



Diego Siebra de Brito

**Forecasting Large Realized Covariance
Matrices: The Benefits of Factor Models and
Shrinkage**

Dissertação de Mestrado

Dissertation presented to the Programa de Pós-graduação em
Economia of PUC-Rio in partial fulfillment of the requirements
for the degree of Mestre em Economia.

Advisor : Prof. Marcelo Cunha Medeiros
Co-advisor: Prof. Ruy Monteiro Ribeiro

Rio de Janeiro
April 2018



Diego Siebra de Brito

**Forecasting Large Realized Covariance
Matrices: The Benefits of Factor Models and
Shrinkage**

Dissertation presented to the Programa de Pós-graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the undersigned Examination Committee.

Prof. Marcelo Cunha Medeiros

Advisor

Departamento de Economia – PUC-Rio

Prof. Ruy Monteiro Ribeiro

Co-advisor

Departamento de Economia – PUC-Rio

Prof. Diogo Abry Guillén

Departamento de Economia – PUC-Rio

Prof. Marcelo Fernandes

Escola de Economia de São Paulo – FGV

Prof. Augusto Cesar Pinheiro da Silva

Vice Dean of Graduate Studies

Centro de Ciências Sociais – PUC-Rio

Rio de Janeiro, April 9th, 2018

All rights reserved.

Diego Siebra de Brito

B.S., Aeronautical Engineering, Instituto Tecnológico de Aeronáutica (ITA), 2013

Bibliographic data

Brito, Diego S

Forecasting Large Realized Covariance Matrices: The Benefits of Factor Models and Shrinkage / Diego Siebra de Brito; advisor: Prof. Marcelo Cunha Medeiros; co-advisor: Prof. Ruy Monteiro Ribeiro. – Rio de Janeiro: PUC-Rio, Departamento de Economia, 2018.

v., 60 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

Inclui bibliografia

1. Economia – Teses. 2. Economia – Teses. 3. Covariância Realizada;. 4. Modelo de Fatores;. 5. LASSO;. 6. Previsão;. 7. Alocação de Portfólio;. 8. Big Data. I. Medeiros, Marcelo Cunha. II. Ribeiro, Ruy Monteiro. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 330

Acknowledgments

I would like to first thank my parents, Giana and Hugo, for all their support throughout the years and all the incentive they gave me to continue my education. To my girlfriend Dianna, thank you for all the support during the last months.

I thank my advisors, Marcelo and Ruy, for all the support, ideas, and complete availability for meetings regarding this project or any other subject.

To my friends from "Favelinha", this was only possible because we were all together!

I thank CNPq, PUC-Rio and Vinci Partners for all the financial support. Finally, I thank ANBIMA for their financial support through the "XIII Prêmio ANBIMA de Mercado de Capitais".

Abstract

Brito, Diego S; Medeiros, Marcelo Cunha (Advisor); Ribeiro, Ruy Monteiro (Co-Advisor). **Forecasting Large Realized Covariance Matrices: The Benefits of Factor Models and Shrinkage**. Rio de Janeiro, 2018. 60p. Dissertação de mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

We propose a model to forecast very large realized covariance matrices of returns, applying it to the constituents of the S&P 500 on a daily basis. To deal with the curse of dimensionality, we decompose the return covariance matrix using standard firm-level factors (e.g. size, value, profitability) and use sectoral restrictions in the residual covariance matrix. This restricted model is then estimated using Vector Heterogeneous Autoregressive (VHAR) models estimated with the Least Absolute Shrinkage and Selection Operator (LASSO). Our methodology improves forecasting precision relative to standard benchmarks and leads to better estimates of the minimum variance portfolios.

Keywords

Realized Covariance; Factor Model; LASSO; Forecasting; Portfolio Allocation; Big Data

Resumo

Brito, Diego S; Medeiros, Marcelo Cunha; Ribeiro, Ruy Monteiro. **Previsão de Matrizes de Covariância Realizada de Alta Dimensão: Os Benefícios de Modelos de Fatores e Shrinkage.** Rio de Janeiro, 2018. 60p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho propõe um modelo de previsão de matrizes de covariância realizada de altíssima dimensão, com aplicação para os componentes do índice S&P 500. Para lidar com o altíssimo número de parâmetros (maldição da dimensionalidade), propõe-se a decomposição da matriz de covariância de retornos por meio do uso de um modelo de fatores padrão (e.g. tamanho, valor, investimento) e uso de restrições setoriais na matriz de covariância residual. O modelo restrito é estimado usando uma especificação de vetores auto regressivos heterogêneos (VHAR) estimados com LASSO (Least Absolute Shrinkage and Selection Operator). O uso da metodologia proposta melhora a precisão de previsão em relação a benchmarks padrões e leva a melhores estimativas de portfólios de menor variância.

Palavras-chave

Covariância Realizada; Modelo de Fatores; LASSO; Previsão; Alocação de Portfólio; Big Data

Table of contents

1	Introduction	10
2	Model	13
2.1	Factor Model	13
2.2	Covariance Decomposition	14
3	Forecasting Methodology	16
3.1	Forecasting $\Sigma_{f,t}$	16
3.1.1	LASSO and adaLASSO	17
3.2	Forecasting B_t	18
3.3	Forecasting $\Sigma_{\varepsilon,t}$	18
4	Data and Descriptive Analysis	20
4.1	Data Cleaning	20
4.2	Factors	20
4.3	Sector Classification	21
4.4	Residual Covariance Matrices	21
5	Forecasting Results	24
5.1	Factor Covariance Matrix	24
5.2	Complete Covariance Matrix	27
6	Portfolio Selection	29
6.1	Global Minimum Variance Portfolios	30
6.2	Restricted Minimum Variance Portfolios	32
6.3	Long-Only Minimum Variance Portfolios	34
7	Conclusions	35
	Bibliography	36
A	Factor Loadings	38
B	Betas' Long Memory	40
C	Factor Construction	43
D	Diagnostic Criterion for Approximate Factor Structure	46
E	Residual Covariance Matrix - Parameter Selection	48
F	Minimum Variance Portfolio: Partial Rebalancing	57

List of figures

Figure 4.1	Realized Raw Correlations	22
Figure 4.2	Realized Residual Correlations	23
Figure 5.1	Average Number of Selected Variables in FHAR Log-matrix LASSO models	25
Figure 5.2	Average Change in the Number of Selected Vari- ables in FHAR Log-matrix models	26
Figure B.1	Estimated Fractional Differencing Parameter (d) - Whittle Method	41
Figure B.2	Estimated Fractional Differencing Parameter (d) - GPH Method	42
Figure D.1	Approximate Factor Structure for the 1F, 3F, 5F, and 7F Models	47
Figure D.2	Approximate Factor Structure for the 7F + 10 Industry Factors Model	47

List of tables

Table 4.1	Number of Stocks per Sector	22
Table 5.1	Forecast Precision for Factor Covariance Matrices	28
Table 5.2	Forecast Precision for Complete Covariance Matrices	28
Table 6.1	S&P Index Performances	30
Table 6.2	Statistics for Daily Portfolios - Global Minimum Variance	33
Table 6.3	Statistics for Daily Portfolios - Restricted Minimum Variance	33
Table 6.4	Statistics for Daily Portfolios - Restricted Minimum Variance (Long Only)	34
Table E.1	Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (1-Factor VHAR).	49
Table E.2	Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (1-Factor VHAR).	50
Table E.3	Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (3-Factor VHAR).	51
Table E.4	Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (3-Factor VHAR).	52
Table E.5	Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (5-Factor VHAR).	53
Table E.6	Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (5-Factor VHAR).	54
Table E.7	Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (7-Factor HVAR).	55
Table E.8	Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (7-Factor HVAR).	56
Table F.1	Statistics for Daily Average Portfolios - Global Minimum Variance.	58
Table F.2	Statistics for Daily Average Portfolios - Restricted Minimum Variance.	59
Table F.3	Statistics for Daily Average Portfolios - Restricted Minimum Variance (Long Only).	60

1 Introduction

The goal of this paper is to construct models based on economically motivated factor decompositions and shrinkage methods to forecast large-dimensional and time-varying realized measures of daily covariance matrices of returns on financial assets. Realized measures of a covariance matrix are estimates, based on intraday returns, of the integrated covariance matrix of a multivariate diffusion process. One example of such an estimator used in this paper is the composite realized kernel method recently introduced by (1). Our proposed model is evaluated in terms of both its forecasting ability and, more importantly, several performance measures in a conditional mean-variance portfolio allocation problem.

Modeling and forecasting the covariance matrix of financial assets are essential for portfolio allocation and risk management. Moreover, it is an established empirical fact that (conditional) covariance matrices vary considerably over time. A natural way to model such dynamics is to either use multivariate generalizations of the ARCH/GARCH family of models, as proposed by (2) or (3), or model directly a given realized measure of the covariance matrices by usual multivariate time-series models, as in (4), (5) or (6). This is motivated by the close connection between the conditional covariance matrix and the integrated covariance of a multivariate diffusion process.

However, when the number of assets increases, the amount of parameters to be estimated becomes very large. For instance, for a covariance matrix of N assets, there are $N(N + 1)/2$ distinct entries to be modeled. If a vector autoregressive specification of order p , $\text{VAR}(p)$, is used, then the total number of parameters will be $(N + 1)(p + 1)/2$. Therefore, the curse of dimensionality precludes the application of the above referenced methods to moderately large covariance matrices, and most of the previous studies in the literature focused on sets of less than ten assets.

More recently, based on the advances of modern statistical tools to handle high-dimensional models, new alternatives have been proposed in the literature in order to address a large number of assets. (7) advocate the use of the least absolute shrinkage and selection operator (LASSO) of (8) and the adaptive LASSO of (9) to model the dynamics of the realized covariance of

the constituents of the Dow Jones index by a large dimensional VAR model. However, their modeling strategy cannot handle sets of assets much larger than the 30 used in their paper. (10) combine nonlinear shrinkage with the DCC-GARCH model and put forward a methodology in which the dynamic of huge latent conditional covariance matrices can be modeled. Their approach makes the estimation of multivariate GARCH-type models feasible even in large-dimensional cases. Nevertheless, the authors did not consider realized measures in their models.

In this paper, we propose a method that can be applied to hundreds or even thousands of assets with the difference that we consider realized covariance measures rather than latent ones, as in (10). We provide empirical evidence that our proposed model drastically improves the forecasting ability and produces portfolios with better performance measures. We apply our model to the constituents of the S&P 500 using daily information on the realized covariance matrix. The models based on our framework are able to improve the quality of the forecasts when compared to the random walk benchmark, by reducing the ℓ_2 -forecast error in 15%. On the other hand, models based on latent volatility are the worst-performing specification in terms of forecasting ability. Furthermore, our models deliver a significant reduction in the volatility of minimum variance portfolios, as the the annualized standard deviation of returns declines by more than 25% relative to the second best alternative in the literature and to nearly 40% versus the random walk alternative. We also observe a decline in realized portfolio risk when we consider longer investment horizons of up to one month and when we introduce portfolio constraints.

Our model applies a methodology similar to (7) to the elements of the factor decomposition of realized covariance matrices. Hence, the daily covariance matrix of returns is first written into a covariance matrix for a low-dimensional set of factors plus an idiosyncratic covariance matrix that is (almost) block diagonal. We consider economically motivated factors that are widely used in the finance literature, based on firm-level characteristics such as size, value and profitability. The dynamics of the variances and covariances of the factors are modeled by a VHAR model estimated with LASSO. To guarantee positive definiteness of the forecasts, we apply the log matrix transformation of (11). The daily factor loadings (“betas”) are computed with high-frequency data. We show that these loadings are time varying and exhibit a high degree of long-range dependence, and we model their dynamics by an HAR specification as well. Finally, the dynamics of each block of the idiosyncratic covariance matrix is modeled by a restricted autoregressive model

also estimated by LASSO.

To the best of our knowledge, this is the first model that can describe the dynamics of realized covariance matrices of such a large dimension. (12) also considered factors and realized covariances; however, their forecasting model is just a random walk, and their results are more concerned with the estimation of the integrated covariance.

In this paper, we do not address the interesting challenges involved in the construction of realized measures (e.g., how to address microstructure noise), specially in the multivariate case (e.g., asynchronicity in transactions of different assets, which bias covariance estimations toward zero). We focus solely on modeling and forecasting a large realized covariance matrix of returns on hundreds of financial assets, which are estimated elsewhere.

