A Dynamic Semiparametric Characteristics-based Model for Optimal Portfolio Selection

Gregory Connor*[1], Shaoran Li†[2], and Oliver Linton‡[3]

[1]*Maynooth University, Ireland*
[2]*University of Cambridge, UK*
[3]*University of Cambridge, UK*

February 10, 2021

Abstract

This paper develops a two-step semiparametric methodology for portfolio weight selection for characteristics-based factor-tilt and factor-timing investment strategies. We build upon the expected utility maximization framework of Brandt (1999) and Aït-sahalia and Brandt (2001). We assume that asset returns obey a characteristics-based factor model with time-varying factor risk premia as in Li and Linton (2020). We prove under our return-generating assumptions that in a market with a large number of assets, an approximately optimal portfolio can be established using a two-step procedure. The first step finds optimal factor-mimicking sub-portfolios using a quadratic objective function over linear combinations of characteristics-based factor loadings. The second step dynamically combines these factor-mimicking sub-portfolios based on a time-varying signal, using the investor's expected utility as the objective function. We develop and implement a two-stage semiparametric estimator. We apply it to CRSP (Center for Research in Security Prices) and FRED (Federal Reserve Economic Data) data and find excellent in-sample and out-sample performance consistent with investors' risk aversion levels.

KEYWORDS: Portfolio management; Single index; GMM;

JEL CLASSIFICATION: C14; G11.

*Electronic address: gregory.connor@mu.ie
†Electronic address: sl736@cam.ac.uk
‡Electronic address: obl20@cam.ac.uk

1 Introduction

The traditional portfolio choice model proceeds by estimating the parameters of an asset return distribution and then finding the portfolio that maximizes expected payoffs for a given risk level, such as the optimal mean-variance portfolio choice model proposed by Markowitz et al. (1952). This approach can produce biases in portfolio weights since the portfolio selection process ignores the estimation error in the empirically-derived return distribution parameters. Furthermore, as the number of assets increases, the estimation of the high-dimensional covariance matrix becomes intractable. Some notable methods have been proposed to solve this issue, such as linear (Ledoit and Wolf (2004)) and nonlinear shrinkage (Ledoit and Wolf (2017)) of the target covariance matrix or selecting main elements by threshold (Fan et al. (2013)). However, these approaches may cause information loss and lead to unsatisfactory results, as illustrate by Ao et al. (2019). At the same time, Ao et al. (2019) studied a method called MAXSER, which is a sparse regression which sets the optimal sharpe ratio as the regressand. Their method also requires a sparsity assumption and can be problematic when the number of assets n is large. Meanwhile, all aforementioned papers ignore the importance of predictive variables, which have been documented by many researchers, such as Fama and French (1989), who analyzed the forecasting ability of dividend yield, default spread and term spread on asset returns, as well as Keim and Stambaugh (1986), Campbell and Shiller (1988), and Hodrick (1992) among others. The goal of this paper is to construct a two-step optimal portfolio that takes advantage of both a large number of assets and dynamic predictors.

Brandt (1999) used nonparametric tools to directly estimate the portfolio weights that maximize expected utility of the observed data, without first estimating the return distribution. He estimated the dynamic portfolio weights of the assets in a two-asset model as a nonparametric function of the univariate time-series predictor of the future excess returns of the risky assets. Aït-sahalia and Brandt (2001) replaced the univariate time series predictor with an index-based set of predictors: the time-varying portfolio weights in a three-asset model were assumed to be a nonlinear function of a linear fix combination of a vector of predictive variables. However, the number of assets included in their portfolio was quite limited.

Brandt et al. (2009) developed a characteristic-based model for portfolio selection with a large cross-section of assets. They assumed that optimal portfolio weights were linearly related to a small set of observable characteristics, such as book-to-market ratio, momentum and market equity. They found the linear coefficients that maximized expected utility under this assumption.

In this paper, we develop a new semiparametric model of portfolio selection, which combines the advantages of a large cross-section of assets and dynamic predictive variables. This is achieved by a characteristics-based asset pricing model. We generalize the methodologies in the above-mentioned papers since we do not impose the assumption that optimal weights are linear in the characteristics. Furthermore, the firm-specific characteristics included in our model can be significantly broadened. There are 33 characteristics in our empirical study, which provides more potential abnormal return opportunities. Also, as in Aït-sahalia and Brandt (2001), we also allow information-based dynamically-varying portfolio allocation based on a single-index function of predictors. We replace weighting across asset classes in Aït-sahalia and Brandt (2001) with weighting across our optimally-constructed characteristics-based sub-portfolios.

We estimate the model using a new, two-stage semiparametric procedure. The first step involves the estimation of the factor-mimicking sub-portfolios. This is a high-dimensional estimation problem since the number of assets is diverging, but the objective function is quadratic, allowing us to solve it using semiparametric techniques. That step compacts those assets into several sub-portfolios rather than discarding some of them, and also helps to reduce the dimensionality, which simplifies the next step. The second step maximizes the dynamic expected utility of a risk-free asset and those sub-portfolios conditional on a set of predictors, similar to Aït-sahalia and Brandt (2001). Our two-step statistical methodology accounts fully for the estimation error in both semiparametric steps, and we show that it approximates the intractable single-stage, asset-by-asset portfolio weight estimation problem in a well-defined sense.

Our model is not entirely general: we do not allow individual asset selection in response to asset-specific valuation information. We essentially allow for factor-tilt strategies, which means weighting securities according to their factor exposure in response to the associated factor risk premia, and factor timing, which means dynamically varying factor-tilt strategies, accounting for predictability in factor risk premia, but not individual asset selection. This method keeps most of the information contained in individual assets, while benefitting greatly from dimensionality reduction.

We base our model on a dynamic, characteristics-based factor model of returns. This kind of model was first studied by Connor and Linton (2007) and Connor et al. (2012), where they specified their model as:

$$y_{it} = \alpha_i + \sum_{j=1}^{J} g_j(X_{ji}) f_{jt} + \epsilon_{it}, \tag{1}$$

where y_{it} is the excess return on security i at time t; f_{jt} is the j^{th} risk factor's return at time t; X_{ji} is the j^{th} observable characteristic of firm i; α_i represents the intercept (mispricing) part of i^{th} asset return; and ϵ_{it} are the mean zero idiosyncratic shocks. They restricted characteristic-based loading $g_j(\cdot)$ to be a univariate nonparametric function. To extend the dimension of the factor loading function $g_j(\cdot)$, Kelly et al. (2019) and Kim et al. (2019) specify both mispricing and factor loading parts as a parametric linear function of a large set of firm-specific characteristics as:

$$y_{it} = h(\boldsymbol{X_i}) + \sum_{j=1}^{J} g_j(\boldsymbol{X_i}) f_{jt} + \epsilon_{it}. \tag{2}$$

They illustrated the valididity of characteristics-based factor models and provided relevant empirical results. Li and Linton (2020) generalized the parametric part of Equation 2 as semiparametric functions to be consistent with earlier research. They also proposed power enhanced tests to verify their model, concluding that the semiparametric mispricing component $h(\boldsymbol{X_i})$ was only significant during certain rolling windows.

Section two describes the econometric framework for our model. We assume the returns are generated by the asset pricing model in Li and Linton (2020) and that the factor risk premia are predictable based on a single-index function involving a set of both stationary and nonstationary predictors.

Section three presents the general portfolio management problem and our restricted class of portfolio selection rules in which the problem is divided into two steps. In the first step, the investors choose a set of characteristics-based sub-portfolios that are well-diversified and mimic the returns of the underlying unobservable factors. In the second step, the investors choose a dynamic combination of these sub-portfolios and a risk-free asset dependent upon their time-varying information set and utility function. The information is specified as a single-index function, which is well-approximated by orthogonal series, allowing both stationary and nonstationary covariates. We show that under reasonable conditions on risk preference, the two-step selection rule has asymptotically zero impact on an investor's expected utility as the number of assets grows to infinity, relative to the unattainable true optimal choice.

Section four derives estimators for both steps. In the factor-tilt step, the factor-mimicking portfolios are constructed by the linear combination of estimated characteristics-based factor loadings. To diversify the idiosyncratic shocks further, the weight for each factor loading function is estimated through a constrained quadratic objective function. In the second step, called factor timing, the optimization of the expected utility function is solved by the methodology of the continuously-updating GMM, as in Hansen et al. (1996). The weights allocated to the risk-free asset and sub-portfolios are determined by the single-index function approximated through Hermite polynomials, which allows for both stationary and nonstationary predictors, as in Dong et al. (2016b). The coefficients of those orthogonal bases are estimated by solving the sample counterpart under the continuously-updating GMM framework. Section five documents the hypothesis tests on the significance of these predictors included in the single-index function.

Section six presents the empirical findings. We apply our approaches to monthly CRSP and FRED data and reveal some popular predictive variables' nonstationarity and significance. Furthermore, we find our portfolios have different but outstanding performance under various levels of risk aversion. Finally, the results of the in-sample and out-sample are similar and reflect the risk preference of the investor.

Section seven concludes and discusses the paper, while proofs of theorems and supplementary tables are arranged in the Appendix.

2 Econometric Framework

We assume there is a large panel of monthly stocks' excess returns generated by the characteristics-based model:

$$y_{it} = \sum_{j=1}^{J} g_j(\boldsymbol{X_i})(f_{jt} + \phi_{jt}) + \epsilon_{it}, \tag{3}$$

where y_{it} is i^{th} stock's excess return at time t while $\boldsymbol{X_i}$ is a large set of assets' P-vector of characteristics, which is regarded as time-invariant within a short time window; $g_j(\boldsymbol{X_i})$ is the j^{th} characteristics-based factor loading, which is specified as a multivariate additive semiparametric function. The factor returns $\boldsymbol{F_t} = (f_{1t}, \ldots, f_{Jt})^\intercal$ are the common sources of risk in asset returns at time t with associated means $\boldsymbol{\phi_t} = \{\phi_{1t}, \ldots, \phi_{Jt}\}$. The asset-specific return ϵ_{it} is conditional zero mean, i.e., $E(\epsilon_{it}|\boldsymbol{X_i}, \boldsymbol{F_t}) = 0$.

This framework is an extension of Connor and Linton (2007) and Connor et al. (2012), who assumed the factor beta function $g(\cdot)$ to be univariate. This model is a special case of Li and Linton (2020) by replacing the mispricing component with the mean value ϕ_{jt} of the j^{th} risk factor.

We allow for time variation in the characteristics of the assets across rolling windows. We treat the $n \times J$ matrix of characteristics in the t^{th} rolling window $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^\intercal$, as a random draw from a multivariate population distribution. Furthermore, the investor can observe $\boldsymbol{X}$ before rolling block t, and then choose his time t portfolio.

We define the $n \times J$ matrix $G(\boldsymbol{X}) = (g_1(\boldsymbol{X}), \ldots, g_J(\boldsymbol{X}))$ and $g_j(\boldsymbol{X}) = (g_j(\boldsymbol{X_1}), \ldots, g_j(\boldsymbol{X_n}))^\intercal$, and the matrix form of the **demeaned** assets' returns at time t is :

$$\boldsymbol{Y_t} = G(\boldsymbol{X})\boldsymbol{F_t} + \boldsymbol{\epsilon_t}, \qquad t = 1, 2, \ldots, T, \tag{4}$$

where $\boldsymbol{Y_t}$ is a $n \times 1$ matrix of the demeaned assets' excess returns at time t, $G(\boldsymbol{X}) = (g_1(\boldsymbol{X}), \ldots, g_J(\boldsymbol{X}))$ is a $n \times J$ factor loading matrix, and $\boldsymbol{\epsilon_t}$ is a $n \times 1$ vector of asset-specific returns.

Furthermore, we assume that there exists a nonsingular $J \times J$ matrix $\boldsymbol{M^G}$ such that:

$$E(G(\boldsymbol{X_i})^\intercal G(\boldsymbol{X_i})) = \boldsymbol{M^G}, \tag{5}$$

where the expectation is over the multivariate probability density of characteristics and the off-diagonal elements are all zeros. The Equation 5 simply imposes that for any large random population of the assets, the factor-loading matrix is nonsingular with the probability of one.

