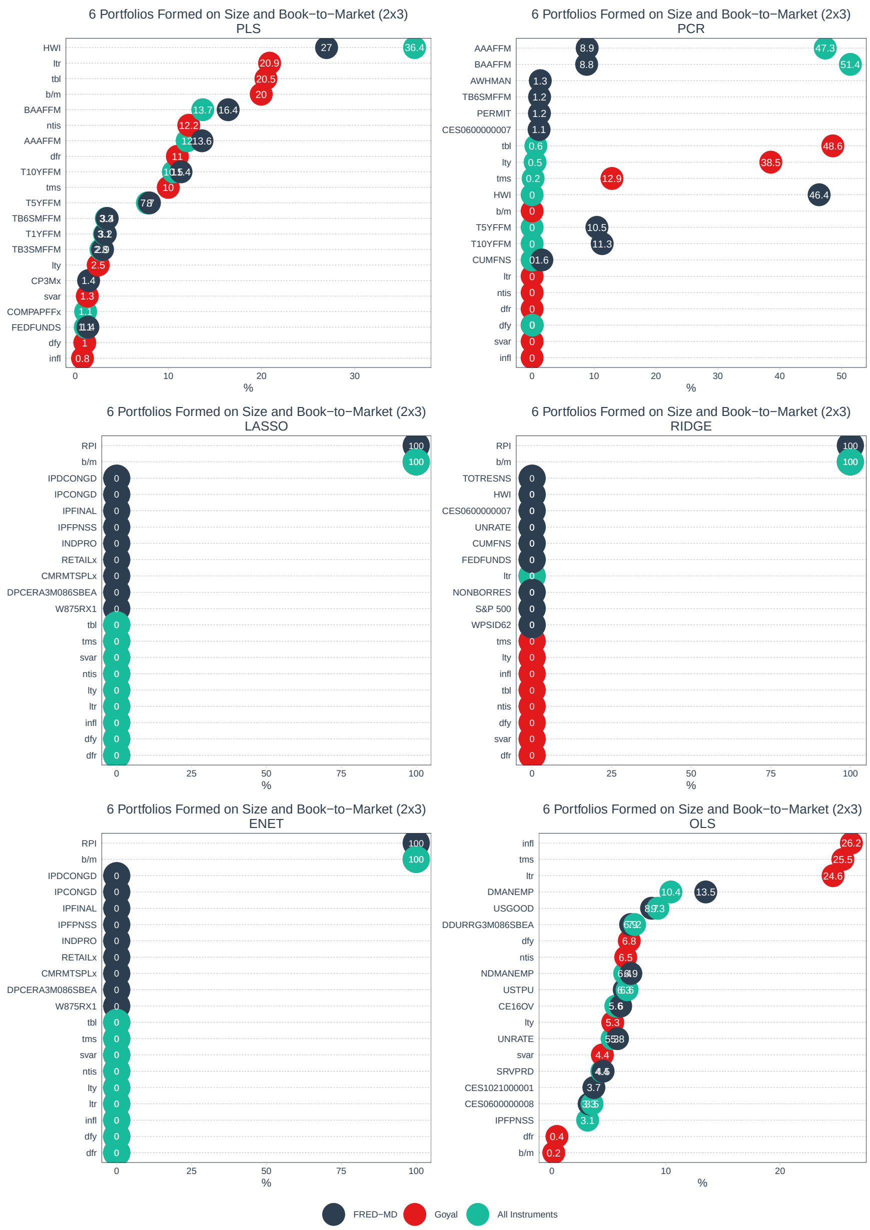

Figure 2.8.3 reports the 10 most influential variables by estimator (panels) and set of conditioning information used (colors) for the for the 6 portfolios formed on Size/BTM. We take a direct approach to obtain the contribution for each variable. At each month t, we compute the absolute values of the estimated coefficients for each asset i and lagged variable k in the conditioning set. We destandardize the variables whenever necessary. Grouping by estimators and $\mathbf{Z}$ used, we pool all the estimates of the N assets and calculate the average of the absolute values. In order to make the comparison clear, we normalize these means to sum one.

Figure 2.8.4: Variable Contribution by Estimator and Set of Conditioning Information - 5 Industry Portfolios

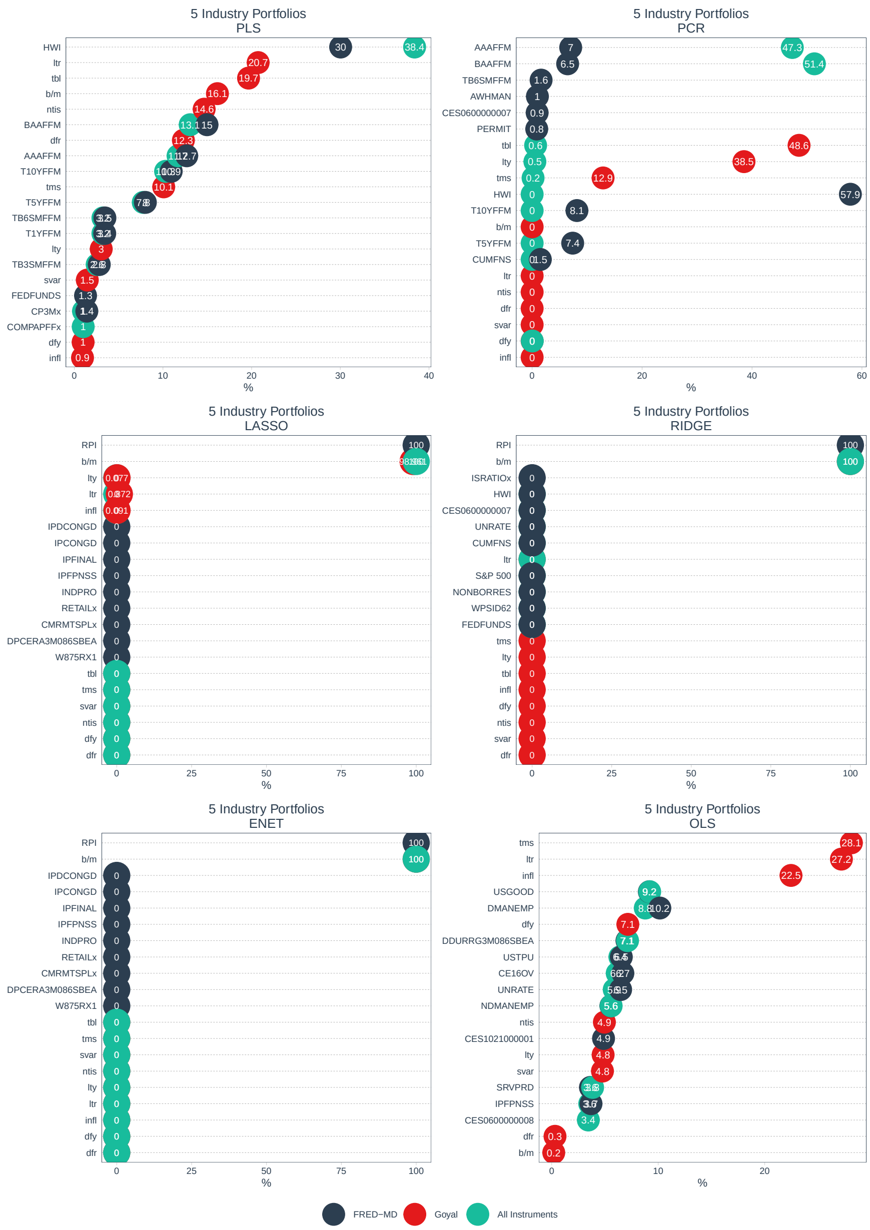

Figure 2.8.4 reports the 10 most influential variables by estimator (panels) and set of conditioning information used (colors) for the for the 5 industry portfolios. We take a direct approach to obtain the contribution for each variable. At each month t, we compute the absolute values of the estimated coefficients for each asset i and lagged variable k in the conditioning set. We destandardize the variables whenever necessary. Grouping by estimators and $\mathbf{Z}$ used, we pool all the estimates of the N assets and calculate the average of the absolute values. In order to make the comparison clear, we normalize these means to sum one.

Figure 2.8.5: Out-of-Sample R^2

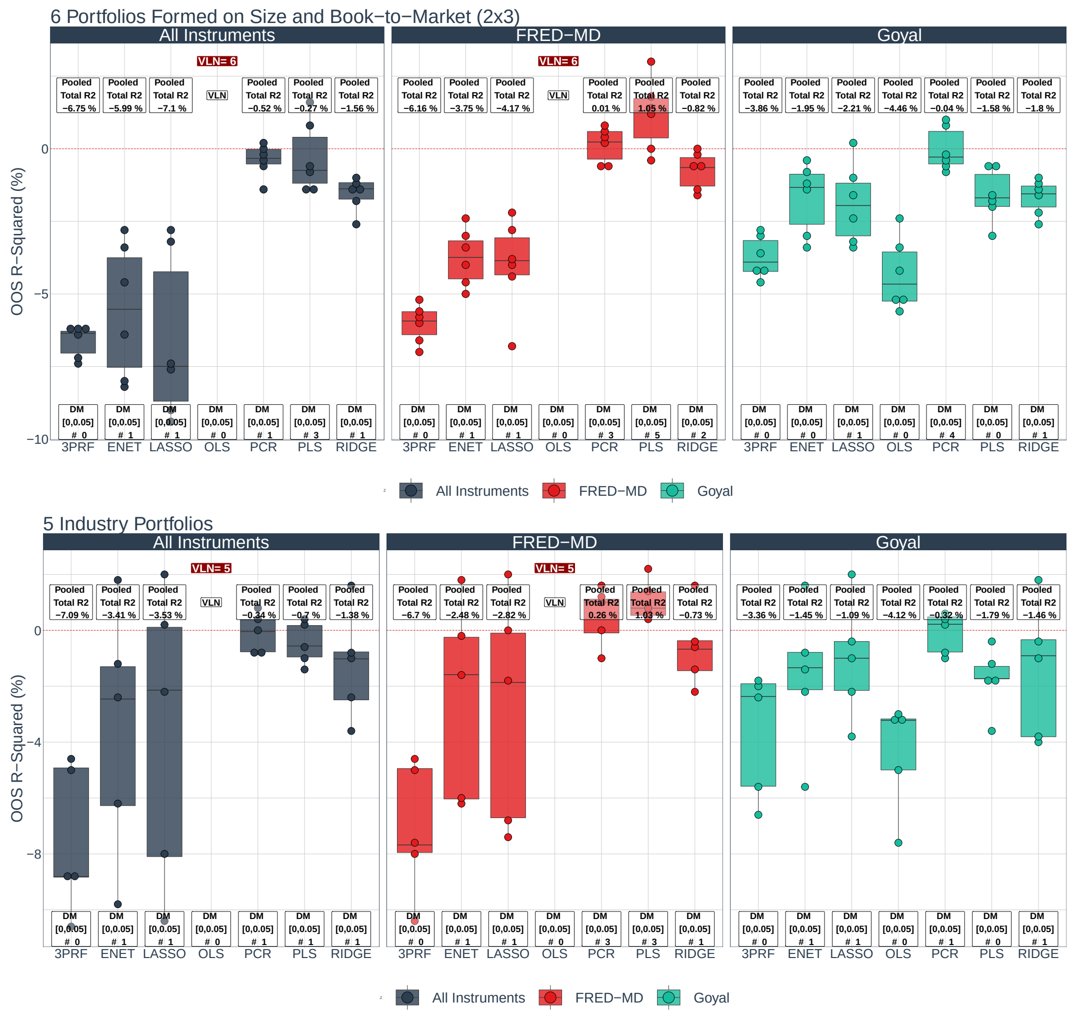

Figure 2.8.5 presents the OOS R^2 for the datasets with 6 portfolios formed on Size/BTM and the 5 industry portfolios by estimator and set of conditioning information used, along with overlapping boxplots. The pooled OOS R^2 across all N risky assets in each dataset is also reported in a white box on the top of each plot. In order to make the plots readable and comparable, we filtered the OOS R^2 larger than -0.5 in absolute values and present the amount of assets that generated these very large numbers (VLN). The amount of VLN per estimator is reported in a red box on top of each plot. We also report at the bottom of each plot the number of significant cases (p-values ≤ 0.05) in terms of a forecast accuracy test. We used the Harvey et al. (1997) modified version of the Diebold and Mariano (1995) t-test among all assets for each estimator and set of conditioning information. There is no significant change when the original Diebold and Mariano (1995) test is used. We assess a one-sided test, in which the alternative hypothesis can be interpreted as our zero forecast benchmark is less accurate than each one of the seven estimators' forecasts.

Figure 2.8.6: Mean Directional Accuracy - 100 Portfolios Formed on Size/BTM

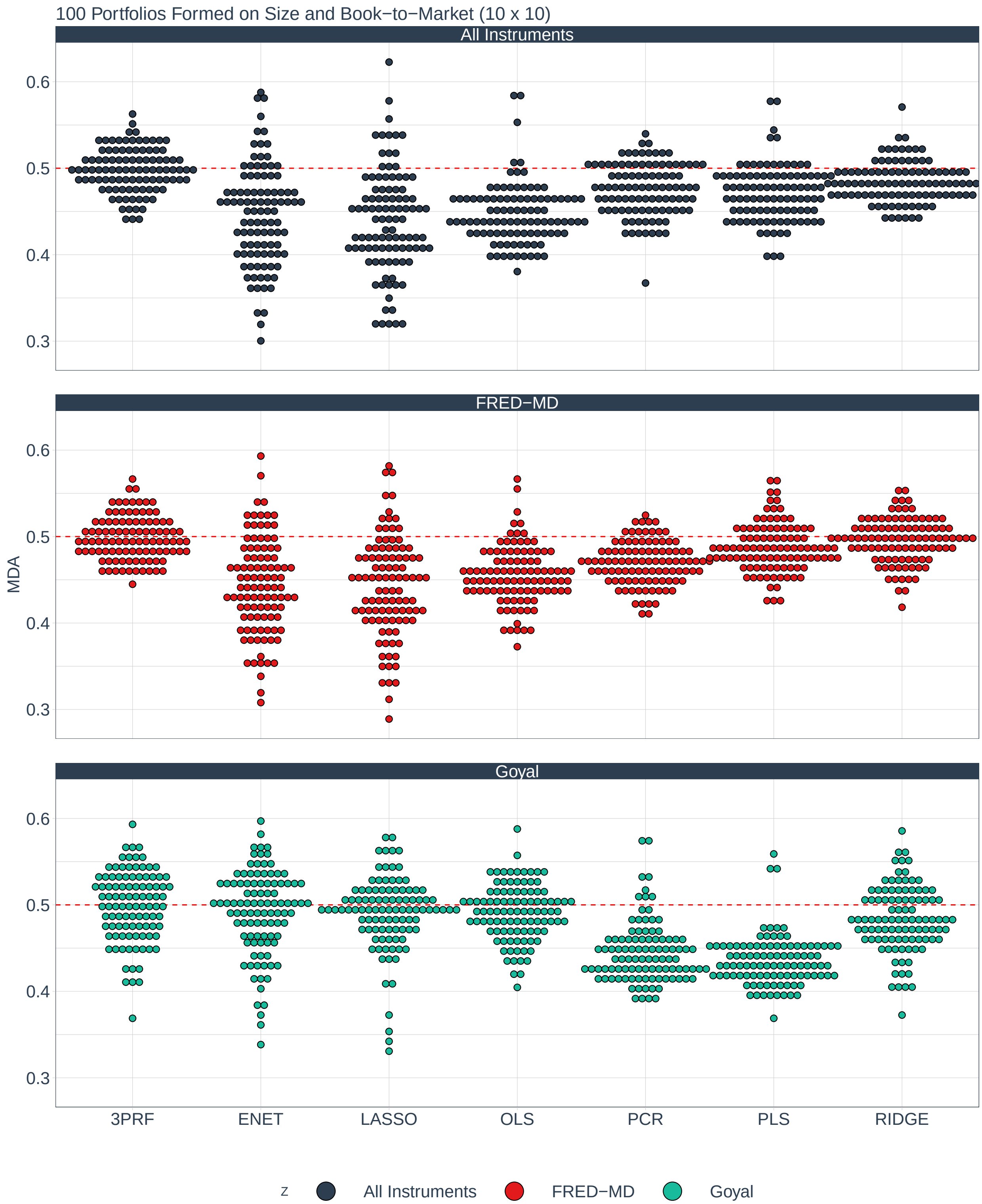

Figure 2.8.6 presents the Mean Directional Accuracy for 100 portfolios formed on Size/BTM by estimator and set of conditioning information used.

Figure 2.8.7: Mean Directional Accuracy - 25 Portfolios Formed on Size/BTM

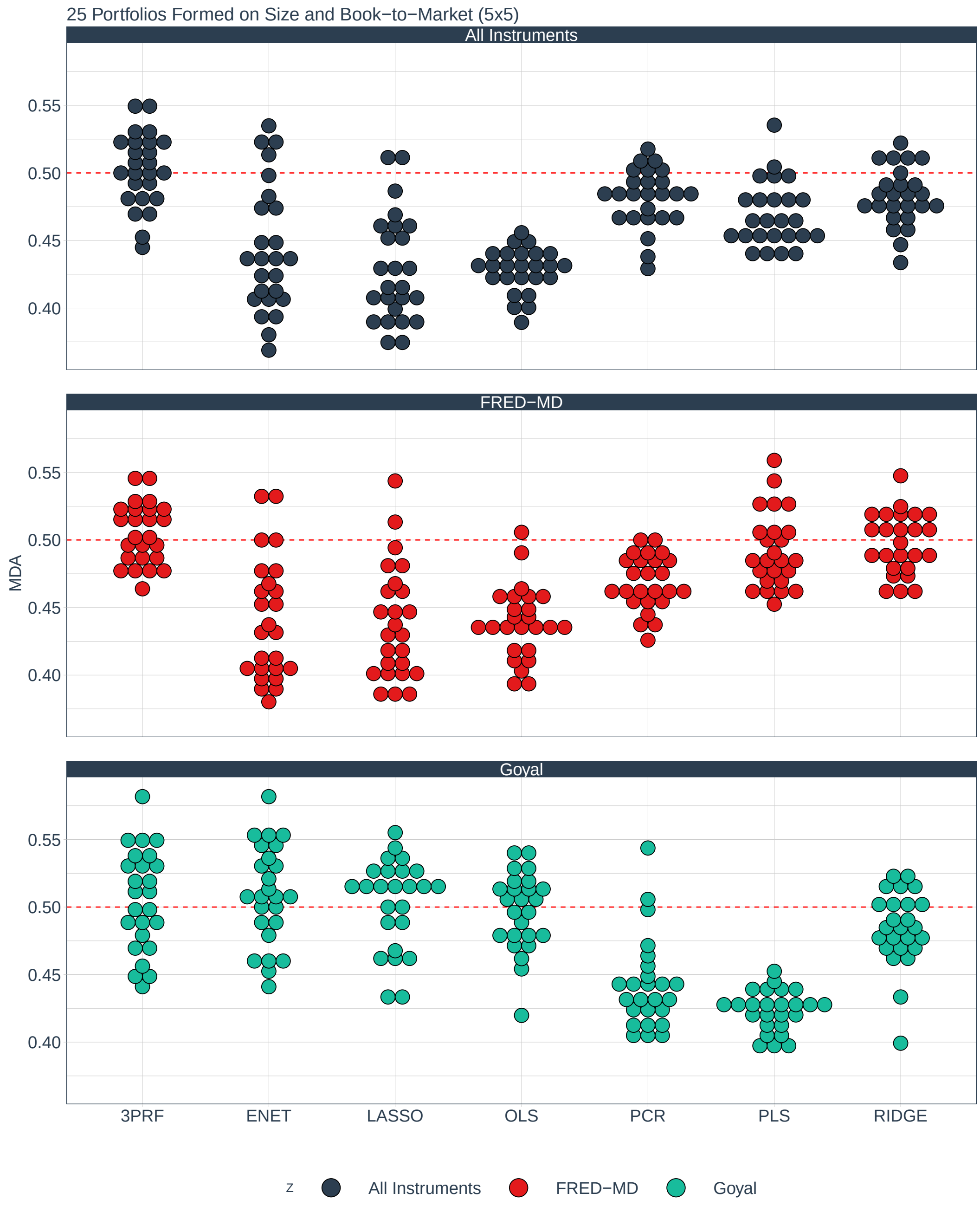

Figure 2.8.7 presents the Mean Directional Accuracy for 25 portfolios formed on Size/BTM by estimator and set of conditioning information used.

Table 2.8.4: Modified Diebold-Mariano (Harvey et al. (1997)) t-test - Proportion for Different Intervals

	25 Portfolios Formed on Size and Book-to-Market					100 Portfolios Formed on Size and Book-to-Market				
	$[0, 0.01]$	$(0.01, 0.05]$	$(0.05, 0.1]$	$(0.1, 0.5]$	$(0.5, 1]$	$[0, 0.01]$	$(0.01, 0.05]$	$(0.05, 0.1]$	$(0.1, 0.5]$	$(0.5, 1]$
Panel A: Goyal										
OLS	0%	0%	0%	84%	16%	0%	0%	0%	64%	36%
3PRF	0%	4%	12%	76%	8%	0%	1%	4%	77%	18%
PLS	0%	4%	28%	64%	4%	1%	10%	12%	59%	18%
PCR	20%	48%	12%	16%	4%	10%	35%	16%	27%	12%
LASSO	4%	0%	8%	84%	4%	1%	8%	9%	58%	24%
RIDGE	4%	4%	28%	60%	4%	1%	8%	19%	58%	14%
ENET	0%	8%	4%	84%	4%	0%	8%	8%	55%	29%
Panel B: FRED-MD										
OLS	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
3PRF	0%	0%	4%	68%	28%	0%	1%	0%	40%	59%
PLS	44%	36%	12%	8%	0%	26%	33%	15%	17%	9%
PCR	12%	40%	24%	24%	0%	5%	27%	24%	35%	9%
LASSO	4%	4%	4%	68%	20%	2%	5%	9%	51%	33%
RIDGE	4%	20%	20%	52%	4%	2%	18%	18%	48%	14%
ENET	4%	4%	4%	68%	20%	0%	8%	8%	62%	22%
Panel C: All Instruments										
OLS	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
3PRF	0%	0%	4%	72%	24%	0%	1%	0%	43%	56%
PLS	4%	40%	8%	44%	4%	2%	30%	15%	40%	13%
PCR	8%	4%	20%	60%	8%	2%	14%	16%	53%	15%
LASSO	4%	0%	0%	64%	32%	0%	3%	6%	44%	47%
RIDGE	0%	12%	12%	68%	8%	2%	8%	12%	58%	20%
ENET	4%	0%	8%	68%	20%	0%	4%	5%	54%	37%

Table 2.8.4 reports the proportion of p-values that falls into five different intervals of the Harvey et al. (1997) t-test of forecast accuracy for each one of the individual assets for the datasets with 100 and 25 portfolios formed on Size/BTM by estimator and set of conditioning information used. Harvey et al. (1997) is modification of the Diebold and Mariano (1995) aimed to improve small-sample properties of the test. There is no significant change when the original Diebold and Mariano (1995) test is used. We assess a one-sided test, in which the alternative hypothesis can be interpreted as our zero forecast benchmark is less accurate than each one of the seven estimators' forecasts. The first interval is for $[0 \leq \mathrm{p} < 0.01]$, the second for $(0.01 \leq \mathrm{p} < 0.05]$, the third for $(0.05 \leq \mathrm{p} < 0.10]$, the fourth for $(0.10 \leq \mathrm{p} < 0.50]$, and the last one for $(0.50 \leq \mathrm{p} < 1]$.

Table 2.8.5: Modified Diebold-Mariano (Harvey et al. (1997)) t-test - Proportion for Different Intervals

	5 Industry Portfolios					6 Portfolios Formed on Size and Book-to-Market				
	$[0, 0.01]$	$(0.01, 0.05]$	$(0.05, 0.1]$	$(0.1, 0.5]$	$(0.5, 1]$	$[0, 0.01]$	$(0.01, 0.05]$	$(0.05, 0.1]$	$(0.1, 0.5]$	$(0.5, 1]$
Panel A: Goyal										
OLS	0%	0%	0%	60%	40%	0%	0%	0%	100%	0%
3PRF	0%	0%	0%	100%	0%	0%	0%	17%	83%	0%
PLS	0%	0%	0%	100%	0%	0%	0%	33%	67%	0%
PCR	0%	20%	60%	20%	0%	17%	50%	0%	33%	0%
LASSO	0%	20%	0%	60%	20%	0%	17%	0%	83%	0%
RIDGE	0%	20%	0%	80%	0%	0%	17%	17%	67%	0%
ENET	0%	20%	0%	60%	20%	0%	0%	17%	83%	0%
Panel B: FRED-MD										
OLS	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
3PRF	0%	0%	0%	20%	80%	0%	0%	0%	67%	33%
PLS	40%	20%	20%	20%	0%	67%	17%	17%	0%	0%
PCR	0%	60%	40%	0%	0%	17%	33%	17%	33%	0%
LASSO	0%	20%	0%	80%	0%	0%	17%	0%	50%	33%
RIDGE	0%	20%	40%	40%	0%	17%	17%	33%	33%	0%
ENET	0%	20%	0%	80%	0%	0%	17%	0%	50%	33%
Panel C: All Instruments										
OLS	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
3PRF	0%	0%	0%	20%	80%	0%	0%	0%	50%	50%
PLS	0%	20%	40%	40%	0%	17%	33%	17%	33%	0%
PCR	0%	20%	40%	40%	0%	17%	0%	17%	50%	17%
LASSO	0%	20%	0%	40%	40%	0%	17%	0%	50%	33%
RIDGE	0%	20%	0%	80%	0%	0%	17%	17%	67%	0%
ENET	0%	20%	0%	60%	20%	0%	17%	0%	67%	17%

Table 2.8.5 reports the proportion of p-values that falls into five different intervals of the Harvey et al. (1997) t-test of forecast accuracy for each one of the individual assets for the datasets with 6 portfolios formed on Size/BTM and the 5 industry portfolios by estimator and set of conditioning information used. Harvey et al. (1997) is modification of the Diebold and Mariano (1995) aimed to improve small-sample properties of the test. There is no significant change when the original Diebold and Mariano (1995) test is used. We assess a one-sided test, in which the alternative hypothesis can be interpreted as our zero forecast benchmark is less accurate than each one of the seven estimators' forecasts. The first interval is for $[0 \leq p < 0.01]$, the second for $(0.01 \leq p < 0.05]$, the third for $(0.05 \leq p < 0.10]$, the fourth for $(0.10 \leq p < 0.50]$, and the last one for $(0.50 \leq p < 1]$.

2.8.4 Economic Value

Table 2.8.6: Implied Sharpe Ratios

	5 Industry Portfolios			6 Portfolios Formed on Size and Book-to-Market		
	CMV	UMV	MVATE	CMV	UMV	MVATE
Panel A: Goyal						
OLS	-	-	-	-	0.055	-
3PRF	-	-	-	0.102	0.164	0.104
PLS	-	-	-	0.211	0.210	0.211
PCR	0.069	0.051	0.063	0.297	0.298	0.299
LASSO	0.181	0.094	0.180	0.076	0.112	0.073
RIDGE	0.051	-	0.051	0.109	0.177	0.110
ENET	0.109	0.048	0.116	0.182	0.217	0.188
Panel B: FRED-MD						
OLS	-	-	-	-	-	-
3PRF	-	-	-	-	-	-
PLS	0.145	0.112	0.144	0.235	0.224	0.226
PCR	0.126	0.055	0.124	0.273	0.255	0.275
LASSO	-	-	-	0.095	-	0.095
RIDGE	0.079	-	0.078	0.224	0.210	0.227
ENET	-	-	-	0.172	0.172	0.172
Panel C: All Instruments						
OLS	-	-	-	-	-	-
3PRF	-	-	-	-	-	-
PLS	-	-	-	0.259	0.258	0.259
PCR	0.119	-	0.115	0.227	0.225	0.228
LASSO	-	-	-	0.049	-	0.062
RIDGE	-	-	-	0.212	0.190	0.212
ENET	-	-	-	0.069	-	0.101

Table 2.8.6 reports the implied SR^* for the 5 Industry portfolios and 6 portfolios formed on Size/BTM as given in equation (2.19) for all seven estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge and ENet), three different sets of conditioning information (Goyal's, FRED-MD, and "All Instruments", wich is the combination of the previous two with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI)) and three different mean-variance approaches (CMV, UMV, MVATE) to build efficient portfolios. Notice that when the OOS R^2 is negative, larger absolute R^2 values than the Sharpe ratio will cause the impossibility of the implied Sharpe ratio to be calculated. These cases are omitted in the table.