The remainder of paper is organized as follows. Chapter 2 describes the proposed model and forecasting framework. In Chapter 4, we present the data and show a descriptive analysis. Forecasting results and portfolio analysis are presented, respectively, in Chapter 5 and 6. Finally, Chapter 7 concludes the paper. Supplementary results are shown in the Appendix.

2 Model

We base our methodology on the fact that realized covariance matrices are highly persistent over time, which suggests the use of an autoregressive model of a large order p , usually larger than 20. Defining $y_t = \text{vech}(\Sigma_t)$, where vech is the half-vectorization operator returning a vector with the unique entries of Σ_t , one possible specification is

$$y_t = \omega + \sum_{i=1}^p \Phi_i y_{t-1} + \epsilon_t, \quad (2-1)$$

where ϵ_t is zero-mean random noise.

While this model is sensible for a small number of assets, the number of parameters grows quadratically as new assets are added (curse of dimensionality). To circumvent this limitation, (7) propose the use of penalized regressions (LASSO) as a way to address the large number of parameters. However, the direct use of penalized regressions is unfeasible for hundreds of assets.

To illustrate, assume that the representation in (2-1) is used to model the covariance matrix for the constituents of the S&P 500 index¹ with 10 lags, that is, $p = 10$. Since each matrix Σ_t has $N(N + 1)/2$ unique entries, this configuration would result in 125,250 equations with $10 \times 125,250$ variables each, plus constants. In this case, estimation is unfeasible even with LASSO.

To reduce the dimensionality to a manageable one, we propose the use of a factor model as well as economic restrictions based on sector classifications and penalized regressions.

2.1 Factor Model

Following the factor model discussed in (13), the excess return on any asset i , $r_{i,t}$, satisfies

$$r_{i,t}^e = \beta_{i1,t} f_{1,t} + \cdots + \beta_{iK,t} f_{K,t} + \varepsilon_{i,t} = \beta_{i,t}' f_t + \varepsilon_{i,t}, \quad (2-2)$$

¹In our application, we restrict the analysis to stocks that remain in the index for the entire period of our sample. This reduces the number of stocks to 430. In terms of the number of equations and potential predicting variables, our problem is still subject to the curse of dimensionality on the same order of magnitude as the illustration.

where $f_{1,t}, \dots, f_{K,t}$ are the excess returns of K factors, $\beta_{ik,t}$, $k = 1, \dots, K$, are factor loadings for asset i , and $\varepsilon_{i,t}$ is the idiosyncratic error term. Note that the factor loadings are variable over time. For N assets, the set of equations can be written in matrix form:

$$r_t^e = B_t' f_t + \varepsilon_t, \quad (2-3)$$

where B_t is a $K \times N$ matrix of loadings, r_t is an $N \times 1$ vector of excess returns, and ε_t is an $N \times 1$ vector of idiosyncratic errors. Throughout, we assume that $(\varepsilon_t | f_t) = 0$. The factors used in this work are linear combinations of returns constructed solely with the assets considered here, i.e., long-short stock portfolios where stocks that are part of our sample are sorted on firm characteristics. In matrix form, for all K factors,

$$\begin{pmatrix} f_{1,t} \\ \vdots \\ f_{K,t} \end{pmatrix} = \begin{pmatrix} w_{1,1} & \dots & w_{1,N} \\ \vdots & \ddots & \vdots \\ w_{K,1} & \dots & w_{K,N} \end{pmatrix} \begin{pmatrix} r_{1,t} \\ \vdots \\ r_{N,t} \end{pmatrix} \quad (2-4)$$

or

$$f_t = W' r_t,$$

where weights are calculated based on accounting and market information, as we describe in Section 4.

2.2

Covariance Decomposition

This section describes how we decompose the realized covariance matrix of returns for all assets into two components: a factor covariance matrix and a residual covariance matrix. Let Σ_t denote the realized covariance matrix of returns at time t , that is, $\Sigma_t = \text{cov}(r_t)$. By using equation 2-3 and the assumption $(\varepsilon_t | f_t) = 0$, we have

$$\Sigma_t = \text{cov}(B_t' f_t) + \text{cov}(\varepsilon_t) = B_t' \Sigma_{f,t} B_t + \Sigma_{\varepsilon,t}, \quad (2-5)$$

where B_t is a $K \times N$ matrix of loadings of N assets on K factors, $\Sigma_{f,t}$ is the $K \times K$ factor covariance matrix and $\Sigma_{\varepsilon,t}$ is the $N \times N$ residual covariance matrix, all at time t . Since each factor is a linear combination of returns, factor covariance matrices can be obtained by using equation 2-4 and the known values of Σ_t , that is,

$$\Sigma_{f,t} = \text{cov}(f_t) = \text{cov}(W' r_t) = W' \Sigma_t W. \quad (2-6)$$

Factor loadings B_t are calculated using a similar procedure (see Appendix A), and the values of $\Sigma_{\varepsilon,t}$ are simply given by the difference $\Sigma_t - B_t' \Sigma_{f,t} B_t$. It

is common to assume that $\Sigma_{\varepsilon,t}$ is diagonal, but we will be less restrictive in this work.

3 Forecasting Methodology

With the decomposition achieved in (2-5), we write the forecasting equation for the complete covariance matrix Σ_t by combining separate forecasts of B_t , $\Sigma_{f,t}$, and $\Sigma_{\varepsilon,t}$. That is,

$$\widehat{\Sigma}_{t+1|t} = \widehat{B}'_{t+1|t} \widehat{\Sigma}_{f,t+1|t} \widehat{B}_{t+1|t} + \widehat{\Sigma}_{\varepsilon,t+1|t}. \quad (3-1)$$

3.1 Forecasting $\Sigma_{f,t}$

Since the number of factors is much smaller than the number of assets, one could propose the use of an unrestricted VAR model for the factor covariance matrix dynamics, as in equation (2-1). Despite the reduction in dimensionality achieved by using a factor model, the number of parameters in this configuration is still quite large (note that each equation would have $K(K+1)/2+1$ parameters). To reduce this number to a more manageable one, we follow the heterogeneous autoregressive (HAR) model proposed by (14). In this model, the predictors are obtained from the simple average of daily data, computed for different horizons (daily, weekly, and monthly). In our case, daily, weekly, and realized covariance matrices of factors are given by

$$\begin{aligned} \Sigma_{f,t}^{day} &= \Sigma_{f,t} \\ \Sigma_{f,t}^{week} &= \frac{1}{5}(\Sigma_{f,t} + \Sigma_{f,t-1} + \cdots + \Sigma_{f,t-4}) \\ \Sigma_{f,t}^{month} &= \frac{1}{22}(\Sigma_{f,t} + \Sigma_{f,t-1} + \cdots + \Sigma_{f,t-21}), \end{aligned} \quad (3-2)$$

where $y_{f,t} = \text{vech}(\Sigma_{f,t})$, $y_{f,t}^{day} = \text{vech}(\Sigma_{f,t}^{day})$, $y_{f,t}^{week} = \text{vech}(\Sigma_{f,t}^{week})$, and $y_{f,t}^{month} = \text{vech}(\Sigma_{f,t}^{month})$. Furthermore, the dynamic process for $y_{f,t}$ is defined as

$$y_{f,t} = \omega + \Phi_{day} y_{f,t-1}^{day} + \Phi_{week} y_{f,t-1}^{week} + \Phi_{month} y_{f,t-1}^{month} + \epsilon_t, \quad (3-3)$$

where Φ_{day} , Φ_{week} , and Φ_{month} are $M \times M$ matrices, where $M = K(K+1)/2$ is the number of unique entries on the factor covariance matrix. ω is an $M \times 1$ vector of constants.

3.1.1

LASSO and adaLASSO

Due to the high number of parameters in equation (3-3), direct estimation with ordinary least squares (OLS) could result in overfitting, harming the precision of model forecasts. LASSO shrinks these estimates by imposing a penalty related to the magnitude of the coefficients. This estimation effectively sets most estimates to zero ((8)). This methodology has been shown to provide a higher out-of-sample forecasting precision, while the reduced number of predictors makes the interpretation of the model easier.

We estimate (3-3) equation by equation. Consider a sample size of T , and let $\mathbf{Z}_t = (1, \mathbf{y}_{f,t-1}^{day'}, \mathbf{y}_{f,t-1}^{week'}, \mathbf{y}_{f,t-1}^{month'})' \in^{3m+1}$ be the vector of explanatory variables and $Z = (Z_T, \dots, Z_1)'$ be the $T \times (3m + 1)$ matrix of covariates. Let $y_{f,i} = (y_{f,T,i}, \dots, y_{f,1,i})' \in^T$ be the vector of observations on the i th equation of (3-3), $i = 1, \dots, M$, and $\epsilon_{f,i} = (\epsilon_{f,T,i}, \dots, \epsilon_{f,1,i})'$ be the corresponding vector of error terms. With $\gamma_{f,i} = (\omega_i, \beta_i)'$ being the $3M + 1$ vector of true parameters for equation i , one can write

$$y_{f,i} = Z\gamma_{f,i} + \epsilon_{f,i}, \quad i = 1, \dots, M, \quad (3-4)$$

where each vector $\gamma_{f,i}$ is then estimated by minimizing

$$L(\gamma_{f,i}) = \frac{1}{T} \|y_{f,i} - Z\gamma_{f,i}\|^2 + 2\lambda_T \|\beta_i\|_{\ell_1}, \quad i = 1, \dots, M. \quad (3-5)$$

The penalty parameter λ_T determines how much penalization is imposed on the size of the coefficients. In our setup, the value of λ_T is determined by minimizing the Bayesian information criterion (BIC). For equation i and penalty parameter λ , the BIC is given by

$$\text{BIC}_i(\lambda) = T \times \log(\hat{\epsilon}'_{\lambda,i} \hat{\epsilon}_{\lambda,i}) + \sum_{j=1}^{3M} 1(\hat{\beta}_{ij}^\lambda \neq 0) \log(T), \quad i = 1, \dots, M, \quad (3-6)$$

where $\hat{\epsilon}_{\lambda,i}$ is the estimated vector of error terms for penalty λ .

After obtaining $\hat{\gamma}_{f,i}$ from equation (3-5), the one-step-ahead forecast is given by

$$\hat{y}_{f,T+1,i} = \hat{\gamma}'_{f,i} \mathbf{Z}_T, \quad i = 1, \dots, M, \quad (3-7)$$

which can be used to provide $\hat{\Sigma}_{f,T+1}$.

We also consider the adaptive version of LASSO (adaLASSO), as in (9). adaLASSO is a two-step procedure that consists of estimating the model via LASSO in the first step and excluding the variables classified as zero from the second step. The second step estimates the parameters with a slightly modified objective function that takes into account the size of the parameters estimated in the first step. To illustrate, consider the set of indices of the

coefficients in the i th equation that are different from zero in the first step: $J(\hat{\beta}_i) = \{j \in R^{3M} : \hat{\beta}_{i,j} \neq 0\}$. adaLASSO estimates the vector of parameters $\gamma_{f,i}$ by minimizing the following objective function:

$$L(\gamma_{f,i}) = \frac{1}{T} \|y_{f,i} - Z\gamma_{f,i}\|^2 + 2\lambda_T \sum_{j \in J(\hat{\beta}_i)} \frac{|\beta_{i,j}|}{|\hat{\beta}_{i,j}|}, \quad i = 1, \dots, M. \quad (3-8)$$

As before, the penalty parameter λ_T is chosen by minimizing the BIC, and the one-step-ahead forecasts are computed as in equation (3-7).

3.2

Forecasting B_t

Instead of using unconditional loadings on factors, we assume that betas change daily and have long-range dependence. In Appendix B, we show the distribution of the estimates of the fractional integration parameter for the beta series. The results show that these series present high persistence, similar to what is observed on realized covariance data. Based on this evidence, and to maintain consistency with the rest of our methodology, we forecast each series of betas using HAR models. For each element of B_t , we have

$$\beta_{k,i,t} = \omega + \Phi_{day}\beta_{k,i,t-1}^{day} + \Phi_{week}\beta_{k,i,t-1}^{week} + \Phi_{month}\beta_{k,i,t-1}^{month} + \epsilon_{k,i,t}, \quad (3-9)$$

where $\beta_{k,i,t}$, the entry $k \times i$ in the matrix B_t , is the loading of stock i on factor k at date t . We estimate the coefficients Φ_{day} , Φ_{week} , and Φ_{month} by OLS.