For a finite value of n and realized characteristics matrix $\boldsymbol{X}$, under weak assumption, the finite sample second moment matrix approaches the population value as:

$$p \lim_{n \to \infty} \frac{1}{n} \sum_{1=1}^{n} G(\boldsymbol{X_i})^\intercal G(\boldsymbol{X_i}) = \boldsymbol{M^G}, \tag{6}$$

There is a scale indeterminacy in Equation 3 since one could multiply the function $g_j(\boldsymbol{X}_i)$ by any non-zero constant c and $(f_{jt} + \phi_{jt})$ by $1/c$, and the model of returns would be identical. We resolve the this indeterminacy by setting $\boldsymbol{M}^G_{jj} = 1, \text{ where } j = 1, \ldots, J.$

The estimation of Equation 5 under Equation 6 is discussed in Li and Linton (2020); it is achieved mainly through "Projected Principal Component Analysis" (PPCA) by Fan et al. (2016). The idea is to project the $n \times T$

asset excess returns onto the B-spline space spanned by $\boldsymbol{X}$, where $\mathbf{X} = (\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n})$, and then, we collect the projected returns $\hat{\mathbf{Y}}$. Furthermore, we perform PCA on $\frac{1}{\mathbf{T}}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\intercal$. Therefore, the $\hat{G}(\boldsymbol{X})$ is estimated as the largest J eigenvectors of $\frac{1}{\mathbf{T}}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\intercal$. Due to the property of the PCA, the assumption Equation 6 is satisfied.

Let $\boldsymbol{M^f} = E(\boldsymbol{F_t}\boldsymbol{F_t}^\intercal)$ denote the nonsingular covariance matrix of the factors and $\boldsymbol{M^\epsilon} = E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\intercal)$ denote the covariance matrix of the idiosyncratic returns; note these two sources of returns are statistically independent. We assume that the asset-specific risks are diversifiable, in particular:

$$\lambda_{max}(\boldsymbol{M^\epsilon}) < c_\lambda < \infty,$$

where the λ_{max} denotes the largest eigenvalue whereas c_λ is a constant. This follows readily from standard assumptions on weak correlations of the asset-specific risks.

We allow dynamic variation in the mean value of factor return premia. At the beginning of each period, a $K \times 1$ vector of random signal $\boldsymbol{z_t} = (z_{1t}, \ldots, z_{Kt})^\intercal$ is observed by the investor before he chooses his portfolio. The expected return on the j^{th} factor in Equation 3 is a nonlinear function of a fixed linear combination of these dynamic signals by coefficients $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)^\intercal$ as:

$$\phi_{jt+1} = \pi_j(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}). \tag{7}$$

The three vectors $\boldsymbol{F_t}$,$\boldsymbol{z_t}$ and $\boldsymbol{\epsilon_t}$ are assumed to be statistically independent. At each time t, the investor observes the characteristics of those assets $\boldsymbol{X}$, which is treated as time-invariant during this time block, and the dynamic signal $\boldsymbol{z_t}$. Then, the investor chooses his time t portfolio based on this information. Finally, at the $(t+1)^{th}$ period, his portfolio return depends upon the realized assets' returns, which in turn depends on the realized factor returns and asset-specific returns $\boldsymbol{F_{t+1}}$ and $\boldsymbol{\epsilon_{t+1}}$ respectively, according to Equation 3.

3 A Two-Step Version of the Portfolio Choice Problem

This section first defines the utility function of a rational decision-maker and then describes how the optimal portfolio weights are chosen through a two-step procedure. In step one, the investor chooses characteristics-based factor-mimicking sub-portfolios based on a linear combination of the beta function$\sum_{j=1}^J g_j(\boldsymbol{X})$ in Equation 3. Step two combines these sub-portfolios optimally using expected utility as the investor's objective function, based on a dynamic index.

3.1 Utility Function of the Investor

The investor in our model is myopic and he chooses his portfolio for time t to maximize one-period expected utility of return. We assume his return at time t is W_t and his risk-averse von Neumann Morgenstern preference is defined over W_t with a lower bound on the second derivative:

$$\frac{d}{dW}u(W) > 0, -c < \frac{d^2}{W^2}u(W) < 0 \tag{8}$$

Additionally, we define the optimal portfolios weights $n \times 1$ vector $\mathbf{w}^*$ such that:

$$\mathbf{w}^* = \arg\max_{\mathbf{w}} E[u(r_{ft} + \mathbf{w}^{*\intercal}\mathbf{r_t})|\boldsymbol{X_t}, \boldsymbol{z_t}], \tag{9}$$

where r_{ft} is the risk-free return at time t and $\mathbf{r_t}$ is a $n \times 1$ vector of stock returns at time t. In practice , the optimal $\mathbf{w}^*$ is hard to determine and unstable when n is large or the trading frequency is high, as discussed in the Introduction. Therefore, we consider optimal portfolio choice under a restriction on portfolio weights. Rather than choosing asset weights directly, the investor chooses a set of J characteristics-based portfolios to approximately mimic the factors. Then, in the second step, the investor combines these factor-mimicking subportfolios optimally using his expected utility function conditional on a group of predictors.

3.2 Step 1: Factor-mimicking Sub-portfolios

In this subsection, we propose a method to construct factor-mimicking sub-portfolios based on Equation 3 and discuss the properties of these sub-portfolios.

We propose a semiparametric weighting function to mimic the risk factors $\boldsymbol{F_t}$, which is in the form of a linear combination of characteristics-based factor loadings as in Equation 3:

$$b_j(\boldsymbol{X}_i) = \gamma_{j1}g_1(\boldsymbol{X}_i) + \cdots + \gamma_{jJ}g_J(\boldsymbol{X}_i), \tag{10}$$

therefore, the portfolio weight of i^{th} asset to construct the j^{th} subportfolio is $\frac{1}{n}b_j(\boldsymbol{X}_i)$.

The weighting matrix of assets to mimic all J factors is as follows:

$$B(\boldsymbol{X}_i) = \frac{1}{n}\mathbf{\Gamma}G(\boldsymbol{X}_i)^\intercal, \tag{11}$$

where

$$B(\boldsymbol{X}_i) = \frac{1}{n}\begin{Bmatrix} b_1(\boldsymbol{X}_i) \\ b_2(\boldsymbol{X}_i) \\ \vdots \\ b_J(\boldsymbol{X}_i) \end{Bmatrix},$$

$$\mathbf{\Gamma} = \begin{Bmatrix} \gamma_{11} & \cdots & \gamma_{1J} \\ \gamma_{21} & \cdots & \gamma_{2J} \\ \cdots & \cdots & \cdots \\ \gamma_{J1} & \cdots & \gamma_{JJ} \end{Bmatrix}, \quad G(\boldsymbol{X}_i)^\intercal = \begin{Bmatrix} G_1(\boldsymbol{X}_i) \\ G_2(\boldsymbol{X}_i) \\ \vdots \\ G_J(\boldsymbol{X}_i) \end{Bmatrix}.$$

Thus, the $J \times 1$ factor-mimicking portfolio return vector at time t is calculated:

$$\boldsymbol{Q_t}(\boldsymbol{X}) = \begin{Bmatrix} q_{1t}(\boldsymbol{X}) \\ q_{2t}(\boldsymbol{X}) \\ \vdots \\ q_{Jt}(\boldsymbol{X}) \end{Bmatrix} = \sum_{i=1}^{n} B(\boldsymbol{X}_i)y_{it}. \tag{12}$$

The factor-mimicking portfolio vector has at least two attractive properties, which are listed as theorems.

Theorem 1. *Each subportfolio in the $J \times 1$ factor-mimicking vector defined by Equation 12 is a linear combination of risk factors f_{jt} directly.*

Theorem1 implies that we can control the similarity between subportfolios and risk factors by adjusting coefficients matrix Γ, which provides us with considerable flexibility.

Theorem 2. *The returns of portfolio defined by Equation 12 have asymptotically zero idiosyncratic variance.*

Theorem 2 illustrates that portfolio returns of factor-mimick sub-portfolios can diversify the asset-specific returns completely as the number of assets goes to infinity.

An investor who uses a semiparametric characteristics-based weight function to choose sub-portfolios rather than individual assets i sacrifices the flexibility to weight assets differently based on the properties of their asset-specific returns ϵ_{it}, since the sub-portfolio weight function Equation 12 only differentiates assets by their characteristic vectors. However, for both hedge fund managers and researchers, there are no satisfactory rules for choosing thousands of assets robustly. Furthermore, some weighting strategies have to be rebalanced once per trading day, and even more frequently for some strategies. This high-speed decision-making problem is intractable without some simplifying applicable rules like Equation 12.

3.3 Step 2: Factor-timing Portfolio Based on Dynamic Signals

This subsection describes how to approximate the dynamic signal function $\pi_j(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})$ in Equation 7, and how to use this function as dynamic weights assigned to those factor-mimicking sub-portfolios in subsection 3.2, to reflect information about their overperformance/underperformance on a risk-adjusted basis. This subsection captures the particular "factor-timing" strategy used by the investor.

Here, we define the objective function as:

$$\arg\max_{\boldsymbol{\theta}} E[u(\alpha r_{ft} + \boldsymbol{\Pi}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})^\intercal \boldsymbol{Q_{t+1}}(\boldsymbol{X}))], \tag{13}$$

subject to

$$\|\boldsymbol{\theta}\|_2 = 1 \text{ and } \theta_1 > 0$$

and

$$\alpha + \sum_{j=1}^{J} \boldsymbol{\pi_j}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) = 1$$

where r_{ft} is the risk-free return at time t and α is its portfolio weight, and

$$\boldsymbol{\Pi}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) = (\pi_1(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}), \ldots, \pi_J(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}))^\intercal.$$

The first restriction is for identification purposes while the second is for unit investment. We do not restrict short selling and leverage. Equation 13 is a transformation of the objective function Equation 9. The $n \times 1$ vector of assets' returns $\boldsymbol{r_t}$ is replaced by the vector of sub-portfolios' returns $\boldsymbol{Q_{t+1}}$ conditional on $\boldsymbol{X}$, which compacts the information of $\mathbf{r_t}$ through observed characteristics $\boldsymbol{X}$. Similarly, the dynamic weights of each asset $\mathbf{w}^*$ is substituted by the dynamic information function $\boldsymbol{\Pi}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})$, which is the mean function for the risky factors $\boldsymbol{\phi_t}$ as in Equation 3. In other words, the objective function Equation 9 is a transformation of the utility function Equation 13 by incorporating conditional variables $\boldsymbol{z_t}, \boldsymbol{X}$.

Our purpose is to maximize the conditional expectation of the investor's utility function. The investment allocation to the j^{th} factor-tilt sub-portfolios is determined by the j^{th} information indicator $\pi_j(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})$, which is a single-index function, to avoid the problem of "curse of dimensionality" caused by fully nonparametric methods. We specify fixed linear combinations as information input in an unknown function $\pi_j(\cdot)$, as stated by Aït-sahalia and Brandt (2001), for at least two reasons. Statistically, this can achieve a better convergence rate for estimates, and economically, a univariate index value provides meaningful and convenient descriptions of current investment opportunities. Meanwhile, these index functions' effects on each sub-portfolio can be highly nonlinear, as documented by Aït-sahalia and Brandt (2001). Therefore, we do not specify the functional form of $\pi_j(\cdot)$, allowing a parametric index function to influence each sub-portfolio's weight nonparametrically.

To facilitate our estimation procedures, we approximate those unknown functions $\pi_j(\cdot)$ by orthonormal bases similar to Dong et al. (2016b). Their methods can allow the elements of information vector $\boldsymbol{z_t}$ to be nonstationary. As pointed by Gao et al. (2013) and Gao and Phillips (2013), conventional kernel estimation as in the Brandt (1999) and Aït-sahalia and Brandt (2001) method may not be workable due to the breakdown of limit theory, when $\boldsymbol{z_t}$ is a multivariate $I(1)$ process. In practice, some time series predictors are likely to be nonstationary, like the unemployment rate, inflation and exchange rates, among other economic indicators. Therefore, we apply a similar method as in the Dong et al. (2016b) to validate a more comprehensive application of our model.

Suppose all the link functions π_j belong to $L^2(\mathbb{R}) = \{f(x) : \int f^2(x)dx < \infty\}$. The Hermite function sequence $\{\mathcal{H}_i\}$ is an orthonormal basis in $L^2(\mathbb{R})$:

$$\mathcal{H}_i(x) = (\sqrt{\pi}2^i i!)^{-1/2} H_i(x) \exp(-\frac{x^2}{2}), \quad i \geqslant 0, \tag{14}$$

where $H_i(x)$ are Hermite polynomials orthogonal with density $\exp(-x^2)$. The orthogonality reads $\int H_i(x)H_j(x)dx = \delta_{ij}$, the Kronecker delta.