Economic Gains

Table 2.8.7: Certainty Equivalent Excess Returns (CER) (Monthly %)

	5 Industry Portfolios						6 Portfolios Formed on Size and Book-to-Market					
	CMV		UMV		MVATE		CMV		UMV		MVATE	
Panel A: Goyal												
OLS	0.15	[5]	0.10	[4]	0.18	[3]	0.21	[4]	0.19	[6]	0.32	[5]
3PRF	0.15	[4]	0.05	[5]	0.09	[5]	0.20	[6]	0.25	[3]	0.34	[4]
PLS	0.13	[6]	-0.04	[7]	-0.22	[6]	0.32	[2]	0.35	[2]	0.50	[2]
PCR	0.01	[7]	-0.03	[6]	-0.57	[7]	0.42	[1]	0.41	[1]	0.69	[1]
LASSO	0.28	[1]	0.16	[1]	0.42	[1]	0.18	[7]	0.15	[7]	0.26	[7]
RIDGE	0.16	[3]	0.12	[3]	0.13	[4]	0.20	[5]	0.22	[5]	0.29	[6]
ENET	0.22	[2]	0.15	[2]	0.28	[2]	0.25	[3]	0.22	[4]	0.42	[3]
Panel B: FRED-MD												
OLS	-0.01	[6]	0.00	[2]	-0.12	[5]	0.07	[7]	0.06	[7]	0.00	[7]
3PRF	0.04	[5]	-0.06	[4]	-0.09	[4]	0.21	[6]	0.20	[5]	0.33	[6]
PLS	-0.11	[7]	-0.14	[6]	-1.39	[7]	0.27	[5]	0.25	[4]	0.39	[5]
PCR	0.11	[4]	-0.40	[7]	-0.28	[6]	0.40	[1]	0.36	[1]	0.64	[1]
LASSO	0.23	[1]	0.02	[1]	0.19	[1]	0.30	[4]	0.18	[6]	0.46	[4]
RIDGE	0.13	[3]	-0.09	[5]	-0.06	[3]	0.32	[3]	0.27	[3]	0.52	[3]
ENET	0.16	[2]	-0.04	[3]	0.07	[2]	0.34	[2]	0.27	[2]	0.55	[2]
Panel C: All Instruments												
OLS	-0.04	[6]	-0.01	[3]	-0.22	[5]	0.06	[7]	0.06	[7]	-0.09	[7]
3PRF	0.03	[4]	-0.03	[4]	-0.12	[3]	0.21	[6]	0.22	[4]	0.32	[6]
PLS	-0.27	[7]	-0.05	[5]	-1.05	[7]	0.32	[3]	0.32	[2]	0.54	[3]
PCR	0.14	[2]	-0.37	[7]	-0.43	[6]	0.37	[1]	0.33	[1]	0.55	[2]
LASSO	0.15	[1]	0.01	[1]	0.04	[1]	0.33	[2]	0.19	[6]	0.55	[1]
RIDGE	0.12	[3]	-0.10	[6]	-0.08	[2]	0.32	[4]	0.27	[3]	0.51	[5]
ENET	0.03	[5]	0.00	[2]	-0.17	[4]	0.30	[5]	0.19	[5]	0.52	[4]

Table 2.8.7 summarises the CER (monthly %) by estimator and optimal portfolio framework (CMV, UMV and MVATE) for the 5 Industry portfolios and 6 portfolios formed on Size/BTM. Panel A reports the Sharpe ratios generated when the variables from Goyal's website are used as **Z**. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. Panel B presents the Sharpe ratios obtained using the FRED-MD variables. The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, panel C shows the Sharpe rations when all variables are used as conditioning information. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). Inside brackets, we also report the ranking among all 7 estimators within strategy and set of conditioning information employed

Figure 2.8.8: Management Fee

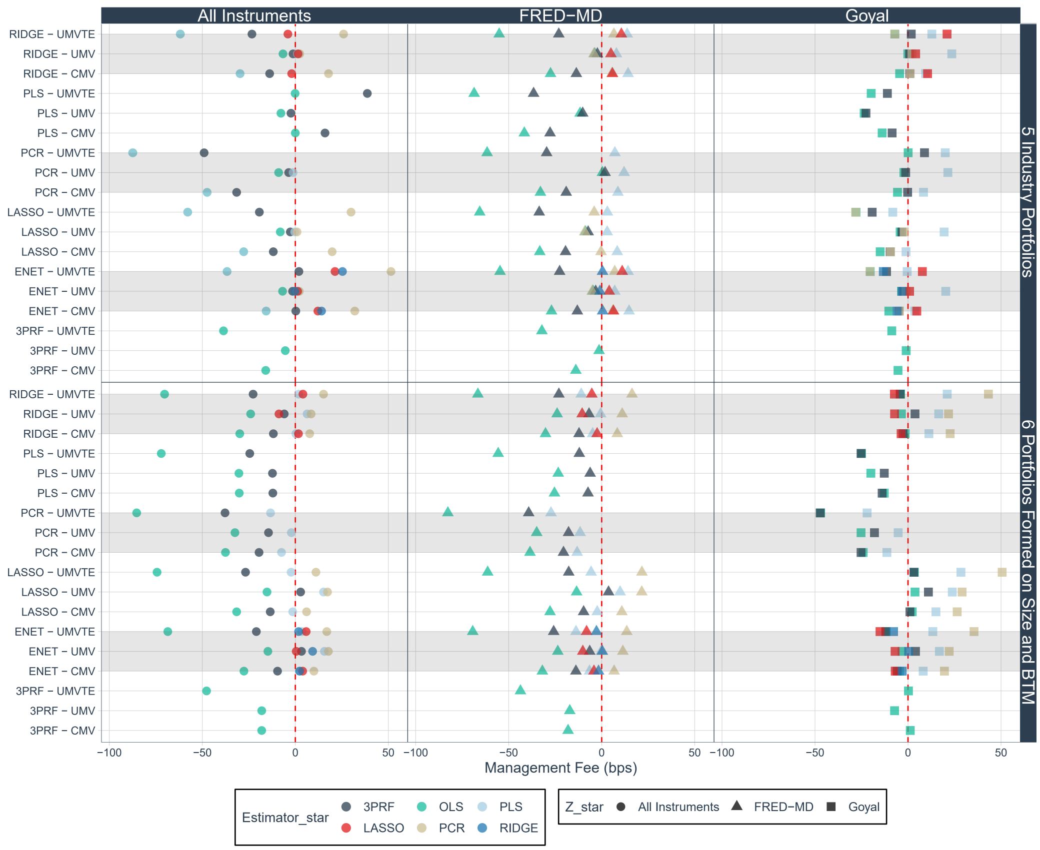

Figure 2.8.8 presents the management fee in bps for the 5 Industry portfolios and 6 portfolios formed on Size/BTM, computed as the solution for $\mathcal{F}$ in equation (2.22), from an investor switching from an optimal portfolio formed by the estimator and mean-variance framework given in the left axis to another portfolio plotted. The comparison is done in pairs of optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and sets of conditioning information.

2.8.5 Out-of-Sample Analysis - Portfolio Efficiency

Figure 2.8.9: p-values

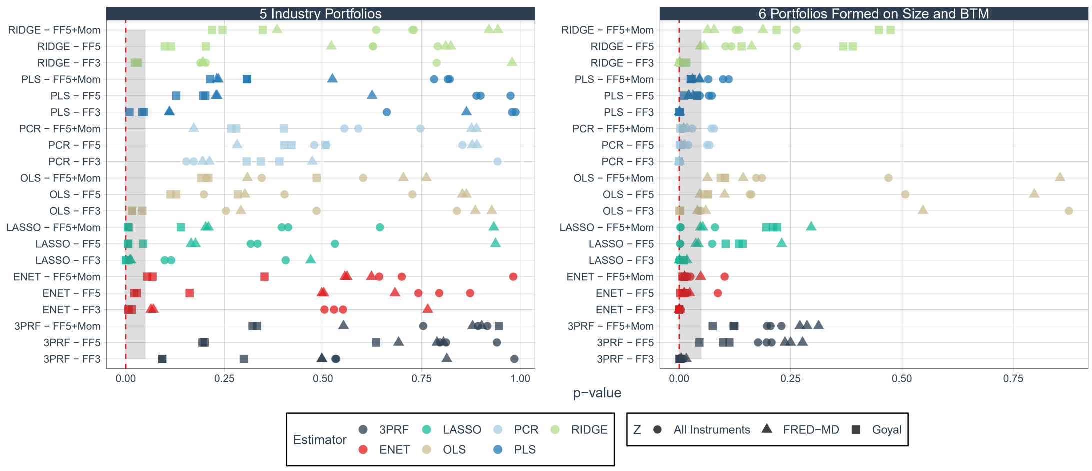

Figure 2.8.9 reports for the 5 Industry portfolios and 6 portfolios formed on Size/BTM the p-values of alphas of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and set of conditioning information used on the Fama-French 3, 5, and 5 + momentum factor models. The p-values are calculated from Newey–West t-statistics computed with one lag.

Figure 2.8.10: R-Squared

Figure 2.8.10 reports for the 5 Industry portfolios and 6 portfolios formed on Size/BTM the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and set of conditioning information used on the Fama-French 3, 5, and 5 + momentum factor models.

Table 2.8.8: Alphas (Monthly %) - 25 and 100 Portfolios Formed on Size and Book-to-Market - "All Instruments"

		25 Portfolios Formed on Size and Book-to-Market									100 Portfolios Formed on Size and Book-to-Market								
		CMV			UMV			MVATE			CMV			UMV			MVATE		
		FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom
Z: All Instruments																			
OLS	α (%)	0.034	0.027	0.025	0.034	0.027	0.025	-0.218	-0.246	-0.237	0.018	0.016	0.015	0.018	0.016	0.015	-0.420	-0.226	-0.181
	t-statistics	[3.07]	[2.23]	[2.39]	[3.11]	[2.3]	[2.4]	[-1.1]	[-1.27]	[-1.18]	[1.67]	[1.67]	[1.72]	[1.66]	[1.67]	[1.72]	[-0.79]	[-0.5]	[-0.38]
	p-val	0.002**	0.027*	0.018*	0.002**	0.022*	0.017*	0.273	0.204	0.238	0.097	0.096	0.087	0.098	0.097	0.088	0.433	0.618	0.706
	R^2	0.02	0.04	0.12	0.02	0.04	0.12	0.07	0.07	0.08	0.01	0.01	0.01	0.01	0.01	0.01	0.06	0.06	0.08
3PRF	α (%)	0.175	0.150	0.141	0.167	0.140	0.131	0.295	0.242	0.229	0.163	0.147	0.147	0.159	0.143	0.143	0.245	0.209	0.221
	t-statistics	[4.01]	[3.3]	[3.4]	[3.77]	[2.96]	[3.24]	[3.3]	[2.56]	[2.66]	[3.7]	[2.72]	[3.01]	[3.57]	[2.64]	[2.91]	[2.18]	[1.83]	[1.84]
	p-val	0**	0.001**	0.001**	0**	0.003**	0.001**	0.001**	0.011*	0.008**	0**	0.007**	0.003**	0**	0.009**	0.004**	0.03*	0.068	0.067
	R^2	0.09	0.11	0.16	0.07	0.09	0.15	0.12	0.15	0.17	0.04	0.05	0.05	0.04	0.05	0.05	0.07	0.08	0.08
PLS	α (%)	0.270	0.204	0.200	0.254	0.175	0.172	0.507	0.375	0.369	0.176	0.148	0.142	0.163	0.136	0.130	0.331	0.257	0.256
	t-statistics	[4]	[3.12]	[2.93]	[3.59]	[2.66]	[2.55]	[3.7]	[2.87]	[2.93]	[4.36]	[3.72]	[3.74]	[4.36]	[3.64]	[3.63]	[2.68]	[2.92]	[2.75]
	p-val	0**	0.002**	0.004**	0**	0.008**	0.011*	0**	0.005**	0.004**	0**	0**	0**	0**	0**	0**	0.008**	0.004**	0.006**
	R^2	0.08	0.16	0.18	0.04	0.15	0.16	0.08	0.17	0.18	0.03	0.07	0.13	0.04	0.08	0.15	0.07	0.11	0.11
PCR	α (%)	0.332	0.278	0.259	0.301	0.265	0.247	0.652	0.542	0.506	0.267	0.251	0.246	0.249	0.232	0.227	0.533	0.498	0.489
	t-statistics	[4.44]	[3.87]	[3.77]	[4.1]	[3.7]	[3.85]	[4.47]	[4.03]	[3.77]	[3.57]	[3.3]	[3.27]	[3.62]	[3.36]	[3.3]	[3.64]	[3.36]	[3.3]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.001**	0.001**	0**	0.001**	0.001**	0**	0.001**	0.001**
	R^2	0.24	0.25	0.38	0.25	0.25	0.37	0.24	0.26	0.38	0.00	0.01	0.02	0.00	0.01	0.03	0.00	0.01	0.02
LASSO	α (%)	0.178	0.139	0.132	0.139	0.114	0.108	0.336	0.260	0.248	0.134	0.121	0.119	0.128	0.116	0.114	0.244	0.221	0.215
	t-statistics	[3.81]	[3.22]	[3.18]	[3.65]	[3.21]	[2.98]	[3.84]	[3.21]	[3.21]	[3.75]	[3.18]	[3.51]	[3.72]	[3.07]	[3.3]	[3.3]	[2.74]	[2.8]
	p-val	0**	0.001**	0.002**	0**	0.002**	0.003**	0**	0.002**	0.002**	0**	0.002**	0.001**	0**	0.002**	0.001**	0.001**	0.007**	0.006**
	R^2	0.08	0.11	0.14	0.08	0.10	0.14	0.08	0.11	0.15	0.04	0.05	0.06	0.05	0.06	0.07	0.03	0.04	0.06
RIDGE	α (%)	0.277	0.207	0.193	0.271	0.208	0.193	0.513	0.380	0.353	0.119	0.099	0.094	0.118	0.099	0.094	0.144	0.108	0.103
	t-statistics	[4.31]	[2.92]	[3.32]	[3.89]	[2.87]	[3.26]	[4.37]	[2.9]	[2.95]	[3.4]	[2.83]	[2.75]	[3.38]	[2.87]	[2.76]	[1.34]	[1.02]	[1.03]
	p-val	0**	0.004**	0.001**	0**	0.004**	0.001**	0**	0.004**	0.004**	0.001**	0.005**	0.006**	0.001**	0.005**	0.006**	0.180	0.307	0.305
	R^2	0.12	0.17	0.28	0.11	0.16	0.29	0.13	0.18	0.28	0.09	0.11	0.15	0.09	0.11	0.15	0.11	0.11	0.12
ENET	α (%)	0.213	0.182	0.177	0.165	0.141	0.137	0.423	0.350	0.344	0.140	0.130	0.128	0.134	0.124	0.122	0.226	0.190	0.192
	t-statistics	[4.79]	[4.19]	[4.05]	[4.13]	[3.48]	[3.48]	[4.64]	[4]	[3.98]	[4.03]	[3.79]	[3.99]	[3.9]	[3.67]	[3.85]	[2.73]	[2.14]	[1.83]
	p-val	0**	0**	0**	0**	0.001**	0.001**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.007**	0.033*	0.069
	R^2	0.10	0.12	0.15	0.08	0.10	0.13	0.06	0.09	0.11	0.05	0.06	0.08	0.06	0.07	0.09	0.01	0.02	0.03

Table 2.8.8 reports the alphas, t-statistics, p-values and the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) using the FRED-MD as conditioning information on the Fama-French 3, 5, and 5 + momentum factor models. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. FRED-MD is a large dataset containing 128 macroeconomic and financial variables.

Table 2.8.9: Alphas (Monthly %) - 25 and 100 Portfolios Formed on Size and Book-to-Market - Goyal Variables

		25 Portfolios Formed on Size and Book-to-Market									100 Portfolios Formed on Size and Book-to-Market								
		CMV			UMV			MVATE			CMV			UMV			MVATE		
		FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom
Z: Goyal																			
OLS	α (%)	0.199	0.184	0.169	0.193	0.176	0.163	0.329	0.309	0.287	0.076	0.072	0.065	0.074	0.070	0.064	-0.023	0.002	0.004
	t-statistics	[5.06]	[4.68]	[4.71]	[4.7]	[4.42]	[4.25]	[4.18]	[3.92]	[4.03]	[1.59]	[1.82]	[1.54]	[1.56]	[1.8]	[1.52]	[-0.16]	[0.01]	[0.03]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.112	0.070	0.124	0.120	0.073	0.128	0.873	0.989	0.974
	R^2	0.01	0.05	0.18	0.02	0.06	0.18	0.01	0.05	0.14	0.01	0.03	0.04	0.01	0.03	0.04	0.01	0.02	0.02
3PRF	α (%)	0.226	0.210	0.194	0.244	0.224	0.210	0.403	0.381	0.357	0.143	0.128	0.121	0.146	0.129	0.122	0.134	0.102	0.101
	t-statistics	[4.99]	[5.06]	[4.97]	[4.89]	[4.71]	[4.53]	[4.63]	[4.54]	[4.61]	[2.46]	[3.04]	[2.74]	[2.68]	[3.29]	[2.99]	[0.98]	[0.92]	[0.91]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.015*	0.003**	0.007**	0.008**	0.001**	0.003**	0.328	0.358	0.363
	R^2	0.03	0.06	0.19	0.03	0.07	0.18	0.02	0.05	0.14	0.01	0.02	0.03	0.01	0.02	0.03	0.01	0.02	0.02
PLS	α (%)	0.292	0.254	0.235	0.352	0.303	0.285	0.537	0.472	0.435	0.242	0.217	0.206	0.243	0.217	0.207	0.462	0.421	0.403
	t-statistics	[4.35]	[3.78]	[3.51]	[3.73]	[3.52]	[3.15]	[4.64]	[3.87]	[3.78]	[3.9]	[3.97]	[3.67]	[3.85]	[3.96]	[3.62]	[4.38]	[3.97]	[4.14]
	p-val	0**	0**	0.001**	0**	0.001**	0.002**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**
	R^2	0.11	0.13	0.22	0.09	0.12	0.20	0.12	0.13	0.23	0.01	0.02	0.05	0.01	0.02	0.05	0.02	0.02	0.05
PCR	α (%)	0.367	0.326	0.311	0.381	0.329	0.318	0.700	0.624	0.594	0.277	0.267	0.260	0.272	0.259	0.252	0.529	0.510	0.496
	t-statistics	[5.38]	[4.64]	[4.52]	[4.62]	[4.07]	[3.98]	[5.36]	[4.65]	[4.56]	[4.21]	[3.98]	[3.73]	[4.21]	[3.78]	[3.77]	[4.23]	[3.83]	[3.83]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**	0**
	R^2	0.17	0.19	0.24	0.13	0.16	0.19	0.18	0.20	0.25	0.01	0.02	0.03	0.01	0.02	0.04	0.02	0.03	0.04
LASSO	α (%)	0.169	0.159	0.151	0.151	0.140	0.133	0.326	0.305	0.291	0.089	0.079	0.079	0.092	0.081	0.080	0.164	0.136	0.131
	t-statistics	[4.09]	[4.29]	[3.75]	[4.43]	[4.45]	[4.07]	[4.04]	[4.51]	[3.8]	[1.71]	[1.71]	[1.76]	[1.83]	[1.94]	[1.94]	[1.57]	[1.61]	[1.55]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.088	0.089	0.080	0.069	0.054	0.053	0.118	0.108	0.123
	R^2	0.06	0.07	0.11	0.07	0.08	0.13	0.07	0.08	0.12	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.02	0.02
RIDGE	α (%)	0.258	0.232	0.214	0.239	0.213	0.199	0.456	0.408	0.375	0.092	0.092	0.089	0.104	0.103	0.100	0.120	0.119	0.117
	t-statistics	[5.67]	[5.29]	[5.25]	[5.3]	[5.16]	[4.86]	[5.3]	[4.95]	[5.31]	[1.21]	[1.38]	[1.39]	[1.58]	[1.81]	[1.77]	[0.77]	[0.9]	[0.87]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.225	0.169	0.166	0.116	0.072	0.077	0.444	0.366	0.385
	R^2	0.05	0.07	0.19	0.08	0.11	0.22	0.05	0.07	0.19	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01
ENET	α (%)	0.192	0.179	0.168	0.176	0.157	0.148	0.329	0.300	0.283	0.094	0.085	0.085	0.095	0.084	0.085	0.076	0.052	0.060
	t-statistics	[5.49]	[5.47]	[5]	[5.42]	[5.61]	[5.17]	[4.1]	[4.02]	[3.9]	[1.66]	[1.8]	[1.88]	[1.73]	[1.9]	[1.99]	[0.53]	[0.46]	[0.53]
	p-val	0**	0**	0**	0**	0**	0**	0**	0**	0**	0.099	0.073	0.061	0.084	0.059	0.047*	0.596	0.646	0.595
	R^2	0.04	0.06	0.14	0.06	0.09	0.17	0.04	0.06	0.11	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01

Table 2.8.9 reports the alphas, t-statistics, p-values and the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) using Goyal's variables as conditioning information on the Fama-French 3, 5, and 5 + momentum factor models. Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*.

Table 2.8.10: Alphas from Factor Models (Monthly %) - 5 Industry Portfolios and 6 Portfolios Formed on Size and Book-to-Market - FRED-MD

		5 Industry Portfolios									6 Portfolios Formed on Size and Book-to-Market								
		CMV			UMV			MVATE			CMV			UMV			MVATE		
		FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom
Z: FRED-MD																			
OLS	α (%)	-0.006	-0.007	-0.012	-0.003	-0.006	-0.011	-0.090	-0.092	-0.095	0.070	0.062	0.058	0.062	0.048	0.044	0.039	0.018	0.013
	t-statistics	[-0.14]	[-0.17]	[-0.3]	[-0.09]	[-0.19]	[-0.38]	[-1.06]	[-1.03]	[-1.02]	[2.05]	[2]	[1.86]	[1.89]	[1.64]	[1.47]	[0.6]	[0.26]	[0.18]
	p-val	0.886	0.864	0.762	0.928	0.853	0.704	0.292	0.302	0.308	0.041*	0.047*	0.063	0.060	0.102	0.144	0.547	0.797	0.855
	R^2	0.04	0.05	0.07	0.04	0.05	0.08	0.03	0.04	0.04	0.05	0.06	0.07	0.06	0.08	0.10	0.04	0.04	0.04
3PRF	α (%)	0.091	0.039	0.019	-0.025	-0.045	-0.066	0.178	0.083	0.046	0.210	0.074	0.067	0.203	0.082	0.071	0.399	0.143	0.130
	t-statistics	[0.68]	[0.25]	[0.12]	[-0.24]	[-0.4]	[-0.6]	[0.68]	[0.27]	[0.15]	[2.84]	[1.15]	[1.07]	[2.42]	[1.09]	[1.01]	[2.91]	[1.18]	[1.1]
	p-val	0.497	0.805	0.902	0.814	0.692	0.552	0.497	0.789	0.879	0.005**	0.250	0.287	0.016*	0.277	0.313	0.004**	0.238	0.271
	R^2	0.10	0.12	0.15	0.10	0.10	0.14	0.10	0.12	0.14	0.35	0.49	0.50	0.24	0.38	0.40	0.38	0.52	0.53
PLS	α (%)	0.484	0.495	0.434	0.035	-0.111	-0.128	0.943	0.974	0.856	0.275	0.153	0.141	0.235	0.133	0.113	0.520	0.285	0.260
	t-statistics	[1.6]	[1.2]	[1.2]	[0.17]	[-0.49]	[-0.64]	[1.6]	[1.21]	[1.19]	[3.74]	[2.3]	[2.01]	[3.25]	[2.32]	[2.18]	[3.73]	[2.17]	[2.01]
	p-val	0.110	0.232	0.231	0.864	0.624	0.524	0.111	0.229	0.234	0**	0.022*	0.046*	0.001**	0.021*	0.03*	0**	0.031*	0.046*
	R^2	0.10	0.13	0.17	0.09	0.11	0.12	0.10	0.13	0.17	0.54	0.62	0.63	0.55	0.60	0.65	0.55	0.63	0.64
PCR	α (%)	0.251	0.028	-0.024	-0.150	-0.196	-0.248	0.468	0.050	-0.053	0.393	0.219	0.203	0.324	0.186	0.171	0.771	0.434	0.403
	t-statistics	[1.3]	[0.15]	[-0.14]	[-0.72]	[-1.08]	[-1.37]	[1.25]	[0.14]	[-0.16]	[4.3]	[2.9]	[2.58]	[4.37]	[2.52]	[2.38]	[4.27]	[2.92]	[2.59]
	p-val	0.195	0.878	0.889	0.473	0.282	0.172	0.212	0.890	0.877	0**	0.004**	0.011*	0**	0.012*	0.018*	0**	0.004**	0.01*
	R^2	0.12	0.19	0.26	0.24	0.24	0.28	0.12	0.18	0.25	0.42	0.55	0.57	0.47	0.56	0.58	0.42	0.55	0.57
LASSO	α (%)	0.310	0.201	0.176	0.081	0.010	-0.010	0.575	0.377	0.330	0.283	0.165	0.154	0.168	0.074	0.061	0.548	0.325	0.303
	t-statistics	[2.59]	[1.39]	[1.28]	[0.73]	[0.08]	[-0.08]	[2.51]	[1.35]	[1.26]	[3.47]	[2.04]	[1.94]	[2.39]	[1.2]	[1.05]	[3.49]	[2.09]	[1.99]
	p-val	0.01*	0.166	0.203	0.469	0.937	0.933	0.013*	0.177	0.210	0.001**	0.043*	0.053	0.017*	0.230	0.296	0.001**	0.037*	0.048*
	R^2	0.07	0.10	0.13	0.02	0.04	0.07	0.06	0.10	0.13	0.28	0.37	0.38	0.17	0.29	0.34	0.28	0.37	0.38
RIDGE	α (%)	0.218	0.043	0.012	-0.004	-0.095	-0.131	0.417	0.090	0.034	0.305	0.133	0.122	0.250	0.098	0.086	0.596	0.272	0.254
	t-statistics	[1.3]	[0.22]	[0.07]	[-0.03]	[-0.64]	[-0.87]	[1.3]	[0.24]	[0.1]	[3.97]	[1.92]	[1.77]	[3.15]	[1.4]	[1.32]	[3.96]	[1.99]	[1.86]
	p-val	0.196	0.824	0.943	0.979	0.521	0.383	0.194	0.811	0.921	0**	0.056	0.078	0.002**	0.163	0.188	0**	0.048*	0.064
	R^2	0.12	0.19	0.22	0.09	0.11	0.17	0.12	0.18	0.21	0.41	0.57	0.59	0.38	0.55	0.57	0.42	0.58	0.59
ENET	α (%)	0.246	0.119	0.099	0.037	-0.056	-0.070	0.466	0.235	0.196	0.313	0.203	0.194	0.254	0.151	0.139	0.606	0.399	0.383
	t-statistics	[1.86]	[0.67]	[0.58]	[0.3]	[-0.41]	[-0.49]	[1.82]	[0.68]	[0.59]	[3.82]	[2.57]	[2.53]	[3.61]	[2.27]	[1.98]	[3.72]	[2.55]	[2.53]
	p-val	0.064	0.502	0.561	0.766	0.682	0.623	0.070	0.497	0.556	0**	0.011*	0.012*	0**	0.024*	0.048*	0**	0.011*	0.012*
	R^2	0.06	0.11	0.13	0.02	0.05	0.07	0.06	0.10	0.12	0.33	0.41	0.42	0.27	0.37	0.40	0.32	0.40	0.40

Table 2.8.10 reports the alphas, t-statistics, p-values and the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) using the FRED-MD as conditioning information on the Fama-French 3, 5, and 5 + momentum factor models. FRED-MD is a large dataset containing 128 macroeconomic and financial variables.

Table 2.8.11: Alphas (Monthly %) - 5 Industry Portfolios and 6 Portfolios Formed on Size and Book-to-Market - “All Instruments”

		5 Industry Portfolios									6 Portfolios Formed on Size and Book-to-Market								
		CMV			UMV			MVATE			CMV			UMV			MVATE		
		FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom
Z: All Instruments																			
OLS	α (%)	-0.032	-0.042	-0.047	-0.006	-0.011	-0.015	-0.162	-0.176	-0.178	0.064	0.046	0.044	0.063	0.044	0.041	-0.016	-0.070	-0.070
	t-statistics	[-0.7]	[-0.84]	[-0.95]	[-0.2]	[-0.35]	[-0.52]	[-1.14]	[-1.29]	[-1.27]	[1.98]	[1.41]	[1.37]	[2.02]	[1.4]	[1.33]	[-0.16]	[-0.66]	[-0.72]
	p-val	0.484	0.402	0.345	0.840	0.727	0.602	0.254	0.198	0.205	0.049*	0.159	0.173	0.044*	0.162	0.186	0.875	0.508	0.470
	R^2	0.01	0.03	0.05	0.02	0.04	0.08	0.01	0.02	0.02	0.08	0.11	0.13	0.11	0.15	0.18	0.04	0.05	0.05
3PRF	α (%)	0.086	0.037	0.016	0.002	-0.009	-0.037	0.168	0.080	0.041	0.213	0.086	0.079	0.219	0.096	0.084	0.393	0.154	0.143
	t-statistics	[0.63]	[0.24]	[0.1]	[0.02]	[-0.07]	[-0.31]	[0.62]	[0.26]	[0.13]	[3.05]	[1.35]	[1.29]	[2.6]	[1.26]	[1.2]	[3.05]	[1.3]	[1.27]
	p-val	0.531	0.812	0.917	0.985	0.941	0.754	0.534	0.796	0.893	0.003**	0.177	0.197	0.01**	0.207	0.230	0.003**	0.196	0.205
	R^2	0.09	0.10	0.13	0.13	0.13	0.20	0.09	0.10	0.13	0.35	0.49	0.50	0.23	0.38	0.41	0.39	0.52	0.53
PLS	α (%)	0.003	-0.025	-0.042	0.063	-0.005	-0.041	-0.010	-0.055	-0.086	0.321	0.174	0.164	0.314	0.166	0.151	0.633	0.348	0.329
	t-statistics	[0.01]	[-0.13]	[-0.23]	[0.44]	[-0.03]	[-0.28]	[-0.03]	[-0.14]	[-0.23]	[3.3]	[1.8]	[1.6]	[3.4]	[2.01]	[1.85]	[3.34]	[1.84]	[1.66]
	p-val	0.988	0.900	0.821	0.662	0.976	0.782	0.979	0.889	0.816	0.001**	0.073	0.111	0.001**	0.046*	0.065	0.001**	0.067	0.099
	R^2	0.04	0.05	0.07	0.05	0.09	0.23	0.04	0.05	0.06	0.30	0.43	0.46	0.29	0.41	0.47	0.31	0.44	0.46
PCR	α (%)	0.376	0.168	0.131	0.020	-0.046	-0.083	0.697	0.302	0.231	0.359	0.171	0.160	0.298	0.187	0.178	0.708	0.343	0.321
	t-statistics	[1.43]	[0.71]	[0.59]	[0.07]	[-0.18]	[-0.32]	[1.37]	[0.66]	[0.54]	[3.23]	[1.83]	[1.77]	[3.88]	[2.34]	[2.19]	[3.24]	[1.87]	[1.81]
	p-val	0.153	0.478	0.554	0.943	0.854	0.747	0.172	0.508	0.589	0.001**	0.068	0.078	0**	0.02*	0.029*	0.001**	0.063	0.072
	R^2	0.11	0.16	0.21	0.16	0.17	0.21	0.11	0.16	0.21	0.44	0.55	0.57	0.53	0.58	0.59	0.44	0.56	0.58
LASSO	α (%)	0.227	0.152	0.127	0.085	0.075	0.052	0.419	0.284	0.236	0.336	0.227	0.216	0.199	0.112	0.100	0.652	0.446	0.426
	t-statistics	[1.66]	[1]	[0.85]	[0.83]	[0.63]	[0.46]	[1.58]	[0.97]	[0.82]	[4.42]	[3.09]	[2.96]	[3.15]	[1.79]	[1.76]	[4.45]	[3.11]	[3.04]
	p-val	0.099	0.317	0.396	0.405	0.530	0.644	0.115	0.334	0.412	0**	0.002**	0.003**	0.002**	0.074	0.080	0**	0.002**	0.003**
	R^2	0.11	0.12	0.19	0.08	0.09	0.17	0.10	0.12	0.19	0.21	0.30	0.32	0.12	0.24	0.30	0.21	0.30	0.33
RIDGE	α (%)	0.228	0.088	0.060	0.043	-0.046	-0.082	0.435	0.173	0.121	0.313	0.132	0.121	0.260	0.096	0.083	0.603	0.258	0.239
	t-statistics	[1.32]	[0.49]	[0.34]	[0.27]	[-0.27]	[-0.48]	[1.28]	[0.49]	[0.35]	[3.62]	[1.63]	[1.53]	[2.97]	[1.12]	[1.12]	[3.6]	[1.57]	[1.5]
	p-val	0.189	0.627	0.730	0.788	0.791	0.635	0.203	0.627	0.727	0**	0.104	0.127	0.003**	0.265	0.263	0**	0.117	0.135
	R^2	0.12	0.16	0.22	0.09	0.11	0.20	0.12	0.16	0.21	0.36	0.54	0.57	0.32	0.49	0.54	0.37	0.54	0.57
ENET	α (%)	0.086	-0.045	-0.067	0.062	0.017	-0.002	0.177	-0.071	-0.112	0.286	0.190	0.181	0.185	0.094	0.086	0.581	0.392	0.378
	t-statistics	[0.63]	[-0.33]	[-0.46]	[0.6]	[0.16]	[-0.02]	[0.67]	[-0.26]	[-0.39]	[3.33]	[2.37]	[2.25]	[2.93]	[1.72]	[1.64]	[3.39]	[2.47]	[2.37]
	p-val	0.528	0.742	0.643	0.551	0.873	0.982	0.504	0.795	0.700	0.001**	0.018*	0.025*	0.004**	0.087	0.102	0.001**	0.014*	0.018*
	R^2	0.10	0.14	0.20	0.06	0.07	0.14	0.10	0.13	0.18	0.24	0.32	0.33	0.16	0.28	0.31	0.21	0.29	0.30

Table 2.8.11 reports the alphas, t-statistics, p-values and the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) using the FRED-MD as conditioning information on the Fama-French 3, 5, and 5 + momentum factor models. “All Instruments” is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. FRED-MD is a large dataset containing 128 macroeconomic and financial variables.