3.3

Forecasting $\Sigma_{\varepsilon,t}$

Since the residual covariance matrix dimension is $N \times N$, the curse of dimensionality remains a concern when forecasting $\Sigma_{\varepsilon,t}$. A factor model may not explain all the covariances between assets depending on the number and choice of factors, implying the the covariance matrix of the residuals is not diagonal. Instead of imposing the common hypothesis that $\Sigma_{\varepsilon,t}$ is diagonal (as in (15)), we consider a less restrictive assumption that stocks in the same sector may still co-vary even after controlling for the standard factors used in the finance literature.

We assume that $\Sigma_{\varepsilon,t}$ is block diagonal, where blocks are defined by industry classification. We assume that there is no contemporaneous correlation between assets in different sectors after controlling for factor exposure. We argue that this assumption is not too strong in Section 4.4, where we analyze residual correlations between stocks in different sectors.

We assume that each asset belongs to an industrial sector s , where S is the total number of sectors and $S \ll N$. Without loss of generality, we can order assets so that the residual covariance matrix has sector covariance matrices near its diagonal:

$$\Sigma_{\varepsilon,t} = \begin{pmatrix} \Sigma_{\varepsilon,t}^1 & & \\ & \ddots & \\ & & \Sigma_{\varepsilon,t}^S \end{pmatrix}. \quad (3-10)$$

Depending on the selected industry classification, the number of assets on each group s (called N^s) can still be quite large, as each block has $M^s = N^s(N^s + 1)/2$ unique elements.

Our second assumption is that the dynamics of each block $\Sigma_{\varepsilon,t}^s$ depend only on the elements of the same block at $t-1$, i.e., only $\Sigma_{\varepsilon,t-1}^s$ predict values in $\Sigma_{\varepsilon,t}^s$. To further simplify this model, we also consider an additional restriction that only past variances matter in this prediction. This last assumption relies on the previous evidence in (7) that past variances are more frequently selected by LASSO as good predictors of covariance and variance terms.

With the notation $y_{\varepsilon,t}^s = \text{vech}(\Sigma_{\varepsilon,t}^s)$, we have

$$y_{\varepsilon,t}^s = \omega_{\varepsilon}^s + \Phi^s \Lambda_{\varepsilon,t-1}^s + u_{\varepsilon,t}^s, \quad s = 1, \dots, S, \quad (3-11)$$

where ω_{ε}^s is an $M^s \times 1$ vector of intercepts, Φ^s is an $M^s \times N^s$ matrix of coefficients, $\Lambda_{\varepsilon,t-1}^s = \text{diag}(\Sigma_{\varepsilon,t-1}^s)$ is an $N^s \times 1$ vector of past variances, and $u_{\varepsilon,t}^s$ is the vector of errors.

The parameters estimation is done block by block, that is, each equation in (3-11) is estimated separately. The procedure is the same as the one used for the factor covariance matrix model: LASSO/adaLASSO regression equation by equation. We then regroup the one-day ahead forecast for each group, $\hat{\Sigma}_{\varepsilon,T+1}^s$, to form the full residual covariance matrix forecast $\hat{\Sigma}_{\varepsilon,T+1}$.

4 Data and Descriptive Analysis

The data consist of daily realized covariance matrices of returns for constituents of the S&P 500 index. These matrices were constructed using 5-minute returns by the composite realized kernel method (discussed in (1) and provided by the authors). The full sample comprises all business days between January 2006 and December 2011. We consider companies that remained in the index and had balance sheet data available for the full sample period, resulting in a total of 430 stocks. With these considerations, the dataset consists of 1,495 daily 430×430 realized covariance matrices of returns.

4.1 Data Cleaning

Our sample spans the 2008 financial crisis as well as the flash crashes from 2010 and 2011. To mitigate the effect of these extreme events, we perform a light cleaning on the realized covariance matrices, as in (7). For each day, we verify which (unique) entries are more than 4 standard errors (of the series corresponding to that entry) away from their sample average up to then. If more than 25% of the unique entries have extreme values according to this criterion, we flag that specific day. We replace the matrices corresponding to the flagged days by an average of the nearest five preceding and five following non-flagged matrices.

4.2 Factors

We construct each factor as the return time series for a long-short stock portfolio derived from individual sorts of the underlying stocks on different signals. Since our approach uses the loading matrix W as an input to calculate factor covariance matrices and factor loadings, we could not use the widely available data on financial factors series of returns (such as on Kenneth French's website). Instead, we construct our own versions of the standard factors, ranking our universe of stocks into the different signals to calculate the matrices W for our sample.

In total, we use seven factors that have been widely used in the finance literature. Besides the market factor, we also consider size (SMB) and value (HML) (16), gross profitability (17), investment (18), asset growth (19), and accruals (20). We report a detailed description of the construction of each factor in Appendix C. We use four different combinations of factors with one, three, five, and seven factors. They are denoted by 1F (market), 3F (market, size, and value), 5F (3F + gross profitability and investment), and 7F (5F + asset growth and accruals), respectively.

4.3 Sector Classification

Each stock is classified into one of 10 sectors, following the Standard Industrial Classification (SIC). Table 4.1 shows the number of companies from our sample in each sector. Note that some groups are quite large (the group Others, for instance, has more than 100 stocks). This motivates the use of additional restrictions we discussed in subsection 3.3.

4.4 Residual Covariance Matrices

In this section, we analyze the sequence of residual covariance matrices for the four different combinations of factors. We use the previous notation, $\Sigma_{\epsilon,t} = \Sigma_t - B_t' \Sigma_{f,t} B_t$, for each of the 1,495 days in the sample. To analyze whether these matrices are approximately block diagonal, we follow the procedure discussed in (21). First, we transform the covariance matrices into correlation matrices. We classify a correlation as significant if it is higher than 0.15 in absolute value in at least 1/3 of the sample. We then plot the significant relations as dots, while the rest of the points are left blank. Sectors in Table 4.1 are represented as red squares (in the same order). Figure 4.1 shows the results when we apply the criterion to the full covariance matrices Σ_t , before the factor decomposition. It is clear that most correlations are classified as significant.

When we follow this procedure for the residual covariance matrices (Figure 4.2), we obtain plots that are much more sparse than previously. This means that most of the correlations are not significant according to the criterion. Furthermore, we can see that most of the dots are contained inside the blocks, meaning that the majority of significant correlations are between stocks in the same sector. We interpret the three plots as evidence that we would not lose much information by assuming block diagonality for the residual covariance matrices. Note that the results are robust across different factor configurations.

Table 4.1: **Number of Stocks per Sector**

The table reports the number of stocks in each of the ten sectors considered. The sector classification follows the Standard Industrial Classification (SIC).

Sector	Number of Stocks
Consumer Non-Durables	31
Consumer Durables	8
Manufacturing	65
Oil, Gas, and Coal Extraction	32
Business Equipment	61
Telecommunications	10
Wholesale and Retail	45
Health Care, Medical Equipments, and Drugs	26
Utilities	36
Others	116

Figure 4.1: **Realized Raw Correlations**

The blue dots represent the correlations between stocks that are larger than 0.15 in absolute value in at least 1/3 of the sample days. Red squares represent the groups defined by sector industrial classification (SIC). The axes have indices that correspond to the 430 stocks in our sample. The correlations are computed using the sequence of realized covariance matrices before any factor decomposition.

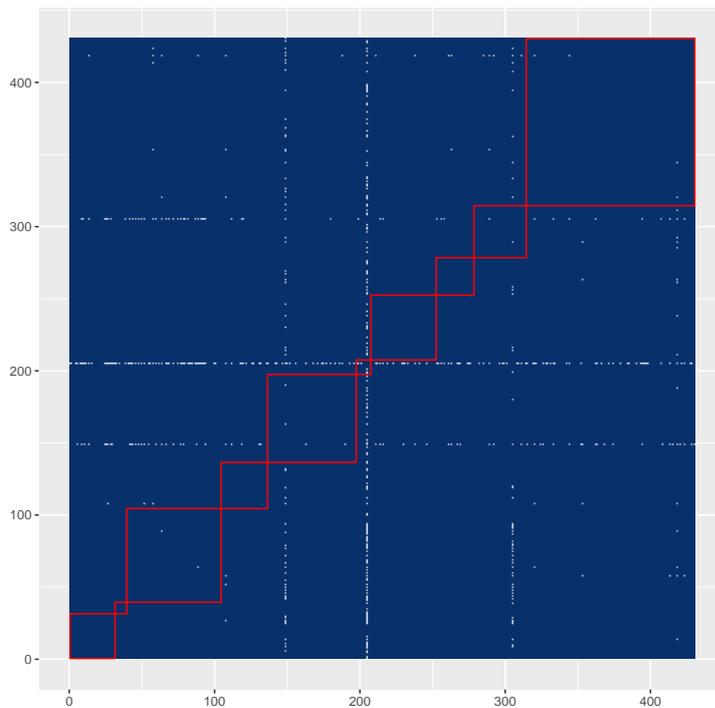
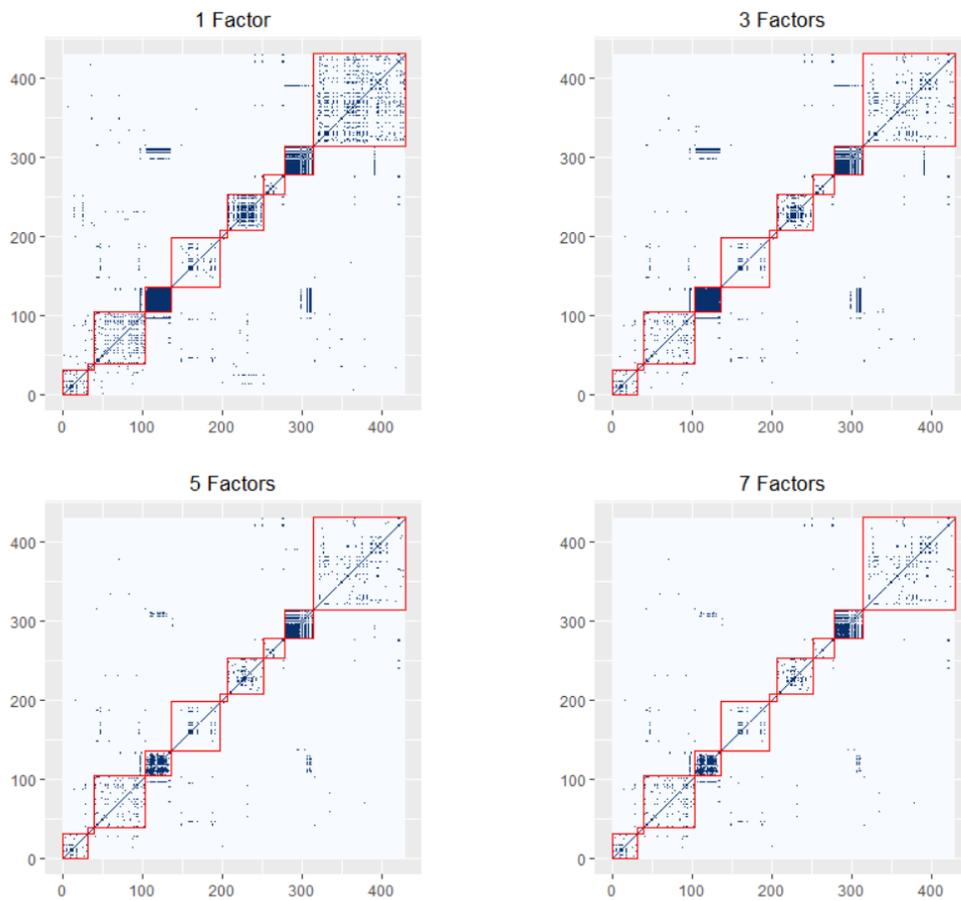


Figure 4.2: Realized Residual Correlations

The blue dots represent the correlations between stocks that are larger than 0.15 in absolute value in at least 1/3 of the sample days. Red squares represent the groups defined by sector industrial classification (SIC). The axes have indices that correspond to the 430 stocks in our sample. This plot shows the results for four series of residual covariance matrices obtained after using four different factor decompositions.



5 Forecasting Results

This section reports the forecasting results obtained using the methodology described in Section 3. Subsection 5.1 shows the forecasting results for Σ_f , while Subsection 5.2 shows the results for the complete covariance matrix Σ . In both cases, daily forecasts are computed using rolling windows of 1,000 observations. Since we need 22 days to compute the first monthly regressor, $\Sigma_{f,t}^{month}$, our out-of-sample forecasts comprise days 1,023 to 1,495 (T_1 to T_2), totaling 473 daily forecasts.