Therefore, any countinuous function $\pi_j(\cdot) \in L^2(\mathbb{R})$ can be expanded into a linear combination of orthogonal series:

$$\pi_j(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) = \sum_{l=0}^{\infty} \beta_{jl} \mathcal{H}_l(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}). \tag{15}$$

We keep the first $L-1$ terms and leave the rest as approximate residues:

$$\pi_j(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) = \sum_{l=0}^{L-1} \beta_{jl}\mathcal{H}_l(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) + \psi(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}), \tag{16}$$

where $\psi(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})$ is the approximation residues.

Furthermore, all J dynamic indicator functions can be approximated (we assume the same truncation parameters for all functions for purposed of notation simplicity only):

$$\boldsymbol{\Pi}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) = \boldsymbol{B}\boldsymbol{\mathcal{H}_L}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) + \Psi(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}), \tag{17}$$

where

$$\boldsymbol{B} = \begin{Bmatrix} \beta_{10} & \dots & \beta_{1(L-1)} \\ \dots & \dots & \dots \\ \beta_{J0} & \dots & \beta_{J(L-1)} \end{Bmatrix}, \quad \boldsymbol{\mathcal{H}_L}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) = \begin{Bmatrix} \mathcal{H}_0(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) \\ \dots \\ \mathcal{H}_{L-1}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t}) \end{Bmatrix},$$

and $\Psi(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})$ is the approximation error.

Therefore, the objective function Equation 13 is transformed through replacing $\boldsymbol{\Pi}(\boldsymbol{\theta}^\intercal \boldsymbol{z_t})$ by $\boldsymbol{B}\boldsymbol{\mathcal{H}_L}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}})$ as:

$$\arg\max_{\alpha,\boldsymbol{B},\boldsymbol{\Theta}} E[u(\alpha r_{ft} + (\boldsymbol{B}\boldsymbol{\mathcal{H}_L}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}))^\intercal \boldsymbol{Q_t}(\boldsymbol{X}))], \tag{18}$$

subject to

$$\|\boldsymbol{\theta}\|_2 = 1 \text{ and } \theta_1 > 0$$

$$\alpha + \sum_{j=1}^{J} \beta_{\mathbf{j}}^\intercal \mathcal{H}_{\mathbf{L}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}) = 1$$

Theorem 3. *The restricted optimal portfolio weight function chosen by Equation 18 gives an approximately optimal portfolio.*

Theorem 3 demonstrates that, as the number of assets $n \to \infty$, our two-step procedure is approximately equivalent to Equation 9, which is the completely unrestricted asset-by-asset portfolio optimization because these two methods give the same expected utility asymptotically.

4 Methodology

This section illustrates procedures for estimating $\boldsymbol{\Gamma}$ in Equation 11, and $\boldsymbol{B}, \boldsymbol{\theta}$ in Equation 18.

We assume that the investor chooses Γ based on the following objective:

$$\hat{\Gamma} = \arg\min_{\Gamma} \sum_{j=1}^{J} E(b_j(\boldsymbol{X_i})^2) \tag{19}$$

subject to

$$E[Q_t(\boldsymbol{X})Q_t(\boldsymbol{X})^\intercal] = \boldsymbol{I}_J,$$

where $\boldsymbol{I}_J$ is a $J \times J$ identity matrix.

In words, we choose the linear combination coefficients $J \times J$ matrix to maximize the spread of the portfolio weights, specifically by minimizing the expected sum of squared portfolio weights, in the class of semiparametric functions of the characteristics, subject to an orthogonality constraint on the vector of sub-portfolios' returns. These portfolios are an econometrically-derived variant of the widely popular Small-Minus-Big (SMB) and High-minus-Low (HML) portfolios designed by Fama and French (1993) to capture the size-related and value-related return factors. Fama and French (1993) did not minimize the sum of squared portfolio weights as was done in Equation 19, but they instead set the portfolio weights using capitalization weight, which, in the highly diversified US equity market, have a very low sum-of-squared relative to the number of assets. Fama and French (1993) did not explicitly impose the orthogonality condition applied in Equation 19, but, as they noted, they chose their size and value breakpoints so that the portfolio returns would have very low correlation. The reason that we set the orthogonal constraint here is to diversify idiosyncratic risks further.

Next we show that Equation 19 can have a Langrangian solution. After expanding the constraint and under the independence between $\boldsymbol{F_t}$ and $\boldsymbol{\epsilon_t}$, we have:

$$\begin{aligned} E[Q_t(\boldsymbol{X})Q_t(\boldsymbol{X})^\intercal] &= E[(\tfrac{1}{n}\boldsymbol{\Gamma G(X)^\intercal G(X) F_t} + \tfrac{1}{n}\boldsymbol{\Gamma G(X)^\intercal \epsilon_t})(\tfrac{1}{n}\boldsymbol{F_t^\intercal G(X)^\intercal G(X)\Gamma^\intercal} + \tfrac{1}{n}\boldsymbol{\epsilon_t^\intercal G(X)\Gamma^\intercal})] \\ &= E(\boldsymbol{\Gamma}\tfrac{\boldsymbol{G(X)^\intercal G(X)}}{\boldsymbol{n}}\boldsymbol{F_t F_t^\intercal}\tfrac{\boldsymbol{G(X)^\intercal G(X)}}{\boldsymbol{n}}\boldsymbol{\Gamma^\intercal}) + \tfrac{1}{n^2}E(\boldsymbol{\Gamma G(X)^\intercal \epsilon_t \epsilon_t^\intercal G(X)\Gamma^\intercal}) \\ &\to \boldsymbol{\Gamma M^G E(F_t F_t^\intercal) M^G \Gamma^\intercal} + \tfrac{1}{n^2}\boldsymbol{\Gamma G(X)^\intercal E(\epsilon_t \epsilon_t^\intercal) G(X)\Gamma^\intercal} \end{aligned},$$

which is a quadratic form in $\boldsymbol{\Gamma}$.

As for the objective function, we have:

$$\sum_{j=1}^{J} E(b_j(\boldsymbol{X_i})^2) = \sum_{j=1}^{J} E(\boldsymbol{\Gamma^2 G(X_i)^2}),$$

which is linear in $\boldsymbol{\Gamma^2}$.

Therefore, we write this constrained optimization problem of sample analogues in Lagrangian form:

$$L(\boldsymbol{\Gamma}) = \frac{1}{n}\sum_{1}^{J}\sum_{i=1}^{n} \boldsymbol{\Gamma} G(\boldsymbol{X_i})^\intercal G(\boldsymbol{X_i})\boldsymbol{\Gamma}^\intercal - \boldsymbol{\Lambda}^\intercal vec((\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Q_t}(\boldsymbol{X})\boldsymbol{Q_t}(\boldsymbol{X})^\intercal) - \boldsymbol{I}_J), \tag{20}$$

where Λ is the $\frac{1}{2}J(J+1)$ vector of Lagrangian multipliers, and vec is the vecterization of a matrix.

The optimal $\boldsymbol{\Gamma}$ and associated Lagrangian multipliers will solve the first order conditions:

$$\frac{\partial L}{\partial \boldsymbol{\Gamma}} = \boldsymbol{0}^{J \times J},$$

$$\frac{\partial L}{\partial \boldsymbol{\Lambda}} = \boldsymbol{0}^{\frac{1}{2}J(J+1)}.$$

Meanwhile, we collect the estimate $\hat{\boldsymbol{\Gamma}}$ to obtain the factor-mimicking sub-portfolios' returns as:

$$\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}) = \begin{Bmatrix} \hat{q}_{1t}(\boldsymbol{X}) \\ \hat{q}_{2t}(\boldsymbol{X}) \\ \vdots \\ \hat{q}_{Jt}(\boldsymbol{X}) \end{Bmatrix} = \sum_{i=1}^{n} \hat{B}(\boldsymbol{X}_i) y_{it} = \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\Gamma}} \hat{G}(\boldsymbol{X}_{\boldsymbol{i}})^{\intercal} y_{it}, \tag{21}$$

where $\hat{G}(\boldsymbol{X}_{\boldsymbol{i}})$ is the consistent estimate of Equation 3 as in Li and Linton (2020), where they specify $\boldsymbol{G}(\boldsymbol{X}_{\boldsymbol{i}})$ as an additive semiparametric function of asset-specified characteristics.

The next step derives an estimator for the dynamic portfolio allocation weighting functions $\boldsymbol{\Pi}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}})$. The portfolio weight estimation problem is to find:

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}) = \arg\max_{\alpha, \boldsymbol{B}, \boldsymbol{\theta}} E[u(\alpha r_{ft} + (\boldsymbol{B} \boldsymbol{\mathcal{H}}_{\boldsymbol{L}}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}}))^{\intercal} \hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}))], \tag{22}$$

subject to

$$\|\boldsymbol{\theta}\|_2 = 1 \text{ and } \theta_1 > 0$$

$$\alpha + \sum_{j=1}^{J} \beta_{\mathbf{j}}^{\intercal} \mathcal{H}_{\mathbf{L}}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}}) = 1$$

where $\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X})$ is the estimate of sub-portfolios from Equation 21. This is essentially the same semiparametric estimation problem analyzed by Aït-sahalia and Brandt (2001). The procedure relies on the profile estimation of the single-index function. We iterate the first order condition to convergence after choosing initial values arbitrarily.

With respect to identification issues, we need to solve another constrained optimization problem as:

$$\arg\max_{\alpha, \boldsymbol{B}, \boldsymbol{\theta}} E[u(\alpha r_{ft} + (\boldsymbol{B} \boldsymbol{\mathcal{H}}_{\boldsymbol{L}}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}}))^{\intercal} \hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}))], \tag{23}$$

subject to

$$\|\boldsymbol{\theta}\|_2 = 1 \text{ and } \theta_1 > 0$$

$$\alpha + \sum_{j=1}^{J} \beta_{\mathbf{j}}^{\intercal} \mathcal{H}_{\mathbf{L}}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}}) = 1$$

Therefore, the first order condition of the maximization with respect to $\boldsymbol{B}, \boldsymbol{\theta}$ is:

$$E[\boldsymbol{M}_t] = \begin{Bmatrix} u'(\cdot) \hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}) \otimes \boldsymbol{\mathcal{H}}_{\boldsymbol{L}}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}})] = \boldsymbol{0}_{\boldsymbol{JL \times 1}} \\ u'(\cdot) \hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}) \boldsymbol{I}_{\boldsymbol{L \times L}} (\boldsymbol{B} \boldsymbol{\mathcal{H}'}_{\boldsymbol{L}}(\boldsymbol{\theta}^{\intercal} \boldsymbol{z}_{\boldsymbol{t-1}}) \otimes \boldsymbol{z}_{\boldsymbol{t-1}})] = \boldsymbol{0}_{\boldsymbol{JL \times 1}} \\ u'(\cdot) r_{ft} = 0 \end{Bmatrix},$$

where $\boldsymbol{I}_{\boldsymbol{L \times L}}$ is a $L \times L$ identity matrix while $\boldsymbol{\mathcal{H}'}_{\boldsymbol{L}}$ and $u'(\cdot)$ are the first derivatives of the truncated orthonormal series and the investor's utility function respectively.

As we can see, there are $2 \times \boldsymbol{JL} + 1$ moment conditions to maximize the objective function.

These moment conditions can be used to construct standard GMM problem as was done in Hansen (1982):

$$(\alpha, \hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}) = \min_{\alpha, \boldsymbol{B}, \boldsymbol{\theta}} E[\boldsymbol{M_t}]^\intercal \boldsymbol{S} E[\boldsymbol{M_t}]$$

subject to

$$\|\boldsymbol{\theta}\|_2 = 1 \text{ and } \theta_1 > 0$$
$$\alpha + \sum_{j=1}^{J} \beta_{\mathrm{j}}^\intercal \mathcal{H}_{\mathbf{L}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}) = 1$$

where $\boldsymbol{S}$ is the optimal weighting positive definite matrix as $\boldsymbol{S} = cov(\boldsymbol{M_t})^{-1}$.