Table 2.8.12: Alphas (Monthly %) - 5 Industry Portfolios and 6 Portfolios Formed on Size and Book-to-Market - Goyal Variables

		5 Industry Portfolios									6 Portfolios Formed on Size and Book-to-Market								
		CMV			UMV			MVATE			CMV			UMV			MVATE		
		FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom	FF3	FF5	FF5 + Mom
Z: Goyal																			
OLS	α (%)	0.202	0.138	0.109	0.197	0.122	0.073	0.394	0.277	0.222	0.220	0.125	0.107	0.205	0.117	0.101	0.406	0.238	0.206
	t-statistics	[2.43]	[1.53]	[1.26]	[2.04]	[1.07]	[0.7]	[2.43]	[1.59]	[1.31]	[3.25]	[1.86]	[1.64]	[3.45]	[1.91]	[1.68]	[3.17]	[1.86]	[1.64]
	p-val	0.016*	0.127	0.209	0.042*	0.284	0.484	0.016*	0.113	0.191	0.001**	0.064	0.103	0.001**	0.057	0.093	0.002**	0.064	0.103
	R^2	0.09	0.12	0.22	0.14	0.17	0.35	0.09	0.11	0.20	0.11	0.21	0.27	0.09	0.21	0.29	0.13	0.21	0.27
3PRF	α (%)	0.203	0.153	0.109	0.153	0.072	0.010	0.391	0.304	0.219	0.209	0.113	0.103	0.272	0.147	0.134	0.401	0.222	0.207
	t-statistics	[1.69]	[1.28]	[0.97]	[1.04]	[0.48]	[0.07]	[1.69]	[1.3]	[0.99]	[3.04]	[1.66]	[1.54]	[3.46]	[2.01]	[1.79]	[3.14]	[1.59]	[1.55]
	p-val	0.093	0.201	0.333	0.299	0.635	0.946	0.093	0.195	0.321	0.003**	0.098	0.124	0.001**	0.045*	0.075	0.002**	0.113	0.122
	R^2	0.22	0.23	0.35	0.16	0.22	0.38	0.22	0.23	0.34	0.12	0.24	0.27	0.09	0.26	0.30	0.13	0.24	0.26
PLS	α (%)	0.352	0.236	0.185	0.528	0.408	0.289	0.676	0.462	0.362	0.310	0.204	0.190	0.363	0.219	0.193	0.595	0.399	0.373
	t-statistics	[2.04]	[1.3]	[1.02]	[2.62]	[1.53]	[1.25]	[2]	[1.28]	[1.02]	[3.35]	[2.06]	[2.23]	[3.64]	[2.58]	[2.21]	[3.31]	[2.06]	[2.24]
	p-val	0.042*	0.196	0.307	0.009**	0.128	0.214	0.046*	0.202	0.308	0.001**	0.04*	0.026*	0**	0.01*	0.028*	0.001**	0.04*	0.026*
	R^2	0.13	0.16	0.23	0.15	0.21	0.39	0.13	0.15	0.22	0.34	0.41	0.43	0.25	0.34	0.40	0.34	0.41	0.43
PCR	α (%)	0.184	-0.141	-0.198	0.164	-0.126	-0.164	0.329	-0.285	-0.396	0.405	0.245	0.236	0.401	0.233	0.224	0.781	0.480	0.462
	t-statistics	[1.02]	[-0.81]	[-1.08]	[0.86]	[-0.67]	[-0.84]	[0.95]	[-0.84]	[-1.11]	[4.68]	[3.12]	[2.97]	[4.42]	[2.83]	[2.75]	[4.7]	[3.24]	[3]
	p-val	0.307	0.420	0.280	0.390	0.506	0.400	0.343	0.402	0.267	0**	0.002**	0.003**	0**	0.005**	0.006**	0**	0.001**	0.003**
	R^2	0.11	0.22	0.30	0.12	0.22	0.26	0.11	0.21	0.29	0.41	0.53	0.54	0.39	0.53	0.54	0.42	0.54	0.55
LASSO	α (%)	0.341	0.300	0.260	0.241	0.178	0.139	0.659	0.588	0.510	0.186	0.107	0.092	0.161	0.080	0.069	0.356	0.206	0.179
	t-statistics	[3.43]	[2.76]	[2.73]	[2.72]	[2.02]	[1.48]	[3.41]	[2.78]	[2.81]	[2.6]	[1.47]	[1.23]	[3.05]	[1.63]	[1.3]	[2.63]	[1.5]	[1.25]
	p-val	0.001**	0.006**	0.007**	0.007**	0.044*	0.140	0.001**	0.006**	0.005**	0.01**	0.142	0.220	0.003**	0.104	0.196	0.009**	0.134	0.211
	R^2	0.11	0.11	0.26	0.08	0.10	0.25	0.10	0.11	0.26	0.24	0.29	0.33	0.22	0.32	0.36	0.26	0.31	0.34
RIDGE	α (%)	0.231	0.154	0.118	0.198	0.116	0.078	0.447	0.311	0.241	0.223	0.069	0.058	0.230	0.100	0.086	0.434	0.138	0.117
	t-statistics	[2.2]	[1.58]	[1.17]	[2.29]	[1.28]	[0.94]	[2.17]	[1.66]	[1.23]	[2.42]	[0.86]	[0.72]	[3.18]	[1.48]	[1.23]	[2.47]	[0.9]	[0.76]
	p-val	0.029*	0.114	0.245	0.023*	0.203	0.347	0.031*	0.099	0.218	0.016*	0.390	0.475	0.002**	0.140	0.219	0.014*	0.369	0.448
	R^2	0.13	0.15	0.24	0.16	0.21	0.34	0.12	0.14	0.23	0.26	0.40	0.42	0.13	0.31	0.35	0.27	0.42	0.43
ENET	α (%)	0.296	0.245	0.204	0.225	0.120	0.084	0.594	0.500	0.421	0.264	0.163	0.152	0.229	0.146	0.131	0.525	0.329	0.311
	t-statistics	[2.76]	[2.2]	[1.83]	[2.44]	[1.4]	[0.93]	[2.7]	[2.31]	[1.94]	[3.92]	[2.6]	[2.53]	[4.14]	[3.08]	[2.53]	[4.04]	[2.73]	[2.69]
	p-val	0.006**	0.029*	0.068	0.015*	0.162	0.352	0.007**	0.021*	0.054	0**	0.01**	0.012*	0**	0.002**	0.012*	0**	0.007**	0.008**
	R^2	0.13	0.14	0.26	0.06	0.11	0.23	0.13	0.13	0.25	0.21	0.30	0.32	0.15	0.26	0.33	0.22	0.31	0.32

Table 2.8.12 reports the alphas, t-statistics, p-values and the R^2 of the regressions of the optimal portfolios generated by each mean-variance strategy (CMV, UMV, and MVATE), estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) using Goyal's variables as conditioning information on the Fama-French 3, 5, and 5 + momentum factor models. Goyal's variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*.

2.8.6 Financial Metrics

Table 2.8.13: Turnover and Financial Metrics (Monthly %)

	5 Industry Portfolios			6 Portfolios Formed on Size and Book-to-Market		
	Turnover (%)	Max DD (%)	Pain Ratio	Turnover (%)	Max DD (%)	Pain Ratio
Panel A: Goyal						
OLS	59.49	10.16	0.073	122.09	7.02	0.146
3PRF	70.34	15.30	0.097	69.41	8.34	0.145
PLS	121.80	20.64	0.057	80.18	17.08	0.110
PCR	148.52	39.55	0.021	60.73	11.69	0.262
LASSO	96.34	8.70	0.199	176.69	14.29	0.109
RIDGE	96.63	13.93	0.080	126.70	18.74	0.096
ENET	95.01	10.47	0.125	176.96	9.84	0.193
Panel B: FRED-MD						
OLS	48.03	10.87	0.000	83.78	5.77	0.067
3PRF	139.96	20.32	0.013	113.50	14.10	0.134
PLS	143.27	32.41	0.051	91.14	22.45	0.120
PCR	150.09	36.17	0.038	60.61	12.37	0.238
LASSO	170.81	17.67	0.075	129.19	7.09	0.281
RIDGE	170.70	24.34	0.038	103.08	13.92	0.198
ENET	179.47	19.11	0.054	129.43	8.19	0.276
Panel C: All Instruments						
OLS	45.56	11.73	-0.004	68.81	5.08	0.084
3PRF	123.67	24.75	0.013	100.53	13.27	0.138
PLS	131.46	48.28	-0.002	73.32	15.26	0.198
PCR	194.51	44.80	0.033	78.35	11.85	0.187
LASSO	160.82	27.37	0.024	129.23	7.94	0.310
RIDGE	173.12	29.60	0.027	120.69	13.23	0.210
ENET	172.31	30.93	0.009	138.13	8.02	0.257

Table 2.8.13 reports several standard financial metrics for the 5 Industry portfolios and 6 portfolios formed on Size/BTM to evaluate portfolios by estimator for the CMV framework. The Turnover is computed following equation (2.24). MaxDD is maximum drawdown as presented in equation (2.25). Pain Ratio is the standard metric, as shown in equation (2.23). Panel A reports the Sharpe ratios generated when the variables from Goyal's website are used as **Z**. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl*. Panel B presents the Sharpe ratios obtained using the FRED-MD variables. The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, panel C shows the Sharpe rations when all variables are used as conditioning information. "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI).

2.9 Appendix - Data

Table 2.9.1: FRED-MD

				Full Sample		In-Sample		Out-of-Sample	
Fred Code	**tcode**	**Group**	**Description**	**Mean**	**sd**	**Mean**	**sd**	**Mean**	**sd**
RPI	5	Output and income	Real Personal Income	0.0020	0.0070	0.0020	0.0090	0.0020	0.0070
W875RX1	5	Output and income	Real personal income ex transfer receipts	0.0020	0.0070	0.0020	0.0110	0.0020	0.0070
INDPRO	5	Output and income	IP Index	0.0020	0.0060	0.0040	0.0040	0.0010	0.0070
IPFPNSS	5	Output and income	IP: Final Products and Nonindustrial Supplies	0.0010	0.0060	0.0030	0.0040	0.0010	0.0060
IPFINAL	5	Output and income	IP: Final Products (Market Group)	0.0010	0.0070	0.0030	0.0050	0.0010	0.0070
IPCONGD	5	Output and income	IP: Consumer Goods	0.0010	0.0070	0.0030	0.0060	0.0000	0.0070
IPDCONGD	5	Output and income	IP: Durable Consumer Goods	0.0020	0.0190	0.0060	0.0120	0.0010	0.0200
IPNCONGD	5	Output and income	IP: Nondurable Consumer Goods	0.0000	0.0070	0.0020	0.0050	0.0000	0.0070
IPBUSEQ	5	Output and income	IP: Business Equipment	0.0030	0.0120	0.0060	0.0070	0.0020	0.0130
IPMAT	5	Output and income	IP: Materials	0.0020	0.0080	0.0040	0.0040	0.0020	0.0090
IPDMAT	5	Output and income	IP: Durable Materials	0.0040	0.0110	0.0070	0.0070	0.0030	0.0110
IPNMAT	5	Output and income	IP: Nondurable Materials	0.0000	0.0110	0.0010	0.0070	0.0000	0.0120
IPMANSICS	5	Output and income	IP: Manufacturing (SIC)	0.0020	0.0070	0.0040	0.0050	0.0010	0.0070
IPB51222s	5	Output and income	IP: Residential Utilities	0.0010	0.0400	0.0030	0.0330	0.0010	0.0410
IPFUELS	5	Output and income	IP: Fuels	0.0010	0.0190	0.0010	0.0160	0.0010	0.0190
CUMFNS	2	Output and income	Capacity Utilization: Manufacturing	-0.0090	0.5000	0.0780	0.3790	-0.0240	0.5170
HWI	2	Labor market	Help-Wanted Index for United States	10.9680	220.7630	33.7610	123.4940	6.9960	233.5480
HWIURATIO	2	Labor market	Ratio of Help Wanted/No. Unemployed	0.0020	0.0330	0.0060	0.0180	0.0010	0.0350
CLF16OV	5	Labor market	Civilian Labor Force	0.0010	0.0020	0.0010	0.0020	0.0010	0.0020
CE16OV	5	Labor market	Civilian Employment	0.0010	0.0020	0.0010	0.0020	0.0010	0.0020
UNRATE	2	Labor market	Civilian Unemployment Rate	-0.0110	0.1560	-0.0390	0.1320	-0.0060	0.1590
UEMPMEAN	2	Labor market	Average Duration of Unemployment (Weeks)	0.0230	0.6950	-0.0090	0.4930	0.0280	0.7250
UEMPLT5	5	Labor market	Civilians Unemployed - Less Than 5 Weeks	-0.0010	0.0600	-0.0050	0.0520	-0.0010	0.0620
UEMP5TO14	5	Labor market	Civilians Unemployed for 43599 Weeks	-0.0010	0.0520	-0.0040	0.0500	-0.0010	0.0520
UEMP15OV	5	Labor market	Civilians Unemployed - 15 Weeks & Over	-0.0010	0.0450	-0.0070	0.0460	0.0000	0.0440
UEMP15T26	5	Labor market	Civilians Unemployed for 15-26 Weeks	-0.0020	0.0710	-0.0060	0.0820	-0.0010	0.0700
UEMP27OV	5	Labor market	Civilians Unemployed for 27 Weeks and Over	0.0000	0.0570	-0.0070	0.0500	0.0010	0.0590
CLAIMSx	5	Labor market	Initial Claims	-0.0020	0.0410	-0.0040	0.0380	-0.0020	0.0420
PAYEMS	5	Labor market	All Employees: Total Nonfarm	0.0010	0.0020	0.0020	0.0010	0.0010	0.0020
USGOOD	5	Labor market	All Employees: Goods-Producing Industries	0.0000	0.0040	0.0010	0.0020	-0.0010	0.0040
CES1021000001	5	Labor market	All Employees: Mining and Logging: Mining	0.0000	0.0090	-0.0030	0.0060	0.0010	0.0090
USCONS	5	Labor market	All Employees: Construction	0.0010	0.0060	0.0030	0.0050	0.0010	0.0060
MANEMP	5	Labor market	All Employees: Manufacturing	-0.0010	0.0030	0.0010	0.0010	-0.0010	0.0040
DMANEMP	5	Labor market	All Employees: Durable goods	-0.0010	0.0040	0.0010	0.0020	-0.0010	0.0050
NDMANEMP	5	Labor market	All Employees: Nondurable goods	-0.0010	0.0020	0.0000	0.0010	-0.0010	0.0030
SRVPRD	5	Labor market	All Employees: Service-Providing Industries	0.0010	0.0010	0.0020	0.0010	0.0010	0.0010
USTPU	5	Labor market	All Employees: Trade, Transportation & Utilities	0.0010	0.0020	0.0020	0.0020	0.0010	0.0020
USWTRADE	5	Labor market	All Employees: Wholesale Trade	0.0000	0.0020	0.0010	0.0020	0.0000	0.0020
USTRADE	5	Labor market	All Employees: Retail Trade	0.0010	0.0020	0.0020	0.0020	0.0000	0.0020
USFIRE	5	Labor market	All Employees: Financial Activities	0.0010	0.0020	0.0010	0.0020	0.0010	0.0020
USGOVT	5	Labor market	All Employees: Government	0.0010	0.0020	0.0010	0.0010	0.0010	0.0030
CES0600000007	1	Labor market	Avg Weekly Hours : Goods-Producing	40.6080	0.5840	40.7170	0.3270	40.5890	0.6170
AWOTMAN	2	Labor market	Avg Weekly Overtime Hours : Manufacturing	0.0020	0.1130	0.0150	0.1560	-0.0010	0.1040
AWHMAN	1	Labor market	Avg Weekly Hours : Manufacturing	41.1990	0.5940	41.2390	0.3630	41.1920	0.6250
CES0600000008	6	Labor market	Avg Hourly Earnings : Goods-Producing	0.0000	0.0030	0.0000	0.0020	0.0000	0.0030
CES2000000008	6	Labor market	Avg Hourly Earnings : Construction	0.0000	0.0060	0.0000	0.0060	0.0000	0.0050
CES3000000008	6	Labor market	Avg Hourly Earnings : Manufacturing	0.0000	0.0030	0.0000	0.0020	0.0000	0.0030
HOUST	4	Housing	Housing Starts: Total New Privately Owned	7.1220	0.3740	7.1890	0.0920	7.1110	0.4020
HOUSTNE	4	Housing	Housing Starts, Northeast	4.8070	0.3400	4.8340	0.1200	4.8020	0.3640
HOUSTMW	4	Housing	Housing Starts, Midwest	5.4140	0.4800	5.6980	0.1040	5.3640	0.5020
HOUSTS	4	Housing	Housing Starts, South	6.3780	0.3470	6.3630	0.1180	6.3800	0.3730
HOUSTW	4	Housing	Housing Starts, West	5.6850	0.4370	5.7590	0.1180	5.6720	0.4700
PERMIT	4	Housing	New Private Housing Permits (SAAR)	7.1370	0.3670	7.1320	0.1050	7.1380	0.3950
PERMITNE	4	Housing	New Private Housing Permits, Northeast (SAAR)	4.8700	0.3340	4.8650	0.1040	4.8710	0.3600
PERMITMW	4	Housing	New Private Housing Permits, Midwest (SAAR)	5.4310	0.4410	5.6480	0.0880	5.3930	0.4660
PERMITS	4	Housing	New Private Housing Permits, South (SAAR)	6.3790	0.3400	6.2720	0.1320	6.3980	0.3610
PERMITW	4	Housing	New Private Housing Permits, West (SAAR)	5.7120	0.4350	5.7260	0.1240	5.7090	0.4680
DPCERA3M086SBEA	5	Consumption, orders, and inventories	Real Personal Consumption Expenditures	0.0020	0.0030	0.0030	0.0030	0.0020	0.0030
CMRMTSPLx	5	Consumption, orders, and inventories	Real Manu. and Trade Industries Sales	0.0020	0.0080	0.0040	0.0080	0.0020	0.0080
RETAILx	5	Consumption, orders, and inventories	Retail and Food Services Sales	0.0040	0.0090	0.0050	0.0080	0.0030	0.0100
ACOGNO	5	Consumption, orders, and inventories	New Orders for Consumer Goods	0.0030	0.0190	0.0050	0.0150	0.0020	0.0190
AMDMNOx	5	Consumption, orders, and inventories	New Orders for Durable Goods	0.0020	0.0410	0.0080	0.0250	0.0020	0.0440
ANDENOx	5	Consumption, orders, and inventories	New Orders for Nondefense Capital Goods	0.0020	0.0810	0.0100	0.0480	0.0010	0.0850
AMDMUOx	5	Consumption, orders, and inventories	Unfilled Orders for Durable Goods	0.0030	0.0100	-0.0010	0.0050	0.0040	0.0100
BUSINVx	5	Consumption, orders, and inventories	Total Business Inventories	0.0030	0.0050	0.0040	0.0030	0.0020	0.0050
ISRATIOx	2	Consumption, orders, and inventories	Total Business: Inventories to Sales Ratio	-0.0010	0.0140	-0.0020	0.0130	0.0000	0.0150
UMCSENTx	2	Consumption, orders, and inventories	Consumer Sentiment Index	0.0870	3.8890	0.4830	3.5940	0.0190	3.9410

(Continued)

Table 2.9.1: FRED-MD *(Continued)*

				Full Sample		In-Sample		Out-of-Sample	
Fred Code	**tcode**	**Group**	**Description**	**Mean**	**sd**	**Mean**	**sd**	**Mean**	**sd**
M1SL	6	Money and credit	M1 Money Stock	0.000	0.012	0.000	0.003	0.000	0.013
M2SL	6	Money and credit	M2 Money Stock	0.000	0.004	0.000	0.003	0.000	0.004
M2REAL	5	Money and credit	Real M2 Money Stock	0.003	0.005	-0.001	0.003	0.003	0.005
AMBSL	6	Money and credit	St. Louis Adjusted Monetary Base	0.000	0.025	0.000	0.003	0.000	0.027
TOTRESNS	6	Money and credit	Total Reserves of Depository Institutions	0.000	0.089	0.001	0.037	0.000	0.096
NONBORRES	7	Money and credit	Reserves Of Depository Institutions	0.000	1.618	0.001	0.036	0.000	1.754
BUSLOANS	6	Money and credit	Commercial and Industrial Loans	0.000	0.006	0.000	0.005	0.000	0.007
REALLN	6	Money and credit	Real Estate Loans at All Commercial Banks	0.000	0.007	0.000	0.003	0.000	0.007
NONREVSL	6	Money and credit	Total Nonrevolving Credit	0.000	0.008	0.000	0.007	0.000	0.008
CONSPI	2	Money and credit	Nonrevolving Consumer Credit to Personal Income	0.000	0.001	0.000	0.001	0.000	0.001
MZMSL	6	Money and credit	MZM Money Stock	0.000	0.005	0.000	0.004	0.000	0.005
DTCOLNVHFNM	6	Money and credit	Consumer Motor Vehicle Loans Outstanding	0.000	0.029	0.000	0.024	0.000	0.030
DTCTHFNM	6	Money and credit	Total Consumer Loans and Leases Outstanding	0.000	0.028	0.000	0.011	0.000	0.030
INVEST	6	Money and credit	Securities in Bank Credit at All Commercial Banks	0.000	0.012	0.000	0.012	0.000	0.012
FEDFUNDS	2	Interest and exchange rates	Effective Federal Funds Rate	-0.009	0.165	0.033	0.178	-0.016	0.162
CP3Mx	2	Interest and exchange rates	3-Month AA Financial Commercial Paper Rate	-0.008	0.194	0.033	0.207	-0.016	0.192
TB3MS	2	Interest and exchange rates	3-Month Treasury Bill	-0.008	0.177	0.028	0.180	-0.014	0.176
TB6MS	2	Interest and exchange rates	6-Month Treasury Bill	-0.008	0.175	0.026	0.209	-0.014	0.168
GS1	2	Interest and exchange rates	1-Year Treasury Rate	-0.008	0.195	0.022	0.260	-0.014	0.181
GS5	2	Interest and exchange rates	5-Year Treasury Rate	-0.014	0.236	-0.023	0.279	-0.013	0.228
GS10	2	Interest and exchange rates	10-Year Treasury Rate	-0.016	0.220	-0.035	0.239	-0.013	0.217
AAA	2	Interest and exchange rates	Moody's Seasoned Aaa Corporate Bond	-0.015	0.177	-0.032	0.179	-0.013	0.176
BAA	2	Interest and exchange rates	Moody's Seasoned Baa Corporate Bond	-0.016	0.201	-0.038	0.191	-0.012	0.203
COMPAPFFx	1	Interest and exchange rates	3-Month Commercial Paper Minus	0.144	0.242	0.255	0.229	0.125	0.239
TB3SMFFM	1	Interest and exchange rates	3-Month Treasury C Minus	-0.196	0.259	-0.099	0.210	-0.213	0.264
TB6SMFFM	1	Interest and exchange rates	6-Month Treasury C Minus	-0.090	0.307	0.081	0.341	-0.120	0.292
T1YFFM	1	Interest and exchange rates	1-Year Treasury C Minus	0.140	0.409	0.514	0.512	0.075	0.351
T5YFFM	1	Interest and exchange rates	5-Year Treasury C Minus	1.110	0.965	1.939	0.961	0.965	0.891
T10YFFM	1	Interest and exchange rates	10-Year Treasury C Minus	1.701	1.251	2.462	1.143	1.568	1.223
AAAFFM	1	Interest and exchange rates	Moody's Aaa Corporate Bond Minus	3.167	1.451	3.555	1.195	3.100	1.483
BAAFFM	1	Interest and exchange rates	Moody's Baa Corporate Bond Minus	4.115	1.658	4.250	1.256	4.092	1.720
TWEXMMTH	5	Interest and exchange rates	Trade Weighted U.S. Dollar Index: Major Currencies	0.000	0.016	-0.001	0.016	0.000	0.017
EXSZUSx	5	Interest and exchange rates	Switzerland / U.S. Foreign Exchange Rate	-0.001	0.025	-0.005	0.031	-0.001	0.024
EXJPUSx	5	Interest and exchange rates	Japan / U.S. Foreign Exchange Rate	0.000	0.026	-0.005	0.030	0.000	0.025
EXUSUKx	5	Interest and exchange rates	U.S. / U.K. Foreign Exchange Rate	-0.001	0.022	-0.003	0.029	-0.001	0.021
EXCAUSx	5	Interest and exchange rates	Canada / U.S. Foreign Exchange Rate	0.000	0.018	0.003	0.011	0.000	0.018
WPSFD49207	6	Prices	PPI: Finished Goods	0.000	0.008	0.000	0.003	0.000	0.008
WPSFD49502	6	Prices	PPI: Finished Consumer Goods	0.000	0.010	0.000	0.004	0.000	0.011
WPSID61	6	Prices	PPI: Intermediate Materials	0.000	0.009	0.000	0.003	0.000	0.009
WPSID62	6	Prices	PPI: Crude Materials	0.000	0.051	0.000	0.020	0.000	0.055
OILPRICEx	6	Prices	Crude Oil, Spliced WTI and Cushing	0.000	0.099	0.001	0.062	0.000	0.105
PPICMM	6	Prices	PPI: Metals and Metal Products:	0.000	0.039	-0.001	0.022	0.000	0.042
CPIAUCSL	6	Prices	CPI : All Items	0.000	0.003	0.000	0.001	0.000	0.003
CPIAPPSL	6	Prices	CPI : Apparel	0.000	0.006	0.000	0.006	0.000	0.006
CPITRNSL	6	Prices	CPI : Transportation	0.000	0.015	0.000	0.005	0.000	0.016
CPIMEDSL	6	Prices	CPI : Medical Care	0.000	0.002	0.000	0.001	0.000	0.002
CUSR0000SAC	6	Prices	CPI : Commodities	0.000	0.007	0.000	0.003	0.000	0.007
CUSR0000SAD	6	Prices	CPI : Durables	0.000	0.002	0.000	0.002	0.000	0.002
CUSR0000SAS	6	Prices	CPI : Services	0.000	0.001	0.000	0.001	0.000	0.001
CPIULFSL	6	Prices	CPI : All Items Less Food	0.000	0.003	0.000	0.001	0.000	0.003
CUSR0000SA0L2	6	Prices	CPI : All Items Less Shelter	0.000	0.004	0.000	0.002	0.000	0.004
CUSR0000SA0L5	6	Prices	CPI : All Items Less Medical Care	0.000	0.003	0.000	0.001	0.000	0.003
PCEPI	6	Prices	Personal Cons. Expend.: Chain Index	0.000	0.002	0.000	0.001	0.000	0.002
DDURRG3M086SBEA	6	Prices	Personal Cons. Exp: Durable goods	0.000	0.003	0.000	0.003	0.000	0.003
DNDGRG3M086SBEA	6	Prices	Personal Cons. Exp: Nondurable goods	0.000	0.008	0.000	0.003	0.000	0.008
DSERRG3M086SBEA	6	Prices	Personal Cons. Exp: Services	0.000	0.002	0.000	0.001	0.000	0.002
S&P 500	5	Stock market	S&P's Common Stock Price Index: Composite	0.006	0.035	0.009	0.017	0.006	0.038
S&P div yield	2	Stock market	S&P's Composite Common Stock: Dividend Yield	-0.004	0.077	-0.016	0.046	-0.002	0.081
S&P PE ratio	5	Stock market	S&P's Composite Common Stock: Price-Earnings Ratio	0.001	0.053	-0.004	0.018	0.001	0.057
VXOCLSx	1	Stock market	VXO	19.650	8.723	13.341	1.497	20.750	8.992

As in McCracken and Ng (2016) we transform the variables following the code presented in column 'tcode'. The transformation for a series x are: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$, (5) $\Delta log(x_t)$, (6) $\Delta^2 log(x_t)$, and (7) $\Delta(x_t/x_{t-1} - 1)$. The column 'gsi' and 'gsi:description' present the comparable series in Global Insight. The Full Sample ranges from to Jul-1963 to Dec-2017, the In-Sample from Jul-1963 to Dec-1995, and the Out-of-Sample from Jan-1996 to Dec-2017.