We evaluate our forecasts by using the ℓ_2 -norm for the vector of errors, that is, $\|\hat{e}_{T+1}\| = \|\text{vech}(\hat{\Sigma}_{T+1} - \Sigma_{T+1})\|$. We compare different methods by using the average ℓ_2 -forecast error:

$$\text{average } \ell_2\text{-forecast error} = \frac{1}{T_2 - T_1 + 1} \sum_{T=T_1}^{T=T_2} \|\hat{e}_{T+1}\|. \quad (5-1)$$

In all cases, our reference forecast is a random walk model: $\hat{\Sigma}_{f,T+1} = \Sigma_{f,T}$ and $\hat{\Sigma}_{T+1} = \Sigma_T$ for the complete covariance matrix.

5.1 Factor Covariance Matrix

Table 5.1 shows the forecast results for the factor covariance matrices, following the method described in Section 3.1 (we refer to this method as FHAR, from factor HAR, hereafter). The results are uniformly stronger when we apply a log-matrix transformation, as proposed by (11). Prior to estimation, we apply the log-matrix transformation to all data, that is, $\Omega_{f,t} = \log(\Sigma_{f,t})$. We then use the method FHAR to compute $\hat{\Omega}_{f,t+1}$. Finally, we revert the transformation by applying the exponential-matrix transformation and obtain our forecasts, that is, $\hat{\Sigma}_{f,t+1} = \exp(\hat{\Omega}_{f,t+1})$. We also report the results without applying the log-matrix transformation.

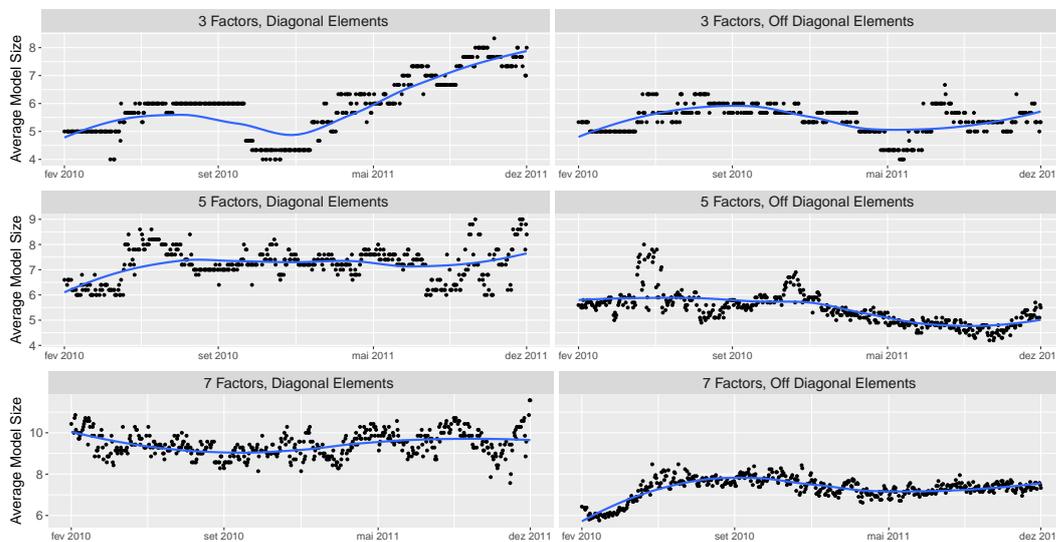
From Table 5.1, we see that the use of the log-matrix transformation considerably improves the forecasts. After this transformation, variance and covariance series become smoother than in the original data. This reduces the weights of outliers when fitting the model and improves forecasting precision. Another advantage is that the exponential matrix is positive definite by

construction. This consideration is important, as we will need the inverse of the covariance matrix Σ when solving the minimum variance problems in Section 6. Finally, applying adaLASSO to the best performing models does not improve the results up to the second decimal place.

Since we estimate all equations on a daily basis, we now investigate their evolution over time¹. Figure 5.1 shows the average number of variables selected by LASSO over the 473 days for the FHAR Log-matrix models. To illustrate, consider the upper-left panel in this figure. It shows the average model size for the diagonal equations in the FHAR Log-matrix model with three factors. In other words, it displays the mean size for the equations of σ_{market}^2 , σ_{SMB}^2 , and σ_{HML}^2 . Panels on the right show the same results for the covariance equations. From top to bottom, we vary the number of factors².

Figure 5.1: Average Number of Selected Variables in FHAR Log-matrix LASSO models

These panels show the daily average number of variables selected by the LASSO method for variance and covariance equations on a given day. We sort the panels by number of factors in the model (increasing from top to bottom) and by variable class (variances on the left and covariances on the right). Blue lines are local polynomial regressions. The maximum number of predictive variables for each configuration are 18 (3 factors), 45 (5 factors), and 76 (7 factors).



In all cases, LASSO reduces the number of selected variables substan-

¹The results for average equation size and average change in the number of selected variables obtained with the adaLASSO estimator are quite similar to the ones obtained via LASSO estimation. These results are readily available upon request.

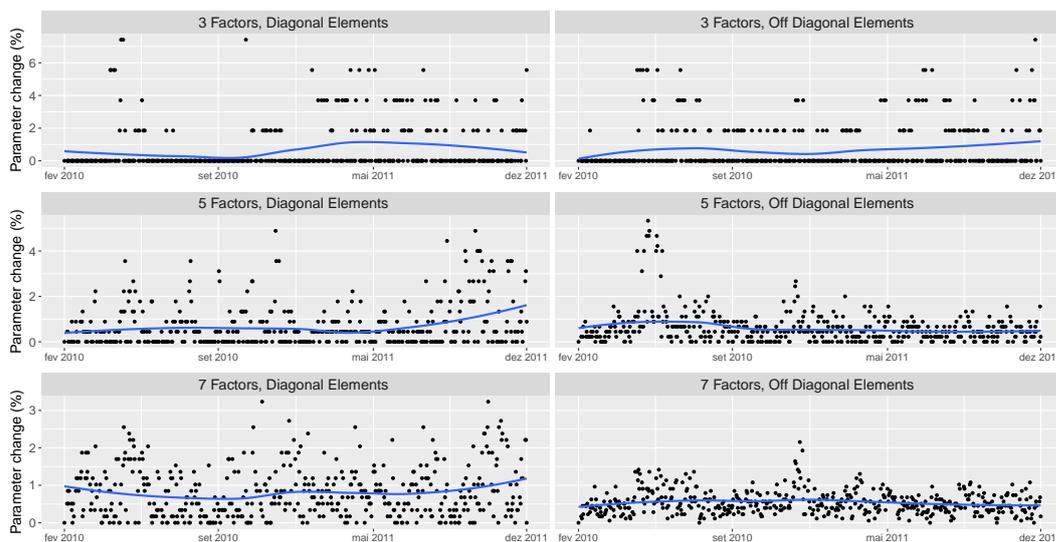
²For the 1F configuration, we only have one element to be forecast: σ_{market}^2 . In this case, the equations estimated by LASSO and adaLASSO have only three potential predictors, which are, in this case, always selected. For this reason, we do not present the average number of selected variables (constant) or average change in the number of selected variables (zero) for this case.

tially. This reduction is most noticeable for the FHAR Log-matrix model with 7 factors, in which the number of selected variables fluctuates around 10 (out of 76 potential predictors) for the variance equations and even less for the covariance equations.

Another interesting feature is that average equation sizes are stable over time. Figure 5.2 shows the average change in the number of selected variables from one day to the next. We consider parameters changing from zero to non-zero values or the opposite direction. All results are reported in percentage (relative to the maximum number of variables in each equation). For the three-factor models presented, both variance and covariance equations remain unchanged on most days. Despite the stronger variation in the 5- and 7-factor configurations, the percentage changes are smaller than 5% on almost all days. In general, variance equations are slightly more stable. We present results for the average equation size and average change in the number of selected variables for the residual covariance matrix forecasts in Appendix E.

Figure 5.2: Average Change in the Number of Selected Variables in FHAR Log-matrix models

These panels show the daily average change in the number of variables selected by LASSO for variance and covariance equations. We count a change as when a variable goes from being zero to non-zero or from being non-zero to zero. The results are presented as a percentage relative to the maximum number of variables in each equation. We sort the panels by the number of factors in the model (increasing from top to bottom) and by variable class (variances on the left and covariances on the right). Blue lines are local polynomial regressions.



5.2

Complete Covariance Matrix

Table 5.2 shows the forecast results for the complete covariance matrix (430 stocks). In this case, VHAR denotes the complete methodology described in Section 3. We compute all models using Log-matrix transformation and estimate these models using LASSO and adaLASSO. Table 5.2 also reports alternative models for comparison purposes:

1. Random walk (RW): the standard reference model throughout this work.
2. Exponentially weighted moving average (EWMA) with smoothing parameter $\lambda = 0.96$, as recommended by the RiskMetrics methodology.
3. The dynamic conditional correlation and BEKK models, both with nonlinear shrinkage, as proposed by (10). These approaches will be called DCC-NL and BEKK-NL hereafter.
4. Modified versions of the random walk model (Blocks 1F, 3F, 5F, and 7F). In these cases, we decompose the original realized covariance matrices using factors and impose block diagonality on the residual covariance matrices. Instead of using VHAR models to forecast future realizations, we simply use the random walk model with the adjusted covariance matrices as forecasts. By doing this, we try to approximate the methodology used by (12)³.

Some of the the alternative models use daily returns data, instead of using realized covariance matrices as inputs. We compare our results to these models to see whether the use of high-frequency data brings gains in terms of portfolio performance.

Models based on daily returns do not outperform the benchmark in any case, with a particularly bad performance for the EWMA. Despite having performed much better, BEKK-NL and DCC-NL have errors that are almost 2 times larger than the random walk model. Imposing block diagonality on the residual covariance matrices (Blocks 1F, 3F, 5F, and 7F) improves the results, relative to the random walk, by 3%. Our methodology is able to beat the random walk model by up to 15%. In terms of the average ℓ_2 -forecast error, there is no difference when estimating the models with LASSO or adaLASSO.

³(12) impose block diagonality when using the high-frequency data for factors and returns to estimate the integrated covariance. Another difference is that the authors use the GICS system to classify stocks in sectors. The model used for forecasting and portfolio allocation is the random walk.

Table 5.1: **Forecast Precision for Factor Covariance Matrices**
 ℓ_2 represents the average ℓ_2 -forecast error over the 473 out-of-sample days, that is,

$$\text{average } \ell_2\text{-forecast error} = \frac{1}{T_2 - T_1 + 1} \sum_{T=T_1}^{T=T_2} \|\hat{\epsilon}_{T+1}\|.$$

$\ell_2/\ell_{2,RW}$ represents the ratio of the ℓ_2 -forecast error for other methods to the random walk value. FHAR stands for the factor heterogeneous autoregressive model described in Section 3.1. Each line represents a different factor configuration. Numbers shown between parenthesis represent the results obtained with the adaLASSO method.

Model	ℓ_2	$\ell_2 / \ell_{2,RW}$	
	Random Walk	FHAR	FHAR, Log-matrix
1F	0.40	0.96 (0.96)	0.92 (0.92)
3F	0.44	0.98 (0.97)	0.90 (0.90)
5F	0.51	0.95 (0.95)	0.89 (0.89)
7F	0.62	0.99 (1.04)	0.86 (0.87)

Table 5.2: **Forecast Precision for Complete Covariance Matrices**
 ℓ_2 represents the average ℓ_2 -forecast error, computed for 473 days, that is,

$$\text{average } \ell_2\text{-forecast error} = \frac{1}{T_2 - T_1 + 1} \sum_{T=T_1}^{T=T_2} \|\hat{\epsilon}_{T+1}\|.$$

$\ell_{2,RW}$ is the error for the random walk model. $\ell_2/\ell_{2,RW}$ represents the ratio between the error values. The column “Latent Covariance Models” displays the results for models based on daily return data. The columns “VHAR” and “VHAR, Log-matrix” show our results (without and with log-matrix transformation, respectively). 1F, 3F, 5F, and 7F stand for the 1-factor, 3-factor, 5-factor, and 7-factor configurations for the factor covariance matrix. For the VHAR (Log-matrix) forecasts, the results are presented for the LASSO and adaLASSO estimation.