Then, we substitute these moment conditions $E[\boldsymbol{M_t}]$ with corresponding sample counterparts as:

$$\boldsymbol{m_t} = \left\{ \begin{array}{c} \frac{1}{T}\sum_{t=1}^{T} u'(\cdot)\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}) \otimes \boldsymbol{\mathcal{H}_{L-1}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}})] = \boldsymbol{0_{JL\times 1}} \\ \frac{1}{T}\sum_{t=1}^{T} u'(\cdot)\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X})\boldsymbol{I_{L\times L}}(\boldsymbol{B\mathcal{H}'_{L-1}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}) \otimes \boldsymbol{z_{t-1}})] = \boldsymbol{0_{JL\times 1}} \\ u'(\cdot) r_{ft} = 0 \end{array} \right\}.$$

Similarly, we obtain the estimate of weighting as:

$$\hat{\boldsymbol{S}} = \frac{1}{T}\sum_{t=1}^{T} \left\{ \begin{array}{c} u'(\cdot)\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}) \otimes \boldsymbol{\mathcal{H}_{L-1}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}})] = \boldsymbol{0_{JL\times 1}} \\ u'(\cdot)\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X})\boldsymbol{I_{L\times L}}(\boldsymbol{B\mathcal{H}'_{L-1}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}) \otimes \boldsymbol{z_{t-1}})] = \boldsymbol{0_{JL\times 1}} \\ u'(\cdot) r_{ft} = 0 \end{array} \right\} \left\{ \begin{array}{c} u'(\cdot)\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X}) \otimes \boldsymbol{\mathcal{H}_{L-1}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}})] = \boldsymbol{0_{JL\times 1}} \\ u'(\cdot)\hat{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X})\boldsymbol{I_{L\times L}}(\boldsymbol{B\mathcal{H}'_{L-1}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}) \otimes \boldsymbol{z_{t-1}})] = \boldsymbol{0_{JL\times 1}} \\ u'(\cdot) r_{ft} = 0 \end{array} \right\}^\intercal.$$

Finally, we substitute the sample analogues and $\hat{\boldsymbol{S}}$ into the objective function, and estimate $\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}$:

$$(\hat{\alpha}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}) = \min_{\alpha, \boldsymbol{B}, \boldsymbol{\theta}} \boldsymbol{m}_{\boldsymbol{t}}^\intercal \hat{\boldsymbol{S}} \boldsymbol{m_t}, \tag{24}$$

subject to

$$\|\boldsymbol{\theta}\|_2 = 1 \text{ and } \theta_1 > 0$$
$$\alpha + \sum_{j=1}^{J} \beta_{\mathrm{j}}^\intercal \mathcal{H}_{\mathbf{L}}(\boldsymbol{\theta}^\intercal \boldsymbol{z_{t-1}}) = 1$$

Furthermore, we substitute Equation 24 into the optimization iteration, which is called the continuously-updating estimator; details can be found in Dong et al. (2018).

5 Hypothesis Tests

This section introduces the hypothesis tests that help us to understand which index variables are important to guide the construction of factor-timing portfolios. We apply a Wald test to infer the significance of θ_j. We have the null and alternative hypotheses as follows:

$$\boldsymbol{\mathcal{H}_0} : \boldsymbol{C\theta} = \mathbf{0}_{\boldsymbol{D}\times\mathbf{1}}, \quad \text{against} \quad \boldsymbol{\mathcal{H}_1} : \boldsymbol{C\theta} \neq \mathbf{0}_{\boldsymbol{D}\times\mathbf{1}},$$

where $\boldsymbol{C}$ is a $D \times K$ fix matrix indicating the number of constraints $\boldsymbol{D}$.

We denote the value of the objective function Equation 24 under $\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}$ as $\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})$ while under the null hypothesis $\boldsymbol{\mathcal{H}_0}$ as $\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}^*)$.

Therefore, if the null hypothesis is correct, we have:

$$T(\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}^*) - \boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})) \sim \boldsymbol{\chi}^2(\boldsymbol{D}), \tag{25}$$

where $\boldsymbol{\chi}^2(\boldsymbol{D})$ is the chi-square distribution with degree of freedom D. This method is a minimum-χ^2 test, the purpose of which is to check the minimized values of objective function Equation 24 after imposing some restrictions.

We reject the null hypothesis if the test statistic exceeds the critical value.

6 Empirical Study

6.1 Data Description

6.1.1 Index Variables

We use the same index variable set as Aït-sahalia and Brandt (2001). These variables are all at a monthly frequency:

- The **Default Spread** is the yield difference between Moody's Baa and Aaa rated bonds, observed from 1967-07-01 to 2017-06-01 (600 months in total) denoted as DS.
- The **Term Spread** is the yield difference between 10 and 1 year government bonds, observed from 1967-07-01 to 2017-06-01 (600 months in total) denoted as TS.
- The **Trend** is the difference between the log of the current S&P 500 index level and the log of the average index level over the previous 12 months, observed from 1967-07-01 to 2017-06-01 (600 months in total).
- The **Dividend Yield**, also called Dividend-to-Price, is the sum of dividends paid on the S&P 500 index over the past 12 months divided by the current level of the index observed from 1967-07-01 to 2017-06-01 (600 months in total). We use the percentage natural logarithm form of Dividend Yield, denoted as Ln(DY%).
- The **Risk Free** rate is obtained from the Fama-French factor model's risk-free rate, observed from 1967-07-01 to 2017-06-01, denoted in the percentage form as RF%.

In Table 1, both "Trend" and "RF%" have small variation while "RF%" has some strong correlation with "TS" and " Ln(DY%)". Apart from these, we also find that all of the index variables are not symmetrically distributed, which is shown by the non-zero skewness. As for the kurtosis, the table indicates that outliers are quite common among these variables.

In Table 2, we conclude the results of the unit root test and autocorrelation. After the Dickey-Fuller tests, we fail to reject the null hypotheses that there are no unit roots among all index variables, especially for the " Ln(DY%)". That can also be found from Figure 4. In terms of autocorrelation, almost all of the index series present persistent autocorrelation even for lag nine, " Ln(DY%)" showing a strong signal of autocorrelation coefficient of 0.94, as shown in Figure 4, Figure 5. However, "Trend" is an exception, where the autocorrelation decays to zero and is negative after lag ten as shown in Figure 3. These test results verify the necessity of applying orthogonal series to approximate the single index function with nonstationary covariates, as in subsection 3.3.

The data above was collected from the websites of FRED and Multpl.

6.2 Monthly Stock Data

We collected monthly stock returns from CRSP and firms' characteristics from Compustat, from 1965 to 2017. We constructed 33 characteristics following the methods of Freyberger et al. (2020). Details of these characteristics can be found in the appendix of Li and Linton (2020). We construct characteristics from fiscal year $t-1$ to explain stock returns between July of year t to June of year $t+1$. Following Hou et al. (2015), we adjust returns of delisted stocks. The method that we apply to estimate the Equation 4 is similar to Li and Linton (2020). We only include firms with at least three years of data in Compustat. The values of firm-specific characteristics are updated annually as most characteristic data are reported every year. We use rolling windows to accommodate these characteristics-based loadings and the risk factors are estimated correspondingly.

The time span of our in-sample analysis is 50 years, from July 1967 to June 2017 (600 months).

6.3 In-sample Factor-mimicking Portfolios

This section presents portfolios that mimic the annually updated risk factors estimated through Equation 5. In this study, we choose the number of unobservable factors in Equation 3 to be three. In Li and Linton (2020), they compared the effects of the number of factors through a simulation study, concluding that underestimating the number of factors can be problematic. However, their discussions mainly focused on the estimation of mispricing functions. We only have four dynamic index variables, and therefore, we follow the renowned research of Fama and French (1993) to set the number of factors to be three. According to the literature, three factors can capture the most important common variation in asset excess returns.

The methods are introduced in subsection 3.2, and we utilize all 600 months of data and construct three such portfolios every year, assuming the number of risk factors in Equation 3 to be three. Then, we conclude the descriptive statistics of these three sub-portfolios in Table 3. The zero mean and unit variance are determined by the constraints in estimation. As for the correlation between risk factors and those sub-portfolios using observations of all 600 months, our first step works very well because the diagonal elements of correlation are quite high while the off-diagonal elements are negligible. That demonstrates that each sub-portfolio imitates only the target risk factor's variation accurately and leaves the rest uncorrelated. The weights put on each asset for these sub-portfolios are calculated through a constrained optimization, which restricts the similarity between sub-portfolios and risk factors. Furthermore, during certain years, the sub-portfolios behaved in the opposite direction of the imitated factor, which can also influence the average correlation over 600 months. The annual correlation can be found Table 7, where some negatively correlated periods are presented.

6.4 Utility Function

We utilitze the classic Constant Relative Risk Aversion (CRRA) utility function to model function $u(W)$ in Equation 8:

$$u(W) = \begin{cases} \frac{W^{1-\xi}}{1-\xi} & \text{if } \xi > 1; \\ \ln(W) & \text{if } x = 1, \end{cases}$$

where ξ is an integer and $\xi = W\frac{\partial^2 u(W)/\partial W^2}{\partial u(W)/\partial W}$, measuring the level of risk aversion. Therefore, under this setting, the investor is risk-averse and tries to maximize his expected utility function through factor-mimicking and factor-timing portfolio strategy. The CRRA utility function is twice differentiable, which can further facilitate our optimization algorithm.

6.5 Selection of Truncation Number

The value of L in Equation 16, which is the truncation number in polynomials, needs to be determined here. Unfortunately, to the best of our knowledge, there is no rule of thumb for the best choice of L. We refer to Dong et al. (2015) and Dong et al. (2016a), where the authors determined L according to the number of observations n. However, the n in this study ranged from 468 to 2928. After trading off the computation burden and approximation accuracy, we choose L to be four throughout the empirical study.

Table 1: Index Variable Summary

		Descriptive Statistics							Correlation Matrix				
Index name	T	Mean	Variance	Median	Max	Min	Skewness	Kurtosis	DS	TS	Trend	Ln(DP%)	RF%
DS	600	1.08	0.2	0.94	3.38	0.55	1.82	7.34	1.00				
TS	600	1.12	1.39	1.23	3.40	-3.07	-0.32	2.72	0.09	1.00			
Trend	600	0.03	0.01	0.05	0.22	-0.4	-1.24	5.63	-0.27	0.07	1.00		
Ln(DY%)	600	1.00	0.17	1.05	1.83	0.10	-0.09	2.07	0.46	-0.26	-0.12	1.00	
RF%	600	0.39	0.08	0.41	1.35	0.00	0.51	3.44	0.23	-0.68	-0.01	0.65	1.00

This table documents the descriptive statistics of the index variables that are used in this empirical study as well as the correlations among them. To be consistent with most of the literature, we use the percentage values of DP and RF.

Table 2: Tests Summary

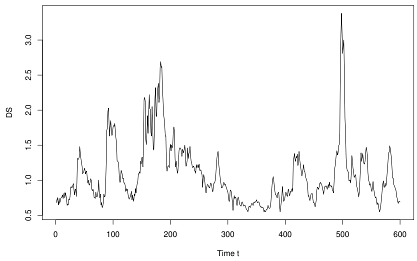
		Unit root test				Autocorrelation		
Index name	T	time trend	p-value (Trend)	ΔY_t	p-value (ΔY_t)	ρ_3	ρ_6	ρ_9
DS	600	-2.3×10^{-5}	-0.81	-3.82×10^{-2}	-3.43	0.85	0.68	0.54
		(2.8×10^{-5})		(1.11×10^{-2})				
TS	600	9.27×10^{-5}	1.23	-3.6×10^{-2}	-3.26	0.88	0.77	0.69
		(7.51×10^{-5})		(1.1×10^{-2})				
Trend	600	4.1×10^{-6}	0.47	-7.59×10^{-2}	-4.86	0.70	0.37	0.12
		(8.75×10^{-6})		(1.56×10^{-2})				
Ln(DY%)	600	-1.8×10^{-5}	-1.44	-8.86×10^{-3}	-1.72	0.98	0.96	0.94
		(1.25×10^{-5})		(5.16×10^{-3})				
RF%	600	-6.6×10^{-5}	-3.22	-5.48×10^{-2}	-4.23	0.94	0.90	0.87
		(2.05×10^{-5})		(5.16×10^{-3})				

This table summarizes the results of unit root tests and autocorrelations of those index variables. It reports the estimates, standard errors (in parentheses) and t-statistics of Dickey-Fuller test with trend individually. Autocorrelation column illustrates the correlation between the series and lag 3, lag 6 and lag 9 respectively, dennoted as ρ_3, ρ_6, ρ_9.