Table 2.9.2: Descriptive Statistics for the Portfolios

5 Industry Portfolios

	Full Sample				In-Sample				Out-of-sample			
Name	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$
Cnsmr	0.009	0.045	0.130	-0.033	0.009	0.049	0.144	-0.003	0.008	0.038	0.095	-0.106
HiTec	0.008	0.054	0.059	-0.017	0.008	0.045	0.079	-0.018	0.007	0.065	0.045	-0.015
Hlth	0.010	0.048	0.003	0.016	0.010	0.051	0.027	0.019	0.008	0.042	-0.057	0.006
Manuf	0.008	0.043	0.027	-0.029	0.009	0.042	-0.014	-0.073	0.008	0.044	0.082	0.030
Other	0.008	0.053	0.133	-0.041	0.009	0.052	0.146	-0.026	0.007	0.053	0.115	-0.064

6 Portfolios Formed on Size and Book-to-Market (2x3)

	Full Sample				In-Sample				Out-of-sample			
Name	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$
BIG.HiBM	0.010	0.048	0.097	-0.035	0.012	0.044	0.024	-0.030	0.007	0.054	0.165	-0.044
BIG.LoBM	0.008	0.045	0.070	-0.028	0.008	0.047	0.078	-0.024	0.008	0.044	0.055	-0.035
ME1.BM2	0.011	0.054	0.152	-0.035	0.012	0.054	0.193	-0.036	0.010	0.053	0.088	-0.036
ME2.BM2	0.008	0.042	0.052	-0.051	0.009	0.041	-0.002	-0.056	0.007	0.044	0.121	-0.048
SMALL.HiBM	0.012	0.056	0.171	-0.049	0.014	0.055	0.180	-0.055	0.010	0.057	0.157	-0.043
SMALL.LoBM	0.007	0.068	0.148	-0.030	0.008	0.067	0.190	-0.006	0.005	0.070	0.090	-0.063

25 Portfolios Formed on Size and Book-to-Market (5x5)

	Full Sample				In-Sample				Out-of-sample			
Name	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$
BIG.HiBM	0.009	0.053	0.068	-0.015	0.011	0.047	0.028	-0.011	0.007	0.062	0.099	-0.023
BIG.LoBM	0.008	0.046	0.059	-0.012	0.008	0.047	0.071	-0.004	0.008	0.044	0.039	-0.027
ME1.BM2	0.009	0.069	0.164	-0.033	0.010	0.066	0.216	-0.010	0.008	0.072	0.099	-0.063
ME1.BM3	0.010	0.060	0.168	-0.020	0.011	0.060	0.211	-0.017	0.009	0.059	0.102	-0.024
ME1.BM4	0.012	0.056	0.170	-0.024	0.013	0.056	0.209	-0.022	0.012	0.056	0.111	-0.027
ME2.BM1	0.006	0.072	0.139	-0.037	0.007	0.071	0.184	-0.016	0.006	0.073	0.076	-0.067
ME2.BM2	0.010	0.060	0.124	-0.046	0.010	0.061	0.171	-0.037	0.009	0.059	0.051	-0.062
ME2.BM3	0.011	0.054	0.134	-0.037	0.012	0.055	0.175	-0.044	0.009	0.053	0.068	-0.028
ME2.BM4	0.012	0.052	0.126	-0.060	0.013	0.051	0.161	-0.051	0.009	0.054	0.077	-0.076
ME2.BM5	0.012	0.060	0.147	-0.070	0.014	0.058	0.147	-0.076	0.009	0.064	0.143	-0.067
ME3.BM1	0.007	0.066	0.123	-0.043	0.008	0.065	0.152	-0.021	0.005	0.068	0.083	-0.072
ME3.BM2	0.010	0.054	0.135	-0.031	0.011	0.055	0.156	-0.023	0.010	0.054	0.101	-0.043
ME3.BM3	0.010	0.050	0.116	-0.060	0.011	0.049	0.150	-0.059	0.009	0.050	0.067	-0.061
ME3.BM4	0.011	0.049	0.145	-0.027	0.013	0.047	0.160	-0.030	0.010	0.051	0.124	-0.025
ME3.BM5	0.013	0.056	0.125	-0.101	0.014	0.055	0.133	-0.111	0.011	0.058	0.113	-0.091
ME4.BM1	0.008	0.059	0.106	-0.022	0.008	0.057	0.116	-0.022	0.008	0.061	0.093	-0.021
ME4.BM2	0.009	0.051	0.127	-0.056	0.008	0.052	0.123	-0.043	0.010	0.049	0.133	-0.078
ME4.BM3	0.010	0.050	0.123	-0.040	0.011	0.048	0.077	-0.049	0.009	0.051	0.181	-0.028
ME4.BM4	0.011	0.047	0.101	-0.034	0.012	0.046	0.077	-0.023	0.010	0.049	0.130	-0.050
ME4.BM5	0.011	0.057	0.111	-0.037	0.013	0.054	0.049	-0.046	0.008	0.060	0.181	-0.030
ME5.BM2	0.008	0.044	0.048	-0.054	0.008	0.045	0.046	-0.061	0.008	0.042	0.050	-0.043
ME5.BM3	0.009	0.042	0.021	-0.075	0.009	0.042	-0.044	-0.071	0.008	0.043	0.109	-0.082
ME5.BM4	0.008	0.046	0.082	0.021	0.010	0.041	-0.067	0.010	0.005	0.053	0.212	0.022
SMALL.HiBM	0.013	0.060	0.222	-0.016	0.014	0.060	0.228	-0.020	0.011	0.059	0.213	-0.010
SMALL.LoBM	0.003	0.079	0.192	0.002	0.005	0.076	0.231	0.024	0.001	0.084	0.144	-0.025

(Continued)

Table 2.9.2: Descriptive Statistics for the Portfolios

100 Portfolios Formed on Size and Book-to-Market (10 x 10)

	Full Sample				In-Sample				Out-of-sample			
Name	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$
BIG.HiBM	0.007	0.071	0.031	-0.043	0.008	0.066	0.023	-0.002	0.005	0.078	0.039	-0.092
BIG.LoBM	0.008	0.048	0.063	0.002	0.007	0.048	0.070	0.006	0.008	0.047	0.052	-0.006
ME1.BM2	0.005	0.077	0.182	0.017	0.007	0.073	0.199	0.020	0.002	0.082	0.160	0.011
ME1.BM3	0.009	0.076	0.184	-0.033	0.009	0.072	0.215	0.006	0.008	0.081	0.147	-0.080
ME1.BM4	0.011	0.069	0.230	-0.020	0.013	0.069	0.263	-0.003	0.007	0.070	0.179	-0.050
ME1.BM5	0.010	0.064	0.174	-0.014	0.011	0.064	0.210	-0.041	0.008	0.064	0.118	0.025
ME1.BM6	0.011	0.062	0.193	-0.002	0.011	0.064	0.222	-0.004	0.011	0.058	0.141	0.002
ME1.BM7	0.012	0.060	0.197	0.001	0.013	0.061	0.243	0.012	0.011	0.059	0.124	-0.016
ME1.BM8	0.012	0.058	0.211	-0.026	0.014	0.060	0.215	-0.032	0.010	0.055	0.201	-0.017
ME1.BM9	0.013	0.058	0.234	-0.002	0.014	0.060	0.227	-0.023	0.012	0.054	0.245	0.032
ME10.BM2	0.008	0.045	0.034	-0.042	0.009	0.046	0.046	-0.042	0.008	0.043	0.011	-0.043
ME10.BM3	0.008	0.046	0.026	-0.056	0.008	0.048	0.024	-0.063	0.009	0.043	0.030	-0.044
ME10.BM4	0.008	0.047	0.032	-0.019	0.009	0.048	0.027	-0.015	0.006	0.047	0.037	-0.027
ME10.BM5	0.007	0.047	0.033	-0.101	0.007	0.045	-0.010	-0.090	0.007	0.049	0.084	-0.116
ME10.BM6	0.009	0.046	-0.038	-0.039	0.009	0.045	-0.127	-0.028	0.009	0.047	0.081	-0.056
ME10.BM7	0.006	0.052	0.050	0.038	0.010	0.045	-0.122	0.031	0.001	0.061	0.176	0.029
ME10.BM8	0.008	0.058	0.087	-0.003	0.009	0.045	-0.050	0.011	0.006	0.072	0.161	-0.013
ME10.BM9	0.009	0.062	0.070	0.000	0.011	0.056	0.043	-0.043	0.005	0.071	0.088	0.032
ME2.BM1	0.002	0.087	0.163	-0.013	0.004	0.083	0.209	0.012	0.001	0.094	0.109	-0.043
ME2.BM10	0.013	0.070	0.161	-0.031	0.015	0.066	0.191	-0.029	0.009	0.076	0.126	-0.037
ME2.BM2	0.008	0.077	0.138	0.005	0.009	0.073	0.195	-0.005	0.006	0.082	0.070	0.016
ME2.BM3	0.009	0.072	0.100	-0.040	0.010	0.069	0.177	-0.016	0.008	0.077	0.011	-0.067
ME2.BM4	0.008	0.069	0.106	-0.015	0.009	0.064	0.167	-0.013	0.008	0.076	0.042	-0.018
ME2.BM5	0.010	0.062	0.116	-0.035	0.011	0.062	0.155	-0.003	0.010	0.064	0.064	-0.080
ME2.BM6	0.009	0.059	0.135	-0.014	0.010	0.058	0.192	-0.006	0.007	0.061	0.056	-0.025
ME2.BM7	0.012	0.058	0.108	-0.018	0.012	0.057	0.150	-0.010	0.012	0.059	0.051	-0.029
ME2.BM8	0.011	0.059	0.111	-0.041	0.011	0.057	0.161	-0.049	0.012	0.062	0.049	-0.030
ME2.BM9	0.012	0.060	0.145	-0.050	0.014	0.059	0.175	-0.043	0.010	0.060	0.099	-0.064
ME3.BM1	0.004	0.083	0.111	-0.050	0.004	0.080	0.144	-0.024	0.005	0.086	0.069	-0.085
ME3.BM10	0.011	0.070	0.157	-0.020	0.013	0.063	0.135	-0.048	0.008	0.079	0.176	0.004
ME3.BM2	0.008	0.072	0.155	-0.053	0.010	0.071	0.176	-0.009	0.006	0.075	0.126	-0.113
ME3.BM3	0.010	0.067	0.111	-0.043	0.010	0.066	0.137	-0.014	0.010	0.068	0.076	-0.083
ME3.BM4	0.008	0.062	0.106	-0.077	0.009	0.065	0.132	-0.103	0.008	0.059	0.064	-0.030
ME3.BM5	0.010	0.059	0.150	-0.025	0.011	0.060	0.185	-0.011	0.009	0.057	0.091	-0.050
ME3.BM6	0.013	0.059	0.099	-0.057	0.014	0.060	0.112	-0.052	0.010	0.058	0.076	-0.069
ME3.BM7	0.011	0.055	0.151	-0.034	0.013	0.054	0.169	-0.026	0.009	0.057	0.126	-0.047
ME3.BM8	0.011	0.056	0.103	-0.047	0.013	0.054	0.127	-0.019	0.009	0.058	0.070	-0.087
ME3.BM9	0.014	0.059	0.084	-0.051	0.016	0.057	0.091	-0.083	0.012	0.062	0.073	-0.013
ME4.BM1	0.005	0.078	0.131	-0.014	0.006	0.077	0.188	0.022	0.005	0.079	0.050	-0.067
ME4.BM10	0.007	0.076	0.177	-0.101	0.011	0.068	0.170	-0.063	0.001	0.087	0.177	-0.144
ME4.BM2	0.007	0.067	0.131	-0.019	0.008	0.067	0.166	-0.035	0.007	0.066	0.077	0.005
ME4.BM3	0.009	0.063	0.114	-0.020	0.009	0.065	0.141	0.001	0.008	0.060	0.068	-0.058
ME4.BM4	0.010	0.062	0.101	-0.051	0.011	0.061	0.172	-0.039	0.009	0.062	-0.002	-0.071
ME4.BM5	0.010	0.057	0.117	-0.043	0.011	0.057	0.161	-0.044	0.009	0.057	0.052	-0.041
ME4.BM6	0.011	0.055	0.123	-0.011	0.013	0.055	0.161	-0.020	0.009	0.055	0.065	0.000
ME4.BM7	0.012	0.054	0.096	-0.047	0.013	0.055	0.142	-0.056	0.010	0.052	0.020	-0.032
ME4.BM8	0.011	0.056	0.140	-0.064	0.014	0.053	0.177	-0.056	0.008	0.060	0.092	-0.080
ME4.BM9	0.013	0.060	0.113	-0.079	0.015	0.058	0.134	-0.082	0.010	0.063	0.082	-0.080
ME5.BM1	0.005	0.075	0.115	-0.026	0.007	0.073	0.147	0.020	0.002	0.079	0.071	-0.089
ME5.BM10	0.011	0.071	0.109	-0.084	0.013	0.065	0.121	-0.072	0.008	0.079	0.094	-0.099

(Continued)

Table 2.9.2: Descriptive Statistics for the Portfolios

100 Portfolios Formed on Size and Book-to-Market (10 x 10)

	Full Sample				In-Sample				Out-of-sample			
Name	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$
ME5.BM2	0.008	0.066	0.089	-0.048	0.009	0.064	0.121	-0.036	0.007	0.068	0.048	-0.065
ME5.BM3	0.010	0.061	0.108	-0.004	0.011	0.060	0.153	0.018	0.007	0.062	0.044	-0.035
ME5.BM4	0.013	0.059	0.121	-0.047	0.014	0.060	0.162	-0.048	0.013	0.059	0.059	-0.043
ME5.BM5	0.010	0.056	0.124	-0.045	0.010	0.057	0.133	-0.060	0.010	0.054	0.109	-0.021
ME5.BM6	0.012	0.055	0.111	-0.044	0.014	0.053	0.118	-0.035	0.009	0.057	0.098	-0.058
ME5.BM7	0.011	0.052	0.132	-0.013	0.013	0.051	0.186	-0.010	0.009	0.054	0.058	-0.023
ME5.BM8	0.012	0.052	0.122	-0.007	0.013	0.050	0.136	-0.009	0.011	0.055	0.103	-0.004
ME5.BM9	0.013	0.055	0.132	-0.080	0.014	0.053	0.101	-0.119	0.011	0.058	0.171	-0.032
ME6.BM1	0.005	0.074	0.130	-0.044	0.006	0.071	0.153	-0.024	0.004	0.077	0.102	-0.068
ME6.BM10	0.013	0.068	0.110	-0.084	0.014	0.067	0.122	-0.098	0.011	0.069	0.093	-0.065
ME6.BM2	0.008	0.062	0.107	-0.055	0.009	0.063	0.132	-0.050	0.008	0.062	0.067	-0.062
ME6.BM3	0.010	0.058	0.129	-0.020	0.010	0.059	0.138	-0.012	0.009	0.055	0.116	-0.035
ME6.BM4	0.009	0.054	0.115	-0.041	0.010	0.055	0.118	-0.025	0.009	0.053	0.109	-0.069
ME6.BM5	0.009	0.052	0.079	-0.045	0.010	0.053	0.114	-0.036	0.008	0.051	0.022	-0.062
ME6.BM6	0.010	0.051	0.125	-0.064	0.010	0.049	0.178	-0.045	0.010	0.055	0.064	-0.085
ME6.BM7	0.012	0.052	0.127	-0.035	0.013	0.050	0.123	-0.047	0.010	0.054	0.131	-0.019
ME6.BM8	0.010	0.052	0.118	-0.038	0.012	0.049	0.119	-0.048	0.008	0.056	0.114	-0.032
ME6.BM9	0.012	0.058	0.062	-0.075	0.014	0.057	0.090	-0.070	0.010	0.060	0.021	-0.083
ME7.BM1	0.009	0.066	0.091	0.013	0.009	0.065	0.123	-0.025	0.009	0.068	0.050	0.064
ME7.BM10	0.010	0.073	0.040	-0.062	0.013	0.067	0.047	-0.108	0.006	0.082	0.028	-0.018
ME7.BM2	0.008	0.058	0.126	-0.042	0.008	0.059	0.122	-0.023	0.007	0.056	0.133	-0.074
ME7.BM3	0.009	0.054	0.126	-0.042	0.009	0.056	0.151	-0.031	0.009	0.052	0.086	-0.062
ME7.BM4	0.009	0.055	0.157	-0.031	0.009	0.056	0.151	-0.036	0.009	0.055	0.168	-0.023
ME7.BM5	0.008	0.054	0.120	-0.038	0.008	0.054	0.092	-0.054	0.008	0.055	0.159	-0.017
ME7.BM6	0.010	0.052	0.111	-0.025	0.011	0.053	0.076	-0.030	0.009	0.051	0.169	-0.017
ME7.BM7	0.012	0.052	0.073	-0.064	0.011	0.051	0.077	-0.071	0.012	0.055	0.068	-0.054
ME7.BM8	0.010	0.053	0.089	-0.027	0.011	0.051	0.038	-0.055	0.009	0.055	0.155	0.009
ME7.BM9	0.011	0.060	0.107	-0.020	0.013	0.058	0.044	-0.019	0.007	0.063	0.180	-0.028
ME8.BM1	0.008	0.065	0.082	-0.028	0.008	0.062	0.097	-0.016	0.008	0.070	0.067	-0.041
ME8.BM10	0.010	0.073	0.035	0.031	0.010	0.066	0.050	0.009	0.009	0.083	0.024	0.049
ME8.BM2	0.008	0.056	0.105	-0.017	0.008	0.055	0.092	-0.004	0.008	0.056	0.126	-0.035
ME8.BM3	0.009	0.054	0.076	-0.041	0.008	0.056	0.071	-0.028	0.011	0.051	0.084	-0.067
ME8.BM4	0.008	0.054	0.105	-0.065	0.007	0.054	0.078	-0.057	0.008	0.055	0.144	-0.078
ME8.BM5	0.011	0.051	0.065	-0.069	0.012	0.050	0.037	-0.085	0.010	0.054	0.099	-0.051
ME8.BM6	0.008	0.056	0.110	-0.041	0.010	0.052	0.044	-0.035	0.006	0.061	0.179	-0.050
ME8.BM7	0.011	0.051	0.089	-0.024	0.012	0.048	0.066	-0.004	0.010	0.056	0.114	-0.046
ME8.BM8	0.011	0.050	0.102	-0.022	0.013	0.050	0.047	0.005	0.007	0.052	0.172	-0.066
ME8.BM9	0.010	0.056	0.095	-0.018	0.012	0.053	-0.020	-0.029	0.008	0.060	0.221	-0.009
ME9.BM1	0.007	0.058	0.128	-0.033	0.007	0.058	0.129	-0.022	0.007	0.059	0.127	-0.050
ME9.BM10	0.009	0.076	0.139	-0.042	0.013	0.058	0.094	0.006	0.004	0.096	0.158	-0.075
ME9.BM2	0.008	0.050	0.092	-0.014	0.008	0.050	0.074	-0.042	0.008	0.051	0.119	0.025
ME9.BM3	0.009	0.049	0.105	-0.032	0.008	0.051	0.117	-0.026	0.009	0.047	0.085	-0.043
ME9.BM4	0.009	0.051	0.052	-0.065	0.010	0.051	0.042	-0.083	0.008	0.051	0.065	-0.041
ME9.BM5	0.010	0.048	0.071	-0.043	0.009	0.048	0.059	-0.058	0.011	0.048	0.088	-0.022
ME9.BM6	0.009	0.046	0.083	-0.035	0.010	0.047	0.057	-0.083	0.009	0.044	0.125	0.041
ME9.BM7	0.010	0.051	0.109	0.013	0.009	0.049	0.063	-0.022	0.011	0.052	0.169	0.057
ME9.BM8	0.008	0.045	0.048	-0.009	0.009	0.045	-0.007	-0.024	0.007	0.046	0.121	0.008
ME9.BM9	0.010	0.055	0.063	-0.108	0.011	0.049	-0.028	-0.098	0.009	0.063	0.143	-0.118
SMALL.HiBM	0.013	0.063	0.262	0.015	0.015	0.063	0.232	0.003	0.011	0.062	0.306	0.031
SMALL.LoBM	0.000	0.084	0.222	0.021	0.002	0.081	0.243	0.058	-0.002	0.089	0.197	-0.025

Table 2.9.2 reports some summary statistics for the 5 Industry portfolios, 6 portfolios formed on size and book-to-market, 25 portfolios formed on size and book-to-market, and the 100 portfolios formed on size and book-to-market. The table reports the monthly returns ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), first and second autocorrelation ($\hat{\rho}_1$ and $\hat{\rho}_2$) of the monthly returns. The Full Sample ranges from to Jul-1963 to Dec-2017, the In-Sample from Jul-1963 to Dec-1995, and the Out-of-Sample from Jan-1996 to Dec-2017. Data is from Ken French's website (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`)

Table 2.9.3: Results for the 5 Industry Portfolios

A: CMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p-val	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.002	0.015	-0.108	0.018	0.133	0.062	0.462	0.214	2.157	0.032	0.038	12.087	0.000
3PRF	0.003	0.021	-0.106	0.089	0.125	0.062	0.434	0.214	2.028	0.044	0.035	12.386	0.000
PLS	0.004	0.032	-0.108	0.116	0.121	0.062	0.419	0.214	1.959	0.051	0.040	10.391	0.000
PCR	0.003	0.035	-0.037	-0.003	0.089	0.062	0.309	0.214	1.449	0.149	0.049	6.355	0.000
LASSO	0.004	0.017	-0.080	0.180	0.209	0.062	0.725	0.216	3.366	0.001	0.032	22.390	0.000
RIDGE	0.003	0.020	-0.185	0.239	0.131	0.062	0.454	0.214	2.123	0.035	0.040	11.321	0.000
ENET	0.003	0.019	0.004	0.216	0.163	0.062	0.565	0.215	2.633	0.009	0.042	13.361	0.000
FRED-MD													
OLS	0.000	0.006	0.003	0.102	0.005	0.062	0.018	0.213	0.083	0.934	0.046	0.384	0.702
3PRF	0.001	0.020	0.070	0.074	0.072	0.062	0.248	0.213	1.164	0.246	0.051	4.915	0.000
PLS	0.005	0.050	-0.027	-0.037	0.102	0.062	0.354	0.214	1.657	0.099	0.032	11.223	0.000
PCR	0.004	0.033	-0.106	0.218	0.115	0.062	0.398	0.214	1.861	0.064	0.043	9.335	0.000
LASSO	0.004	0.023	-0.084	0.077	0.156	0.062	0.540	0.214	2.517	0.012	0.041	13.134	0.000
RIDGE	0.003	0.026	0.019	0.013	0.116	0.062	0.402	0.214	1.881	0.061	0.049	8.250	0.000
ENET	0.003	0.023	-0.048	0.051	0.127	0.062	0.438	0.214	2.048	0.042	0.042	10.477	0.000
All Instruments													
OLS	0.000	0.007	-0.032	0.077	-0.035	0.066	-0.123	0.230	-0.534	0.594	0.046	-2.682	0.008
3PRF	0.001	0.020	0.081	0.081	0.067	0.062	0.231	0.213	1.081	0.281	0.053	4.380	0.000
PLS	0.000	0.033	0.001	-0.034	0.002	0.066	0.007	0.230	0.030	0.976	0.054	0.128	0.898
PCR	0.005	0.039	-0.038	0.157	0.133	0.067	0.461	0.231	1.998	0.047	0.058	7.949	0.000
LASSO	0.003	0.022	0.000	0.166	0.122	0.067	0.424	0.231	1.837	0.067	0.052	8.127	0.000
RIDGE	0.003	0.026	-0.002	0.035	0.113	0.067	0.392	0.231	1.698	0.091	0.056	6.944	0.000
ENET	0.002	0.023	-0.053	0.146	0.069	0.066	0.239	0.230	1.037	0.301	0.051	4.638	0.000

B: UMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.002	0.019	-0.149	0.086	0.100	0.062	0.345	0.214	1.616	0.107	0.038	9.176	0.000
3PRF	0.002	0.026	-0.051	0.096	0.083	0.062	0.286	0.214	1.341	0.181	0.043	6.590	0.000
PLS	0.005	0.047	-0.308	0.220	0.109	0.062	0.378	0.214	1.767	0.078	0.030	12.410	0.000
PCR	0.003	0.034	0.091	-0.072	0.076	0.062	0.264	0.214	1.239	0.217	0.055	4.795	0.000
LASSO	0.002	0.016	-0.132	0.061	0.141	0.062	0.488	0.214	2.278	0.024	0.036	13.381	0.000
RIDGE	0.002	0.018	-0.148	0.082	0.112	0.062	0.387	0.214	1.809	0.072	0.040	9.716	0.000
ENET	0.002	0.017	-0.150	0.192	0.130	0.062	0.449	0.214	2.098	0.037	0.041	11.050	0.000
FRED-MD													
OLS	0.000	0.005	-0.045	0.100	0.009	0.062	0.032	0.213	0.150	0.881	0.044	0.735	0.463
3PRF	0.000	0.017	0.007	0.088	0.009	0.062	0.031	0.213	0.147	0.883	0.046	0.676	0.499
PLS	0.002	0.034	0.079	-0.103	0.046	0.062	0.160	0.213	0.749	0.455	0.055	2.915	0.004
PCR	0.001	0.044	-0.109	0.190	0.019	0.062	0.064	0.213	0.301	0.764	0.043	1.480	0.140
LASSO	0.001	0.019	-0.125	0.012	0.056	0.062	0.194	0.213	0.911	0.363	0.040	4.893	0.000
RIDGE	0.001	0.025	-0.038	-0.011	0.028	0.062	0.095	0.213	0.448	0.655	0.046	2.056	0.041
ENET	0.001	0.021	-0.104	0.025	0.032	0.062	0.112	0.213	0.524	0.601	0.040	2.772	0.006
All Instruments													
OLS	0.000	0.005	-0.101	0.101	-0.002	0.066	-0.006	0.230	-0.027	0.979	0.049	-0.126	0.900
3PRF	0.001	0.018	0.020	0.162	0.029	0.062	0.100	0.213	0.468	0.640	0.046	2.151	0.032
PLS	0.001	0.023	-0.028	-0.049	0.037	0.066	0.127	0.230	0.552	0.581	0.055	2.303	0.022
PCR	0.002	0.047	-0.075	0.171	0.040	0.066	0.137	0.230	0.596	0.552	0.053	2.586	0.010
LASSO	0.001	0.019	-0.110	0.167	0.052	0.066	0.180	0.230	0.781	0.435	0.048	3.713	0.000
RIDGE	0.001	0.028	-0.097	0.003	0.035	0.066	0.120	0.230	0.522	0.602	0.049	2.463	0.015
ENET	0.001	0.018	-0.105	0.051	0.047	0.066	0.162	0.230	0.702	0.483	0.047	3.417	0.001

(Continued)

Table 2.9.3: Results for the 5 Industry Portfolios *(Continued)*

C: MVATE

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.004	0.030	-0.121	0.015	0.135	0.062	0.467	0.214	2.181	0.030	0.037	12.513	0.000
3PRF	0.005	0.042	-0.107	0.091	0.125	0.062	0.435	0.214	2.030	0.043	0.035	12.492	0.000
PLS	0.007	0.062	-0.107	0.114	0.119	0.062	0.414	0.214	1.934	0.054	0.040	10.274	0.000
PCR	0.006	0.068	-0.038	-0.007	0.085	0.062	0.295	0.214	1.383	0.168	0.049	6.076	0.000
LASSO	0.007	0.033	-0.082	0.184	0.209	0.062	0.723	0.216	3.355	0.001	0.031	23.150	0.000
RIDGE	0.005	0.039	-0.192	0.243	0.131	0.062	0.455	0.214	2.125	0.034	0.040	11.503	0.000
ENET	0.006	0.037	-0.001	0.222	0.168	0.062	0.581	0.215	2.704	0.007	0.037	15.738	0.000
FRED-MD													
OLS	-0.001	0.014	0.019	0.115	-0.046	0.062	-0.160	0.213	-0.748	0.455	0.044	-3.608	0.000
3PRF	0.003	0.038	0.071	0.080	0.073	0.062	0.252	0.213	1.180	0.239	0.050	5.011	0.000
PLS	0.010	0.097	-0.026	-0.038	0.101	0.062	0.351	0.214	1.640	0.102	0.032	11.038	0.000
PCR	0.007	0.063	-0.104	0.221	0.112	0.062	0.389	0.214	1.821	0.070	0.043	9.116	0.000
LASSO	0.007	0.044	-0.081	0.074	0.152	0.062	0.528	0.214	2.461	0.014	0.041	12.803	0.000
RIDGE	0.006	0.051	0.022	0.013	0.116	0.062	0.401	0.214	1.873	0.062	0.049	8.174	0.000
ENET	0.005	0.044	-0.036	0.044	0.125	0.062	0.434	0.214	2.026	0.044	0.043	10.166	0.000
All Instruments													
OLS	-0.001	0.018	0.022	0.136	-0.078	0.066	-0.271	0.230	-1.179	0.240	0.042	-6.499	0.000
3PRF	0.003	0.039	0.084	0.082	0.068	0.062	0.235	0.213	1.102	0.271	0.053	4.459	0.000
PLS	0.000	0.065	0.002	-0.038	-0.001	0.066	-0.003	0.230	-0.011	0.991	0.053	-0.048	0.962
PCR	0.010	0.075	-0.035	0.154	0.129	0.067	0.447	0.231	1.936	0.054	0.058	7.669	0.000
LASSO	0.005	0.043	0.003	0.172	0.118	0.067	0.408	0.231	1.768	0.078	0.052	7.811	0.000
RIDGE	0.006	0.051	0.003	0.036	0.112	0.067	0.388	0.231	1.680	0.094	0.056	6.868	0.000
ENET	0.003	0.044	-0.045	0.147	0.069	0.066	0.239	0.230	1.038	0.300	0.052	4.586	0.000

Table 2.9.3 reports some summary statistics of the OOS (Jan-1996 to Dec-2017) for the 5 Industry portfolios by estimator and three different set of conditioning information (Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios (UMV). Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). The table reports the monthly returns ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), first and second autocorrelation ($\hat{\rho}_1$ and $\hat{\rho}_2$) of the monthly returns. The Sharpe ratios (SR) are presented monthly and annualized, as well as the standard error (se). For the sake of completeness, it is also reported the t-statistic (t), the standard p-value (p), as well as the HAC standard error ($se(HAC)$), the t-statistic obtained using HAC ($t(HAC)$) and its p-value ($p(HAC)$).

Table 2.9.4: Results for the 6 Portfolios Formed on Size and Book-to-Market

A: CMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p-val	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.002	0.012	0.057	-0.024	0.207	0.062	0.719	0.215	3.335	0.001	0.049	14.647	0.000
3PRF	0.002	0.010	0.160	0.031	0.222	0.062	0.769	0.216	3.563	0.000	0.052	14.904	0.000
PLS	0.004	0.015	0.158	-0.016	0.247	0.062	0.856	0.216	3.954	0.000	0.059	14.417	0.000
PCR	0.005	0.016	0.132	0.024	0.298	0.063	1.032	0.218	4.738	0.000	0.048	21.518	0.000
LASSO	0.002	0.013	0.030	-0.002	0.168	0.062	0.580	0.215	2.704	0.007	0.046	12.726	0.000
RIDGE	0.003	0.015	0.131	-0.024	0.173	0.062	0.600	0.215	2.793	0.006	0.059	10.197	0.000
ENET	0.003	0.012	0.041	-0.097	0.231	0.062	0.799	0.216	3.698	0.000	0.045	17.752	0.000
FRED-MD													
OLS	0.001	0.005	0.050	0.024	0.154	0.062	0.534	0.214	2.491	0.013	0.049	10.802	0.000
3PRF	0.003	0.014	-0.018	-0.066	0.186	0.062	0.646	0.215	3.003	0.003	0.048	13.522	0.000
PLS	0.003	0.016	0.142	0.046	0.210	0.062	0.726	0.216	3.369	0.001	0.052	13.883	0.000
PCR	0.005	0.017	0.127	0.098	0.272	0.063	0.944	0.217	4.346	0.000	0.044	21.651	0.000
LASSO	0.004	0.016	-0.039	-0.057	0.226	0.062	0.783	0.216	3.627	0.000	0.033	23.946	0.000
RIDGE	0.004	0.016	0.037	-0.029	0.242	0.062	0.840	0.216	3.881	0.000	0.047	17.773	0.000
ENET	0.004	0.015	0.054	-0.027	0.261	0.063	0.905	0.217	4.172	0.000	0.040	22.576	0.000
All Instruments													
OLS	0.001	0.004	0.097	-0.017	0.160	0.067	0.554	0.231	2.395	0.017	0.065	8.465	0.000
3PRF	0.003	0.013	0.001	-0.086	0.193	0.062	0.670	0.215	3.113	0.002	0.047	14.340	0.000
PLS	0.004	0.014	0.122	0.087	0.264	0.068	0.916	0.234	3.915	0.000	0.063	14.545	0.000
PCR	0.005	0.019	0.136	0.104	0.238	0.067	0.825	0.233	3.539	0.000	0.054	15.290	0.000
LASSO	0.004	0.014	-0.019	-0.044	0.271	0.063	0.940	0.219	4.296	0.000	0.038	24.768	0.000
RIDGE	0.004	0.015	0.034	-0.016	0.248	0.067	0.858	0.233	3.674	0.000	0.054	15.940	0.000
ENET	0.003	0.013	-0.011	-0.035	0.255	0.063	0.883	0.218	4.052	0.000	0.040	21.857	0.000

B: UMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.002	0.010	0.055	-0.110	0.218	0.062	0.757	0.216	3.508	0.001	0.046	16.626	0.000
3PRF	0.003	0.011	0.144	0.000	0.258	0.063	0.893	0.217	4.122	0.000	0.046	19.347	0.000
PLS	0.004	0.017	0.139	0.111	0.246	0.062	0.853	0.216	3.941	0.000	0.054	15.805	0.000
PCR	0.005	0.016	0.167	0.083	0.299	0.063	1.035	0.218	4.749	0.000	0.052	19.905	0.000
LASSO	0.002	0.009	0.041	-0.071	0.187	0.062	0.648	0.215	3.011	0.003	0.049	13.293	0.000
RIDGE	0.002	0.011	0.077	-0.040	0.224	0.062	0.774	0.216	3.587	0.000	0.048	16.120	0.000
ENET	0.002	0.009	0.029	-0.074	0.260	0.063	0.901	0.217	4.157	0.000	0.044	20.464	0.000
FRED-MD													
OLS	0.001	0.004	0.116	0.043	0.158	0.062	0.546	0.215	2.547	0.011	0.057	9.636	0.000
3PRF	0.002	0.012	0.135	-0.021	0.197	0.062	0.681	0.215	3.163	0.002	0.060	11.418	0.000
PLS	0.003	0.016	0.150	0.076	0.198	0.062	0.687	0.215	3.189	0.002	0.058	11.751	0.000
PCR	0.004	0.017	0.118	0.055	0.255	0.063	0.882	0.217	4.070	0.000	0.059	14.887	0.000
LASSO	0.002	0.010	0.098	0.017	0.201	0.062	0.697	0.215	3.239	0.001	0.046	15.132	0.000
RIDGE	0.003	0.014	0.142	-0.021	0.230	0.062	0.797	0.216	3.688	0.000	0.054	14.801	0.000
ENET	0.003	0.012	0.138	0.047	0.261	0.063	0.905	0.217	4.173	0.000	0.044	20.408	0.000
All Instruments													
OLS	0.001	0.004	0.125	-0.007	0.168	0.067	0.583	0.232	2.519	0.012	0.070	8.374	0.000
3PRF	0.003	0.012	0.167	-0.027	0.213	0.062	0.738	0.216	3.421	0.001	0.059	12.574	0.000
PLS	0.004	0.014	0.104	0.032	0.263	0.068	0.911	0.234	3.897	0.000	0.060	15.097	0.000
PCR	0.004	0.017	0.133	0.053	0.237	0.067	0.820	0.233	3.517	0.001	0.069	11.858	0.000
LASSO	0.002	0.009	0.069	-0.026	0.228	0.063	0.790	0.218	3.628	0.000	0.043	18.550	0.000
RIDGE	0.003	0.014	0.147	-0.021	0.229	0.067	0.794	0.233	3.407	0.001	0.060	13.331	0.000
ENET	0.002	0.009	0.038	-0.045	0.234	0.063	0.812	0.217	3.736	0.000	0.043	19.023	0.000

(Continued)

Table 2.9.4: Results for the 6 Portfolios Formed on Size and Book-to-Market *(Continued)*

C: MVATE

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.004	0.022	0.059	-0.025	0.203	0.062	0.703	0.215	3.266	0.001	0.050	14.175	0.000
3PRF	0.004	0.020	0.161	0.034	0.223	0.062	0.773	0.216	3.583	0.000	0.052	14.856	0.000
PLS	0.007	0.029	0.169	-0.011	0.247	0.062	0.857	0.216	3.958	0.000	0.061	14.037	0.000
PCR	0.009	0.031	0.144	0.034	0.300	0.063	1.039	0.218	4.768	0.000	0.048	21.704	0.000
LASSO	0.004	0.025	0.027	-0.010	0.166	0.062	0.575	0.215	2.678	0.008	0.046	12.624	0.000
RIDGE	0.005	0.029	0.132	-0.022	0.174	0.062	0.604	0.215	2.811	0.005	0.060	10.138	0.000
ENET	0.006	0.024	0.033	-0.098	0.236	0.062	0.816	0.216	3.774	0.000	0.044	18.609	0.000
FRED-MD													
OLS	0.001	0.014	-0.175	-0.006	0.037	0.062	0.128	0.213	0.599	0.550	0.035	3.648	0.000
3PRF	0.005	0.027	-0.019	-0.064	0.188	0.062	0.652	0.215	3.031	0.003	0.048	13.619	0.000
PLS	0.006	0.032	0.138	0.037	0.201	0.062	0.695	0.215	3.226	0.001	0.051	13.732	0.000
PCR	0.009	0.033	0.131	0.105	0.275	0.063	0.951	0.217	4.378	0.000	0.043	22.089	0.000
LASSO	0.007	0.030	-0.039	-0.053	0.226	0.062	0.783	0.216	3.625	0.000	0.032	24.708	0.000
RIDGE	0.007	0.030	0.036	-0.023	0.246	0.062	0.850	0.216	3.930	0.000	0.047	18.052	0.000
ENET	0.008	0.029	0.050	-0.018	0.261	0.063	0.905	0.217	4.173	0.000	0.039	23.380	0.000
All Instruments													
OLS	0.000	0.018	-0.130	-0.057	-0.002	0.066	-0.008	0.230	-0.034	0.973	0.042	-0.184	0.854
3PRF	0.005	0.025	-0.001	-0.087	0.191	0.062	0.661	0.215	3.072	0.002	0.048	13.744	0.000
PLS	0.007	0.028	0.132	0.095	0.265	0.068	0.917	0.234	3.921	0.000	0.064	14.399	0.000
PCR	0.009	0.037	0.143	0.106	0.240	0.067	0.831	0.233	3.565	0.000	0.054	15.490	0.000
LASSO	0.007	0.026	-0.019	-0.038	0.274	0.063	0.950	0.219	4.339	0.000	0.037	25.655	0.000
RIDGE	0.007	0.029	0.044	-0.016	0.247	0.067	0.856	0.233	3.669	0.000	0.054	15.731	0.000
ENET	0.007	0.026	-0.026	-0.033	0.266	0.063	0.921	0.219	4.206	0.000	0.039	23.441	0.000

Table 2.9.4 reports some summary statistics of the OOS (Jan-1996 to Dec-2017) for the 6 portfolios formed on size and book-to-market by estimator and three different set of conditioning information (Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios (UMV). Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). The table reports the monthly returns ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), first and second autocorrelation ($\hat{\rho}_1$ and $\hat{\rho}_2$) of the monthly returns. The Sharpe ratios (SR) are presented monthly and annualized, as well as the standard error (se). For the sake of completeness, it is also reported the t-statistic (t), the standard p-value (p), as well as the HAC standard error ($se(HAC)$), the t-statistic obtained using HAC ($t(HAC)$) and its p-value ($p(HAC)$).