Model (Benchmarks)	$\ell_2/\ell_{2,RW}$	VHAR (Log-matrix)	$\ell_2/\ell_{2,RW}$
RW	1.00	1F, LASSO	0.86
EWMA (Returns)	6.93	3F, LASSO	0.85
BEKK-NL	1.71	5F, LASSO	0.85
DCC-NL	1.71	7F, LASSO	0.85
Block 1F	0.97	1F, adaLASSO	0.86
Block 3F	0.97	3F, adaLASSO	0.85
Block 5F	0.97	5F, adaLASSO	0.85
Block 7F	0.97	7F, adaLASSO	0.85
Random Walk (RW) $\ell_{2,RW}$	341.57		

6

Portfolio Selection

Despite the encouraging results obtained when forecasting covariance matrices, one could argue that our measure of forecast precision is too aggregate, specially when taking into account the fact that we have more than 90,000 unique entries being forecasted simultaneously. For instance, these estimates may produce large mistakes in only a few entries but with large implications when applied to practical problems. To evaluate the economic benefits provided by our forecasts, we use them to construct daily investment portfolios¹. We construct minimum variance portfolios using the models described in the previous section. We consider the global minimum variance portfolio (without any restrictions on weights), the restricted portfolio (limiting long positions and short-selling) and the long-only portfolios (without any short-selling).

We focus our analysis on variations of the minimum variance problem, as it is the sole portfolio problem that requires only information on the covariance matrix. Other portfolio optimization problems, such as finding tangency portfolios that maximize Sharpe ratios, also require estimates of expected returns for every single eligible asset. When we analyze the outcomes, we also focus on risk measures, such as the portfolio standard deviation, as our aim is to minimize the risk of the portfolio. In finance theory, all efficient portfolios have higher expected returns and Sharpe ratios than the minimum variance portfolio. Minimum variance optimizations that are less precise or biased may, by chance, find portfolios that have higher realized returns and higher Sharpe ratios. While we should base our assessment of the models on the risk measures, other measures related to the concentration, trading volume and even realized return may also provide relevant information.

First, we compute various statistics for indices directly related to our portfolio exercise. Table 6.1 shows the standard deviation, kurtosis, skewness, average excess return over the risk-free rate and Sharpe ratio for the daily series of returns for three S&P 500 indices: minimum volatility, low volatility, and equal weight. We provide these values as reference points to evaluate our

¹In this section, we use daily covariance matrix forecasts to build portfolios for the next day. These portfolios are rebalanced daily, using the most recent forecasts. We investigate portfolios formed for longer holding periods in Appendix F. Similar results to this section hold even with infrequent rebalancing.

results, as these indices are commonly used by practitioners. According to the index provider, the first index “is designed to reflect a managed-volatility equity strategy that seeks to achieve lower total risk, measured by standard deviation, than the S&P 500 while maintaining similar characteristics.” It is based on the Northfield Open Optimizer and rebalanced every six months. The second index is based on risk measures for each asset individually, focusing on stocks with low volatility.

Table 6.1: **S&P Index Performances**

Average excess returns are defined as the annualized average excess returns over the risk-free rate for the index total return series. Standard deviations are calculated for these series and are also annualized. The Sharpe ratio is computed using these two quantities. The sample is from 2010/02/12 to 2011/12/29, which corresponds to 473 days considered in the portfolio exercises.

	S&P 500	S&P 500 Min. Vol.	S&P 500 Low Vol.	S&P 500 Equal Weight
Standard Deviation (%)	21.00	16.41	14.84	23.48
Kurtosis	2.88	3.37	3.54	2.67
Skewness	-0.35	-0.36	-0.38	-0.34
Average Excess Return (%)	10.51	13.00	13.50	12.78
Cumulative Return (%)	17.11	24.70	26.47	20.93
Sharpe Ratio	0.50	0.79	0.91	0.54

6.1 Global Minimum Variance Portfolios

Consider the problem of an investor at time $t = t_0, \dots, T - 1$ who wishes to construct a minimum variance portfolio to be held in time $t + 1$. For this minimization problem, the investor needs to forecast the future covariance matrix, $\hat{\Sigma}_{t+1}$. The optimization problem consists of choosing a vector of weights \hat{w}_{t+1} (dimension $N \times 1$):

$$\begin{aligned} \hat{w}_{t+1} = \arg \min_{w_{t+1}} \quad & w'_{t+1} \hat{\Sigma}_{t+1} w_{t+1} \\ \text{subject to} \quad & w'_{t+1} \mathbf{1} = 1. \end{aligned} \quad (6-1)$$

We use our VHAR method with log matrix transformation, as it provides the best results in terms of forecasting precision. Estimation is done via LASSO and adaLASSO. We compare our results against the random walk models (RW and Block models), the EWMA, BEKK-NL, and DCC-NL. We evaluate ex-post portfolio performance; that is, we use our time t estimated weights, \hat{w}_{t+1} , with data from $t + 1$. In the following, \hat{w}_{it} is the i -th component of \hat{w}_t . We also define $r_{pt}^* = (r_{pt} - \frac{1}{(T-t_0)} \sum_{t=t_0+1}^T r_{pt})$. We show the following statistics:

1. Standard deviation: $\sigma_p = \sqrt{\frac{1}{(T-t_0)} \sum_{t_0+1}^T (r_{pt}^*)^2}$.

2. Lower partial standard deviation: $\sqrt{\frac{1}{\sum_{t=t_0+1}^T I(r_{pt}^* < 0)} \sum_{t=t_0+1}^T (r_{pt}^*)^2 * I(r_{pt}^* < 0)}$.
3. Kurtosis: $\frac{\sum_{t=t_0+1}^T (r_{pt}^*)^4 / (T-t_0)}{\sigma_p^4} - 3$.
4. Skewness: $\frac{\sqrt{(T-t_0)(T-t_0-1)} \sum_{t=t_0+1}^T (r_{pt}^*)^3 / (T-t_0)}{T-t_0-2 \sigma_p^3}$.
5. Average diversification ratio: $\frac{1}{(T-t_0)} \sum_{t=t_0+1}^T \frac{\sum_{i=1}^N \hat{w}_{it} \sigma_{it}}{\sigma_{pt}}$, where $\sigma_{pt} = \hat{w}'_t \Sigma_t \hat{w}_t$.
6. Average max. weight: $\frac{1}{(T-t_0)} \sum_{t=t_0+1}^T \max_{1 \leq i \leq N} (\hat{w}_{it})$ for $i = 1, \dots, N$.
7. Average min. weight: $\frac{1}{(T-t_0)} \sum_{t=t_0+1}^T \min_{1 \leq i \leq N} (\hat{w}_{it})$ for $i = 1, \dots, N$.
8. Average gross leverage: $\frac{1}{N(T-t_0)} \sum_{t=t_0+1}^T \sum_{i=1}^N |\hat{w}_{it}|$.
9. Proportion of leverage: $\frac{1}{N(T-t_0)} \sum_{t=t_0+1}^T \sum_{i=1}^N I(\hat{w}_{it} < 0)$.
10. Average turnover: $\frac{1}{N(T-t_0)} \sum_{t=t_0+1}^T \sum_{i=1}^N |\hat{w}_{it} - \hat{w}_{it}^{hold}|$, where $\hat{w}_{it}^{hold} = \hat{w}_{it-1} \frac{1+r_{it}-1}{1+r_{pt-1}}$. r_{pt} is the portfolio return at time t , r_{it} is the stock i return at time t , and \hat{w}_{it}^{hold} is the weight of stock i in the hold portfolio. The hold portfolio at time $t+1$ is defined as the resulting portfolio from keeping all the stocks from period t .
11. Average excess return: $\mu_p^e = \frac{1}{(T-t_0)} \sum_{t=t_0+1}^T (r_{pt}^e) = \frac{1}{(T-t_0)} \sum_{t=t_0+1}^T (\hat{w}'_t r_t - r_{f,t})$, where $r_{f,t}$ is the risk free rate.
12. Cumulative Return: $\prod_{t=t_0+1}^T (1 + r_{pt})$.
13. Sharpe ratio: $\frac{\mu_p^e}{\sigma_p}$.

Table 6.2 shows these results for the minimum variance portfolio optimization problem, presented in equation (6-1). In terms of standard deviations, the VHAR class of models perform better than all others. For these models, the standard deviations go from a high of 10.39 % (1 factor and LASSO) to a low of 8.58 % (5 factors and adaLASSO). Note that the lowest value is almost half of the standard deviation for the S&P 500 Minimum Volatility Index, presented in Table 6.1. Benchmark models (upper results) also have smaller standard deviation than the S&P 500 indices (except for the EWMA, which has the worst performance in this sense). Similar results hold for lower partial standard deviations.

Table 6.2 shows that some portfolios have extreme short positions. In most of the cases, the proportion of leverage is close to 50%. Average gross leverage is also quite high for many of the alternative models. Since shorting stocks may not be feasible and can be costly, we consider these positions

potentially extreme. The values for average maximum and minimum weights are also quite large in some cases. Our models appear to generate more balanced portfolios.

In terms of average excess returns and cumulative returns, our models provide mixed results: slightly inferior performance compared to BEKK-NL and DCC-NL; similar performance compared to the RW and Block models; and much better performance relative to the EWMA. Since the optimization problem focuses on minimizing the portfolio variance, it is not obvious which models should achieve higher returns. For this reason, we do not see the slightly superior performance of the BEKK-NL and DCC-NL models as a problem, as they are choosing riskier portfolios that are likely to deliver higher returns. Moreover, the small difference in realized returns is more than compensated by the substantial reduction in standard deviation (as evidenced by the higher Sharpe ratios for our models).

6.2 Restricted Minimum Variance Portfolios

We now resort to what we consider a more realistic investor problem. In this section, we solve a problem similar to equation (6-1), except that now we impose two additional restrictions. First, we allow the maximum leverage to be 30% (in some sense, consistent with a 130-30 fund concept in the mutual fund industry). Second, we restrict the maximum weights on individual stocks to be 20% (in absolute value). The problem for an investor at time $t = t_0, \dots, T - 1$ is then given by

$$\begin{aligned} \hat{w}_{t+1} = \arg \min_{w_{t+1}} \quad & w'_{t+1} \hat{\Sigma}_{t+1} w_{t+1} \\ \text{subject to} \quad & w'_{t+1} \mathbf{1} = 1, \end{aligned} \tag{6-2}$$

$$\sum_{i=1}^N |w_{it+1}| I(w_{it} < 0) \leq 0.30 \quad \text{and} \quad |w_{it+1}| \leq 0.20.$$

In Table 6.3 we report the same performance statistics as in the last subsection. Again, our methodology generates lower standard deviations than the competing models (our VHAR 1-factor model performs closely to the RW and Block benchmarks). In terms of returns, the results are, again, mixed. While our models perform better against the EWMA, BEKK-NL, and DCC-NL, the RW and Block models are much more competitive this time. Again, we do not consider this to be a problem, as we were able to obtain smaller standard deviations with our models. Finally, the restriction on leverage drastically reduces the proportion of leverage in all cases.

Table 6.2: Statistics for Daily Portfolios - Global Minimum Variance
 The table reports the results for portfolios constructed according to the optimization problem as in (6-1). Different models are used to provide the one-step-ahead forecasts for the realized covariance matrix: RW is the random walk model, and Blocks 1F, 3F, 5F, and 7F are random walks applied to the residual covariance matrix after the four different factor decompositions.