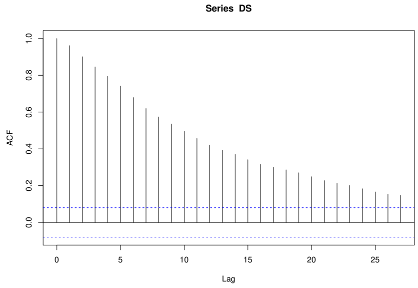
Table 3: Factor-mimicking Portfolios Summary

		Descriptive Statistics							Average Correlation		
Index name	T	Mean	Variance	Median	Max	Min	Skewness	Kurtosis	$\hat{\boldsymbol{f}}_{\boldsymbol{1}}$	$\hat{\boldsymbol{f}}_{\boldsymbol{2}}$	$\hat{\boldsymbol{f}}_{\boldsymbol{3}}$
$\hat{\boldsymbol{q}}_{\boldsymbol{1}}$	600	0.00	1.00	0.08	2.81	-3.10	-0.38	3.33	0.48	-0.01	0.13
$\hat{\boldsymbol{q}}_{\boldsymbol{2}}$	600	0.00	1.00	0.03	2.70	-3.20	-0.15	3.42	-0.11	0.59	0.05
$\hat{\boldsymbol{q}}_{\boldsymbol{3}}$	600	0.00	1.00	0.05	2.60	-2.79	-0.10	2.82	0.06	-0.03	0.63

This table presents the descriptive statistics of factor-mimicking portfolios and their correlations with estimated risk factors. $\hat{\boldsymbol{q}}_{\boldsymbol{1}}$, $\hat{\boldsymbol{q}}_{\boldsymbol{2}}$, and $\hat{\boldsymbol{q}}_{\boldsymbol{3}}$ are constructed portfolios through all 600 months' data while $\boldsymbol{f}_{\boldsymbol{1}}$, $\boldsymbol{f}_{\boldsymbol{2}}$, and $\boldsymbol{f}_{\boldsymbol{3}}$ are three $T \times 1$ factors estimated by rolling windows.

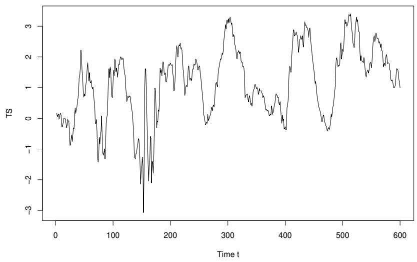


(a) Series Plot of DS

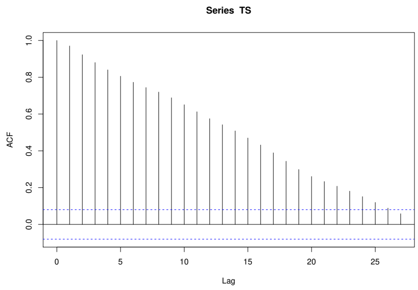


(b) Autocorrelation Plot of DS

Figure 1: The Plot of DS

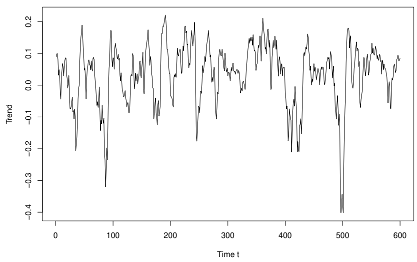


(a) Series Plot of TS

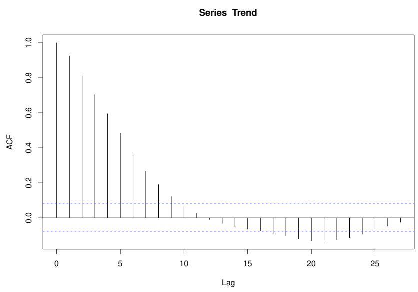


(b) Autocorrelation Plot of TS

Figure 2: The Plot of DS

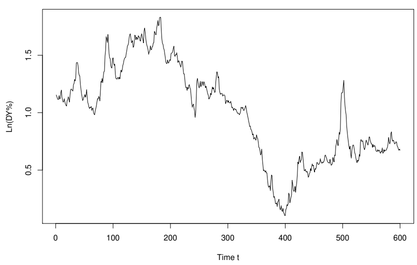


(a) Series Plot of Trend

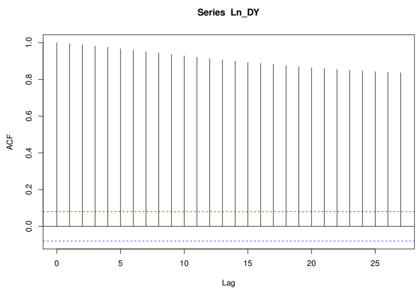


(b) Autocorrelation Plot of Trend

Figure 3: The Plot of Trend

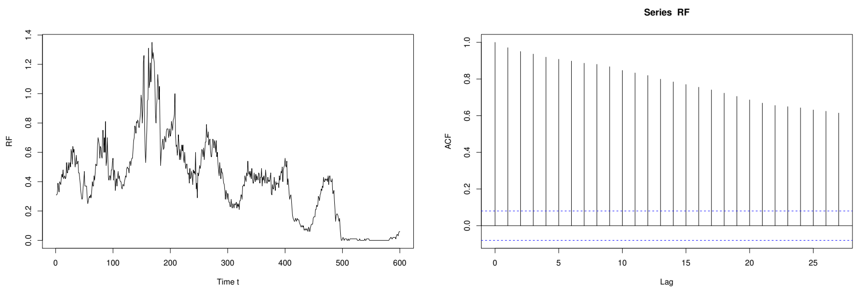


(a) Series Plot of Ln(DY%)



(b) Autocorrelation Plot of Ln(DY%)

Figure 4: The Plot of Ln(DY%)



(a) Series Plot of RF (b) Autocorrelation Plot of RF

Figure 5: The Plot of RF

6.6 Estimation of Dynamic Signals

This section presents the estimates of single-index coefficient vector $\boldsymbol{\theta}$ and the results of the corresponding hypothesis tests. We use the CRRA utility function with various risk aversion levels ξ. Meanwhile, we also test the null hypothesis in section 5 to examine whether some of the coefficients of dynamic variables are significantly different from 0. These tests are used to show the importance of this dynamic information during the second step of portfolio management, namely factor-timing.

During our estimation, all the optimization processes converged, and the optimized values are reported. The in-sample results are based on the data for all 600 months and the estimation procedures are repeated under different risk-aversion levels $\xi = 2, \xi = 5$, and $\xi = 10$. We then obtain the values of the objective function Equation 24, denoted as $\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})$. The hypothesis tests are conducted by setting $\theta_i = 0$, where i indicates the i^{th} index variable. We denote the value of the objective function Equation 24 under $\mathcal{H}_0 : \theta_i = 0$ as $\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\theta}_i = 0)$. In addition, χ^2 statistics are calculated as $T\Delta\boldsymbol{V} = T(\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\theta}_i = 0) - \boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}}))$.

Table 4: Index Variable Summary

		$\xi = 2$				$\xi = 5$				$\xi = 10$			
Index name	T	$\hat{\theta}_i$	$\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\theta}_i = 0)$	$\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})$	$T\Delta\boldsymbol{V}$	$\hat{\theta}_i$	$\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\theta}_i = 0)$	$\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})$	$T\Delta\boldsymbol{V}$	$\hat{\theta}_i$	$\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\theta}_i = 0)$	$\boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})$	$T\Delta\boldsymbol{V}$
DS	600	0.15	0.23	0.006	134.4	0.06	0.0001	2.5×10^{-8}	0.066	0.08	0.0026	4.9×10^{-10}	1.56
TS	600	0.19	0.064	0.006	34.8	-0.34	0.0008	2.5×10^{-8}	0.481	-0.06	5.2×10^{-9}	4.9×10^{-10}	0
Trend	600	0.03	0.01	0.006	2.4	0.06	1×10^{-6}	2.5×10^{-8}	0.006	0.03	8.1×10^{-10}	4.9×10^{-10}	0
Ln(DY%)	600	-0.97	0.06	0.006	32.4	-0.93	3.2×10^{-6}	2.5×10^{-8}	0.002	-0.995	2×10^{-8}	4.9×10^{-10}	0

This table reports the estimates and hypothesis test of dynamic index variables. $\hat{\theta}_i$ is the estimate of the coefficient of the i^{th} index variable while $\boldsymbol{V}$ represent the value of the objective function. $\Delta\boldsymbol{V} = \boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\theta}_i = 0) - \boldsymbol{\mathcal{V}}(\hat{\boldsymbol{B}}, \hat{\boldsymbol{\theta}})$.

The findings in Table 4 differ across the risk-aversion levels. When the magnitude of risk aversion is low, the influence of the dynamic index variables is significant. With $\xi = 2$, nearly all of the values of $T\Delta\boldsymbol{V}$ exceed the 95% critical value of $\chi^2(1)$, which is 3.84, except for DS. As the risk-aversion becomes larger, the importance of these dynamic variables declines. This can be confirmed when $\xi = 5$ and $\xi = 10$, where all the four variables are insignificant. We compare the values of the objective function Equation 24, and most of them are quite similar, both close to zero. That means the moment conditions in Equation 24 can be satisfied even if we restrict the coefficient of the i^{th} index variable to zero. Nevertheless, we cannot reject their joint significance.

6.7 In-sample Performance of Factor-timing Portfolios

This section presents portfolio performance estimated using in-sample data. As mentioned previously, there are two steps in the construction of our dynamic portfolio, namely, factor-tilt and factor-timing steps. In subsection 6.3, we describe how to build the sub-portfolios that mimic the behavior of risk factors. This section solves the second step, factor-timing: choosing the time-varying weights for the risk-free asset and risky sub-portfolios. The dynamic weights are determined by a single-index function with a set of index variables. These variables capture investment opportunities. We standardize the amount of investment to be 1 unit and take the monthly returns as the wealth gleaned by the investor. We do not restrict leverage or short-selling in order to check the influence of the risk-aversion level.

As we have 600 months in total, we record the average returns every year and annual standard deviations in Table 5 to save the space, and we calculate the Sharpe-ratio directly through $mean(Return_t)/SD_{annual}$. Table 5 shows the in-sample results from July 1967 to June 2017 under all three risk-aversion levels defined in subsection 6.4, and these results are compared with monthly S&P 500 returns.

Some findings here are significant and worth discussing. Firstly, for investors who have relatively lower risk-aversion, the average portfolio returns are more rewarding, with some extremely high returns appearing as well. For example, when the risk-aversion level $\xi = 2$, the twelve month average monthly returns can be 10.61and 8. 65. As for $\xi = 10$, the average monthly returns are more normal. Most monthly returns are around 5% except for some outliers. Secondly, a higher risk-aversion level corresponds to more volatile returns, such as losing -2.33 monthly during the whole year when $\xi = 2$, provided the standard deviation of the monthly return is 6.44. But the circumstances can be much more favourable when ξ increases to 5 and 10, with the standard deviation of the monthly return being 3.98 and 2.81, respectively. Especially under $\xi = 10$, the returns are quite acceptable and stable. Thirdly, all of the portfolios under various ξ have a relatively low Sharpe-ratio, compared with S&P 500 returns, which may be due to the high volatility.

In this empirical study, we optimize over three risky sub-portfolios and one risk-free asset without restricting leveraging or short-selling, and the weights for each asset are plotted in Figure 6. As we can see in (a), when the relative risk aversion level is low, $\xi = 2$, the weights for each asset are variable, while the scale of the vertical axis here is wider than (b) and (c). As we increase ξ, the weights become more stable. Specifically, when the ξ increases to 10, the only substantial volatility in the weights appeared around the stock crash in March 2000.

6.8 Out-sample Performance of Factor-timing Portfolios

This section examines the out-sample performance of our two-step portfolio selection procedure. We test the last six months of the last ten years in our data set for various risk aversion levels. The coefficients of the dynamic information function are estimated using all of the past information while the sub-portfolios are estimated using the first six months each year. The "Return" in Table 6 is calculated by substituting the predictors observed at the beginning of time $t + 1$. The sub-portfolios are constructed at time t, based on all the available data at the target year before time $t + 1$. Table 6 also lists the assigned weights to each sub-portfolios and the risk-free asset using 1 unit of investment, represented by c_1,c_2,c_3 and c_0. To summarize each column, we also provide the mean and standard deviation values at the end of the table, indicated by "ColMeans" and "ColStd".

As we can see from Table 6, most of the out-sample performance is quite similar to the in-sample performance Table 5. When the risk-aversion level is low such as $\xi = 2$, the variation of assets' weights is the largest and with an extensive range. Correspondingly, the realized monthly returns are also variable and high on average. The mean return of all 60 months is 0.36, which is very similar to that of the in-sample result which is 0.37. Not surprisingly, the out-sample standard deviation 8.88 is bigger than that of the in-sample result (6.44).