Table 2.9.5: Results for the 25 Portfolios Formed on Size and Book-to-Market

A: CMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p-val	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.002	0.007	-0.108	0.124	0.307	0.063	1.062	0.218	4.868	0.000	0.038	27.935	0.000
3PRF	0.002	0.007	0.005	0.044	0.327	0.063	1.132	0.219	5.171	0.000	0.044	25.789	0.000
PLS	0.003	0.010	0.017	0.046	0.307	0.063	1.063	0.218	4.870	0.000	0.046	23.139	0.000
PCR	0.004	0.011	0.092	-0.042	0.383	0.064	1.325	0.221	6.000	0.000	0.054	24.516	0.000
LASSO	0.002	0.006	0.016	0.167	0.276	0.063	0.957	0.217	4.407	0.000	0.047	20.181	0.000
RIDGE	0.003	0.008	-0.130	0.147	0.313	0.063	1.083	0.218	4.961	0.000	0.039	27.911	0.000
ENET	0.002	0.006	-0.123	0.155	0.315	0.063	1.092	0.218	4.998	0.000	0.036	30.736	0.000
FRED-MD													
OLS	0.000	0.002	-0.061	-0.031	0.161	0.062	0.559	0.215	2.605	0.010	0.045	12.313	0.000
3PRF	0.002	0.007	0.096	-0.006	0.282	0.063	0.978	0.217	4.497	0.000	0.052	18.951	0.000
PLS	0.003	0.009	0.026	0.057	0.296	0.063	1.024	0.218	4.701	0.000	0.043	23.716	0.000
PCR	0.004	0.011	0.079	0.048	0.371	0.064	1.286	0.220	5.834	0.000	0.052	24.585	0.000
LASSO	0.002	0.008	-0.071	-0.076	0.215	0.062	0.743	0.216	3.448	0.001	0.044	17.027	0.000
RIDGE	0.003	0.011	0.053	-0.014	0.274	0.063	0.949	0.217	4.370	0.000	0.049	19.392	0.000
ENET	0.002	0.008	-0.025	-0.053	0.280	0.063	0.970	0.217	4.465	0.000	0.048	20.323	0.000
All Instruments													
OLS	0.000	0.002	-0.005	0.009	0.201	0.067	0.696	0.232	2.996	0.003	0.053	13.137	0.000
3PRF	0.002	0.007	0.076	-0.005	0.290	0.063	1.006	0.218	4.621	0.000	0.050	19.946	0.000
PLS	0.003	0.009	0.080	0.050	0.331	0.068	1.145	0.236	4.849	0.000	0.053	21.547	0.000
PCR	0.004	0.013	0.007	0.029	0.307	0.068	1.063	0.235	4.518	0.000	0.055	19.254	0.000
LASSO	0.002	0.008	-0.098	0.018	0.231	0.063	0.799	0.220	3.637	0.000	0.040	20.170	0.000
RIDGE	0.003	0.010	0.006	0.017	0.305	0.068	1.057	0.235	4.493	0.000	0.054	19.399	0.000
ENET	0.002	0.008	-0.118	-0.018	0.302	0.064	1.045	0.222	4.701	0.000	0.040	25.901	0.000

B: UMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.002	0.006	-0.078	0.107	0.293	0.063	1.017	0.218	4.668	0.000	0.037	27.165	0.000
3PRF	0.002	0.007	0.033	-0.005	0.330	0.063	1.143	0.219	5.219	0.000	0.042	27.225	0.000
PLS	0.003	0.011	0.099	0.074	0.322	0.063	1.114	0.219	5.094	0.000	0.041	27.417	0.000
PCR	0.004	0.011	0.121	0.024	0.386	0.064	1.337	0.221	6.049	0.000	0.050	26.591	0.000
LASSO	0.001	0.005	0.016	0.102	0.279	0.063	0.968	0.217	4.453	0.000	0.046	20.973	0.000
RIDGE	0.002	0.007	-0.075	0.042	0.311	0.063	1.076	0.218	4.928	0.000	0.040	26.583	0.000
ENET	0.002	0.005	-0.066	0.088	0.316	0.063	1.095	0.218	5.014	0.000	0.037	29.458	0.000
FRED-MD													
OLS	0.000	0.002	-0.058	-0.023	0.166	0.062	0.576	0.215	2.683	0.008	0.045	12.696	0.000
3PRF	0.002	0.007	0.113	-0.001	0.262	0.063	0.907	0.217	4.182	0.000	0.059	15.475	0.000
PLS	0.003	0.009	0.089	0.057	0.299	0.063	1.036	0.218	4.756	0.000	0.043	24.224	0.000
PCR	0.004	0.011	0.072	0.042	0.361	0.064	1.252	0.220	5.687	0.000	0.051	24.710	0.000
LASSO	0.002	0.007	-0.062	-0.106	0.209	0.062	0.724	0.216	3.357	0.001	0.048	15.114	0.000
RIDGE	0.003	0.010	0.135	-0.047	0.277	0.063	0.960	0.217	4.417	0.000	0.051	18.862	0.000
ENET	0.002	0.007	-0.017	-0.123	0.267	0.063	0.924	0.217	4.259	0.000	0.052	17.747	0.000
All Instruments													
OLS	0.000	0.002	0.001	0.011	0.204	0.067	0.706	0.232	3.039	0.003	0.054	13.132	0.000
3PRF	0.002	0.007	0.105	-0.007	0.278	0.063	0.964	0.217	4.436	0.000	0.054	17.895	0.000
PLS	0.003	0.009	0.115	-0.014	0.303	0.068	1.051	0.235	4.469	0.000	0.053	19.951	0.000
PCR	0.004	0.012	0.046	0.019	0.293	0.068	1.015	0.235	4.322	0.000	0.062	16.363	0.000
LASSO	0.001	0.006	-0.070	0.014	0.230	0.063	0.797	0.220	3.628	0.000	0.043	18.443	0.000
RIDGE	0.003	0.010	0.101	-0.020	0.306	0.068	1.061	0.235	4.511	0.000	0.057	18.572	0.000
ENET	0.002	0.006	-0.066	-0.057	0.280	0.064	0.970	0.222	4.376	0.000	0.047	20.478	0.000

(Continued)

Table 2.9.5: Results for the 25 Portfolios Formed on Size and Book-to-Market *(Continued)*

C: MVATE

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.003	0.013	-0.127	0.142	0.270	0.063	0.934	0.217	4.302	0.000	0.037	25.109	0.000
3PRF	0.004	0.014	-0.019	0.027	0.307	0.063	1.064	0.218	4.876	0.000	0.043	25.024	0.000
PLS	0.006	0.019	-0.011	0.071	0.300	0.063	1.039	0.218	4.766	0.000	0.045	23.133	0.000
PCR	0.008	0.020	0.098	-0.040	0.380	0.064	1.318	0.221	5.967	0.000	0.052	25.274	0.000
LASSO	0.003	0.012	0.020	0.177	0.278	0.063	0.962	0.217	4.427	0.000	0.045	21.145	0.000
RIDGE	0.005	0.015	-0.121	0.169	0.299	0.063	1.035	0.218	4.750	0.000	0.038	27.334	0.000
ENET	0.003	0.013	-0.128	0.199	0.259	0.063	0.898	0.217	4.143	0.000	0.049	18.277	0.000
FRED-MD													
OLS	0.000	0.039	-0.010	0.045	0.006	0.062	0.020	0.213	0.096	0.924	0.043	0.472	0.637
3PRF	0.004	0.015	0.142	0.018	0.261	0.063	0.904	0.217	4.171	0.000	0.050	17.999	0.000
PLS	0.005	0.019	0.041	0.012	0.285	0.063	0.989	0.218	4.546	0.000	0.045	21.889	0.000
PCR	0.008	0.022	0.082	0.058	0.374	0.064	1.295	0.221	5.872	0.000	0.051	25.189	0.000
LASSO	0.003	0.016	-0.074	-0.074	0.212	0.062	0.735	0.216	3.410	0.001	0.043	16.970	0.000
RIDGE	0.006	0.020	0.069	-0.001	0.275	0.063	0.951	0.217	4.380	0.000	0.049	19.403	0.000
ENET	0.004	0.016	-0.011	-0.089	0.286	0.063	0.990	0.218	4.553	0.000	0.045	21.977	0.000
All Instruments													
OLS	-0.001	0.025	-0.007	0.035	-0.054	0.066	-0.186	0.230	-0.810	0.419	0.046	-4.040	0.000
3PRF	0.004	0.014	0.105	0.013	0.254	0.063	0.880	0.217	4.063	0.000	0.050	17.726	0.000
PLS	0.005	0.017	0.068	0.063	0.314	0.068	1.089	0.236	4.621	0.000	0.053	20.460	0.000
PCR	0.008	0.024	0.009	0.031	0.310	0.068	1.074	0.235	4.562	0.000	0.055	19.574	0.000
LASSO	0.004	0.015	-0.100	0.019	0.232	0.063	0.802	0.220	3.649	0.000	0.039	20.450	0.000
RIDGE	0.006	0.020	0.040	0.024	0.299	0.068	1.035	0.235	4.405	0.000	0.054	18.999	0.000
ENET	0.005	0.015	-0.097	-0.106	0.302	0.064	1.045	0.223	4.681	0.000	0.040	26.200	0.000

Table 2.9.5 reports some summary statistics of the OOS (Jan-1996 to Dec-2017) for the 25 portfolios formed on size and book-to-market by estimator and three different set of conditioning information (Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios (UMV). Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). The table reports the monthly returns ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), first and second autocorrelation ($\hat{\rho}_1$ and $\hat{\rho}_2$) of the monthly returns. The Sharpe ratios (SR) are presented monthly and annualized, as well as the standard error (se). For the sake of completeness, it is also reported the t-statistic (t), the standard p-value (p), as well as the HAC standard error ($se(HAC)$), the t-statistic obtained using HAC ($t(HAC)$) and its p-value ($p(HAC)$).

Table 2.9.6: Results for the 100 Portfolios Formed on Size and Book-to-Market

A: CMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p-val	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.001	0.009	-0.175	0.106	0.076	0.062	0.262	0.214	1.227	0.221	0.072	3.635	0.000
3PRF	0.001	0.012	-0.208	0.065	0.117	0.062	0.404	0.214	1.890	0.060	0.095	4.269	0.000
PLS	0.002	0.010	-0.002	0.009	0.245	0.062	0.848	0.216	3.920	0.000	0.151	5.624	0.000
PCR	0.003	0.009	0.104	0.090	0.303	0.063	1.048	0.218	4.808	0.000	0.127	8.270	0.000
LASSO	0.001	0.009	-0.051	-0.002	0.087	0.062	0.302	0.214	1.415	0.158	0.096	3.140	0.002
RIDGE	0.001	0.013	-0.027	0.010	0.060	0.062	0.207	0.213	0.970	0.333	0.082	2.512	0.013
ENET	0.001	0.009	0.047	0.022	0.097	0.062	0.336	0.214	1.573	0.117	0.101	3.314	0.001
FRED-MD													
OLS	0.000	0.001	0.185	0.005	0.128	0.062	0.443	0.214	2.068	0.040	0.065	6.830	0.000
3PRF	0.002	0.007	-0.255	0.134	0.210	0.062	0.728	0.216	3.376	0.001	0.042	17.141	0.000
PLS	0.002	0.007	0.081	0.283	0.350	0.063	1.211	0.220	5.513	0.000	0.060	20.049	0.000
PCR	0.002	0.011	0.072	0.081	0.203	0.062	0.703	0.215	3.266	0.001	0.144	4.898	0.000
LASSO	0.001	0.003	0.049	0.115	0.312	0.063	1.079	0.218	4.943	0.000	0.041	26.408	0.000
RIDGE	0.001	0.005	-0.027	0.082	0.274	0.063	0.950	0.217	4.375	0.000	0.062	15.315	0.000
ENET	0.001	0.004	-0.027	0.139	0.307	0.063	1.064	0.218	4.878	0.000	0.037	28.613	0.000
All Instruments													
OLS	0.000	0.001	0.211	-0.001	0.118	0.067	0.409	0.231	1.773	0.078	0.082	4.972	0.000
3PRF	0.002	0.007	-0.188	0.220	0.240	0.062	0.832	0.216	3.846	0.000	0.049	16.828	0.000
PLS	0.002	0.006	0.106	0.049	0.309	0.068	1.072	0.235	4.554	0.000	0.074	14.532	0.000
PCR	0.003	0.010	0.119	0.031	0.269	0.068	0.932	0.234	3.983	0.000	0.154	6.069	0.000
LASSO	0.001	0.004	0.179	0.223	0.323	0.068	1.120	0.236	4.747	0.000	0.042	26.651	0.000
RIDGE	0.001	0.005	-0.042	0.147	0.243	0.067	0.842	0.233	3.608	0.000	0.038	21.963	0.000
ENET	0.001	0.004	0.212	0.261	0.359	0.068	1.243	0.237	5.241	0.000	0.040	31.397	0.000

B: UMV

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.001	0.009	-0.175	0.102	0.072	0.062	0.248	0.213	1.164	0.245	0.072	3.471	0.001
3PRF	0.001	0.010	-0.165	0.086	0.133	0.062	0.462	0.214	2.157	0.032	0.104	4.445	0.000
PLS	0.002	0.010	0.002	0.017	0.241	0.062	0.836	0.216	3.867	0.000	0.154	5.439	0.000
PCR	0.003	0.009	0.132	0.109	0.315	0.063	1.090	0.218	4.992	0.000	0.110	9.935	0.000
LASSO	0.001	0.008	-0.026	0.014	0.100	0.062	0.346	0.214	1.617	0.107	0.104	3.327	0.001
RIDGE	0.001	0.011	-0.010	0.036	0.083	0.062	0.289	0.214	1.351	0.178	0.095	3.031	0.003
ENET	0.001	0.008	0.058	0.040	0.103	0.062	0.357	0.214	1.669	0.096	0.104	3.444	0.001
FRED-MD													
OLS	0.000	0.001	0.186	0.006	0.128	0.062	0.442	0.214	2.064	0.040	0.065	6.802	0.000
3PRF	0.001	0.008	-0.266	0.128	0.197	0.062	0.684	0.215	3.177	0.002	0.039	17.320	0.000
PLS	0.002	0.007	0.065	0.297	0.346	0.063	1.199	0.220	5.463	0.000	0.056	21.277	0.000
PCR	0.002	0.010	0.070	0.082	0.209	0.062	0.723	0.216	3.356	0.001	0.140	5.185	0.000
LASSO	0.001	0.003	0.055	0.123	0.305	0.063	1.057	0.218	4.846	0.000	0.040	26.500	0.000
RIDGE	0.001	0.005	-0.033	0.087	0.265	0.063	0.918	0.217	4.232	0.000	0.059	15.576	0.000
ENET	0.001	0.004	-0.037	0.141	0.301	0.063	1.044	0.218	4.789	0.000	0.036	28.759	0.000
All Instruments													
OLS	0.000	0.001	0.211	-0.001	0.118	0.067	0.409	0.231	1.771	0.078	0.082	4.959	0.000
3PRF	0.002	0.007	-0.194	0.210	0.228	0.062	0.791	0.216	3.663	0.000	0.047	16.654	0.000
PLS	0.002	0.005	0.134	0.062	0.315	0.068	1.092	0.236	4.634	0.000	0.066	16.579	0.000
PCR	0.002	0.009	0.131	0.007	0.272	0.068	0.944	0.234	4.031	0.000	0.136	6.925	0.000
LASSO	0.001	0.004	0.184	0.234	0.315	0.068	1.090	0.236	4.626	0.000	0.042	25.906	0.000
RIDGE	0.001	0.005	-0.021	0.165	0.247	0.067	0.857	0.233	3.670	0.000	0.037	23.256	0.000
ENET	0.001	0.004	0.222	0.280	0.352	0.068	1.218	0.237	5.142	0.000	0.040	30.241	0.000

(Continued)

Table 2.9.6: Results for the 100 Portfolios Formed on Size and Book-to-Market *(Continued)*

C: MVATE

					Monthly		Annual						
Estimator	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	SR	se	SR	se	t	p	$se(HAC)$	$t(HAC)$	$p(HAC)$
Goyal													
OLS	0.000	0.024	-0.094	0.078	0.003	0.062	0.010	0.213	0.047	0.962	0.042	0.239	0.811
3PRF	0.002	0.026	-0.175	0.064	0.059	0.062	0.204	0.213	0.955	0.341	0.060	3.374	0.001
PLS	0.005	0.017	0.042	0.035	0.279	0.063	0.965	0.217	4.440	0.000	0.101	9.541	0.000
PCR	0.005	0.017	0.103	0.108	0.312	0.063	1.082	0.218	4.957	0.000	0.098	11.065	0.000
LASSO	0.001	0.017	-0.016	0.048	0.084	0.062	0.292	0.214	1.365	0.173	0.089	3.263	0.001
RIDGE	0.001	0.026	-0.010	0.060	0.044	0.062	0.154	0.213	0.720	0.472	0.070	2.197	0.029
ENET	0.001	0.021	0.014	0.026	0.043	0.062	0.150	0.213	0.703	0.483	0.067	2.231	0.027
FRED-MD													
OLS	0.003	0.069	-0.136	0.022	0.047	0.062	0.163	0.213	0.763	0.446	0.025	6.613	0.000
3PRF	0.003	0.023	-0.220	0.009	0.135	0.062	0.469	0.214	2.190	0.029	0.015	31.078	0.000
PLS	0.004	0.019	0.094	0.210	0.230	0.062	0.796	0.216	3.683	0.000	0.101	7.858	0.000
PCR	0.004	0.022	0.082	0.084	0.198	0.062	0.687	0.215	3.189	0.002	0.139	4.931	0.000
LASSO	0.002	0.008	0.074	0.113	0.240	0.062	0.833	0.216	3.851	0.000	0.037	22.748	0.000
RIDGE	0.003	0.015	-0.034	0.082	0.188	0.062	0.652	0.215	3.030	0.003	0.029	22.333	0.000
ENET	0.001	0.012	0.192	0.065	0.129	0.062	0.448	0.214	2.092	0.037	0.069	6.465	0.000
All Instruments													
OLS	-0.002	0.085	-0.103	-0.020	-0.026	0.066	-0.091	0.230	-0.394	0.694	0.040	-2.284	0.023
3PRF	0.003	0.020	-0.187	0.084	0.145	0.062	0.502	0.214	2.344	0.020	0.026	19.497	0.000
PLS	0.003	0.013	0.110	-0.064	0.263	0.068	0.912	0.234	3.898	0.000	0.046	19.874	0.000
PCR	0.005	0.020	0.125	0.031	0.271	0.068	0.937	0.234	4.003	0.000	0.140	6.698	0.000
LASSO	0.003	0.010	0.117	0.200	0.253	0.067	0.875	0.234	3.748	0.000	0.036	24.412	0.000
RIDGE	0.002	0.015	-0.049	0.083	0.117	0.067	0.407	0.231	1.762	0.079	0.036	11.442	0.000
ENET	0.002	0.012	-0.021	-0.093	0.179	0.067	0.620	0.232	2.675	0.008	0.060	10.313	0.000

Table 2.9.6 reports some summary statistics of the OOS (Jan-1996 to Dec-2017) for the 100 portfolios formed on size and book-to-market by estimator and three different set of conditioning information (Goyal, FRED-MD and "All Instruments"). Each panel shows the test of pairs of Sharpe ratios for three different framework to build efficient portfolios. Panel A reports conditionally mean-variance (CMV) efficient portfolios. Panel B reports unconditionally mean-variance efficient portfolios (UMV). Panel C presents the mean-variance tracking error (MVATE) portfolios. We split the results depending on conditioning information set used. Goyal variables comprises: *b/m, dfr, dfy, infl, ltr, lty, ntis, svar, tms* and *tbl* . The FRED-MD is a large dataset containing 128 macroeconomic and financial variables. Finally, "All Instruments" is the combination of Goyal and FRED-MD datasets with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI). The table reports the monthly returns ($\hat{\mu}$), standard deviation ($\hat{\sigma}$), first and second autocorrelation ($\hat{\rho}_1$ and $\hat{\rho}_2$) of the monthly returns. The Sharpe ratios (SR) are presented monthly and annualized, as well as the standard error (se). For the sake of completeness, it is also reported the t-statistic (t), the standard p-value (p), as well as the HAC standard error ($se(HAC)$), the t-statistic obtained using HAC ($t(HAC)$) and its p-value ($p(HAC)$).

Robustness - $\alpha = 1\%$, $\alpha = 1.5\%$, and $\alpha = 2\%$

Table 2.9.7: Implied Sharpe Ratios - $\alpha = 1\%$, $\alpha = 1.5\%$, and $\alpha = 2\%$

	25 Portfolios Formed on Size and Book-to-Market			100 Portfolios Formed on Size and Book-to-Market		
	1%	1.50%	2%	1%	1.50%	2%
Panel A: Goyal						
OLS	0.307	0.288	0.276	-	-	-
3PRF	0.342	0.323	0.310	0.160	0.159	0.158
PLS	0.345	0.330	0.322	0.454	0.437	0.427
PCR	0.439	0.425	0.417	0.457	0.442	0.433
LASSO	0.331	0.305	0.289	0.178	0.164	0.155
RIDGE	0.355	0.338	0.327	0.077	0.065	0.058
ENET	0.343	0.332	0.323	0.125	0.120	0.117
Panel B: FRED-MD						
OLS	-	-	-	-	-	-
3PRF	0.246	0.226	0.212	0.195	0.197	0.197
PLS	0.370	0.357	0.349	0.391	0.397	0.401
PCR	0.427	0.413	0.405	0.329	0.319	0.313
LASSO	0.213	0.188	0.172	0.246	0.220	0.199
RIDGE	0.320	0.305	0.296	-	-	-
ENET	0.309	0.286	0.271	0.137	0.167	0.177
Panel C: All Instruments						
OLS	-	-	-	-	-	-
3PRF	0.236	0.218	0.204	-	-	-
PLS	0.398	0.381	0.371	0.470	0.463	0.453
PCR	0.357	0.343	0.334	0.427	0.413	0.404
LASSO	0.171	0.133	0.106	0.421	0.405	0.395
RIDGE	0.349	0.333	0.322	-	-	-
ENET	0.305	0.280	0.263	0.402	0.419	0.424

Table 2.9.7 reports the implied SR^* as given in equation (2.19) for all seven estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge and ENet), three different sets of conditioning information (Goyal's, FRED-MD, and "All Instruments", wich is the combination of the previous two with the Economic Policy Uncertainty (EPU) index and the Financial Stress Indicator (FSI)), for the **MVATE** mean-variance strategy. Notice that when the OOS R^2 is negative, larger absolute R^2 values than the Sharpe ratio will cause the impossibility of the implied Sharpe ratio to be calculated. These cases are omitted in the table.

Figure 2.9.1: Management Fee - $\alpha = 1\%$

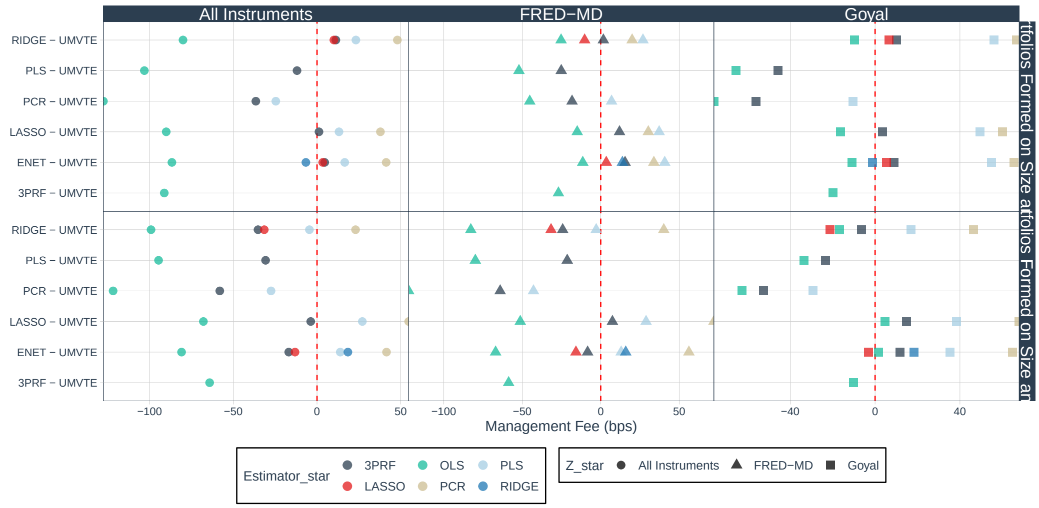

Figure 2.9.1 presents the management fee in bps for the 25 and 100 portfolios formed on Size/BTM, computed as the solution for $\mathcal{F}$ in equation (2.22), from an investor switching from an optimal portfolio formed by the estimator and mean-variance framework given in the left axis to another portfolio plotted. The comparison is done in pairs of optimal portfolios generated by each estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and sets of conditioning information for the **MVATE** mean-variance strategy setting $\alpha = 1\%$.

Figure 2.9.2: Management Fee - $\alpha = 1.5\%$

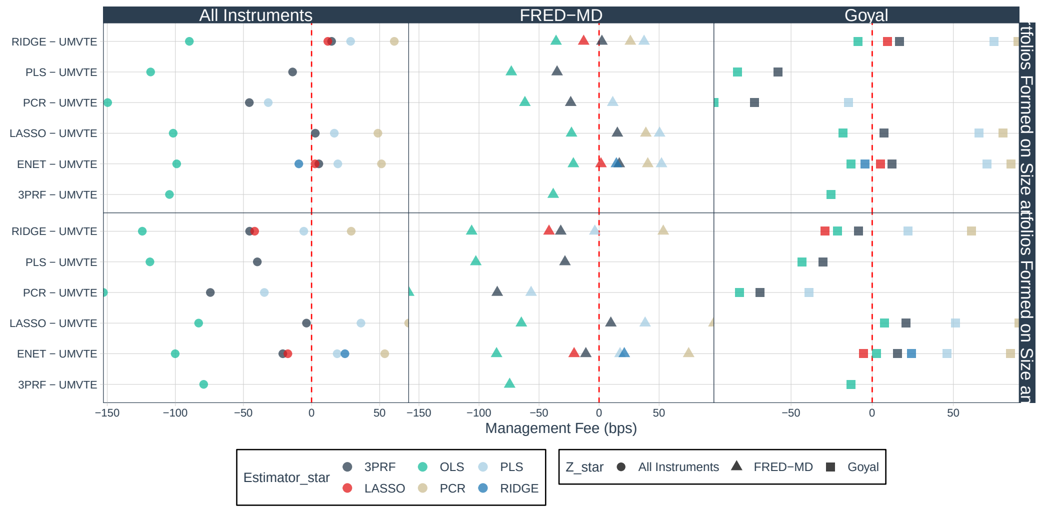

Figure 2.9.2 presents the management fee in bps for the 25 and 100 portfolios formed on Size/BTM, computed as the solution for $\mathcal{F}$ in equation (2.22), from an investor switching from an optimal portfolio formed by the estimator and mean-variance framework given in the left axis to another portfolio plotted. The comparison is done in pairs of optimal portfolios generated by each estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and sets of conditioning information for the **MVATE** mean-variance strategy setting $\alpha = 1.5\%$.

Figure 2.9.3: Management Fee - $\alpha = 2\%$

Figure 2.9.3 presents the management fee in bps for the 25 and 100 portfolios formed on Size/BTM, computed as the solution for $\mathcal{F}$ in equation (2.22), from an investor switching from an optimal portfolio formed by the estimator and mean-variance framework given in the left axis to another portfolio plotted. The comparison is done in pairs of optimal portfolios generated by each estimators (OLS, 3PRF, PLS, PCR, LASSO, Ridge, and ENet) and sets of conditioning information for the **MVATE** mean-variance strategy setting $\alpha = 2\%$.