	RW	Block 1F	Block 3F	Block 5F	Block 7F	EWMA (Returns)	BEKK-NL	DCC - NL	
Standard Deviation (%)	12.07	8.21	8.29	8.25	8.25	14.62	9.41	10.65	
Lower Partial Standard Deviation (%)	12.82	8.79	8.94	8.73	8.83	14.90	9.63	11.31	
Kurtosis	1.29	0.69	0.77	0.72	0.73	2.12	1.60	4.60	
Skewness	-0.35	-0.39	-0.40	-0.38	-0.40	-0.15	-0.33	-0.50	
Average Diversification Ratio	3.29	6.05	6.11	6.11	6.12	1.01	3.03	3.53	
Average Max. Weight	0.06	0.10	0.10	0.10	0.10	0.20	0.06	0.12	
Average Min. Weight	-0.06	-0.03	-0.03	-0.03	-0.03	-0.15	-0.04	-0.04	
Average Gross Leverage	5.94	3.08	3.14	3.14	3.19	12.55	5.09	4.11	
Proportion of Leverage (%)	44.30	44.40	44.22	44.10	44.11	49.17	45.11	51.73	
Average Turnover (%)	1.80	0.75	0.78	0.78	0.80	0.27	0.11	0.21	
Average Excess Return (%)	14.20	12.72	14.46	15.37	14.95	3.42	17.98	17.46	
Cumulative Return (%)	29.04	26.42	30.59	32.86	31.82	4.74	39.27	37.58	
Sharpe Ratio	1.18	1.55	1.74	1.86	1.81	0.23	1.91	1.64	
		1 Factor VHAR (Log matrix)	3 Factors VHAR (Log matrix)		5 Factors VHAR (Log matrix)		7 Factors VHAR (Log matrix)		
		LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO
Standard Deviation (%)	8.46	8.42	8.37	8.32	8.29	8.25	8.12	8.09	
Lower Partial Standard Deviation (%)	8.86	8.81	8.78	8.68	8.57	8.53	8.52	8.51	
Kurtosis	0.96	0.98	0.99	0.98	0.96	0.94	1.13	1.06	
Skewness	-0.21	-0.20	-0.18	-0.17	-0.15	-0.13	-0.24	-0.22	
Average Diversification Ratio	4.79	4.81	5.02	5.03	4.87	4.88	4.96	4.96	
Average Max. Weight	0.07	0.08	0.08	0.08	0.08	0.09	0.08	0.09	
Average Min. Weight	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.03	
Average Gross Leverage	2.66	2.67	2.80	2.80	2.82	2.82	2.93	2.93	
Proportion of Leverage (%)	45.89	46.01	44.88	45.03	44.89	45.12	45.26	45.50	
Average Turnover (%)	0.20	0.22	0.20	0.22	0.19	0.21	0.20	0.22	
Average Excess Return (%)	15.24	15.18	17.69	17.45	18.93	18.61	18.09	17.85	
Cumulative Return (%)	32.49	32.35	38.74	38.13	42.01	41.19	39.85	39.21	
Sharpe Ratio	1.80	1.80	2.11	2.10	2.28	2.26	2.23	2.21	

Table 6.3: Statistics for Daily Portfolios - Restricted Minimum Variance

The table reports the results for portfolios constructed according to the optimization problem as in (6-2). Different models are used to provide the one-step-ahead forecasts for the realized covariance matrix: RW is the random walk model, and Blocks 1F, 3F, 5F, and 7F are random walks applied to the residual covariance matrix after the four different factor decompositions.

	RW	Block 1F	Block 3F	Block 5F	Block 7F	EWMA (Returns)	BEKK-NL	DCC - NL	
Standard Deviation (%)	13.29	13.34	13.20	13.17	13.25	15.28	15.49	14.72	
Lower Partial Standard Deviation (%)	14.13	13.91	13.66	13.35	13.68	16.47	16.24	15.28	
Kurtosis	3.40	4.71	4.68	4.79	5.10	4.16	4.36	3.08	
Skewness	-0.27	-0.13	-0.08	-0.10	-0.17	-0.48	-0.36	-0.15	
Average Diversification Ratio	3.55	3.92	4.11	4.12	4.07	2.22	2.24	2.15	
Average Max. Weight	0.17	0.16	0.15	0.15	0.15	0.19	0.19	0.19	
Average Min. Weight	-0.09	-0.07	-0.07	-0.07	-0.08	-0.16	-0.15	-0.11	
Average Gross Leverage	1.60	1.60	1.60	1.60	1.60	1.60	1.60	1.60	
Proportion of Leverage (%)	1.91	3.11	3.08	3.06	2.93	0.71	0.85	1.41	
Average Turnover (%)	0.43	0.40	0.42	0.41	0.42	0.09	0.10	0.11	
Average Excess Return (%)	16.72	18.23	19.01	22.42	21.22	13.68	14.24	16.91	
Cumulative Return (%)	34.88	38.74	40.83	50.14	46.79	26.74	27.99	34.86	
Sharpe Ratio	1.26	1.37	1.44	1.70	1.60	0.90	0.92	1.15	
		1 Factor VHAR (Log matrix)	3 Factors VHAR (Log matrix)		5 Factors VHAR (Log matrix)		7 Factors VHAR (Log matrix)		
		LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO
Standard Deviation (%)	13.20	13.37	12.81	12.86	12.57	12.83	12.63	12.75	
Lower Partial Standard Deviation (%)	13.29	13.64	12.60	12.54	12.54	12.75	12.52	12.62	
Kurtosis	3.89	4.09	4.82	4.84	4.44	5.05	4.47	5.14	
Skewness	-0.03	-0.12	0.14	0.12	0.04	0.10	0.03	0.05	
Average Diversification Ratio	3.41	3.38	3.64	3.65	3.73	3.70	3.68	3.65	
Average Max. Weight	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	
Average Min. Weight	-0.08	-0.08	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	
Average Gross Leverage	1.60	1.60	1.60	1.60	1.60	1.60	1.60	1.60	
Proportion of Leverage (%)	2.46	2.44	2.37	2.38	2.43	2.41	2.27	2.25	
Average Turnover (%)	0.22	0.23	0.24	0.24	0.23	0.24	0.22	0.23	
Average Excess Return (%)	16.07	19.89	19.72	21.04	20.56	18.93	20.74	19.19	
Cumulative Return (%)	33.30	43.13	42.88	46.43	45.22	40.76	45.67	41.48	
Sharpe Ratio	1.22	1.49	1.54	1.64	1.64	1.48	1.64	1.51	

Table 6.4: **Statistics for Daily Portfolios - Restricted Minimum Variance (Long Only)**

The table reports the results for portfolios constructed according to the optimization problem as in (6-3). Different models are used to provide the one-step-ahead forecasts for the realized covariance matrix: RW is the random walk model, and Blocks 1F, 3F, 5F, and 7F are random walks applied to the residual covariance matrix after the four different factor decompositions.

	RW	Block 1F	Block 3F	Block 5F	Block 7F	EWMA (Returns)	BEKK-NL	DCC - NL
Standard Deviation (%)	17.10	17.06	16.96	16.85	16.88	17.74	17.92	17.78
Lower Partial Standard Deviation (%)	17.56	17.83	17.63	17.49	17.58	18.94	19.16	19.13
Kurtosis	3.29	3.04	3.16	3.22	3.17	2.84	2.88	2.60
Skewness	-0.25	-0.30	-0.32	-0.31	-0.31	-0.32	-0.35	-0.33
Average Diversification Ratio	3.10	3.04	3.07	3.09	3.09	2.63	2.60	2.51
Average Max. Weight	0.18	0.19	0.19	0.19	0.19	0.18	0.18	0.18
Average Min. Weight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Gross Leverage	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Proportion of Leverage (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Turnover (%)	0.17	0.16	0.16	0.16	0.16	0.03	0.03	0.04
Average Excess Return (%)	14.29	15.86	16.18	14.98	15.06	20.22	15.85	16.28
Cumulative Return (%)	27.49	31.30	32.15	29.25	29.44	42.18	30.91	32.04
Sharpe Ratio	0.84	0.93	0.95	0.89	0.89	1.14	0.88	0.92

	1 Factor VHAR (Log matrix)		3 Factors VHAR (Log matrix)		5 Factors VHAR (Log matrix)		7 Factors VHAR (Log matrix)	
	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO
	Standard Deviation (%)	16.96	16.98	16.55	16.59	16.34	16.47	16.31
Lower Partial Standard Deviation (%)	17.51	17.64	17.29	17.27	16.88	17.10	16.89	17.03
Kurtosis	4.32	4.27	4.18	3.70	3.37	3.46	3.69	3.56
Skewness	-0.35	-0.35	-0.36	-0.27	-0.19	-0.21	-0.20	-0.23
Average Diversification Ratio	3.10	3.09	3.09	3.10	3.12	3.13	3.11	3.11
Average Max. Weight	0.17	0.18	0.19	0.19	0.19	0.19	0.19	0.19
Average Min. Weight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Gross Leverage	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Proportion of Leverage (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Turnover (%)	0.08	0.08	0.07	0.08	0.07	0.08	0.07	0.07
Average Excess Return (%)	17.60	17.57	17.62	18.04	18.02	18.17	17.13	17.04
Cumulative Return (%)	35.71	35.63	35.95	37.01	37.06	37.38	34.79	34.50
Sharpe Ratio	1.04	1.03	1.06	1.09	1.10	1.10	1.05	1.04

6.3

Long-Only Minimum Variance Portfolios

In this section, we impose a restriction of no short-selling. The problem for an investor at time $t = t_0, \dots, T - 1$ is then given by

$$\begin{aligned} \hat{w}_{t+1} &= \arg \min_{w_{t+1}} w'_{t+1} \hat{\Sigma}_{t+1} w_{t+1} \\ &\text{subject to } w'_{t+1} \mathbf{1} = 1, \\ &0 \leq w_{it+1} \leq 0.20. \end{aligned} \tag{6-3}$$

From Table 6.4, we see that, except for the 1-factor model, our methodology is able to reduce standard deviations compared to the benchmarks. This time, interestingly, average returns and cumulative returns are higher in all cases. These facts are reflected in higher Sharpe ratios.

7 Conclusions

We propose a model to forecast very large realized covariance matrices of returns, applying it to the constituents of the S&P 500 on a daily basis. To address the curse of dimensionality, we decompose the return covariance matrix using standard firm-level factors (e.g., size, value, and profitability) and use sectoral restrictions in the residual covariance matrix. This restricted model is then estimated using vector heterogeneous autoregressive (VHAR) models estimated with the least absolute shrinkage and selection operator (LASSO). Our methodology improves forecasting precision relative to standard benchmarks and leads to better estimates of the minimum variance portfolios.

Bibliography

- [1] LUNDE, A.; SHEPHARD, N. ; SHEPPARD, K.. **Econometric analysis of vast covariance matrices using composite realized kernels and their application to portfolio choice.** *Journal of Business & Economic Statistics*, 34(4):504–518, 2016.
- [2] ENGLE, R.; KRONER, F.. **Multivariate simultaneous generalized arch.** *Econometric Theory*, 11:122–150, 1995.
- [3] ENGLE, R.. **Dynamic conditional correlation.** *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.
- [4] BAUER, G.; VORNIK, K.. **Forecasting multivariate realized stock market volatility.** *Journal of Econometrics*, 160:93–101, 2011.
- [5] CHIRIAC, R.; VOEV, V.. **Modelling and forecasting multivariate realized volatility.** *Journal of Applied Econometrics*, 26:922–947, 2011.
- [6] GOLOSNOY, V.; GRIBISCH, B. ; LIESENFELD, R.. **The conditional autoregressive wishart model for multivariate stock market volatility.** *Journal of Econometrics*, 167(1):211–223, 2012.
- [7] CALLOT, L. A.; KOCK, A. B. ; MEDEIROS, M. C.. **Modeling and forecasting large realized covariance matrices and portfolio choice.** *Journal of Applied Econometrics*, 32(1):140–158, 2017.
- [8] TIBSHIRANI, R.. **Regression shrinkage and selection via the lasso.** *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 267–288, 1996.
- [9] ZOU, H.. **The adaptive lasso and its oracle properties.** *Journal of the American statistical association*, 101(476):1418–1429, 2006.
- [10] ENGLE, R. F.; LEDOIT, O. ; WOLF, M.. **Large dynamic covariance matrices.** *Journal of Business & Economic Statistics*, (forthcoming), 2017.
- [11] CHIU, T. Y.; LEONARD, T. ; TSUI, K.-W.. **The matrix-logarithmic covariance model.** *Journal of the American Statistical Association*, 91(433):198–210, 1996.

- [12] FAN, J.; FURGER, A. ; XIU, D.. **Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high frequency data**". *Journal of Business & Economic Statistics*, 34:489–503, 2016.
- [13] CHAMBERLAIN, G.; ROTHSCILD, M.. **Arbitrage, factor structure, and mean-variance analysis on large asset markets**. *Econometrica*, 51(5):1281, 1983.
- [14] CORSI, F.. **A simple approximate long-memory model of realized volatility**. *Journal of Financial Econometrics*, 7(2):174–196, 2009.
- [15] FAN, J.; FAN, Y. ; LV, J.. **High dimensional covariance matrix estimation using a factor model**. *Journal of Econometrics*, 147(1):186–197, 2008.
- [16] FAMA, E. F.; FRENCH, K. R.. **Common risk factors in the returns on stocks and bonds**. *Journal of Financial Economics*, 33(1):3–56, 1993.
- [17] NOVY-MARX, R.. **The other side of value: The gross profitability premium**. *Journal of Financial Economics*, 108(1):1–28, 2013.
- [18] LYANDRES, E.; SUN, L. ; ZHANG, L.. **The new issues puzzle: Testing the investment-based explanation**. *The Review of Financial Studies*, 21(6), 2008.
- [19] COOPER, M. J.; GULEN, H. ; SCHILL, M. J.. **Asset growth and the cross-section of stock returns**. *The Journal of Finance*, 63(4):1609–1651, 2008.
- [20] SLOAN, R.. **Do stock prices fully reflect information in accruals and cash flows about future earnings?** *The Accounting Review*, 71:289–315, 1996.
- [21] AIT-SAHALIA, Y.; XIU, D.. **Using principal component analysis to estimate a high dimensional factor model with high-frequency data**. *Journal of Econometrics*, forthcoming.
- [22] GAGLIARDINI, P.; OSSOLA, E. ; SCALLET, O.. **A diagnostic criterion for approximate factor structure**. 2018.