When the risk-aversion level increases to $\xi = 5$, the weights' volatility decreases, and the mean return also falls from 0.36 to 0.27, which is similar to the in-sample result (0.29). Compared with the $\xi = 2$ situation, the standard deviation of all the assigned weights and the monthly returns decline.

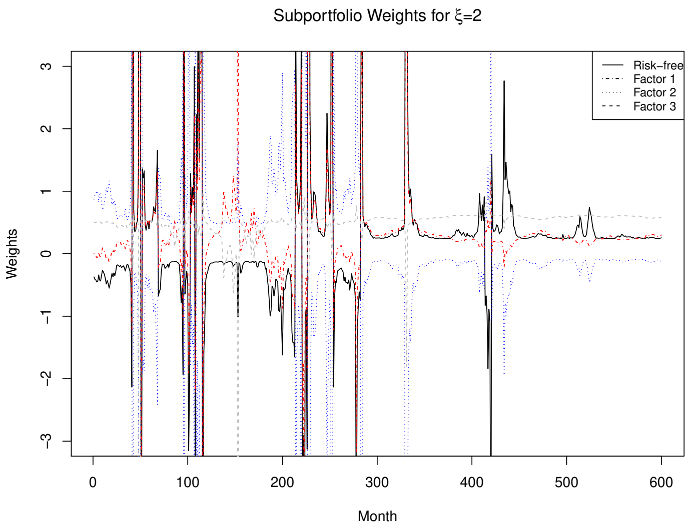
In the case of $\xi = 10$, all of the weights and monthly returns become more stable and less volatile. However, the average monthly return here is much lower than the in-sample result (0.17), with a smaller standard deviation of 1.51.

From the above analysis, we can conclude that our out-sample results are robust and vary according to the risk aversion levels. When the risk-aversion level is low, the investor reassigns his weights broadly and frequently, with an high average monthly return but high volatility. As the risk aversion level increases, the investor adjusts his weights more moderately, and the monthly average return and its standard deviation are reduced.

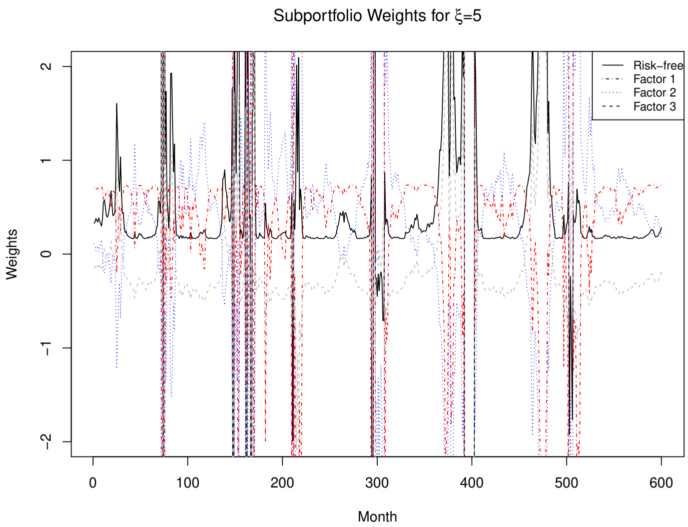
Table 5: Average Annual In-sample Results

n	$\xi = 2$			$\xi = 5$			$\xi = 10$			Average Monthly S&P 500		
	Return	SD	Sharpe-ratio	Return	SD	Sharpe-ratio	Return	SD	Sharpe-ratio	return	SD	Sharpe-ratio
468	-0.13	1.07	-0.12	0.15	0.70	0.22	0.06	0.61	0.10	0.01	0.03	0.01
894	-0.15	1.07	-0.14	0.20	0.47	0.41	0.08	0.63	0.13	-0.00	0.03	-0.00
1108	-0.13	1.11	-0.11	0.45	0.66	0.69	0.06	0.72	0.09	-0.02	0.04	-0.02
1199	-2.33	9.36	-0.25	0.10	0.88	0.12	0.04	0.59	0.08	0.02	0.03	0.02
1333	10.61	37.73	0.28	0.09	0.74	0.12	0.04	0.80	0.04	0.01	0.03	0.01
1409	-0.23	1.38	-0.16	0.10	0.77	0.13	0.06	0.82	0.08	-0.00	0.03	-0.00
1466	-0.16	1.03	-0.16	-1.02	3.50	-0.29	0.08	0.70	0.12	-0.01	0.04	-0.01
1560	-0.10	1.12	-0.09	0.35	1.19	0.30	0.05	0.70	0.07	0.00	0.07	0.00
1494	0.19	2.89	0.06	0.32	1.25	0.26	0.06	0.59	0.10	0.01	0.04	0.01
1292	0.52	7.48	0.07	0.14	1.29	0.11	0.03	0.60	0.05	-0.00	0.02	-0.00
1393	-0.12	0.73	-0.16	0.08	0.92	0.09	0.03	0.58	0.06	-0.00	0.03	-0.00
1340	-0.10	0.84	-0.12	0.38	0.94	0.41	0.05	0.79	0.06	0.00	0.03	0.00
1285	0.39	2.92	0.13	-0.13	5.27	-0.02	0.09	0.61	0.16	0.01	0.04	0.01
1181	-0.12	0.76	-0.16	-0.68	8.15	-0.08	0.03	0.57	0.06	0.01	0.03	0.01
1110	-0.15	0.83	-0.18	0.82	2.08	0.39	0.05	0.75	0.07	-0.01	0.04	-0.01
1044	-0.09	1.34	-0.07	-0.05	1.64	-0.03	-0.07	0.59	-0.11	0.04	0.03	0.04
1125	-0.73	1.72	-0.42	0.12	1.15	0.10	0.05	0.75	0.07	-0.01	0.02	-0.01
2192	-0.88	1.83	-0.48	-1.83	13.38	-0.14	0.05	0.73	0.06	0.02	0.03	0.02
2236	8.65	16.35	0.53	-0.87	2.35	-0.37	0.06	0.77	0.07	0.02	0.03	0.02
2273	0.07	0.53	0.13	0.08	0.87	0.09	0.05	0.95	0.05	0.02	0.03	0.02
2235	0.60	2.48	0.24	0.02	0.63	0.04	0.04	0.73	0.05	-0.01	0.06	-0.01
2270	-0.02	1.15	-0.02	0.28	0.95	0.30	0.09	0.58	0.16	0.02	0.02	0.02
2405	-0.33	1.19	-0.28	0.24	0.82	0.29	0.09	0.61	0.15	0.01	0.02	0.01
2376	2.03	3.87	0.53	0.02	1.00	0.02	0.03	0.63	0.05	0.01	0.05	0.01
2323	0.09	0.62	0.14	3.80	9.31	0.41	0.05	0.66	0.08	0.01	0.02	0.01
2344	0.05	0.68	0.08	2.45	3.21	0.76	0.03	0.61	0.05	0.01	0.01	0.01
2434	0.06	0.69	0.09	0.03	1.07	0.03	0.05	0.63	0.08	0.00	0.01	0.00
2548	-0.87	11.71	-0.07	-0.04	0.95	-0.04	0.10	0.56	0.17	0.01	0.02	0.01
2741	0.26	1.06	0.25	0.15	0.59	0.25	0.15	0.69	0.21	0.02	0.02	0.02
2928	0.10	0.55	0.19	0.14	0.74	0.18	0.25	0.53	0.47	0.02	0.04	0.02
2894	0.10	0.68	0.14	-0.03	1.90	-0.01	0.38	0.67	0.57	0.02	0.03	0.02
2905	0.11	0.73	0.16	-0.14	2.53	-0.05	0.25	0.64	0.39	0.02	0.05	0.02
2804	0.13	0.89	0.15	2.60	7.64	0.34	7.55	18.53	0.41	0.01	0.03	0.01
2570	0.10	0.69	0.14	-1.02	4.43	-0.23	0.95	1.64	0.58	-0.01	0.04	-0.01
2516	0.26	1.33	0.20	0.11	0.94	0.12	0.07	0.61	0.12	-0.02	0.05	-0.02
2491	0.11	0.89	0.13	0.07	1.05	0.07	0.03	0.60	0.05	-0.00	0.05	-0.00
2402	0.20	0.97	0.20	0.07	0.97	0.07	-0.04	0.58	-0.08	0.01	0.02	0.01
2326	0.06	0.90	0.07	0.01	0.50	0.03	0.11	0.58	0.18	0.01	0.02	0.01
2241	0.06	0.69	0.09	0.46	0.96	0.48	0.13	0.59	0.23	0.00	0.02	0.00
2178	0.12	0.73	0.17	7.04	15.93	0.44	0.15	0.60	0.25	0.02	0.02	0.02
2113	0.05	0.68	0.07	0.14	0.66	0.20	0.11	0.57	0.18	-0.01	0.04	-0.01
2023	-0.00	0.57	-0.00	-0.86	2.32	-0.37	0.04	0.85	0.04	-0.03	0.08	-0.03
2007	0.03	0.53	0.05	0.23	1.87	0.13	0.02	0.69	0.03	0.01	0.04	0.01
1924	-0.02	0.87	-0.02	0.02	1.14	0.02	-0.04	0.75	-0.05	0.01	0.02	0.01
1990	0.01	0.88	0.01	0.04	0.66	0.07	0.03	0.80	0.03	0.00	0.04	0.00
1937	0.00	0.89	0.00	-0.02	0.59	-0.03	0.01	0.79	0.01	0.02	0.02	0.02
1909	0.01	0.89	0.01	-0.02	0.94	-0.02	-0.01	0.70	-0.01	0.02	0.01	0.02
1872	0.00	0.69	0.01	-0.03	0.93	-0.04	-0.00	0.63	-0.00	0.01	0.02	0.01
1841	0.00	0.70	0.01	0.05	0.93	0.06	0.00	0.63	0.01	0.00	0.04	0.00
1826	0.01	0.70	0.02	-0.02	0.90	-0.02	-0.01	0.61	-0.01	0.01	0.01	0.01
Total mean	0.37	6.44	0.06	0.29	3.98	0.07	0.23	2.81	0.08	0.01	0.04	0.17

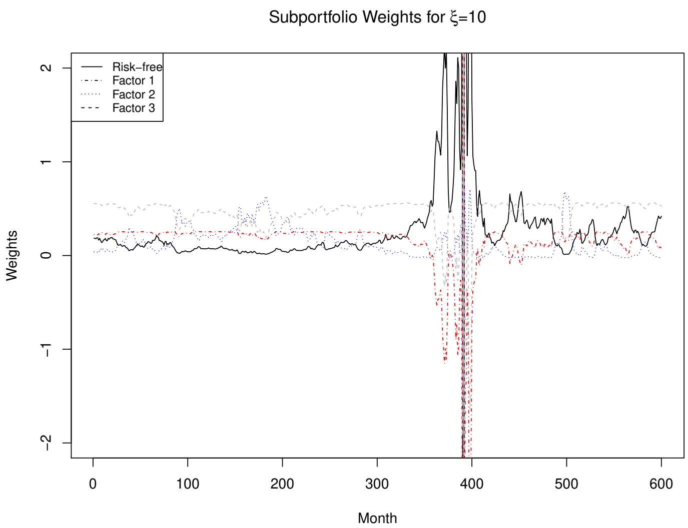
This table illustrates the in-sample results under various risk-aversion levels annually from July 1967- June 2017. n represents the number of stocks included in the portfolio. Both returns and standard deviations are calculated based on each year's results. The results of monthly S&P 500 returns are reported for comparison. No restrictions on leverage or short-selling.