Chapter 3

Identifying Short Lived Signals in Intraday Foreign Exchange Returns

Abstract

Using high-frequency foreign exchange data, we estimate short-lived signals to forecast one-minute-ahead rolling currency prices using other currency pairs, commodities and stock market indices as predictors. With such a large set of covariates, we impose sparsity using Elastic-Net estimator to find unexpected signals in the intraday currency market, what allow us to deal dynamically with multicollinearity, while constraining the size of the estimates of uninformative variables. We investigate the existence of signal patterns for eleven currency pairs among the most liquid ones spanning the years of 2016 to 2019. The paper also shed lights on the time-of-day effects reported in the literature.

JEL classification: G15, G17, F31, F37, C58.

Keywords: Foreign-exchange. Shrinkage. Pennalized Regression. Elastic-Net.

3.1 Introduction

Recent advances in the literature gathered interesting results in intraday patterns for foreign exchange markets. Breedon and Ranaldo (2013) find evidence of time-of-day effects in exchange returns, in such a way that home currencies tend to depreciate during domestic business hours and appreciate during foreign business hours. However, other studies found some contradictory results (Cornett et al., 1995; Cai et al., 2008). A possible explanation for these conflicting results may lie in the fact of the existence of unexpected, significant and temporary signals in the market.

In this paper we investigate whether there are and if it is possible to identify short-lived signals in intraday currency exchange markets, accounting for these patterns studied in the literature. Motivated by Chinco et al. (2019) that documented interesting insights to incorporate high-dimensional data in intraday stock market forecasts, we show that it is possible to incorporate a vast set of sources of movements to capture short-lived and unexpected signals that can improve the forecasts in intraday FX market.

Given the complexity and interconnectedness in the current state of financial markets, movements in many other variables can be informative to a specific currency. In order to try to capture these possibly short-lived and unexpected signaling information, we make use of the Elastic-Net (ENet) estimator to assist us in the process of selecting important sources of predictability, and shrinking and discarding uninformative information that just add noise to the process.

We motivate this idea of using high-dimensional sets of predictors in intraday FX data, along with a statistical rule to filter the signals, in a similar way to Chinco et al. (2019). Next minute informative signals about the state of an asset in the currency exchange market exists at a scale that makes too hard to obtain economic interpretation to explain intraday movements. Thus, making too difficult to a researcher or market practitioner to try to intuit all the sources that may be impacting the movements in FX markets.

Hence, using a large set of covariates, we impose sparsity using the Elastic-Net estimator to find unexpected signals in the intraday currency market, what allow us to deal dynamically with multicollinearity, while constraining the size of the estimates of uninformative variables. We investigate the existence of signal patterns for the five most liquid currencies (British Pound, Canadian Dollar, Euro, Japanese Yen, and Swiss-Franc) for the year of 2018. The process of incorporating short-lived and unexpected signals is done in a rolling 1-minute basis. We use other currency pairs, commodities and stock market indices as predictors. This work also shed lights on the time-of-day effects reported in the literature.

The structure of this paper is as follows. Next section introduces the Elastic-Net estimator and discuss its features. Section 3.3 discuss the data used and the forecasting framework designed to obtain out-of-sample forecasts. Section 3.4 presents the results. Additional results, tables and figures are presented in Appendix 3.6.

3.2 Estimation

We are interested in identifying short-lived predictors in financial markets. Short-lived signals arise quite frequently in financial data, especially when dealing with high-frequency data. However, given the complexity and the exponential increase of the speed in transactions, it makes almost impossible to find these signals uniquely or mainly using economic intuition. We focus in the intraday exchange currency market and dealing with high-dimensional set of predictors, we seek to find sources of predictability in FX markets that last for short periods and are difficult to find based in economic interpretability.

Consider a simple predictive regression such as

$$r_{t+1,j} = \beta_0 + \boldsymbol{\beta_1}^\top \boldsymbol{z}_t + \epsilon_{t+1} \qquad t = 1, 2, \ldots, T \tag{3.1}$$

where $\boldsymbol{z}_t$ is a high-dimensional vector with predictive information. It is well known that the standard OLS faces several issues, widely known as the "curse of dimensionality". When dealing with large number of predictors, whenever we have $K > T$, where K is the dimension of $\boldsymbol{z}_t$, and T the number of observations in the series of $r_{t+1,j}$, OLS is unable to find an unique solution. Additionally, the infinite set of solutions have a tendency to overfit the data instead of extracting informative signals for the conditional mean. This fact, which is more serious when the signal-noise ratio in return predictions is low, causes the OLS estimations to have a poor behavior in forecasting.

To overcome these issues when dealing with FX markets data and, especially being able to use a statistical method to identify unexpected signals that may last only in short horizons in intraday data, we make use of a penalized estimator. Specifically, we use the Elastic-Net, which is one of the most flexible estimators in this class, allowing for a combination of shrinkage and variable selection in high-dimensional settings. In general, penalized estimators share the following common structure:

$$\mathcal{L}(\boldsymbol{\beta}; \cdot\,) = \underbrace{\mathcal{L}(\boldsymbol{\beta})}_{\text{Loss Function}} + \underbrace{\phi(\boldsymbol{\beta}; \cdot\,)}_{\text{Penalty}} \quad . \tag{3.2}$$

In the structure of equation 3.2, the Elastic-Net (ENet) estimator proposed by Zou and

Hastie (2005) chooses as penalty function a convex combination of ℓ_1- and ℓ_2-norm for the size of the betas in the regression 3.1. Following Hastie et al. (2015), the ENet solves the following convex program:

$$\hat{\boldsymbol{\beta}}_i^{ENet} := \left\{ \arg\min_{\beta_{0,i},\boldsymbol{\beta}_i} \frac{1}{T} \sum_{t=1}^{T} \left(r_{i,t+1} - \beta_{0,i} - \sum_{k=1}^{K} Z_{k,t}\beta_{k,i} \right)^2 + \lambda \sum_{k=1}^{K} \left(\alpha \left|\beta_{k,i}\right| + (1-\alpha)\, \beta_{k,i}^2 \right) \right\} \tag{3.3}$$

where $\alpha \in [0,1]$. Setting $\alpha = 1$, this estimator imposes uniquely a ℓ_1-norm penalty. In this case, it reduces to the Least Absolute Selection and Shrinkage Operator (LASSO), proposed originally by Tibshirani (1996). A ℓ_1-norm penalty is related to the idea of variable selection to impose sparsity in the data, in such a way that many of the coefficients from the high-dimensional vector of independent variables are set equal to 0. This estimator regularizes the estimation process constraining the size of the estimates.

On the other hand, setting $\alpha = 0$ reduces to the Ridge regression, which imposes only a ℓ_2-penalty. An important characteristic of Ridge approach is the one-to-one correspondence between the parameters λ and t. This fact makes Ridge estimation to attenuate multicollinearity, when it is present in the data. Since, when K is large, the high amount of regressors may result in high correlation among some of them. With standard "naive" OLS, multicollinearity causes poorly determined coefficients with large variances (and large variance inflator factor - VIF), what Ridge can mitigate.

Among the positive features of the ENet, when adding both penalties, this estimator automatically controls for strong within-group correlations. The ENet is also a sctrictly convex problem; thus, providing a unique solution independently of correlations or duplications in the $Z_{k,t}$. However, in equation 3.3 we have two tuning parameters λ and α that has to be determined, or defined *ad hoc*. The parameter λ controls the amount the shrinkage: the larger it is, the greater is the shrinkage.

3.3 Methodology

3.3.1 Data

We focus our analysis on five of the most liquid[1] currencies: the British Pound (GBP), Canadian Dollar (CAD), Euro (EUR), Japanese Yen (JPY), and the Swiss Franc (CHF)[2]

[1] According to the Bank of International Settlements(BIS).

[2] We plan to extend this to several other currencies, such as Australian Dollar (AUD), Swedish Krona (SEK) and Norwegian Krone (NOK).

against the United States Dollar (USD). Given the large amount of data and the process of obtaining estimates in a rolling structure at a 1-minute scale, which is highly computationally demanding, we only use the intraday data for the year of 2018.

For each currency pair, we use the orderbook information for bid and ask prices to construct our 1-minute tick data. Using the last bid and ask in each minute we calculate the returns based on mid prices. Thus, the price of currency j at the minute t in a given day is given by:

$$P_{t,j} = \frac{\text{Bid}_{t,j}^{\text{last}} + \text{Ask}_{t,j}^{\text{last}}}{2} \tag{3.4}$$

Then, we can compute the 1-minute return of currency j between minutes t and $t-1$ as:

$$r_{t,j} = log(P_{t,j}) - log(P_{t-1,j}) \tag{3.5}$$

For each currency pair above we build a large set of potential predictors. This set is formed with the last minute bid, ask and volume of all other currency pairs we have in our dataset. There are other 60 currency pairs in our dataset for each currency j we are interested in. In order to impose stationarity for the bid and ask variables, we apply the log difference in each one of these variables. Table 3.6.1 in Appendix 3.6 shows the entire list of currencies data we consider in our set of predictors, along with some descriptive statistics for the year of 2018.

Additionally, we also consider other sources of predictability in the 1-minute scale. We consider 11 commodities prices, 2 metals prices, as well as 18 stock indexes. For all of them, we use the 1-minute log difference of their original values. Table 3.6.2 in Appendix 3.6 reports the description and further descriptive statistics of these variables for the full sample of 2018.

As the foreign exchange market operates 24 hours 6 days a week (Monday through Saturday), there are $8,640$ observations per week. We decided to keep Saturday data, i.e., the last day of each week data, even though there is less liquidity in the market, mainly because we also seek to evaluate and identify what periods of the day and week there are more unexpected short-lived lived predictors. We base the timing of the data on Coordinated Universal Time (UTC), as this time zone is more stable.

All our data were extracted from historical data feed of Dukascopy [3], a large and liquid Swiss FX bank. As this data is publicly available, making use of this source allows easy replication of our results. Table 3.3.1 shows the summary statistics of our five currency pairs returns for the full year of 2018. In this table we report the 1-minute mean, standard

[3] Dukascopy Bank SA. https://www.dukascopy.com/swiss/english/marketwatch/historical/

Table 3.3.1: Descriptive Statistics

	USD/CHF	EUR/USD	USD/CAD	USD/JPY	GBP/USD
Mean $\times 10^6$ (%)	2.91	-12.92	24.05	-4.54	-14.92
Mean (annualized %)	1.32	-5.64	11.41	-2.02	-6.48
Standard Deviation $\times 10^4$	1.175	1.261	1.223	1.156	1.449
Skewness	-0.015	0.137	-0.208	0.261	0.527
Kurtosis	18.735	26.092	154.880	27.370	85.951
Normality (p-value)	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001
ADF (p-value)	0.010	0.010	0.010	0.010	0.010
1st-Order Autocorrelation	-0.023	-0.018	-0.019	-0.020	-0.022
BP Heteroskedacity (p-value)	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001

Table 3.3.1 reports the descriptive statistics of our five currency pairs returns. Normality is the Jarque-Bera test p-value. We test stationarity with the ADF test. The heteroskedasticity is the Breusch-Pagan (BP) test. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

deviation, skewness and kurtosis. We also report the Jarque-Bera p-value test of normality, the ADF test of stationarity, the first-order autocorrelation and the Breusch-Pagan test of heteroskedsticity.

Figure 3.3.1: Mean Annualized Return per Hour (2016-2019)

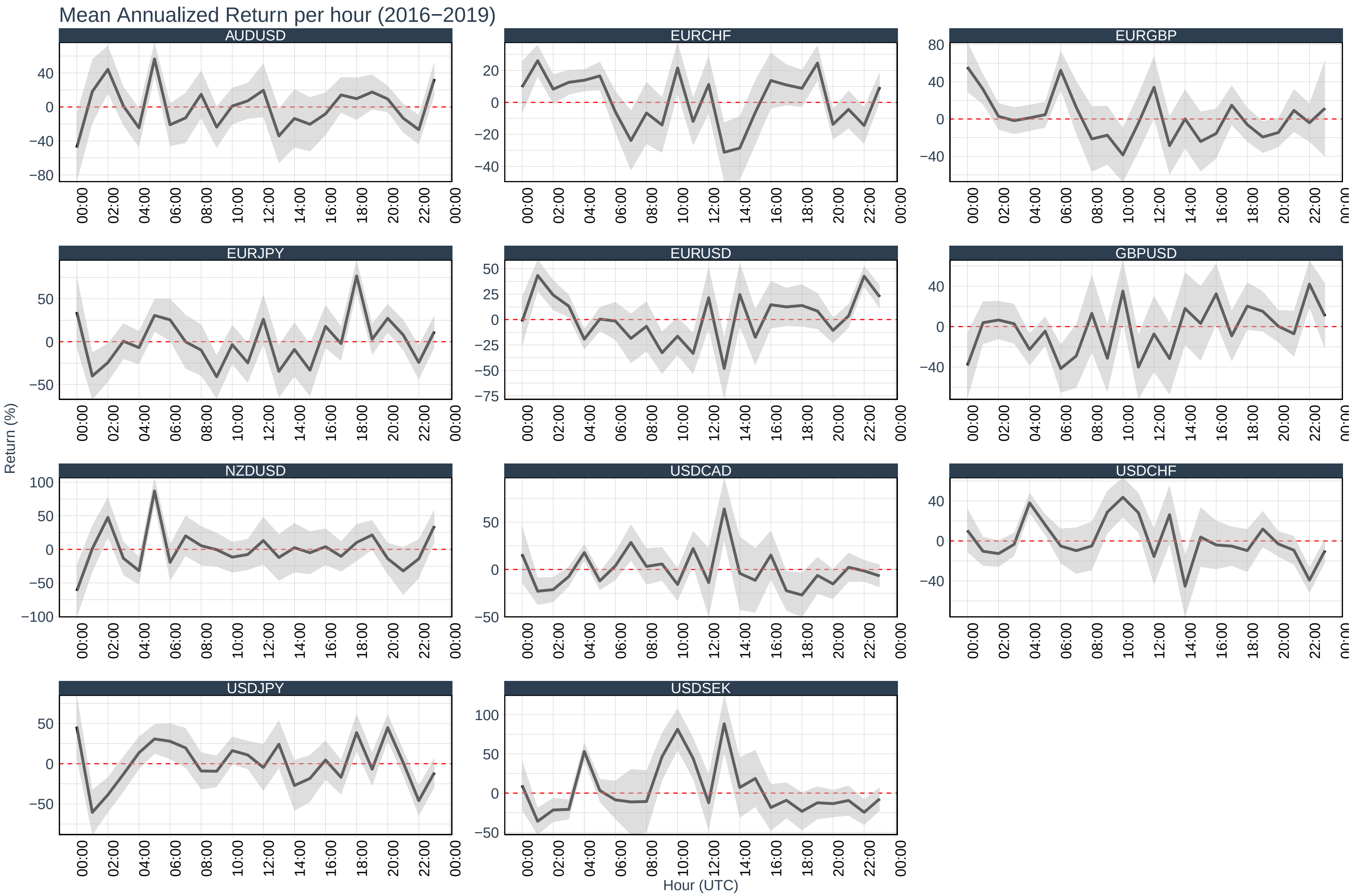

Figure 3.3.1 shows the estimated mean per hour of the annualized return of each one of the eleven currency pairs in analysis. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), and y-axis the the hourly mean of the annualized 5-minute return. The shaded areas denote the 68% confidence bands.

Figure 3.3.2: Mean Annualized Return per Weekday (2016-2019)

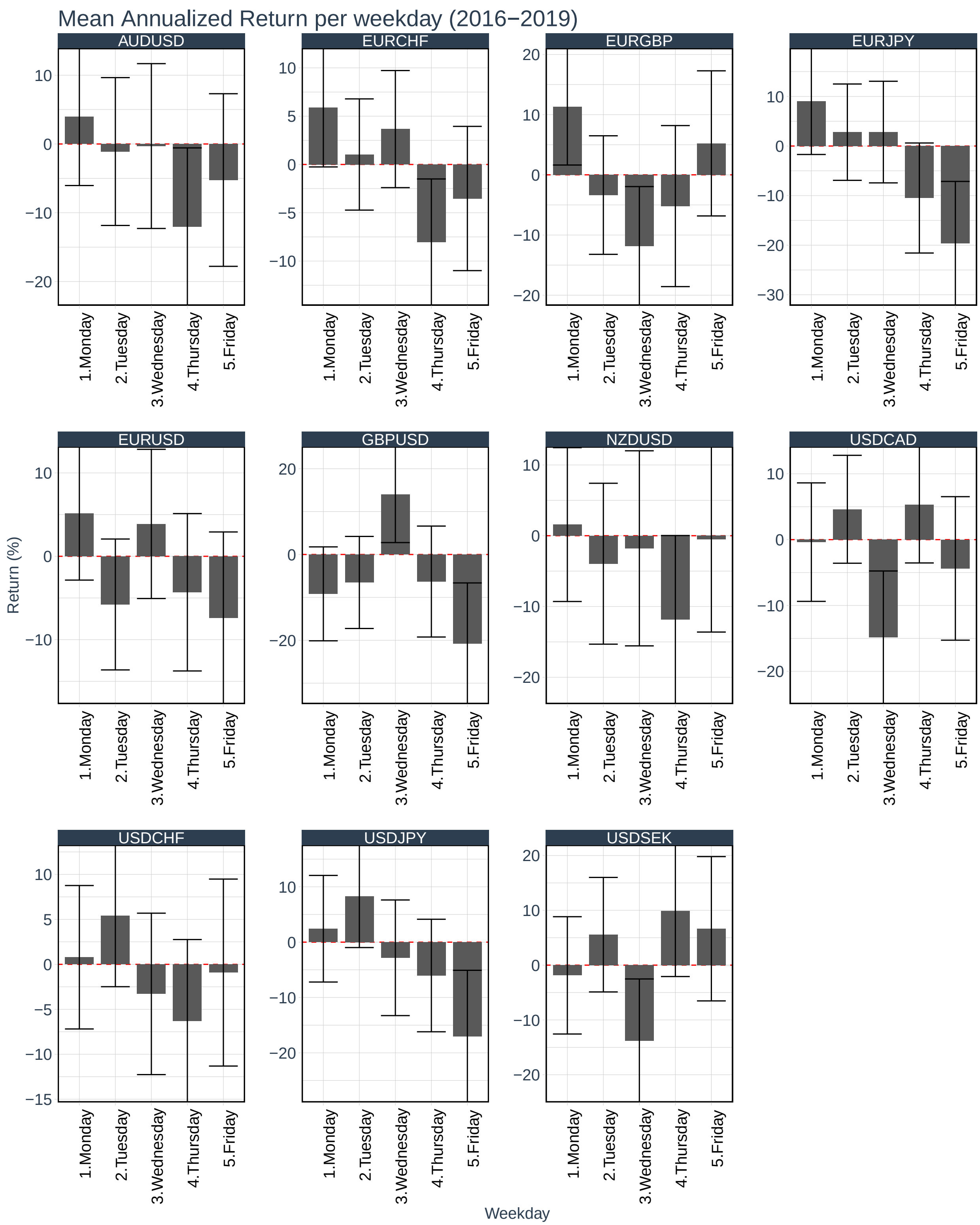

Figure 3.3.2 shows the estimated mean per weekday (Monday-Friday) of the annualized return of each one of the eleven currency pairs in analysis. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), and y-axis the the mean of the annualized 5-minute return for each weekday (bars). The error bars denote the 68% confidence bands.

Figure 3.3.3: Boxplot of 5-Minute Return per Market Hours (2016-2019)

Figure 3.3.3 shows the boxplots of the 5-minute return of each one of the eleven currency pairs in analysis for the four trading hours groups: *U.S. hours* (2:00 PM to 9:00 PM GMT), *Europe hours* (8:00 AM to 5:00 PM GMT), the overlap of U.S. and Europe hours (2:00 PM to 5:00 PM GMT), and the hours outside the previous three groups - *Not U.S. and Europe hours* (8:00 PM to 8:00 AM GMT). The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent each one of the four trading hours groups, and y-axis the the 5-minute return.

Figure 3.3.4: Density of 5-Minute Return per Market Hours (2016-2019)

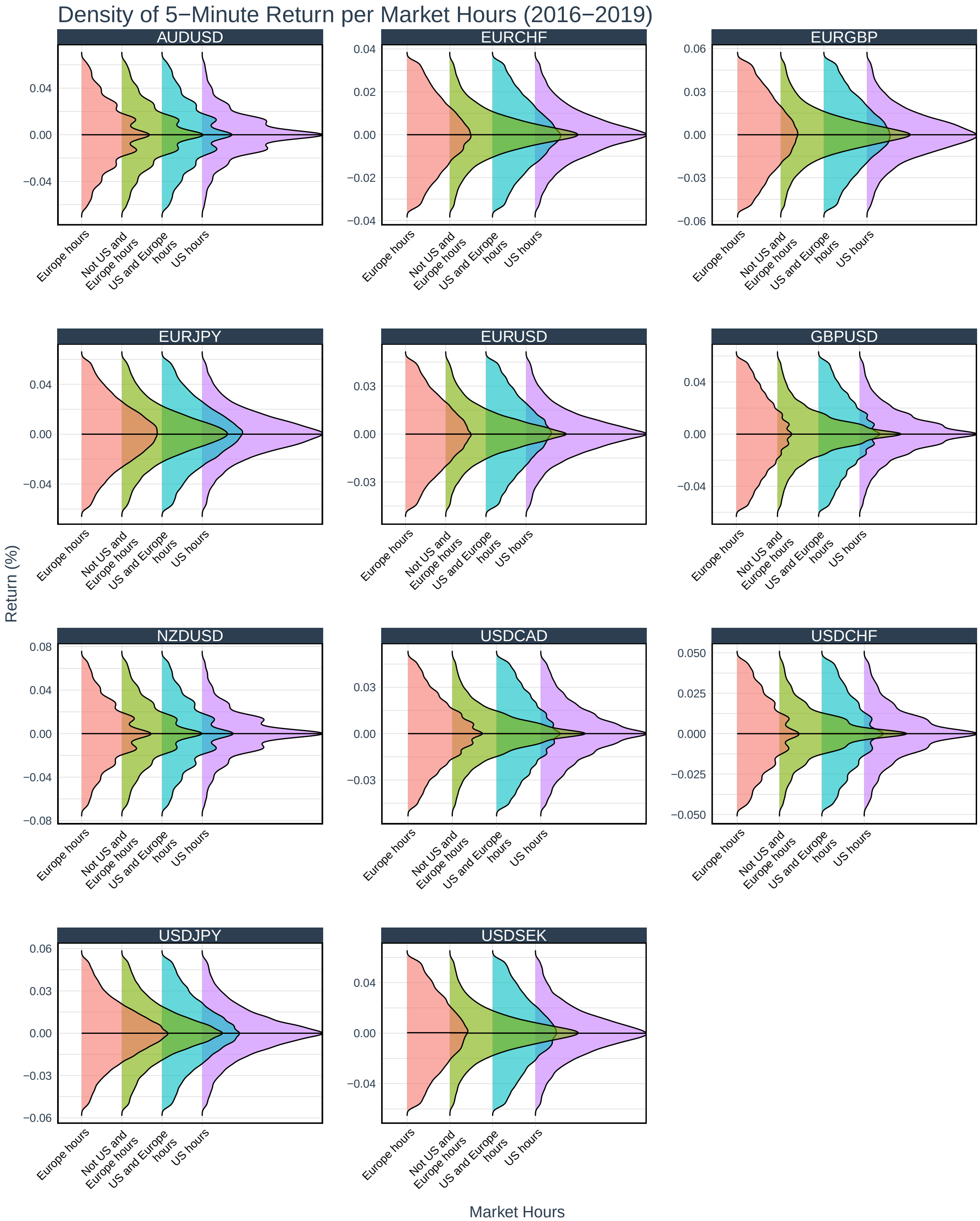

Figure 3.3.4 shows the estimated densities of the 5-minute return of each one of the eleven currency pairs in analysis for the four trading hours groups: *U.S. hours* (2:00 PM to 9:00 PM GMT), *Europe hours* (8:00 AM to 5:00 PM GMT), the overlap of U.S. and Europe hours (2:00 PM to 5:00 PM GMT), and the hours outside the previous three groups - *Not U.S. and Europe hours* (8:00 PM to 8:00 AM GMT). The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent each one of the four trading hours groups, and y-axis the the 5-minute return.

Figure 3.3.5: Boxplot of 5-Minute Return per Hours (2016-2019)

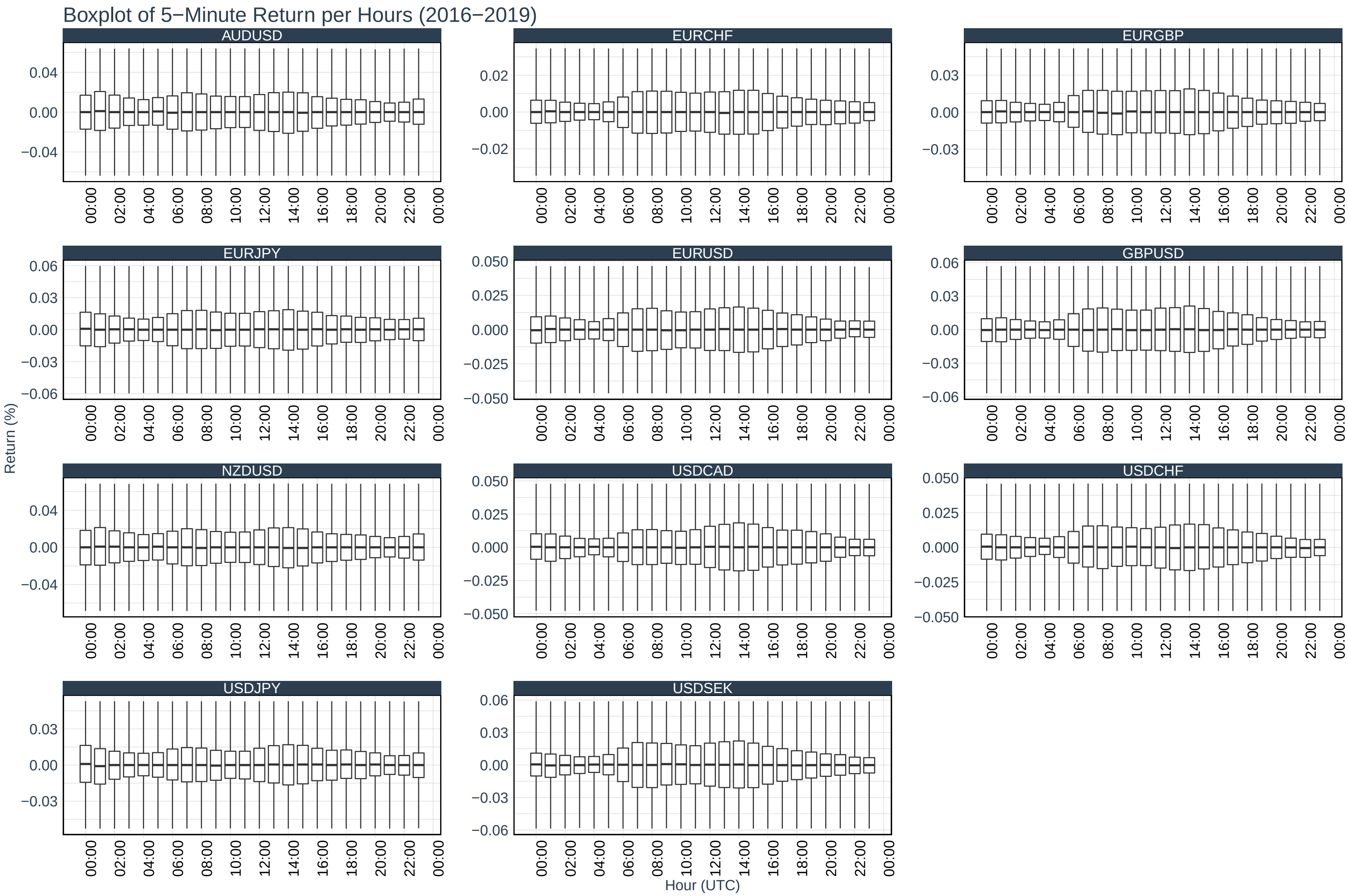

Figure 3.3.5 shows the hourly boxplots of the 5-minute return of each one of the eleven currency pairs in analysis. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), and y-axis the the 5-minute return.

3.3.2 The Forecasting Framework

In this section we discuss our research design to obtain intraday forecasts for the returns of our currency pairs. We make use of the rolling window scheme using intraday data of our set of predictors. For each currency pair we are interested in, we set the length of the window as 30 minutes. Thus, we use the past 30 observations as a cross-section. The list of predictors, as explained in section 3.3.1, contains the log differences of bid prices and ask prices, as well as volume of all other currency pairs, stock indexes, commodities and metals. We also add the last 5 minutes lag of each one of these variables. In total, there are $1,648$ independent variables in each regression window. Hence, our matrix of covariates has $30 \times 1,648$ dimension. Clearly, OLS would fail in such case. We summarize the predictive regression for currency j in each window as:

$$\boldsymbol{r}_{t+1,j} = \boldsymbol{Z}_t \boldsymbol{\beta}^{\text{ENet}} + \boldsymbol{\epsilon}_{t+1} \tag{3.6}$$

where $\boldsymbol{r_{t+1,j}}$ is a 30×1 vector, $\boldsymbol{\beta}^{\text{ENet}}$ is a 30×1 vector, and $\boldsymbol{Z}$ is a $30 \times (1,648 + 1)$ matrix of covariates.

Then, obtained $\widehat{\boldsymbol{\beta}}^{\text{ENet}}$ from a window, we use the last 5 minutes observations data to construct our fitted value for the next minute $t+1$. We denote this forecast as $\widehat{f}_{t,j}^{ENet}$. Next, we move the 30 minutes window one minute ahead and repeat this process for all intraday data, Monday through Friday, for the whole year from 2016 to 2019, for each one of the eleven currency pairs we are interested in.

3.3.3 Validation

As mentioned in section 3.2, penalized regressions do not have a closed form solution mainly because of the penalty λ imposed in the objective function. For ENet especifically, λ and α are hyperparameters that need to be defined in each regression. To tune these parameters, we recursively define the training and validation periods. The prevailing approach in the literature is to use the tunning subsample to adaptively determine the hyperparameters. Since this is a time series, standard approaches such as traditional K-fold cross-validation that randomly selects subsambles may be problematic.

Thus, at each 30 minutes window, we start with a vector of the first 10 minutes as our training. Then, the remaining 20 observations are used as validation, in such a way that we enlarge the training vector with one observation at time. For α and λ, we defined their individual grids as a 15×1 vector equally spaced of possible hyperparameters values in the interval $[0, 1]$[4]. For a list of possible different hyperparameters[5], the mean square error (MSE) is computed in this training plus validation sample. The hyperparameters values that produce the lowest MSE are chosen. In this sense, we say that the training and the validation samples are consequential to preserve the time series dependence[6].

3.4 Empirical Results

3.4.1 Out-of-Sample Forecasting Performance

In order to evaluate our results, we define three benchmarks to compare with the forecasts obtained with the Elastic-Net estimator. The idea is not solely assess if the forecast provides better in- or out-of-sample results. We also seek to know whether we see any gains when using these benchmarks in conjunction with $\widehat{f}_{t,j}^{ENet}$. The first benchmark is the random walk (RW). We also consider autoregressive processes such as AR(1) and ARMA(p,q) with parameters obtained by BIC (denoted by $ARMA_{BIC}$). Each benchmark was built in a similar fashion to the process of estimating $\widehat{f}_{t,j}^{ENet}$. In a 30-minutes rolling window scheme, we used the 30×1 vector of the last 30 minutes returns $r_{t,j}$ from currency j to generate each benchmark forecast for the following minute.