A Factor Loadings

This appendix describes how daily loadings on factors can be calculated using daily realized covariance matrices of returns Σ_t , daily factor covariance matrices $\Sigma_{f,t}$, and factor weights W (which are calculated yearly). To simplify the notation, we drop the subscript t and derive the equation for a daily matrix of loadings \hat{B} (the final equation is applied to daily data, providing 1,495 matrices \hat{B}). Letters with one subscript t represent vectors with cross-sectional data at time t , while letters with one superscript n represent time series data for asset n .

Assume that we have T observations for K factors and N assets (in our case, T can be thought of as intra-daily high-frequency observations). Stacking all observations in matrix form,

$$F = \begin{pmatrix} f_{1,1} & \dots & f_{K,1} \\ \vdots & \ddots & \vdots \\ f_{1,T} & \dots & f_{K,T} \end{pmatrix} = \begin{pmatrix} f'_1 \\ \vdots \\ f'_T \end{pmatrix} \quad (\text{A-1})$$

$$R = \begin{pmatrix} r_{1,1} & \dots & r_{1,T} \\ \vdots & \ddots & \vdots \\ r_{N,1} & \dots & r_{N,T} \end{pmatrix} = \begin{pmatrix} r'_1 \\ \vdots \\ r'_N \end{pmatrix} = (r_1 \dots r_T).$$

From equation (2-4), we can rewrite the matrix F as

$$F = \begin{pmatrix} w'_1 r_1 & \dots & w'_K r_1 \\ \vdots & \ddots & \vdots \\ w'_1 r_T & \dots & w'_K r_T \end{pmatrix} = \begin{pmatrix} r'_1 w_1 & \dots & r'_1 w_K \\ \vdots & \ddots & \vdots \\ r'_T w_1 & \dots & r'_T w_K \end{pmatrix} = \begin{pmatrix} r'_1 \\ \vdots \\ r'_T \end{pmatrix} (w_1 \dots w_K) = R'W. \quad (\text{A-2})$$

With this setup, consider the linear model of asset n , $n \in \{1, \dots, N\}$, on K factors:

$$\begin{pmatrix} r_{n,1} \\ \vdots \\ r_{n,T} \\ r^n \end{pmatrix} = \begin{pmatrix} f'_1 \\ \vdots \\ f'_T \\ F \end{pmatrix} \begin{pmatrix} b_{n,1} \\ \vdots \\ b_{n,K} \\ b_n \end{pmatrix} + \begin{pmatrix} \epsilon_{n,1} \\ \vdots \\ \epsilon_{n,T} \end{pmatrix}. \quad (\text{A-3})$$

The OLS estimator for b_n , \hat{b}_n , is given by

$$\hat{b}_n = (F'F)^{-1}F'r^n \quad \text{for } i = 1, \dots, N. \quad (\text{A-4})$$

In matrix form for all n ,

$$\begin{aligned}
\hat{B} &= (\hat{b}_1 \quad \dots \quad \hat{b}_N) = (F'F)^{-1}F' (r^1 \quad \dots \quad r^N) = \\
&= (F'F)^{-1}F' \begin{pmatrix} r_1' \\ \vdots \\ r_T' \end{pmatrix} = (F'F)^{-1}F'R' = & \quad (A-5) \\
&= (F'F)^{-1}W'RR' \xrightarrow{mean=0} (\Sigma_f)^{-1}W'\Sigma,
\end{aligned}$$

with equation (A-2) being used on the last line. Again, in our setup, \hat{B} is different for every day in the sample.

B Betas' Long Memory

In this section, we investigate the long-range dependency for the series of factor loadings (“betas”). For this, we estimate the fractional differencing parameter (d) distribution for four groups of betas, corresponding to the 1F, 3F, 5F, and 7F factor configurations. For each factor configuration, we pool all series into a single distribution (e.g., for FF3, we plot the distribution of d for the 430 x 3 series). Figure B.1 shows the results obtained using the Whittle method. From the results, we see that, in all four cases, few series have values of d close to zero, which evidences long memory.

Figure B.2 shows similar results for the GPH (Geweke and Porter-Hudak) estimator. Again, in all cases, most series have a value of d different from zero. We can also see that the distributions are centered close to 0.5 and away from 1, which indicates stationarity.

Figure B.1: Estimated Fractional Differencing Parameter (d) - Whittle Method

These panels show the distribution of fractional differencing parameter (d) estimated by the Whittle method. Each panel corresponds to one of the factor configurations (1, 3, 5, or 7 factors). In total, there are $430 \times$ number of factors series of betas for each case. We plot the distribution of d for these series.

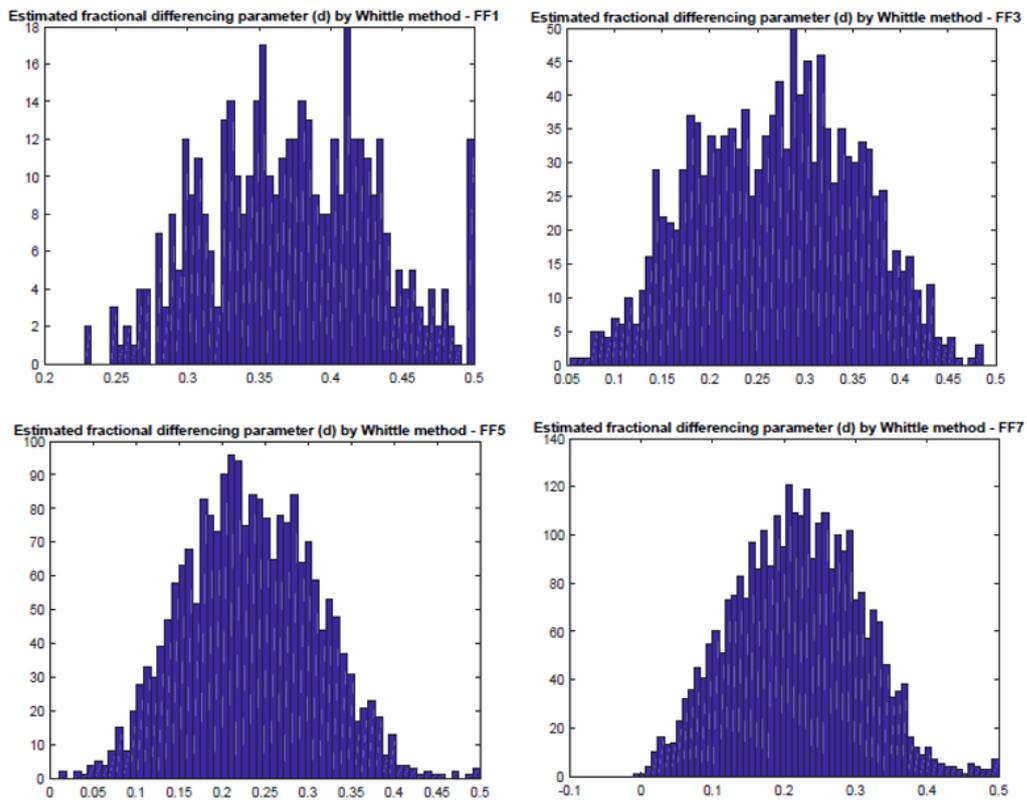
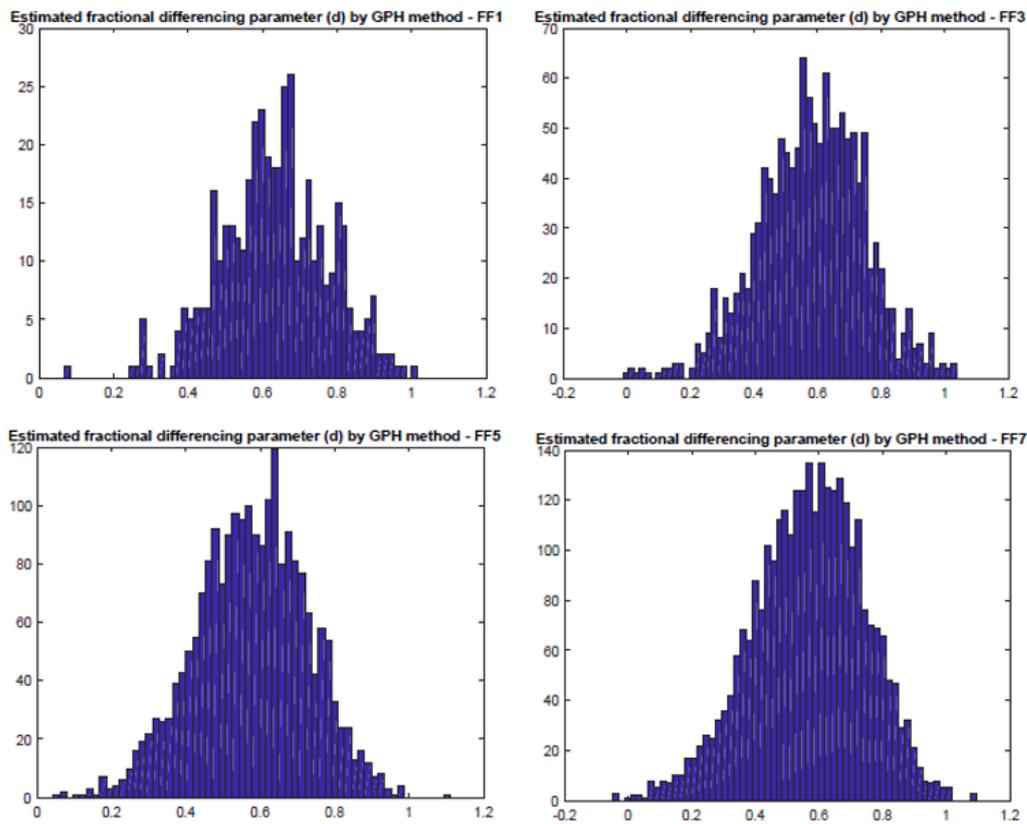


Figure B.2: Estimated Fractional Differencing Parameter (d) - GPH Method

These panels show the distribution of fractional differencing parameter (d) estimated by the GPH method. Each panel corresponds to one of the factor configurations (1, 3, 5, or 7 factors). In total, there are $430 \times$ number of factors series of betas for each case. We plot the distribution of d for these series.



C

Factor Construction

Market: The market portfolio is a value-weighted portfolio of all stocks in our sample. The market factor is the excess return of this portfolio relative to the risk-free rate.