(a) Monthly Weights Change Under $\xi = 2$



(b) Monthly Weights Change Under $\xi = 5$



(c) Monthly Weights Change Under $\xi = 10$

Figure 6: The Plot of Subportfolio Weights

Table 6: Monthly Out-sample Results Comparison

	$\xi = 2$					$\xi = 5$					$\xi = 10$				
Year	c_0	c_1	c_2	c_3	Return	c_0	c_1	c_2	c_3	Return	c_0	c_1	c_2	c_3	Return
2008	1.12	0.04	0.05	-0.21	-0.04	-0.14	0.59	0.22	0.32	0.12	0.07	0.73	0.01	0.20	0.65
	2.03	-0.13	-0.41	-0.49	0.21	-0.11	0.63	0.18	0.30	0.60	0.11	2.12	-0.37	-0.86	2.67
	2.41	-0.21	-0.64	-0.55	0.67	-0.23	0.89	-0.09	0.43	0.22	-0.14	-2.37	1.00	2.51	0.84
	2.31	-0.19	-0.58	-0.54	-0.82	-0.20	0.82	-0.02	0.40	0.69	-0.14	-2.31	0.99	2.46	4.20
	2.24	-0.17	-0.54	-0.53	-1.28	-0.14	0.69	0.12	0.33	1.35	0.62	11.15	-3.11	-7.65	-6.09
	1.85	-0.09	-0.31	-0.45	0.77	-0.11	0.63	0.18	0.30	0.49	0.11	2.02	-0.34	-0.79	3.93
2009	-0.00	1.47	0.79	-1.26	-1.60	0.15	-1.56	2.56	-0.15	6.52	-0.29	0.20	0.24	0.85	2.69
	-0.00	1.45	0.83	-1.28	0.73	0.10	-1.08	2.03	-0.06	-0.30	-0.26	0.31	0.03	0.93	0.31
	-0.00	1.46	0.81	-1.28	-1.44	0.11	-1.15	2.11	-0.07	0.92	-0.26	0.30	0.04	0.92	-0.44
	-0.00	1.49	0.77	-1.25	-0.86	0.06	-0.66	1.60	-0.00	-1.90	-0.25	0.39	-0.13	1.00	-1.87
	-0.00	1.45	0.83	-1.28	-0.05	0.03	-0.25	1.19	0.03	1.21	-0.26	0.67	-0.66	1.25	0.82
	-0.00	1.39	0.94	-1.33	-0.48	0.02	0.03	0.93	0.03	0.33	-0.39	1.86	-2.80	2.33	-0.63
2010	-0.00	0.22	0.02	0.75	-0.21	-0.01	2.06	-0.73	-0.31	-1.17	-0.00	2.17	-0.84	-0.33	-1.23
	-0.00	0.23	0.02	0.75	0.50	-0.01	2.04	-0.72	-0.31	-0.00	-0.00	2.18	-0.84	-0.34	-0.05
	-0.00	0.23	0.02	0.75	-0.72	-0.01	2.07	-0.74	-0.31	-2.10	-0.00	2.17	-0.84	-0.33	-2.17
	-0.00	0.22	0.03	0.75	0.31	-0.01	1.97	-0.67	-0.30	-2.02	-0.00	2.18	-0.84	-0.34	-2.31
	-0.00	0.27	-0.01	0.74	-0.86	-0.01	2.11	-0.78	-0.32	-1.49	-0.00	2.18	-0.83	-0.35	-1.51
	-0.00	0.30	-0.03	0.74	-0.71	-0.01	2.39	-1.03	-0.36	-1.00	-0.00	2.17	-0.84	-0.32	-0.97
2011	1.49	-0.90	0.61	-0.20	1.01	0.00	1.07	0.28	-0.35	-0.11	-0.02	0.37	0.66	0.00	0.36
	1.02	-0.91	1.10	-0.20	1.42	0.00	1.09	0.26	-0.35	-0.35	-0.02	0.37	0.64	0.01	0.12
	1.18	-0.92	0.96	-0.22	-1.19	0.00	1.06	0.28	-0.35	0.45	-0.03	0.39	0.62	0.02	0.09
	1.05	-0.92	1.07	-0.21	-1.56	0.00	1.08	0.27	-0.35	0.54	-0.03	0.39	0.62	0.02	0.05
	1.31	-0.92	0.83	-0.21	-2.11	0.00	1.04	0.31	-0.34	1.24	-0.03	0.40	0.61	0.02	0.73
	1.52	-0.89	0.56	-0.19	-0.78	0.00	1.01	0.33	-0.35	0.78	-0.03	0.43	0.56	0.04	0.70
2012	-33.41	5.01	8.21	21.20	53.27	0.04	0.75	0.17	0.04	1.64	0.19	0.12	0.30	0.39	0.96
	31.14	-4.95	-6.39	-18.80	-37.15	0.05	0.76	0.17	0.02	1.23	0.19	0.12	0.30	0.39	0.77
	22.59	-3.62	-4.46	-13.50	-6.39	0.05	0.76	0.17	0.03	0.30	0.17	0.12	0.33	0.39	0.16
	13.19	-2.17	-2.36	-7.67	13.97	0.05	0.77	0.17	0.01	-1.08	0.18	0.12	0.31	0.39	-0.93
	11.36	-1.88	-1.95	-6.53	9.04	0.07	0.79	0.17	-0.03	-0.80	0.23	0.12	0.26	0.38	-0.72
	14.33	-2.34	-2.61	-8.37	-4.71	0.07	0.79	0.17	-0.04	0.33	0.27	0.12	0.23	0.38	0.26
2013	1.06	-0.71	2.04	-1.40	0.44	0.05	1.18	-0.87	0.64	-0.30	0.09	0.19	0.36	0.36	-0.04
	1.40	-1.02	2.53	-1.91	-0.90	0.05	1.17	-0.85	0.63	0.43	0.09	0.19	0.36	0.36	0.03
	1.12	-0.76	2.12	-1.48	0.89	0.05	1.14	-0.77	0.58	-0.64	0.09	0.18	0.37	0.35	-0.17
	0.43	-0.12	1.10	-0.41	-0.26	0.05	1.12	-0.72	0.55	0.92	0.10	0.18	0.37	0.35	0.32
	1.09	-0.73	2.08	-1.43	-5.69	0.05	1.07	-0.60	0.48	2.90	0.10	0.17	0.40	0.33	0.15
	-0.42	0.65	-0.08	0.85	1.06	0.04	1.05	-0.53	0.44	0.94	0.10	0.17	0.40	0.33	0.46
2014	0.01	0.24	0.63	0.13	-1.20	-0.04	2.39	1.34	-2.68	-4.24	0.08	-0.76	1.23	0.45	-3.96
	0.01	0.25	0.63	0.12	0.02	-0.05	2.39	1.34	-2.68	0.29	0.09	-0.36	0.89	0.38	0.14
	0.01	0.25	0.63	0.12	0.35	-0.05	2.39	1.34	-2.68	0.43	0.09	-0.28	0.82	0.37	0.34
	0.01	0.25	0.63	0.12	0.75	-0.05	2.39	1.34	-2.68	1.40	0.09	-0.37	0.90	0.38	0.97
	0.01	0.26	0.62	0.11	1.29	-0.06	2.40	1.34	-2.69	1.29	0.11	0.00	0.59	0.29	1.15
	0.01	0.25	0.62	0.12	-0.01	-0.05	2.39	1.34	-2.68	0.33	0.10	-0.19	0.75	0.34	0.08
2015	0.37	0.36	1.09	-0.82	-1.66	-0.35	0.80	0.34	0.21	-0.13	0.40	0.20	0.29	0.12	-0.06
	0.50	0.54	1.42	-1.46	2.83	-1.35	4.17	-1.20	-0.62	4.29	0.37	0.20	0.24	0.20	-0.21
	0.51	0.55	1.43	-1.48	-0.58	-1.32	4.08	-1.16	-0.60	-0.60	0.36	0.20	0.22	0.22	-0.05
	0.42	0.43	1.21	-1.05	2.27	-0.56	1.51	0.01	0.04	3.29	0.39	0.19	0.27	0.15	1.19
	2.79	3.58	7.06	-12.44	-3.46	0.33	-1.46	1.36	0.77	-0.37	0.31	0.23	0.15	0.31	-0.09
	-1.06	-1.50	-2.36	5.92	2.20	0.17	-0.91	1.11	0.64	-0.17	0.28	0.27	0.10	0.35	-0.04
2016	1.26	-0.15	-0.86	0.75	-1.11	-0.01	1.45	-1.28	0.84	-0.05	0.00	0.09	-0.11	1.02	-0.86
	1.37	-0.29	-0.40	0.32	0.19	-0.01	1.12	-0.89	0.78	0.50	0.00	0.11	-0.10	0.99	0.37
	1.37	-0.30	-0.35	0.28	1.24	-0.01	1.09	-0.85	0.78	-1.09	0.00	0.11	-0.10	0.99	-0.01
	1.36	-0.28	-0.43	0.34	0.80	-0.01	1.03	-0.78	0.76	0.84	0.00	0.13	-0.10	0.97	1.02
	1.37	-0.29	-0.40	0.32	-0.18	-0.01	0.94	-0.67	0.74	-2.72	0.00	0.17	-0.09	0.92	-1.29
	1.37	-0.33	-0.19	0.15	-0.06	-0.01	0.85	-0.56	0.72	2.38	0.00	0.17	-0.09	0.92	0.18
2017	-0.00	-1.51	2.14	0.37	-0.92	0.60	-1.03	-2.37	3.80	-0.65	0.08	0.25	0.46	0.21	0.04
	-0.00	-1.43	2.06	0.37	0.23	0.59	-1.00	-2.32	3.73	0.01	0.10	0.21	0.51	0.18	0.23
	-0.00	-1.09	1.72	0.37	1.21	0.34	-0.24	-1.03	1.93	-0.14	0.14	0.10	0.65	0.11	0.53
	-0.00	-0.71	1.35	0.37	1.56	0.25	0.10	-0.40	1.06	0.52	0.21	-0.08	0.87	0.00	0.85
	-0.00	-0.60	1.23	0.37	1.18	0.24	0.16	-0.27	0.87	-0.08	0.26	-0.20	1.01	-0.07	1.18
	-0.00	-0.30	0.93	0.37	0.28	0.24	0.25	-0.02	0.54	0.07	0.60	-1.08	2.03	-0.55	0.45
ColMean	1.58	-0.18	0.47	-0.88	0.36	-0.02	0.95	0.03	0.04	0.27	0.08	0.53	0.15	0.25	0.07
Colstd	7.15	1.42	2.03	4.75	8.88	0.30	1.18	1.02	1.21	1.58	0.19	1.67	0.81	1.22	1.51

This table demonstrates the out-sample results under various risk-aversion levels of the last six months from 2008- 2017. c_0 represents the weights of the risk-free asset while c_1 c_2 and c_3 show the weights of three sub-portfolios individually. "Return" represents the monthly return. "ColMean" and "ColStd" show the column means and standard deviations, respectively. No restrictions on leverage or short-selling.

7 Conclusion

This paper develops and tests a two-step portfolio selection procedure which relies on a large universe of investable assets and a set of dynamic predictors of factor-related returns. The first step in the procedure creates a collection of well-diversified mimicking portfolios to approximate the returns of pervasive risk factors. The second step uses a set of predictors including default spread, term spread, price trend, and dividend yield. These predictors are combined into a single-index function, which in turn determines a dynamic allocation of portfolio weights across the factor-mimicking portfolios in order to maximize investor's expected utility. Due to the nonstationarity of some predictive variables, we apply orthogonal series to approximate the single-index function in estimation. We apply the technique to fifty years of monthly U.S. data and find very good performance both in-sample and out-of-sample. We show empirically that the factor mimicking portfolios have high correlation with the targeted factors and low correlation with each other. Our dynamic portfolios perform well, both for high risk-aversion and low risk aversion investors, providing high average returns and also high return volatility for the less risk-averse and correspondingly lower average returns and lower volatility for the more risk-averse investor.

References

Y. Aït-sahalia and M. W. Brandt. Variable selection for portfolio choice. *The Journal of Finance*, 56(4):1297–1351, 2001.

M. Ao, L. Yingying, and X. Zheng. Approaching mean-variance efficiency for large portfolios. *The Review of Financial Studies*, 32(7):2890–2919, 2019.

M. W. Brandt. Estimating portfolio and consumption choice: A conditional euler equations approach. *The Journal of Finance*, 54(5):1609–1645, 1999.

M. W. Brandt, P. Santa-Clara, and R. Valkanov. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies*, 22(9):3411–3447, 2009.

J. Y. Campbell and R. J. Shiller. Stock prices, earnings, and expected dividends. *The Journal of Finance*, 43(3): 661–676, 1988.

G. Connor and O. Linton. Semiparametric estimation of a characteristic-based factor model of common stock returns. *Journal of Empirical Finance*, 14(5):694–717, 2007.

G. Connor, M. Hagmann, and O. Linton. Efficient semiparametric estimation of the fama–french model and extensions. *Econometrica*, 80(2):713–754, 2012.

C. Dong, J. Gao, and B. Peng. Semiparametric single-index panel data models with cross-sectional dependence. *Journal of Econometrics*, 188(1):301–312, 2015.

C. Dong, J. Gao, and B. Peng. Another look at single-index models based on series estimation. *Available at SSRN 2858624*, 2016a.