3.4.2 Predictive Regressions

We are interested to know whether (i) it is important to account for high-dimensional data, (ii) how our forecast based on Elastic-Net that imposes sparsity and variable selection improves our prediction in an out-of-sample analysis, and (iii) how much return variation does our Elastic-Net forecast explain. In order to provide some insights to answer these questions, we estimate the following regression for each currency pair j:

$$\boldsymbol{r}_{t+1,j} = \widehat{\boldsymbol{f}}_{t,j}^{ENet} \boldsymbol{\delta}_j + \boldsymbol{\epsilon}_{t+1} \tag{3.7}$$

where $\boldsymbol{r_{t+1,j}}$ is the vector with 1-minute return of currency j for the entire period of analysis and $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ is the forecast for the 1-minute-ahead obtained through Elastic-Net as shown in equation 3.6. We estimate a separate regression for each currency pair we are interested in. We seek to evaluate if the estimates $\widehat{\boldsymbol{\delta}}_j$ are statistically significant, and how much of the

[4] The larger this vector is, the higher is the computational burden. Taking into consideration that this is done at the 1-minute level with rolling regressions, we made the choice of a vector size of 15 for α and λ. Common choices ranges from vector with 5 to 50 values.

[5] We set a grid of 20 possible values that λ and α could take in the interval $[0, 1]$

[6] An interesting discussion of using cross-validation on time series data was done by Bergmeir et al. (2018)

return variation can $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ explain using the fraction of the variation for currency pair j is explained by our forecasts.

Table 3.4.1 reports the out-of-sample predictive regression as in 3.7 for all eleven currency pairs split by panels sorted by turnover. In equation 3.7 we also replace the ENet forecast by each one of the three benchmarks: RW, AR(1), and $ARMA_{BIC}$. Columns (1) – (4) are univariate regressions either using the Elastic-Net forecasts, or one of the three benchmarks. Columns (5) – (7) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark.

A quick examination show that, over the full sample

Using Newey-West robust standard errors, we notice that the Elastic-Net forecast is vastly statistically significant at the 1% significance level for the eleven currency pairs. The only case that this does not apply is to GBP/USD, which is the currency pair with the third largest turnover. On the other hand, we see that AR(1) is weakly significant or not significant for most of the currencies. We find similar evidence for the parameter associated to the $ARMA_{BIC}$ benchmark. Among the three benchmarks we consider, the random walk show evidence of being a good source of predictability, with a similar significance to the $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$.

Notice as well that either $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ or the random walk out-of-sample positive forecasts are associate with a depreciation for CHF, JPY and CAD, while for EUR and GBP it is the opposite. Another important evidence gathered from tables 3.4.1 and 3.4.1 is that the adjusted R^2, even though, as expected, is small for all cases increases significantly when $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the random walk are used in conjunction, especially for CHF.

To answer whether this behavior maintains if we control for some of the intraday patterns reported in the literature, we use sets of dummies. First, we create 23 to represent each one of the hours of the day. The second set are two dummies, representing the trading hours in US and Europe. US trading hours are restricted to be range from 9 : 00 AM to 4 : 00 PM (EST), while Europe trading hours range from 8 : 00 to 5 : 00 PM (GMT). We consider a wider time range to Europe, to cover the trading hours in UK, France, Germany, Italy and Switzerland.

Tables 3.4.2, 3.4.3, and 3.4.3 report the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for one currency pair. Columns (1) – (7) control for all hours of the days, and columns (8) – (14) control for the US and Europe trading hours, in such there are two dummies, one for each case. We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) and (8) – (11) are univariate regressions, while (5) – (7) and (12) – (14) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark.

Table 3.4.1: Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks

Panel A: **EUR/USD**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.024***				−0.023***	−0.021***	−0.011
	(0.007)				(0.007)	(0.006)	(0.007)
AR(1)		−0.029***			−0.017**		
		(0.008)			(0.008)		
$ARMA_{BIC}$			−0.046**			−0.035*	
			(0.019)			(0.020)	
RW				−0.016***			−0.012***
				(0.003)			(0.004)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.0002	0.0001	0.0001	0.0003	0.0002	0.0003	0.0003
F Statistic	287.657***	75.947***	196.750***	374.348***	155.575***	196.894***	203.682***

Panel B: **USD/JPY**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.050***				−0.041***	−0.036**	−0.046***
	(0.015)				(0.013)	(0.017)	(0.017)
AR(1)		−0.093**			−0.065		
		(0.044)			(0.041)		
$ARMA_{BIC}$			−0.109***			−0.086**	
			(0.041)			(0.043)	
RW				−0.021***			−0.003
				(0.006)			(0.005)
Observations	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920
Adjusted R^2	0.001	0.001	0.001	0.0005	0.001	0.001	0.001
F Statistic	1,209.864***	787.245***	1,455.151***	643.446***	782.054***	1,013.807***	609.608***

Panel C: **GBP/USD**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.004				−0.001	0.011	0.010
	(0.043)				(0.046)	(0.046)	(0.043)
AR(1)		−0.023			−0.022		
		(0.020)			(0.031)		
$ARMA_{BIC}$			−0.088***			−0.095***	
			(0.030)			(0.035)	
RW				−0.009			−0.013
				(0.018)			(0.009)
Observations	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920
Adjusted R^2	0.00000	0.00003	0.001	0.0001	0.00003	0.001	0.0001
F Statistic	7.321***	49.617***	863.845***	105.030***	24.969***	458.883***	66.773***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.1 reports the out-of-sample predictive regression as in equation 3.7 for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for three currency pairs split by panels. We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) are univariate regressions, while (5) – (7) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.1: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks

Panel D: **AUD/USD**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.026***				−0.025***	−0.024***	−0.011**
	(0.004)				(0.004)	(0.004)	(0.005)
AR(1)		−0.021***			−0.008		
		(0.007)			(0.007)		
$ARMA_{BIC}$			−0.031***			−0.020*	
			(0.011)			(0.011)	
RW				−0.017***			−0.013***
				(0.003)			(0.003)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.0002	0.00003	0.0001	0.0003	0.0002	0.0002	0.0003
F Statistic	304.877***	40.093***	87.892***	417.489***	155.212***	169.629***	225.388***

Panel E: **USD/CAD**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.017***				−0.016***	−0.014***	−0.010*
	(0.005)				(0.005)	(0.005)	(0.006)
AR(1)		−0.021***			−0.011		
		(0.008)			(0.008)		
$ARMA_{BIC}$			−0.039***			−0.032***	
			(0.011)			(0.011)	
RW				−0.010***			−0.007**
				(0.002)			(0.003)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.0001	0.00003	0.0001	0.0001	0.0001	0.0002	0.0001
F Statistic	139.308***	37.391***	140.661***	155.013***	75.153***	116.564***	90.094***

Panel F: **USD/CHF**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.027***				−0.025***	−0.025***	−0.013**
	(0.005)				(0.005)	(0.005)	(0.006)
AR(1)		−0.031***			−0.017**		
		(0.008)			(0.008)		
$ARMA_{BIC}$			−0.031**			−0.018	
			(0.013)			(0.013)	
RW				−0.017***			−0.013***
				(0.003)			(0.003)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.0002	0.0001	0.0001	0.0003	0.0002	0.0003	0.0003
F Statistic	328.064***	85.686***	93.401***	425.524***	176.948***	179.404***	234.476***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.1 reports the out-of-sample predictive regression as in equation 3.7 for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for two currency pairs split by panels. We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) − (4) are univariate regressions, while (5) − (7) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.1: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks

Panel G: **EUR/GBP**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.059***				−0.063***	−0.060***	−0.007
	(0.021)				(0.023)	(0.021)	(0.016)
AR(1)		0.010			0.039		
		(0.041)			(0.045)		
$ARMA_{BIC}$			−0.020			0.005	
			(0.026)			(0.027)	
RW				−0.050***			−0.047***
				(0.017)			(0.015)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.001	0.00001	0.00002	0.002	0.001	0.001	0.002
F Statistic	1,745.112***	9.746***	36.273***	3,520.391***	939.943***	873.849***	1,767.772***

Panel H: **EUR/JPY**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.050***				−0.045***	−0.032	−0.049*
	(0.011)				(0.011)	(0.021)	(0.025)
AR(1)		−0.069			−0.040		
		(0.047)			(0.035)		
$ARMA_{BIC}$			−0.145**			−0.124*	
			(0.062)			(0.071)	
RW				−0.020**			−0.001
				(0.008)			(0.017)
Observations	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920
Adjusted R^2	0.001	0.0003	0.002	0.0004	0.001	0.002	0.001
F Statistic	1,249.415***	435.627***	2,326.273***	566.592***	692.188***	1,403.984***	624.951***

Panel I: **NZD/USD**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.028***				−0.028***	−0.025***	−0.016**
	(0.006)				(0.006)	(0.006)	(0.008)
AR(1)		−0.017*			−0.002		
		(0.009)			(0.009)		
$ARMA_{BIC}$			−0.040***			−0.027**	
			(0.013)			(0.014)	
RW				−0.016***			−0.010**
				(0.003)			(0.004)
Observations	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920	1,414,920
Adjusted R^2	0.0002	0.00002	0.0001	0.0003	0.0002	0.0003	0.0003
F Statistic	342.745***	26.323***	146.435***	363.873***	171.511***	202.550***	216.125***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.1 reports the out-of-sample predictive regression as in equation 3.7 for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for two currency pairs split by panels. We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) are univariate regressions, while (5) – (7) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.1: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks

Panel J: **USD/SEK**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.038***				−0.040***	−0.037***	−0.006
	(0.004)				(0.004)	(0.004)	(0.005)
AR(1)		0.001			0.019**		
		(0.008)			(0.009)		
$ARMA_{BIC}$			−0.031***			−0.015	
			(0.012)			(0.012)	
RW				−0.031***			−0.029***
				(0.002)			(0.003)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.0005	-0.00000	0.0001	0.001	0.0005	0.0005	0.001
F Statistic	643.389***	0.025	93.047***	1,336.197***	337.560***	332.389***	672.822***

Panel K: **EUR/CHF**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
f^{ENet}	−0.062***				−0.066***	−0.064***	−0.018**
	(0.008)				(0.008)	(0.008)	(0.009)
AR(1)		0.014			0.039**		
		(0.014)			(0.016)		
$ARMA_{BIC}$			−0.003			0.023	
			(0.012)			(0.014)	
RW				−0.045***			−0.039***
				(0.003)			(0.003)
Observations	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000	1,422,000
Adjusted R^2	0.001	0.00001	-0.00000	0.002	0.001	0.001	0.002
F Statistic	1,608.166***	16.844***	0.870	2,879.774***	870.048***	826.780***	1,484.531***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.1 reports the out-of-sample predictive regression as in equation 3.7 for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for two currency pairs split by panels. We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) − (4) are univariate regressions, while (5) − (7) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.2: Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks - Intraday Patters

Panel A: **EUR/USD**

	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.025**	−0.025**	−0.022**	−0.014	−0.017*	−0.014	−0.013*	−0.005	−0.032**	−0.031**	−0.027	−0.024
	(0.011)	(0.011)	(0.011)	(0.012)	(0.009)	(0.009)	(0.007)	(0.009)	(0.014)	(0.015)	(0.017)	(0.018)
AR(1)		−0.008				−0.028**				−0.007		
		(0.011)				(0.011)				(0.013)		
$ARMA_{BIC}$			−0.034*				−0.039				−0.051	
			(0.018)				(0.026)				(0.038)	
RW				−0.010**				−0.011**				−0.007
				(0.004)				(0.005)				(0.007)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.0002	0.0002	0.0003	0.0003	0.0001	0.0001	0.0002	0.0002	0.0003	0.0003	0.0005	0.0004
F Statistic	105.679***	53.806***	69.552***	67.039***	63.296***	46.377***	58.895***	53.857***	192.383***	96.988***	143.733***	103.845***

Panel B: **USD/JPY**

	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.032***	−0.032***	−0.031***	−0.020**	−0.025***	−0.023***	−0.020***	−0.017**	−0.069**	−0.050*	−0.041	−0.074**
	(0.008)	(0.009)	(0.009)	(0.010)	(0.005)	(0.005)	(0.006)	(0.007)	(0.029)	(0.030)	(0.045)	(0.035)
AR(1)		0.001				−0.019*				−0.133*		
		(0.013)				(0.012)				(0.078)		
$ARMA_{BIC}$			−0.007				−0.047***				−0.126**	
			(0.018)				(0.016)				(0.059)	
RW				−0.010**				−0.007				0.004
				(0.004)				(0.004)				(0.012)
Observations	474,240	474,240	474,240	474,240	603,000	603,000	603,000	603,000	578,880	578,880	578,880	578,880
Adjusted R^2	0.0003	0.0003	0.0003	0.0004	0.0002	0.0002	0.0003	0.0002	0.002	0.003	0.003	0.002
F Statistic	159.949***	79.994***	80.720***	94.369***	124.158***	68.662***	101.838***	69.839***	1,018.289***	806.715***	960.748***	511.888***

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 3.4.2 reports the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for two currency pairs split by panels. Columns (1) – (7) control for all hours of the days, in such a way that there are 23 dummies, one for each hour. Columns (8) – (14) control for the US and Europe trading hours, in such there are two dummies, one for each case. US trading hours are restricted to be range from 9 : 00 AM to 4 : 00 PM (EST), while Europe trading hours range from 8 : 00 to 5 : 00 PM (GMT). We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) and (8) – (11) are univariate regressions, while (5) – (7) and (12) – (14) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.3: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks - Intraday Patters

Panel C:	**GBP/USD**											
	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.054***	−0.052***	−0.052***	−0.032***	−0.031***	−0.028***	−0.026***	−0.019**	0.055	0.064	0.107	0.057
	(0.009)	(0.009)	(0.010)	(0.012)	(0.006)	(0.006)	(0.006)	(0.008)	(0.105)	(0.121)	(0.113)	(0.107)
AR(1)		−0.015				−0.025**				−0.052		
		(0.015)				(0.011)				(0.118)		
$ARMA_{BIC}$			−0.013				−0.043***				−0.225**	
			(0.019)				(0.014)				(0.099)	
RW				−0.019***				−0.010***				−0.002
				(0.006)				(0.004)				(0.029)
Observations	474,240	474,240	474,240	474,240	603,000	603,000	603,000	603,000	578,880	578,880	578,880	578,880
Adjusted R^2	0.001	0.001	0.001	0.001	0.0003	0.0004	0.0004	0.0004	0.001	0.001	0.005	0.001
F Statistic	421.758***	214.325***	213.394***	264.649***	190.320***	106.679***	129.769***	114.378***	694.239***	393.520***	1,557.407***	347.468***

Panel D:	**AUD/USD**											
	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.027***	−0.027***	−0.027***	−0.004	−0.024***	−0.023***	−0.021***	−0.012**	−0.021***	−0.021***	−0.020***	−0.011
	(0.007)	(0.007)	(0.007)	(0.009)	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)	(0.006)	(0.007)	(0.009)
AR(1)		0.001				−0.016*				−0.0004		
		(0.013)				(0.009)				(0.011)		
$ARMA_{BIC}$			−0.004				−0.041***				−0.020	
			(0.016)				(0.014)				(0.020)	
RW				−0.021***				−0.011***				−0.009
				(0.005)				(0.003)				(0.006)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.0002	0.0002	0.0002	0.0005	0.0002	0.0002	0.0003	0.0003	0.0001	0.0001	0.0002	0.0002
F Statistic	109.911***	54.977***	55.223***	116.354***	115.731***	62.302***	84.440***	77.895***	84.023***	42.014***	49.429***	55.556***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.3 reports the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for two currency pairs split by panels. Columns $(1)-(7)$ control for all hours of the days, in such a way that there are 23 dummies, one for each hour. Columns $(8)-(14)$ control for the US and Europe trading hours, in such there are two dummies, one for each case. US trading hours are restricted to be range from $9:00$ AM to $4:00$ PM (EST), while Europe trading hours range from $8:00$ to $5:00$ PM (GMT). We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns $(1)-(4)$ and $(8)-(11)$ are univariate regressions, while $(5)-(7)$ and $(12)-(14)$ are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.3: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks - Intraday Patters

Panel E: **USD/CAD**

	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.025***	−0.025***	−0.021***	−0.021**	−0.015***	−0.013**	−0.011*	−0.008	−0.022***	−0.021***	−0.023***	−0.009
	(0.007)	(0.007)	(0.007)	(0.009)	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
AR(1)		−0.002				−0.017*				−0.003		
		(0.011)				(0.009)				(0.016)		
$ARMA_{BIC}$			−0.042**				−0.045***				0.017	
			(0.017)				(0.014)				(0.019)	
RW				−0.004				−0.006				−0.011***
				(0.004)				(0.004)				(0.004)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.0002	0.0002	0.0003	0.0002	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001	0.0002	0.0002
F Statistic	103.282***	51.715***	79.840***	54.206***	44.848***	27.255***	60.209***	28.295***	79.945***	40.155***	45.281***	63.529***

Panel F: **USD/CHF**

	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.029***	−0.027***	−0.024**	−0.021*	−0.023***	−0.019***	−0.020***	−0.015**	−0.027***	−0.028***	−0.028***	−0.006
	(0.009)	(0.010)	(0.009)	(0.011)	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)	(0.008)	(0.009)	(0.011)
AR(1)		−0.017				−0.031***				0.004		
		(0.012)				(0.009)				(0.017)		
$ARMA_{BIC}$			−0.045*				−0.025*				0.011	
			(0.024)				(0.015)				(0.020)	
RW				−0.007				−0.007*				−0.018***
				(0.004)				(0.004)				(0.007)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.0003	0.0003	0.0004	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0004
F Statistic	130.747***	69.646***	97.508***	71.622***	100.105***	67.307***	61.349***	58.090***	122.737***	61.706***	64.021***	122.089***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.2 reports the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for one currency pair. Columns (1) – (7) control for all hours of the days, in such a way that there are 23 dummies, one for each hour. Columns (8) – (14) control for the US and Europe trading hours, in such there are two dummies, one for each case. US trading hours are restricted to be range from 9 : 00 AM to 4 : 00 PM (EST), while Europe trading hours range from 8 : 00 to 5 : 00 PM (GMT). We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) and (8) – (11) are univariate regressions, while (5) – (7) and (12) – (14) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.3: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks - Intraday Patters

Panel G:	**EUR/GBP**											
	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.069***	−0.072***	−0.070***	−0.019*	−0.043***	−0.042***	−0.039***	−0.012*	−0.069	−0.077	−0.072	0.013
	(0.010)	(0.010)	(0.011)	(0.011)	(0.005)	(0.005)	(0.005)	(0.007)	(0.052)	(0.056)	(0.053)	(0.042)
AR(1)		0.031**				−0.007				0.114		
		(0.015)				(0.010)				(0.112)		
$ARMA_{BIC}$			0.009				−0.038***				0.069	
			(0.020)				(0.014)				(0.082)	
RW				−0.044***				−0.027***				−0.076*
				(0.006)				(0.005)				(0.044)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.001	0.002	0.001	0.003	0.001	0.001	0.001	0.001	0.002	0.003	0.002	0.005
F Statistic	714.543***	371.909***	358.365***	634.229***	362.824***	182.203***	208.387***	314.584***	1,119.428***	786.225***	641.502***	1,470.029***
Panel H:	**EUR/JPY**											
	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.044***	−0.047***	−0.045***	−0.006	−0.032***	−0.032***	−0.029***	−0.007	−0.062***	−0.046**	−0.019	−0.105**
	(0.009)	(0.009)	(0.009)	(0.010)	(0.006)	(0.006)	(0.006)	(0.008)	(0.023)	(0.023)	(0.052)	(0.044)
AR(1)		0.041***				0.003				−0.113		
		(0.011)				(0.009)				(0.070)		
$ARMA_{BIC}$			0.017				−0.028*				−0.226**	
			(0.015)				(0.016)				(0.109)	
RW				−0.034***				−0.023***				0.039
				(0.004)				(0.004)				(0.032)
Observations	474,240	474,240	474,240	474,240	603,000	603,000	603,000	603,000	578,880	578,880	578,880	578,880
Adjusted R^2	0.001	0.001	0.001	0.001	0.0003	0.0003	0.0004	0.001	0.001	0.002	0.006	0.002
F Statistic	297.489***	173.484***	153.000***	309.476***	208.810***	104.532***	118.058***	197.005***	864.866***	639.286***	1,665.379***	670.067***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.2 reports the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for one currency pair. Columns (1) – (7) control for all hours of the days, in such a way that there are 23 dummies, one for each hour. Columns (8) – (14) control for the US and Europe trading hours, in such there are two dummies, one for each case. US trading hours are restricted to be range from 9 : 00 AM to 4 : 00 PM (EST), while Europe trading hours range from 8 : 00 to 5 : 00 PM (GMT). We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) and (8) – (11) are univariate regressions, while (5) – (7) and (12) – (14) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.3: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}^{ENet}_{t,j}$ and Benchmarks - Intraday Patters

Panel I: **NZD/USD**

	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.017*	−0.017*	−0.019**	0.005	−0.020***	−0.018***	−0.015***	−0.014**	−0.040***	−0.042***	−0.035***	−0.033**
	(0.009)	(0.010)	(0.009)	(0.013)	(0.005)	(0.005)	(0.005)	(0.007)	(0.011)	(0.011)	(0.011)	(0.014)
AR(1)		0.004				−0.024***				0.016		
		(0.019)				(0.009)				(0.015)		
$ARMA_{BIC}$			0.015				−0.051***				−0.039*	
			(0.024)				(0.012)				(0.021)	
RW				−0.019**				−0.005				−0.007
				(0.009)				(0.003)				(0.007)
Observations	474,240	474,240	474,240	474,240	603,000	603,000	603,000	603,000	578,880	578,880	578,880	578,880
Adjusted R^2	0.0001	0.0001	0.0001	0.0003	0.0001	0.0002	0.0003	0.0001	0.001	0.001	0.001	0.001
F Statistic	40.185***	20.291***	23.515***	72.527***	80.003***	50.498***	82.754***	44.527***	292.127***	150.681***	174.023***	153.520***

Panel J: **USD/SEK**

	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.041***	−0.043***	−0.041***	−0.009	−0.027***	−0.027***	−0.024***	−0.008	−0.055***	−0.059***	−0.053***	−0.007
	(0.006)	(0.006)	(0.006)	(0.008)	(0.005)	(0.005)	(0.005)	(0.006)	(0.011)	(0.010)	(0.010)	(0.011)
AR(1)		0.021**				−0.001				0.043**		
		(0.010)				(0.008)				(0.020)		
$ARMA_{BIC}$			−0.001				−0.031*				−0.015	
			(0.015)				(0.016)				(0.023)	
RW				−0.029***				−0.017***				−0.041***
				(0.004)				(0.003)				(0.007)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.001	0.001	0.001	0.001	0.0002	0.0002	0.0003	0.0004	0.001	0.001	0.001	0.002
F Statistic	251.897***	132.497***	125.960***	247.551***	140.424***	70.247***	86.867***	122.356***	489.203***	278.030***	249.929***	559.844***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.2 reports the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}^{ENet}_{t,j}$ and the three benchmarks for one currency pair. Columns (1) – (7) control for all hours of the days, in such a way that there are 23 dummies, one for each hour. Columns (8) – (14) control for the US and Europe trading hours, in such there are two dummies, one for each case. US trading hours are restricted to be range from $9:00$ AM to $4:00$ PM (EST), while Europe trading hours range from $8:00$ to $5:00$ PM (GMT). We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) and (8) – (11) are univariate regressions, while (5) – (7) and (12) – (14) are bivariate regressions with $\widehat{\boldsymbol{f}}^{ENet}_{t,j}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.4.3: **(Continued)** Out-of-Sample Regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and Benchmarks - Intraday Patters

Panel I:	**EUR/CHF**											
	US hours				Europe hours				US & Europe hours			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
f^{ENet}	−0.054***	−0.056***	−0.055***	−0.006	−0.039***	−0.040***	−0.036***	−0.005	−0.103***	−0.110***	−0.113***	−0.056**
	(0.007)	(0.007)	(0.007)	(0.008)	(0.005)	(0.005)	(0.005)	(0.006)	(0.024)	(0.025)	(0.025)	(0.028)
AR(1)		0.027**				0.006				0.101**		
		(0.013)				(0.009)				(0.043)		
$ARMA_{BIC}$			0.011				−0.032**				0.115***	
			(0.020)				(0.014)				(0.031)	
RW				−0.042***				−0.030***				−0.042***
				(0.004)				(0.003)				(0.008)
Observations	476,640	476,640	476,640	476,640	606,000	606,000	606,000	606,000	581,760	581,760	581,760	581,760
Adjusted R^2	0.001	0.001	0.001	0.002	0.0005	0.0005	0.001	0.001	0.003	0.004	0.004	0.004
F Statistic	410.384***	215.793***	206.968***	474.517***	285.298***	143.375***	160.523***	315.523***	1,700.886***	1,041.788***	1,111.620***	1,179.624***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3.4.2 reports the out-of-sample predictive regression accounting for the intraday patterns for $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and the three benchmarks for one currency pair. Columns (1) – (7) control for all hours of the days, in such a way that there are 23 dummies, one for each hour. Columns (8) – (14) control for the US and Europe trading hours, in such there are two dummies, one for each case. US trading hours are restricted to be range from 9 : 00 AM to 4 : 00 PM (EST), while Europe trading hours range from 8 : 00 to 5 : 00 PM (GMT). We consider three possible benchmarks: a random walk (RW), an AR(1), and an ARMA(p,q) with the number of legs chosen by BIC. Columns (1) – (4) and (8) – (11) are univariate regressions, while (5) – (7) and (12) – (14) are bivariate regressions with $\widehat{\boldsymbol{f}}_{t,j}^{ENet}$ and a benchmark. We use the Newey-West robust standard errors. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

3.4.3 Behavior of the Penalization

Figure 3.4.1: Mean Alpha and Lambda per Hour (2016-2019)

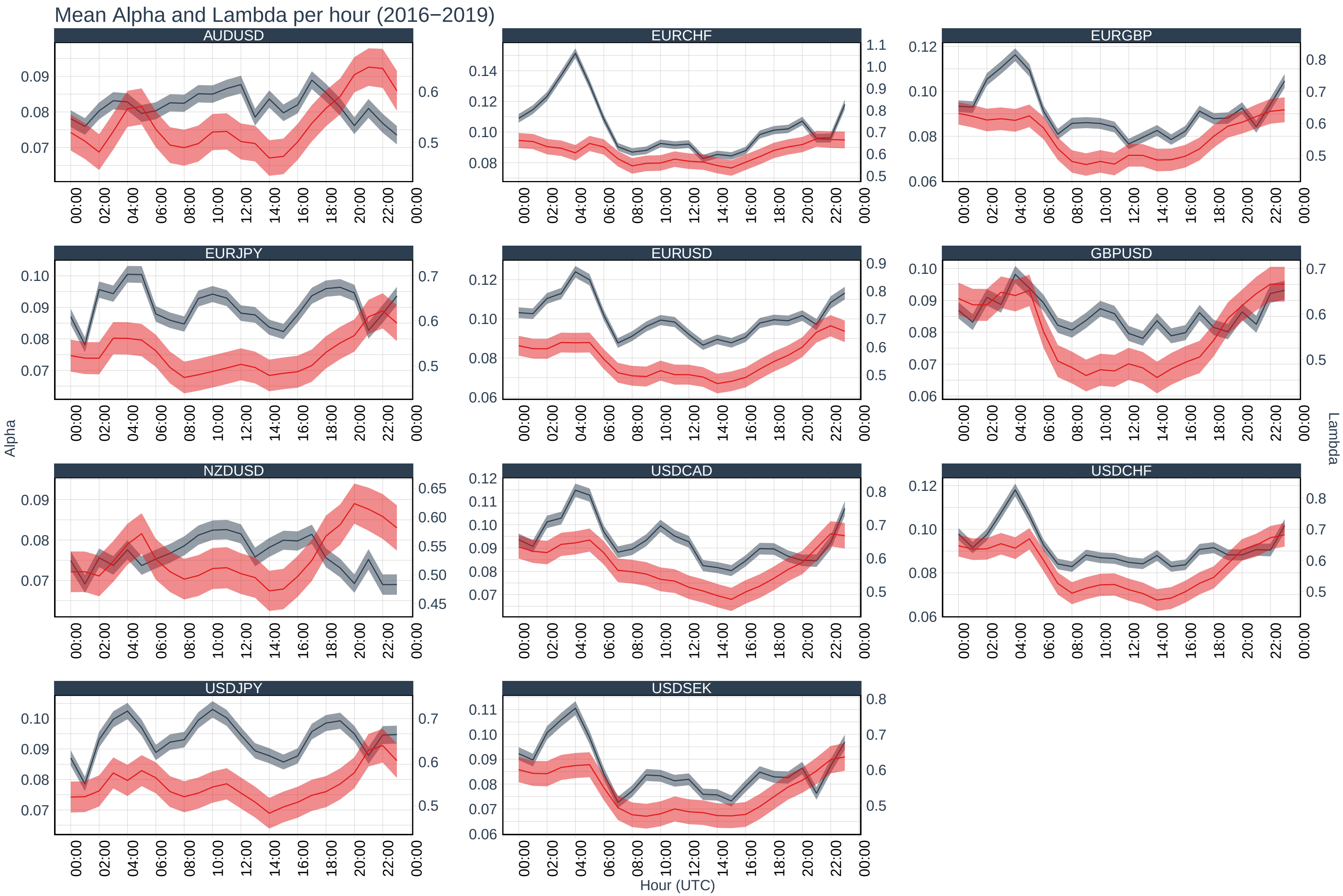

Figure 3.4.1 shows the estimated mean per hour of the tuned hyperparameters: α and λ. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), while the left y-axis the the hourly average of α and the right y-axis the hourly average of λ. The grey solid line represents α, while the the red solid line shows λ. Shaded areas for both parameter represent the 99% confidence interval. The process of tuning the hyperparameters α and λ is done through cross-validation in each time window in the rolling regression process.

Figure 3.4.2: Mean Minutes with Informative Signals per Hour (2016-2019)

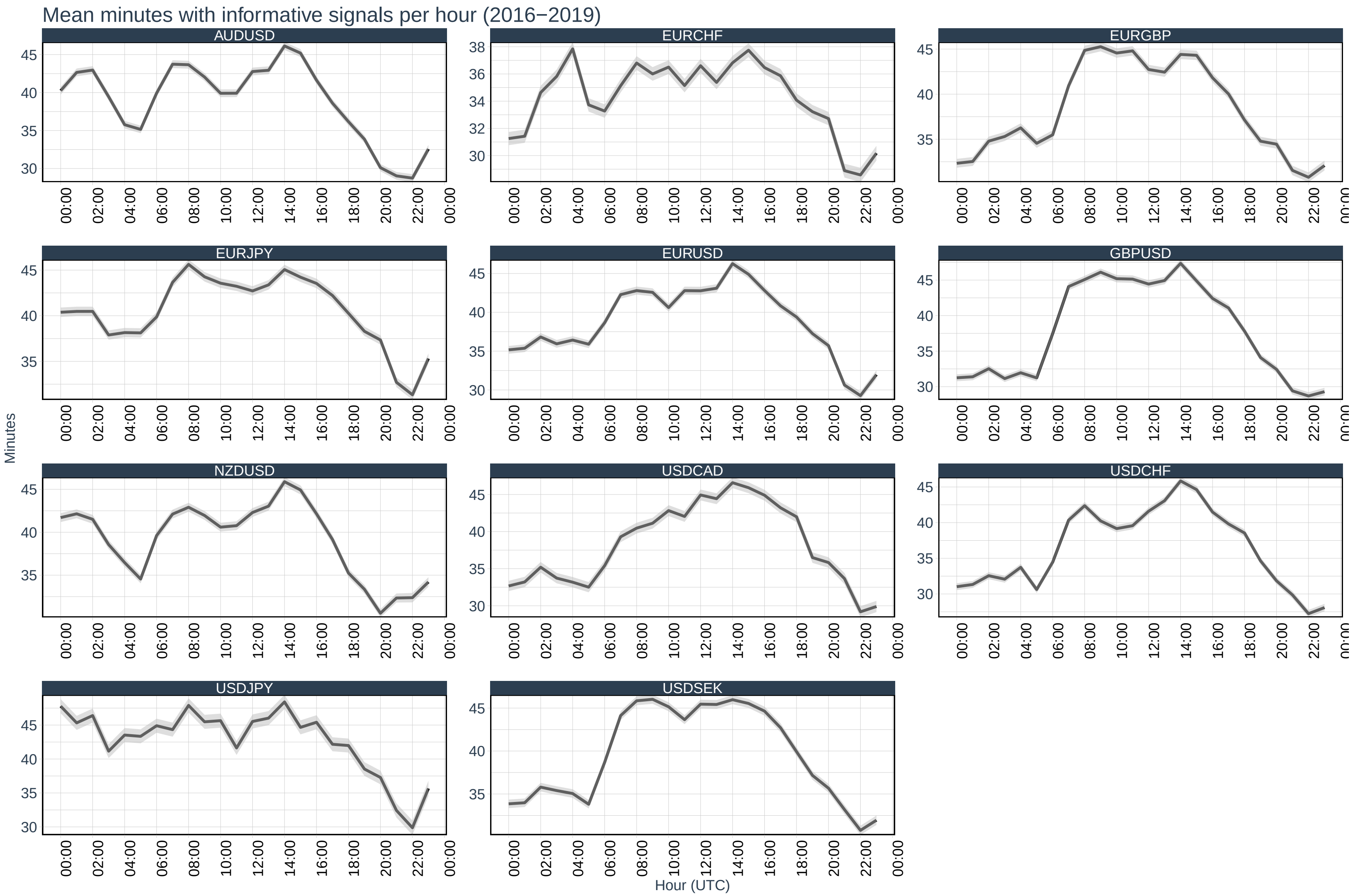

Figure 3.4.2 shows the estimated proportion of minutes with informative signals per hour. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), while in the y-axis we have minutes. The grey solid line is the estimated mean of minutes, while the shaded area represent the 99% confidence interval. For each minute t in the rolling regression process, we consider a minute to have informative signal if the shrinkage and selection estimator ENet yielded at least a coefficient for the high-dimensional set of signals to be not equal to 0.