C. Dong, J. Gao, D. Tjøstheim, et al. Estimation for single-index and partially linear single-index integrated models. *The Annals of Statistics*, 44(1):425–453, 2016b.

C. Dong, J. Gao, and O. B. Linton. High dimensional semiparametric moment restriction models. *Available at SSRN 3045063*, 2018.

E. F. Fama and K. R. French. Business conditions and expected returns on stocks and bonds. *Journal of financial economics*, 25(1):23–49, 1989.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56, 1993.

J. Fan, Y. Liao, and M. Mincheva. Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society. Series B, Statistical methodology*, 75(4), 2013.

J. Fan, Y. Liao, and W. Wang. Projected principal component analysis in factor models. *Annals of statistics*, 44 (1):219, 2016.

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

J. Gao and P. C. Phillips. Semiparametric estimation in triangular system equations with nonstationarity. *Journal of Econometrics*, 176(1):59–79, 2013.

J. Gao, C. Phillips, et al. Functional coefficient nonstationary regression with non-and semi parametric cointegration. 2013.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

L. P. Hansen, J. Heaton, and A. Yaron. Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics*, 14(3):262–280, 1996.

R. J. Hodrick. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies*, 5(3):357–386, 1992.

K. Hou, C. Xue, and L. Zhang. Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705, 2015.

D. B. Keim and R. F. Stambaugh. Predicting returns in the stock and bond markets. *Journal of financial Economics*, 17(2):357–390, 1986.

B. T. Kelly, S. Pruitt, and Y. Su. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 2019.

S. Kim, R. A. Korajczyk, and A. Neuhierl. Arbitrage portfolios. *Georgia Tech Scheller College of Business Research Paper*, (18-43), 2019.

O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of multivariate analysis*, 88(2):365–411, 2004.

O. Ledoit and M. Wolf. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12):4349–4388, 2017.

S. Li and O. B. Linton. A dynamic network of arbitrage characteristics. *Available at SSRN 3638105*, 2020.

H. M. Markowitz et al. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

8 Appendix

8.1 Proof

Proof of Theorem 1 : Write the subportfolio vector $\boldsymbol{Q_t}(\boldsymbol{X})$ as:

$$\boldsymbol{Q_t}(\boldsymbol{X}) = \sum_{i=1}^{n} \boldsymbol{B}(\boldsymbol{X_i}) y_{it}.$$

Because:

$$y_{it} = G(\boldsymbol{X_i})\boldsymbol{F_t} + \epsilon_{it}.$$

Then, substitute y_{it} into $\boldsymbol{Q_t}(\boldsymbol{X})$:

$$\begin{aligned} \boldsymbol{Q_t}(\boldsymbol{X}) &= \textstyle\sum_{i=1}^{n} \boldsymbol{B}(\boldsymbol{X_i})(G(\boldsymbol{X_i})\boldsymbol{F_t} + \epsilon_{it}) \\ &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \Gamma G(\boldsymbol{X_i})^\intercal (G(\boldsymbol{X_i})\boldsymbol{F_t} + \epsilon_{it}) \\ &= \Gamma(\tfrac{1}{n}\textstyle\sum_{i=1}^{n}(G(\boldsymbol{X_i})^\intercal G(\boldsymbol{X_i}))\boldsymbol{F_t} + \tfrac{1}{n}\textstyle\sum_{i=1}^{n}(G(\boldsymbol{X_i})^\intercal \epsilon_{it})) \end{aligned}$$

Given:

$$p \lim_{n\to\infty} \frac{1}{n} \sum_{1=1}^{n} G(\boldsymbol{X_i})^\intercal G(\boldsymbol{X_i}) = \boldsymbol{M^G},$$

where $\boldsymbol{M^G}$ is an identity matrix, and

$$E(\epsilon_{it}|\boldsymbol{X_i}, \boldsymbol{F_t}) = 0.$$

Therefore, we have:

$$p \lim_{n\to\infty} \frac{1}{n} \sum_{1=1}^{n} G(\boldsymbol{X_i})^\intercal \epsilon_{it} = 0.$$

Thus,

$$p \lim_{n\to\infty} \boldsymbol{Q_t}(\boldsymbol{X}) = \Gamma \boldsymbol{M^G} \boldsymbol{F_t} = \Gamma \boldsymbol{F_t}.$$

This shows that the factor-mimicking portfolio is a linear combination of risk factors given Γ is a non-zero matrix. □

Proof of Theorem 2 : Let $\tilde{\boldsymbol{F}}$ represent the demeaned risk factor matrix while

$$\tilde{\boldsymbol{y}}_t = G(\boldsymbol{X})\tilde{\boldsymbol{F}}_t + \boldsymbol{\epsilon_t}.$$

Correspondingly, we have:

$$\begin{aligned} \tilde{\boldsymbol{Q}}_t(\boldsymbol{X}) &= \boldsymbol{B}(\boldsymbol{X}_i)\tilde{\boldsymbol{y}}_t \\ &= \tfrac{1}{n}\Gamma G(\boldsymbol{X})^\intercal \tilde{\boldsymbol{y}}_t. \end{aligned}$$

And then,

$$\begin{aligned} E(\tilde{\boldsymbol{Q}}_t(\boldsymbol{X})\tilde{\boldsymbol{Q}}_t(\boldsymbol{X})^\intercal|\boldsymbol{X}) = \;& \Gamma(\tfrac{1}{n}G(\boldsymbol{X})^\intercal G(\boldsymbol{X}))E(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^\intercal)(\tfrac{1}{n}G(\boldsymbol{X})^\intercal G(\boldsymbol{X}))\Gamma^\intercal + \\ & \Gamma(\tfrac{1}{n}G(\boldsymbol{X})^\intercal \tfrac{1}{n}E(\boldsymbol{\epsilon_t}\boldsymbol{\epsilon_t}^\intercal)(G(\boldsymbol{X}))\Gamma^\intercal. \end{aligned}$$

Given $E(\epsilon_{it}|\boldsymbol{X_i}, \boldsymbol{F_t}) = 0$, the cross terms are $E(\tilde{\boldsymbol{F}}\tilde{\boldsymbol{\epsilon}}_t) = 0$.

Taking the second term and using the Euclidian matrix norm:

$$||\Gamma(\tfrac{1}{n}G(\boldsymbol{X})^\intercal \tfrac{1}{n}E(\tilde{\boldsymbol{\epsilon}}_t\tilde{\boldsymbol{\epsilon}}_t{}^\intercal)G(\boldsymbol{X}))\Gamma^\intercal|| \leqslant$$
$$\tfrac{1}{n}||\Gamma\tfrac{1}{n}(G(\boldsymbol{X})^\intercal G(\boldsymbol{X}))\Gamma^\intercal|| \times ||E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^\intercal)|| \rightarrow_{n\rightarrow\infty}$$
$$\tfrac{1}{n}||\Gamma\Gamma^\intercal|| \times ||E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^\intercal)|| \rightarrow_{n\rightarrow\infty} 0$$

The conclusion of the above formula is due to

$$p \lim_{n\rightarrow\infty} \frac{1}{n}\sum_{1=1}^{n} G(\boldsymbol{X_i})^\intercal G(\boldsymbol{X_i}) = \boldsymbol{M^G},$$

and $||E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t^\intercal)||$ has bounded eigenvalues for all n.

Furthermore, the well-chosen coefficient matrix Γ can give:

$$E(\tilde{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X})\tilde{\boldsymbol{Q}}_{\boldsymbol{t}}(\boldsymbol{X})^\intercal|\boldsymbol{X}) = \boldsymbol{I}_{JJ}$$

□

Proof of Theorem 3 : We decompose the investment returns of optimal asset-by-asset portfolio and risk-free rate as:

$$r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^*},$$

where the R_F is the optimal factor returns since the return generation function states the risk premiums come from risk factors. The $\boldsymbol{\epsilon_t^*}$ is the zero mean idiosyncratic returns.

Since

$$E(\epsilon_{it}|\boldsymbol{X_i}, \boldsymbol{F_t}) = 0,$$

it follows from the second-order stochastic dominance that the expected utility has the following relationship:

$$E(u(r_{ft} + \boldsymbol{R_F})) > E(u(r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^*})).$$

because zero mean $\boldsymbol{\epsilon_t^*}$ only contributes variance rather than returns.

According to Theorem 1 and Equation 17, the restricted portfolio optimally combines the factors' returns. Therefore, our two-stage portfolio's return can be written as:

$$r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^{**}},$$

where the only difference is the idiosyncratic returns. Our goal now is to show that:

$$E(u(r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^{**}})|\boldsymbol{X}, \boldsymbol{z_t}) \rightarrow^{n\rightarrow\infty} E(u(r_{ft} + \boldsymbol{R_F})|\boldsymbol{X}, \boldsymbol{z_t}).$$

Next, we take the Taylor expansion of $u(r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^{**}})$ around $r_{ft} + \boldsymbol{R_F}$:

$$u(r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^{**}}) = u(r_{ft} + \boldsymbol{R_F}) + \frac{d}{d(r_{ft} + \boldsymbol{R_F})} u(r_{ft} + \boldsymbol{R_F})\boldsymbol{\epsilon_t^{**}} + \frac{d^2}{d(r_{ft} + \boldsymbol{R_F})} u(r_{ft} + \boldsymbol{R_F})^2 \boldsymbol{\epsilon_t^{**2}}.$$

We take the expection on both sides, given $E(\boldsymbol{\epsilon_t^{**}}) = 0$ and $\frac{d^2 u(\cdot)}{dW^2} \geqslant -c$. Therefore, we have:

$$E(u(r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^{**}})) \geqslant E(u(r_{ft} + \boldsymbol{R_F})) - cE(\boldsymbol{\epsilon_t^{**2}}),$$

where $p\lim_{n\rightarrow\infty} E(\boldsymbol{\epsilon_t^{**2}}) = 0$, according to Theorem2. Therefore, we have :

$$p\lim_{n\rightarrow\infty} E(u(r_{ft} + \boldsymbol{R_F} + \boldsymbol{\epsilon_t^{**}})|\boldsymbol{X}, \boldsymbol{z_t}) - E(u(r_{ft} + \boldsymbol{R_F})|\boldsymbol{X}, \boldsymbol{z_t}) = 0,$$

which completes the proof. □

Table 7: Annual Correlation Between Subportfolios and Risk Factors 1-20

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
f_1	0.61	-0.49	0.68	0.69	0.66	0.73	-0.10	-0.16	0.72	0.79	0.73	0.69	0.77	-0.17	0.64	0.87	0.65	0.71	0.51	0.75
f_2	0.89	0.64	0.70	0.75	0.65	0.79	0.60	0.91	0.71	0.79	0.90	0.92	0.97	-0.01	0.79	0.74	0.79	0.86	0.88	0.41
f_3	0.80	0.81	0.63	0.40	0.66	0.39	0.75	0.64	0.40	0.72	0.86	0.74	0.96	0.56	0.72	0.86	0.84	0.66	0.78	0.55

This table shows the annual correlation between factor-mimicking subportfolios and corresponding risk factors from Jul. 1967- Jun.1987.

Table 8: Annual Correlation Between Subportfolios and Risk Factors 21-40

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
f_1	0.78	0.86	0.54	0.58	0.22	0.76	0.68	0.16	0.60	0.74	0.75	0.67	0.72	0.73	0.69	0.67	0.96	0.66	0.73	0.76
f_2	0.77	0.79	0.73	-0.24	0.93	0.66	0.86	-0.06	0.87	0.81	0.94	0.76	0.22	-0.06	0.97	0.94	0.92	0.36	0.93	0.93
f_3	-0.40	0.72	0.76	0.64	0.71	0.83	0.70	0.57	0.73	0.67	0.72	0.58	0.66	0.60	0.79	0.77	0.92	0.68	0.79	0.72

This table shows the annual correlation between factor-mimicking subportfolios and corresponding risk factors from Jul. 1987- Jun.2007.

Table 9: Annual Correlation Between Subportfolios and Risk Factors 41-50

	41	42	43	44	45	46	47	48	49	50
f_1	-0.22	0.73	-0.15	0.72	-0.55	-0.08	0.33	0.74	0.63	0.77
f_2	0.63	0.76	0.29	0.73	0.65	0.34	-0.23	0.81	0.83	0.81
f_3	0.64	0.58	0.65	0.61	0.82	0.51	0.87	0.86	0.82	0.53

This table shows the annual correlation between factor-mimicking subportfolios and corresponding risk factors from Jul. 2007- Jun.2017.