Figure 3.4.3: Mean Duration of Signal Blocks in Minutes per Hour (2016-2019)

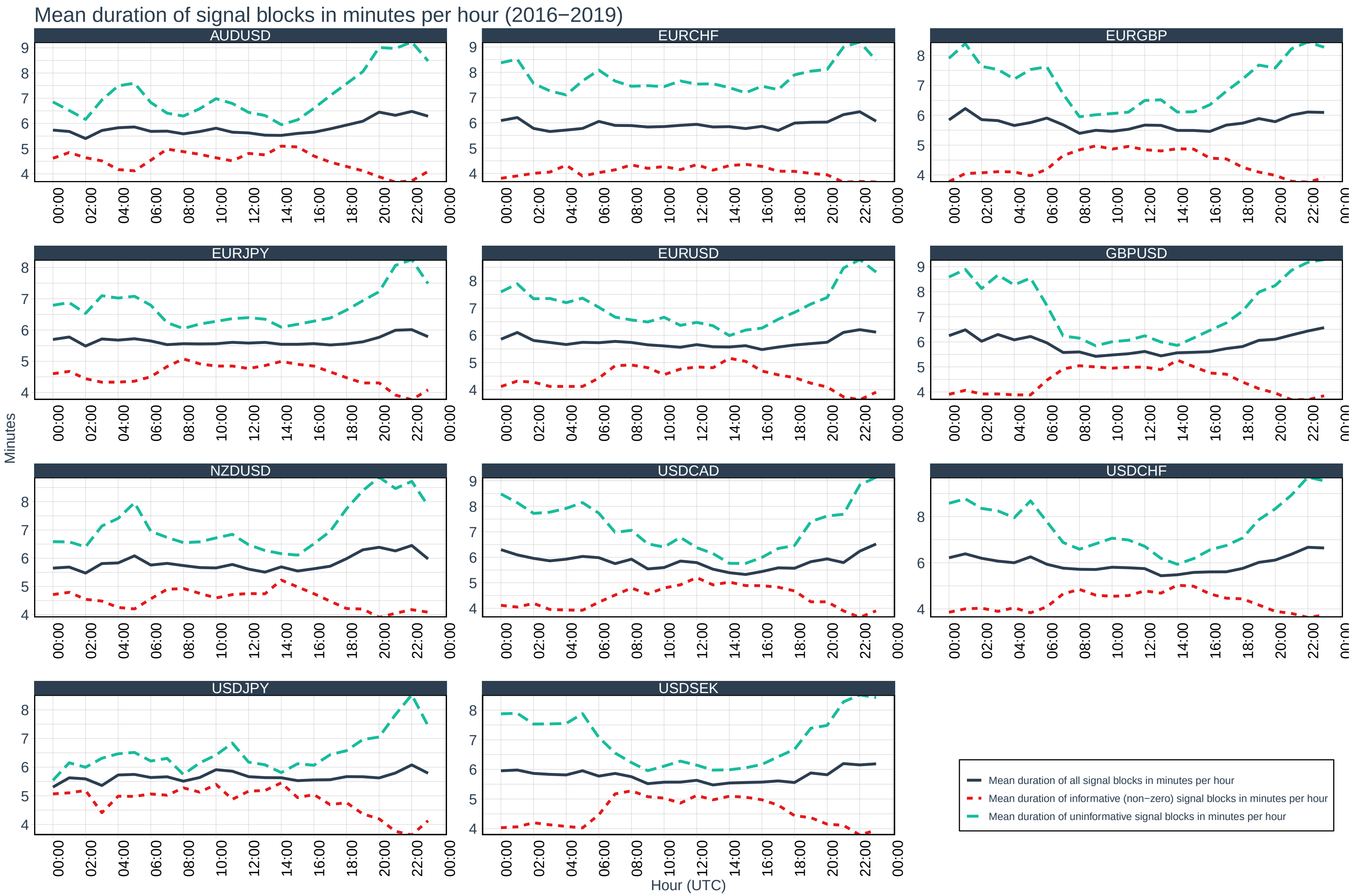

Figure 3.4.3 shows the estimated proportion of minutes with informative signals per hour. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), while in the y-axis we have minutes. The grey solid line is the estimated mean of minutes, while the shaded area represent the 99% confidence interval. For each minute t in the rolling regression process, we consider a minute to have informative signal if the shrinkage and selection estimator ENet yielded at least a coefficient for the high-dimensional set of signals to be not equal to 0.

Figure 3.4.4: Mean Duration of Signal Blocks in Minutes per Hour (2016-2019)

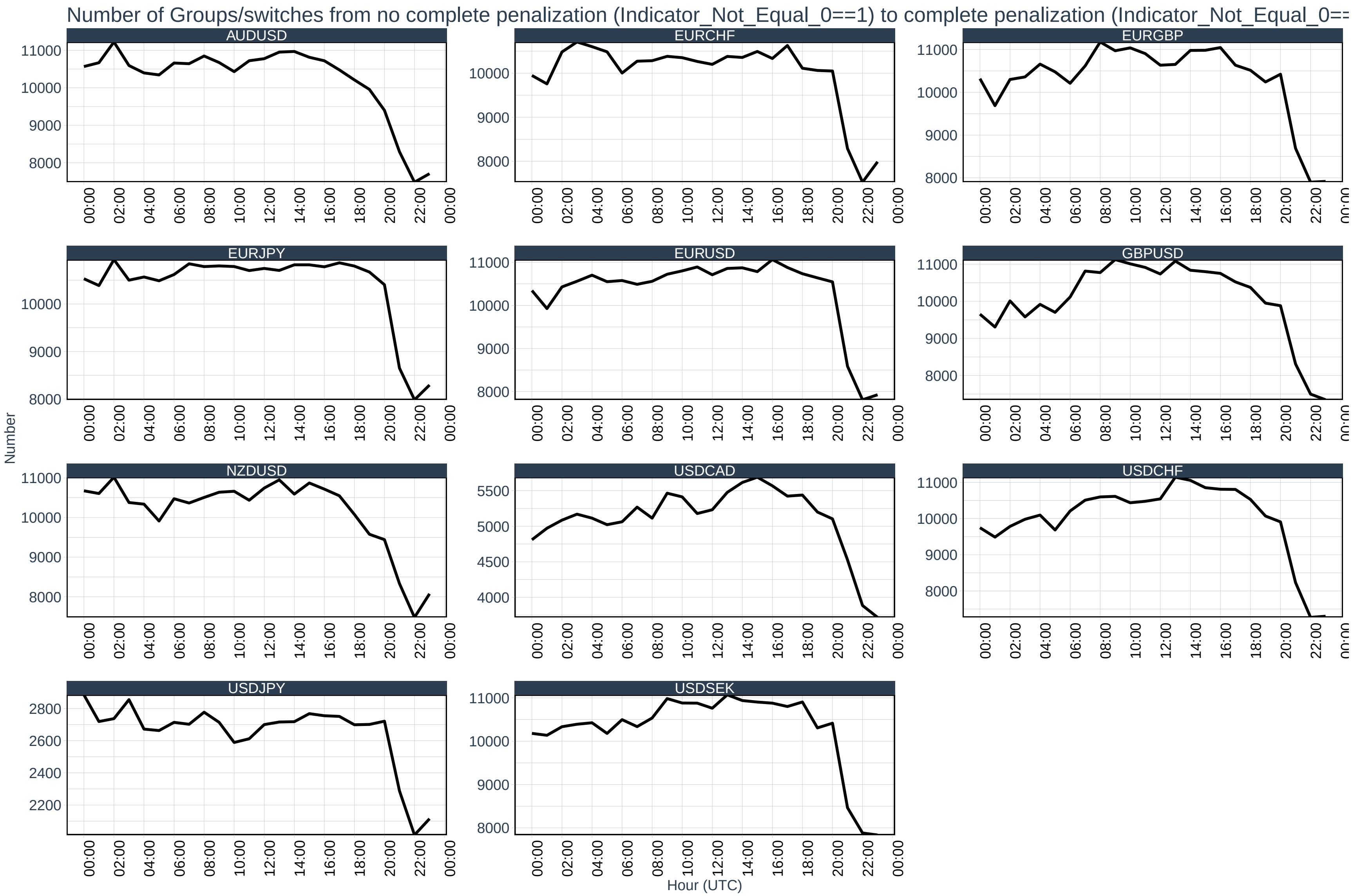

Figure 3.4.4 shows the estimated proportion of minutes with informative signals per hour. The sample period spans the 1-minute data over the whole 24-hour trading period from 2016 to 2019. The x-axis of each panel represent the hours (UTC), while in the y-axis we have minutes. The grey solid line is the estimated mean of minutes, while the shaded area represent the 99% confidence interval. For each minute t in the rolling regression process, we consider a minute to have informative signal if the shrinkage and selection estimator ENet yielded at least a coefficient for the high-dimensional set of signals to be not equal to 0.

3.5 Conclusion

We plan to extend our analysis to other currency pairs. The following currencies are in line to be estimted: Australian Dollar (AUD), Swedish Krona (SEK) and Norwegian Krone (NOK). We believe that, as we move to less traded currencies, the amount of signal extracted by the Elastic-Net using a large set of cross-sectional data will be more relevant. The intuition lies in the fact that the most liquid currencies are affected by economic factors or external shocks first, and these impacts rollover to other less liquid assets.

Many recent studies showed many interesting intraday patterns. We seek to evaluate how Elastic-Net 1-minute ahead behave extracting signals when accounting for these patterns. Additionally, we seek to study the economic origins of these possible sources of short-lived and unexpected signals.

3.6 Appendix

Table 3.6.1: List of Currencies Pairs

Type	Code	Description	Mean $\times 10^6$ (%)	Mean (annualized %)	sd $\times 10^4$	1st-Order Autocorrelation
Currency	AUDCAD	Australian Dollar to Canadian Dollar	5.64	2.57	2.4	-0.017
Currency	AUDCHF	Australian Dollar to Swiss Franc	-11.60	-5.08	3.1	-0.013
Currency	AUDJPY	Australian Dollar to Japanese Yen	-18.60	-8.02	4.4	-0.005
Currency	AUDNZD	Australian Dollar to New Zealand Dollar	-7.08	-3.13	1.7	-0.027
Currency	AUDSGD	Australian Dollar to Singapore Dollar	-4.76	-2.12	3.2	-0.013
Currency	AUDUSD	Australian Dollar to United States Dollar	-1.29	-0.58	4.7	-0.002
Currency	CADCHF	Canadian Dollar to Swiss Franc	-17.30	-7.48	2.1	-0.019
Currency	CADHKD	Canadian Dollar to Hong Kong Dollar	-6.40	-2.83	3.1	-0.006
Currency	CADJPY	Canadian Dollar to Japanese Yen	-24.30	-10.34	3.4	-0.009
Currency	CHFJPY	Swiss Franc to Japanese Yen	-6.97	-3.08	2.2	-0.019
Currency	CHFSGD	Swiss Franc to Singapore Dollar	6.95	3.17	1.5	-0.038
Currency	EURAUD	Euro to Australian Dollar	8.15	3.73	2.2	-0.018
Currency	EURCAD	Euro to Canadian Dollar	13.80	6.40	1.8	-0.026
Currency	EURCHF	Euro to Swiss Franc	-3.51	-1.56	1.5	-0.019
Currency	EURCZK	Euro to Czech Koruna	-1.93	-0.86	1.1	-0.035
Currency	EURDKK	Euro to Danish Krone	0.12	0.05	0.1	-0.078
Currency	EURGBP	Euro to Pound Sterling	1.06	0.48	1.4	-0.044
Currency	EURHKD	Euro to Hong Kong Dollar	7.41	3.39	3.2	-0.002
Currency	EURHUF	Euro to Hungarian Forint	2.59	1.17	2.2	-0.002
Currency	EURJPY	Euro to Japanese Yen	-10.50	-4.61	2.9	-0.010
Currency	EURNOK	Euro to Norwegian Krone	-6.92	-3.06	1.5	-0.051
Currency	EURNZD	Euro to New Zealand Dollar	1.09	0.49	1.9	-0.025
Currency	EURPLN	Euro to Polish Zloty	0.65	0.29	1.8	-0.020
Currency	EURRUB	Euro to Russian Rubble	3.36	1.52	4.7	-0.003
Currency	EURSEK	Euro to Swedish Krona	9.08	4.16	2.8	-0.016
Currency	EURSGD	Euro to Singapore Dollar	3.35	1.52	1.8	-0.021
Currency	EURTRY	Euro to Turkish Lira	7.70	3.52	14.2	0.001
Currency	EURUSD	Euro to United States Dollar	6.85	3.13	3.3	-0.003
Currency	GBPAUD	Pound Sterling to Australian Dollar	7.03	3.21	2.1	-0.026
Currency	GBPCAD	Pound Sterling to Canadian Dollar	12.70	5.87	2.2	-0.022
Currency	GBPCHF	Pound Sterling to Swiss Franc	-4.63	-2.06	2.3	-0.017
Currency	GBPJPY	Pound Sterling to Japanese Yen	-11.60	-5.08	3.5	-0.008
Currency	GBPNZD	Pound Sterling to New Zealand	-0.03	-0.01	2.1	-0.025
Currency	GBPUSD	Pound Sterling to United States Dollar	5.74	2.61	3.9	-0.004
Currency	HKDJPY	Hong Kong Dollar to Japanese Yen	-17.90	-7.73	2.1	-0.007
Currency	NZDCAD	New Zealand Dollar to Canadian Dollar	12.70	5.87	2.2	-0.019
Currency	NZDCHF	New Zealand Dollar to Swiss Franc	-4.57	-2.03	2.4	-0.016
Currency	NZDJPY	New Zealand Dollar to Japanese Yen	-11.50	-5.04	3.5	-0.009
Currency	NZDUSD	New Zealand Dollar to United States Dollar	5.80	2.64	3.8	-0.003
Currency	SGDJPY	Singapore Dollar to Japanese Yen	-13.80	-6.01	2.0	-0.014
Currency	TRYJPY	Turkish Lira to Japanese Yen	-18.30	-7.89	16.2	-0.005
Currency	USDCAD	United States Dollar to Canadian Dollar	6.93	3.16	3.1	-0.004
Currency	USDCHF	United States Dollar to Swiss Franc	-10.40	-4.57	2.5	-0.006
Currency	USDCNH	United States Dollar to Offshore Chinese Renminbi	-7.83	-3.46	3.5	0.001
Currency	USDCZK	United States Dollar to Czech Koruna	-8.81	-3.88	4.1	-0.003
Currency	USDDKK	United States Dollar to Danish Krone	-6.69	-2.96	3.4	-0.007
Currency	USDHKD	United States Dollar to Hong Kong Dollar	0.56	0.25	0.2	-0.002
Currency	USDHUF	United States Dollar to Hungarian Forint	-4.11	-1.83	5.2	0.002
Currency	USDILS	United States Dollar to Israeli Shekel	-3.09	-1.38	3.6	-0.013
Currency	USDJPY	United States Dollar to Japanese Yen	-17.30	-7.48	2.1	-0.006
Currency	USDMXN	United States Dollar to Mexican Peso	-11.90	-5.21	3.2	-0.001
Currency	USDNOK	United States Dollar to Norwegian Krone	-13.70	-5.97	3.8	-0.011
Currency	USDPLN	United States to Polish Zloty	-6.13	-2.72	4.8	-0.001
Currency	USDRON	United States to Romanian Leu	-6.97	-3.08	3.3	-0.004
Currency	USDRUB	United States Dollar to Russian Rubble	-3.76	-1.68	7.3	0.000
Currency	USDSEK	United States to Swedish Krona	2.05	0.93	5.6	-0.005
Currency	USDSGD	United States Dollar to Singapore Dollar	-3.49	-1.56	1.8	0.000
Currency	USDTHB	United States Dollar to Thai Baht	-9.45	-4.16	1.9	-0.003
Currency	USDTRY	United States Dollar to Turkish Lira	0.81	0.36	16.7	-0.002
Currency	USDZAR	United States Dollar to South African Rand	-10.40	-4.57	7.4	-0.004
Currency	ZARJPY	South African Rand to Japanese Yen	-6.78	-3.00	6.7	-0.010

Table 3.6.1 reports the full list of currencies pairs used, along with some descriptive statistics. For all of them, we use the 1-minute log difference of their original values. The sample is the 1-minute data for the year of 2018, from Monday through Saturday.

Table 3.6.2: List of Commodities, Stock Indexes and Metals

Type	Code	Description	Mean $\times 10^6$ (%)	Mean (annualized %)	sd $\times 10^4$ (%)	1st-Order Autocorrelation
Commodity	BRENTCMDUSD	US Brent Crude Oil	-8.55	-3.77	8.5	-0.003
Commodity	COCOACMDUSD	NY Cocoa	56.40	28.84	12.0	-0.002
Commodity	COFFEECMDUSX	Coffee Arabica	-14.40	-6.26	10.5	-0.026
Commodity	COPPERCMDUSD	High Grade Copper	-12.90	-5.63	7.4	-0.015
Commodity	COTTONCMDUSX	Cotton	12.70	5.87	8.9	-0.148
Commodity	DIESELCMDUSD	Gas oil	-7.40	-3.27	7.6	-0.013
Commodity	GASCMDUSD	Natural Gas	-34.60	-14.40	12.6	0.002
Commodity	LIGHTCMDUSD	US Light Crude Oil	5.68	2.58	10.2	-0.002
Commodity	OJUICECMDUSX	Orange Juice	-0.15	-0.07	8.4	-0.047
Commodity	SOYBEANCMDUSX	Soybean	31.40	15.15	6.2	0.004
Commodity	SUGARCMDUSD	Sugar	-24.00	-10.22	19.8	-0.208
Index	AUSIDXAUD	Australian 200 Index	-6.23	-2.76	3.3	-0.034
Index	CHEIDXCHF	Switzerland 20 Index	-20.30	-8.72	4.3	-0.009
Index	CHIIDXUSD	China A50 Index	2.30	1.04	12.3	-0.004
Index	DEUIDXEUR	Germany 30 Index	-18.50	-7.98	6.9	-0.002
Index	DOLLARIDXUSD	US Dollar Index	-6.16	-2.73	3.4	-0.017
Index	ESPIDXEUR	Spain 35 Index	-13.30	-5.80	6.6	-0.009
Index	EUSIDXEUR	Europe 50 Index	-11.20	-4.91	5.8	-0.007
Index	FRAIDXEUR	French 40 Index	-6.16	-2.73	4.6	-0.005
Index	GBRIDXGBP	UK 100 Index	-19.10	-8.22	4.0	-0.005
Index	HKGIDXHKD	Hong Kong 40 Index	2.41	1.09	8.7	0.000
Index	INDIDXUSD	India 50 Index	0.59	0.26	2.9	-0.017
Index	JPNIDXJPY	Japan 225 Index	-19.20	-8.26	5.2	-0.017
Index	NLDIDXEUR	Netherlands 25 Index	-13.20	-5.76	4.4	-0.021
Index	PLNIDXPLN	Poland 20 Index	-16.90	-7.31	7.2	-0.022
Index	SGDIDXSGD	Singapore Blue Chip Cash Index	5.26	2.39	6.3	-0.038
Index	USA30IDXUSD	Dow Jones Industrial Index	-2.48	-1.11	3.9	-0.006
Index	USA500IDXUSD	Standard and Poor 500 Index	1.40	0.63	3.7	-0.003
Index	USATECHIDXUSD	Nasdaq Index	16.70	7.79	5.0	-0.004
Metal	XAGUSD	Spot Silver	-7.98	-3.52	6.6	-0.012
Metal	XAUUSD	Spot Gold	3.20	1.45	4.1	-0.003

Table 3.6.2 reports the list of commodities, stock indexes and metals used, along with some descriptive statistics. For all of them, we use the 1-minute log difference of their original values. The sample is the 1-minute data for the year of 2018.

Bibliography

Abhyankar, A., Basu, D., and Stremme, A. (2012). The optimal use of return predictability: An empirical study. *Journal of Financial and Quantitative Analysis*, 47(5):973–1001.

Anatolyev, S. (2005). GMM, GEL, serial correlation, and asymptotic bias. *Econometrica*, 73(3):983–1002.

Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda*, pages 507–547. University of Chicago Press.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11:685–725.

Bai, J. and Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, 74(4):1133–1150.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Bauer, M. D. and Hamilton, J. D. (2018). Robust bond risk premia. *The Review of Financial Studies*, 31(2):399–448.

Bauer, M. D. and Rudebusch, G. D. (2017). Resolving the spanning puzzle in macro-finance term structure models. *Review of Finance*, 21(2):511–553.

Bergmeir, C., Hyndman, R. J., and Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120:70–83.

Bianchi, D., Büchner, M., and Tamoni, A. (2019). Bond risk premia with machine learning. *USC-INET Research Paper*, (19-11).

Brandt, M. W. (2010). Portfolio choice problems. In *Handbook of financial econometrics: Tools and techniques*, pages 269–336. Elsevier.

Brandt, M. W. and Santa-Clara, P. (2006). Dynamic portfolio selection by augmenting the asset space. *The Journal of Finance*, 61(5):2187–2217.

Breedon, F. and Ranaldo, A. (2013). Intraday patterns in fx returns and order flow. *Journal of Money, Credit and Banking*, 45(5):953–965.

Cai, F., Howorka, E., and Wongswan, J. (2008). Informational linkages across trading regions: Evidence from foreign exchange markets. *Journal of International Money and Finance*, 27(8):1215–1243.

Campbell, J. Y. and Shiller, R. J. (1991). Yield spreads and interest rate movements: A bird's eye view. *The Review of Economic Studies*, 58(3):495–514.

Campbell, J. Y. and Thompson, S. B. (2007). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Chen, L., Pelger, M., and Zhu, J. (2019). Deep learning in asset pricing. *Available at SSRN 3350138*.

Chiang, I.-H. E. (2015). Modern portfolio management with conditioning information. *Journal of Empirical Finance*, 33:114–134.

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

Cieslak, A. and Povala, P. (2015). Expected returns in treasury bonds. *The Review of Financial Studies*, 28(10):2859–2901.

Cochrane, J. H. (2009). *Asset Pricing*. Princeton University Press.

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

Cochrane, J. H. (2015). Comments on "robust bond risk premia" by michael bauer and jim hamilton. *Unpublished working paper. University of Chicago*.

Cochrane, J. H. and Piazzesi, M. (2005). Bond risk premia. *American Economic Review*, 95(1):138–160.

Cooper, I. and Maio, P. (2019). New evidence on conditional factor models. *Journal of Financial and Quantitative Analysis*, 54(5):1975–2016.

Cooper, I. and Priestley, R. (2009). Time-varying risk premiums and the output gap. *The Review of Financial Studies*, 22(7):2801–2833.

Coqueret, G. (2015). Diversified minimum-variance portfolios. *Annals of Finance*, 11(2):221–241.

Cornett, M. M., Schwarz, T. V., and Szakmary, A. C. (1995). Seasonalities and intraday return patterns in the foreign currency futures market. *Journal of Banking & Finance*, 19(5):843–869.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314.

Dai, Z. and Wang, F. (2019). Sparse and robust mean–variance portfolio optimization problems. *Physica A: Statistical Mechanics and its Applications*, 523:1371–1378.

DeMiguel, V., Garlappi, L., Nogales, F. J., and Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management science*, 55(5):798–812.

DeMiguel, V., Gil-Bazo, J., Nogales, F. J., and AP Santos, A. (2021). Can machine learning help to select portfolios of mutual funds? *Can Machine Learning Help to Select Portfolios of Mutual Funds*.

DeMiguel, V., Martín-Utrera, A., Nogales, F. J., and Uppal, R. (2020). A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *The Review of Financial Studies*, 33(5):2180–2222.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–263.

Dubiel-Teleszynski, T., Kalogeropoulos, K., and Karouzakis, N. (2019). Predicting bond risk premia via sequential learning.

Duffee, G. (2013). Forecasting interest rates. In *Handbook of economic forecasting*, volume 2, pages 385–426. Elsevier.

Fama, E. F. and Bliss, R. R. (1987). The information in long-maturity forward rates. *The American Economic Review*, pages 680–692.

Faria, G. and Verona, F. (2017). Forecasting the equity risk premium with frequency-decomposed predictors. *Bank of Finland Research Discussion Paper*, (1).

Farrell, M. H., Liang, T., and Misra, S. (2021). Deep neural networks for estimation and inference. *Econometrica*, 89(1):181–213.

Feng, G., Giglio, S., and Xiu, D. (2017). Taming the factor zoo. *Fama-Miller Working Paper*, 24070.

Feng, G., He, J., and Polson, N. G. (2018a). Deep learning for predicting asset returns. *arXiv preprint arXiv:1804.09314*.

Feng, G., Polson, N. G., and Xu, J. (2018b). Deep learning in characteristics-sorted factor models. *arXiv preprint arXiv:1805.01104*.

Ferson, W. E. and Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *The Journal of Finance*, 56(3):967–982.

Ferson, W. E. and Siegel, A. F. (2009). Testing portfolio efficiency with conditioning information. *Review of Financial Studies*, 22(7):2735–2758.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56(1):329–352.

Fletcher, J. and Basu, D. (2016). An examination of the benefits of dynamic trading strategies in uk closed-end funds. *International Review of Financial Analysis*, 47:109–118.

Frank, L. E. and Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, 35(2):109–135.

Freyberger, J., Neuhierl, A., and Weber, M. (2017). Dissecting characteristics nonparametrically. Technical report, National Bureau of Economic Research.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics New York.

Gargano, A., Pettenuzzo, D., and Timmermann, A. (2019). Bond return predictability: Economic value and links to the macroeconomy. *Management Science*, 65(2):508–540.

Ghysels, E., Horan, C., and Moench, E. (2018). Forecasting through the rearview mirror: Data revisions and bond return predictability. *The Review of Financial Studies*, 31(2):678–714.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

Goto, S. and Xu, Y. (2015). Improving mean variance optimization through sparse hedging restrictions. *Journal of Financial and Quantitative Analysis*, 50(6):1415–1441.

Goyal, A. (2012). Empirical cross-sectional asset pricing: a survey. *Financial Markets and Portfolio Management*, 26(1):3–38.

Green, J., Hand, J. R., and Zhang, X. F. (2013). The supraview of return predictive signals. *Review of Accounting Studies*, 18(3):692–730.

Gu, S., Kelly, B., and Xiu, D. (2018). Empirical asset pricing via machine learning. Technical report, National Bureau of Economic Research.

Gu, S., Kelly, B., and Xiu, D. (2020). Autoencoder asset pricing models. *Journal of Econometrics*.

Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The us treasury yield curve: 1961 to the present. *Journal of monetary Economics*, 54(8):2291–2304.

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

Hansen, L. P. and Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, pages 587–613.

Hansen, P. R. and Timmermann, A. (2012). Choice of sample split in out-of-sample forecast evaluation.

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

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2):281–291.

Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and generalizations*. Chapman and Hall/CRC.

Heaton, J., Polson, N. G., and Witte, J. (2016). Deep portfolio theory. *arXiv preprint arXiv:1605.07230*.

Heaton, J. B., Polson, N. G., and Witte, J. H. (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, 33(1):3–12.

Hornik, K., Stinchcombe, M., White, H., et al. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366.

Huang, J.-z. and Shi, Z. (2011). Determinants of bond risk premia. Technical report, Citeseer.

Huang, J.-Z. and Shi, Z. (2019). Determinants of bond risk premia: A machine-learning-based resolution of the spanning controversy. Working Paper.

Hyndman, R. J. et al. (2006). Another look at forecast-accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting*, 4(4):43–46.

Joslin, S., Priebsch, M., and Singleton, K. J. (2014). Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance*, 69(3):1197–1233.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Kelly, B. and Pruitt, S. (2013). Market expectations in the cross-section of present values. *The Journal of Finance*, 68(5):1721–1756.

Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, 186(2):294–316.

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

Kirby, C. and Ostdiek, B. (2012). It's all in the timing: simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis*, 47(2):437–467.

Kozak, S., Nagel, S., and Santosh, S. (2019). Shrinking the cross-section. *Journal of Financial Economics*.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

Lee, J. (2018). Risk premium information from treasury-bill yields. *Journal of Financial and Quantitative Analysis*, 53(1):437–454.

Light, N., Maslov, D., and Rytchkov, O. (2017). Aggregation of information about the cross section of stock returns: A latent variable approach. *The Review of Financial Studies*, 30(4):1339–1381.

Ludvigson, S. C. and Ng, S. (2009). Macro factors in bond risk premia. *The Review of Financial Studies*, 22(12):5027–5067.

McCracken, M. W. and Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

McLean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32.

Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Newey, W. K. and Smith, R. J. (2004). Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica*, 72(1):219–255.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Peñaranda, F. and Sentana, E. (2016). Duality in mean-variance frontiers with conditioning information. *Journal of Empirical Finance*, 38:762–785. Recent developments in financial econometrics and empirical finance.

Peñaranda, F. (2016). Understanding portfolio efficiency with conditioning information. *Journal of Financial and Quantitative Analysis*, 51(3):985–1011.

Pettenuzzo, D. and Ravazzolo, F. (2016). Optimal portfolio choice under decision-based model combinations. *Journal of Applied Econometrics*, 31(7):1312–1332.

Politis, D. N. and Romano, J. P. (1992). A circular block-resampling procedure for stationary data. *Exploring the limits of bootstrap*, 2635270.

Püttmann, L. (2018). Patterns of panic: Financial crisis language in historical newspapers.

Rossi, B. and Inoue, A. (2012). Out-of-sample forecast tests robust to the choice of window size. *Journal of Business & Economic Statistics*, 30(3):432–453.

Stock, J. H. and Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business & Economic Statistics*, 14(1):11–30.

Stone, M. and Brooks, R. J. (1990). Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(2):237–258.

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

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2):3–28.

Welch, I. and Goyal, A. (2007). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4):1455–1508.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2):301–320.