Size and Value: These factors are based on a double sorting of stocks on market equity and book-to-market equity. Market equity is defined as the last available price in June times the corresponding value of shares outstanding. Book-to-market ratio is calculated using the book value of the previous fiscal year and the market equity from the last day of December of the previous year. Stocks are ranked on market equity and split into two portfolios, small and big (S and B). Book-to-market is used to split the stocks into three book-to-market equity groups based on breakpoints for the bottom 30% (low, or L), middle 40% (medium, or M), and high 30% (high, or H). Six portfolios are formed from intersections of the aforementioned groups: S/L, S/M, S/H, B/L, B/M, and B/H. Daily value-weighted returns are calculated for the six portfolios, for the whole year (portfolios are rebalanced annually). The factor SMB (small minus big) is the difference between the simple average of returns on the three small-stock portfolios and the simple average of returns on the three big-stock portfolios, that is,

$$SMB = \frac{(R_{S/L} + R_{S/M} + R_{S/H})}{3} - \frac{(R_{B/L} + R_{B/M} + R_{B/H})}{3}.$$

Similarly, the factor HML (high minus low) is defined as the daily difference of the simple average of returns on the two high book-to-market portfolios (S/H and B/H) and the simple average of returns on the two low book-to-market portfolios (S/L and B/L), that is,

$$HML = \frac{(R_{S/H} + R_{B/H})}{2} - \frac{(R_{S/L} + R_{B/L})}{2}.$$

Gross Profitability: The factor construction follows (17). The signal gross profitability (GP) is defined as the ratio between gross profits and total assets. Stocks are ranked and split in 10 deciles every year, using current year financial data. The gross profitability portfolio consists of a strategy that is long in the group with the lowest GP and short in the group with the highest GP. For

both groups, value-weighted returns are calculated, and the factor GP is given by

$$GP = R_{lowGP} - R_{highGP}.$$

Accruals: This factor follows (20). The signal accrual is defined as

$$\text{Accruals} = \frac{\Delta ACT - \Delta CHE - \Delta LCT + \Delta DLC + \Delta TXP - DP}{(AT + AT_{-12})/2},$$

where ACT is the annual total current assets, CHE is the annual total cash and short-term investments, LCT is the annual current liabilities, DLC is the annual debt in current liabilities, TXP is the annual income taxes payable, DP is the annual depreciation and amortization, and $(AT + AT_{-12})/2$ represents the average total assets over the last two years. Δ stands for the annual variation in these variables. Stocks are split in 10 deciles every year, and two value-weighted portfolios are formed on the lowest and highest deciles. The accruals portfolio consists of a strategy that is long on the small accrual portfolio and short on the high accrual portfolio. The factor is then given by

$$\text{Accrual} = R_{lowAccrual} - R_{highaccrual}.$$

Asset Growth: This factor follows (19). The asset growth signal is defined as $\text{Asset Growth} = AT/AT_{-12}$, with AT and AT_{-12} previously defined in the accruals section. Analogous to the procedure used for the accruals factor, the asset growth portfolio is simply a strategy that is long on the lowest decile asset growth portfolio and short on the highest decile portfolio (the stocks are sorted into 10 deciles). Following this procedure, the factor is then given by

$$\text{AssetGrowth} = R_{lowAssetGrowth} - R_{highAssetGrowth}.$$

Investment: The factor construction follows (18). The investment signal is defined as

$$\text{Investment} = (\Delta PPEGT + \Delta INVT)/AT_{-12},$$

where $PPEGT$ is the gross total property, plant, and equipment (COMPUSTAT's variable 'ppegt'), and $INVT$ is total inventories (COMPUSTAT's 'invnt'). For the investment portfolio, the stocks are triple sorted in size, value (as in the Fama-French factors), and investment. For each of the three characteristics, each stock is classified into one of three groups: low (30%), medium (30%-70%), or high (70%-100%). This procedure results in 27 different portfolios. The investment portfolio consists of a strategy that is long on the low investment portfolio (the simple average between the 9 groups with low invest-

ment) and short on the high investment portfolio (the simple average between the 9 groups with high investment), that is,

$$Investment = R_{lowInvestment} - R_{highInvestment}.$$

Risk Free: We obtain the daily series of returns from Kenneth French's website.

D Diagnostic Criterion for Approximate Factor Structure

In this appendix, we follow the procedure described in (22) to check for the existence of omitted factors in our models. The diagnostic criterion for omitted factors is based on the following expression:

$$\xi(k) = \mu_{k+1} \left(\frac{1}{NT} \sum_i \epsilon_i \epsilon_i' \right) - g(N, T),$$

where $\mu_{k+1}(\cdot)$ denotes the $(k+1)$ -th largest eigenvalue of a symmetric matrix, N represents the number of assets, T represents the sample size, and ϵ_i is a vector of dimension $T \times 1$ with the residuals from the i -th asset relative to the model being tested. $g(N, T)$ is a penalty function given by $g(N, T) = \left[\frac{N+T}{NT} \log\left(\frac{NT}{N+T}\right) \right] \hat{\sigma}^2$, where $\hat{\sigma}^2$ is the estimated variance of the residuals. The criterion detects k omitted factors in the residuals when $\xi(k) < 0$. To illustrate, for a given specification, if $\xi(2) > 0$ and $\xi(3) < 0$, the residuals present a factor structure with two omitted factors.

In Figure D.1, we show the values of $\xi(k)$ for four factor configurations: 1F, 3F, 5F, and 7F. In all cases, we detect the presence of omitted factors, which suggests that imposing more structure on the residuals is appropriate. Since we used sector classifications to restrict the residual dynamics, we also evaluate the criterion in one of these cases. In Figure D.2, we show similar results for the better performing model of Section 6, that is, the 7F with block diagonal residual covariance matrices. From the results, we verify a reduction in the number of omitted factors from five to three in this case.

Figure D.1: Approximate Factor Structure for the 1F, 3F, 5F, and 7F Models

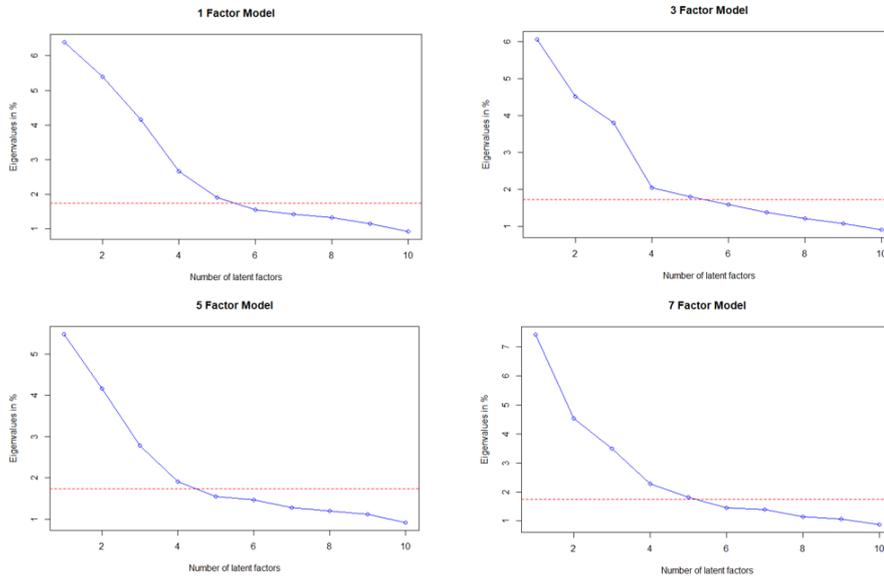
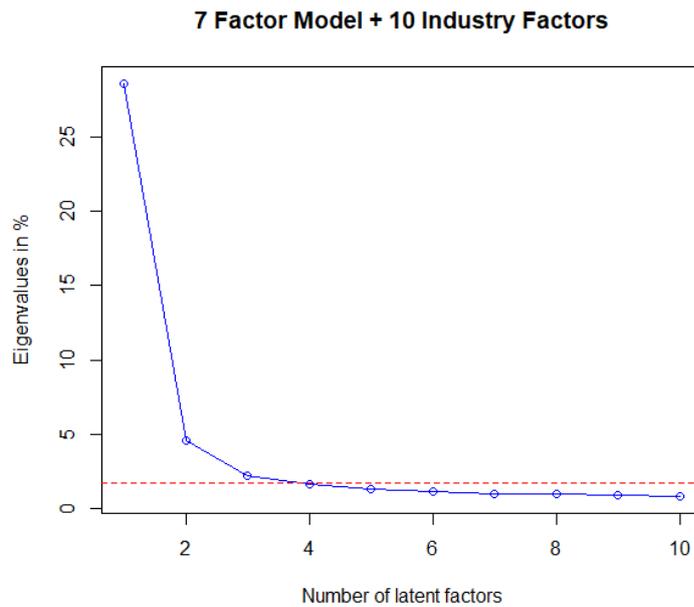


Figure D.2: Approximate Factor Structure for the 7F + 10 Industry Factors Model



E

Residual Covariance Matrix - Parameter Selection

In this appendix, we show the average equation size and average parameter change for models used to forecast the residual covariance matrices. Since the number of categories of equations is much higher than the one used for the factor covariance matrices, we present the results in percentiles of the distribution (instead of individual plots).

Table E.1: Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (1-Factor VHAR).

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0.25	0.5	0.75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	11.29	11.90	12.16	12.48	13.55	12.25
Consumer Durables	6.12	6.38	6.62	6.88	7.62	6.69
Manufacturing	16.05	16.57	16.74	16.92	17.26	16.74
Oil, Gas, and Coal Extraction	9.38	10.16	10.47	10.66	11.38	10.43
Business Equipment	14.70	15.15	15.33	15.52	15.97	15.33
Telecommunications	6.50	6.90	7.40	7.60	7.90	7.28
Wholesale and Retail	12.51	13.27	13.44	13.6	14.00	13.43
Health Care, Medical Equipments, and Drugs	11.69	12.19	12.46	12.85	13.35	12.5
Utilities	11.86	12.56	12.83	13.17	13.72	12.85
Others	15.46	15.88	16.52	16.97	18.1	16.47
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.77	1.07	1.51	1.63	1.75	1.36
Consumer Durables	0.39	0.96	1.50	1.68	1.93	1.35
Manufacturing	0.51	0.60	0.69	0.76	0.82	0.68
Oil, Gas, and Coal Extraction	1.72	2.03	2.43	2.59	3.00	2.33
Business Equipment	0.37	0.45	0.60	0.64	0.69	0.55
Telecommunications	1.44	1.67	1.78	1.89	2.18	1.79
Wholesale and Retail	0.56	0.69	1.03	1.14	1.28	0.95
Health Care, Medical Equipments, and Drugs	0.58	0.85	1.39	1.83	2.02	1.34
Utilities	0.69	0.77	0.84	1.48	1.92	1.09
Others	0.48	0.54	0.64	0.74	0.77	0.63

Table E.2: Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (1-Factor VHAR).

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0.25	0.5	0.75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	0.21	0.94	1.25	1.66	3.75	1.33
Consumer Durables	0.00	0.00	0.00	1.56	4.69	0.52
Manufacturing	0.17	0.78	0.99	1.25	3.48	1.05
Oil, Gas, and Coal Extraction	0.10	0.88	1.37	1.86	4.30	1.43
Business Equipment	0.46	0.96	1.16	1.43	4.03	1.21
Telecommunications	0.00	0.00	0.00	1.00	4.00	0.65
Wholesale and Retail	0.20	0.89	1.14	1.43	2.91	1.19
Health Care, Medical Equipment, and Drugs	0.00	0.74	1.04	1.48	6.07	1.16
Utilities	0.15	0.62	0.93	1.16	4.32	0.94
Others	0.37	0.68	0.79	0.94	1.81	0.82
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.10	0.30	0.39	0.48	0.96	0.39
Consumer Durables	0.00	0.45	0.89	1.34	4.46	0.94
Manufacturing	0.07	0.11	0.13	0.14	0.24	0.13
Oil, Gas, and Coal Extraction	0.26	0.55	0.67	0.79	1.32	0.68
Business Equipment	0.05	0.09	0.10	0.12	0.17	0.1
Telecommunications	0.00	0.44	0.89	1.33	3.56	0.9
Wholesale and Retail	0.07	0.17	0.21	0.26	0.42	0.21
Health Care, Medical Equipment, and Drugs	0.11	0.30	0.43	0.58	1.04	0.45
Utilities	0.06	0.22	0.31	0.41	0.78	0.32
Others	0.04	0.06	0.07	0.07	0.1	0.07

Table E.3: Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (3-Factor VHAR).

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0,25	0,5	0,75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	11.19	11.97	12.19	12.48	13.61	12.29
Consumer Durables	6.25	6.50	6.62	7.00	7.75	6.79
Manufacturing	16.22	16.69	16.85	17.05	17.46	16.86
Oil, Gas, and Coal Extraction	9.00	9.78	10.19	10.59	11.12	10.17
Business Equipment	15.03	15.46	15.64	15.77	16.18	15.62
Telecommunications	6.60	6.90	7.30	7.60	7.90	7.27
Wholesale and Retail	12.80	13.36	13.53	13.71	14.22	13.55
Health Care, Medical Equipment, and Drugs	11.54	12.19	12.42	12.92	13.58	12.53
Utilities	12.03	12.75	13.00	13.25	14.25	13.03
Others	15.41	15.85	16.47	17.13	18.11	16.53
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.66	0.83	1.28	1.39	1.52	1.15
Consumer Durables	0.21	0.36	0.54	0.68	0.89	0.53
Manufacturing	0.47	0.55	0.63	0.68	0.72	0.61
Oil, Gas, and Coal Extraction	1.46	1.76	2.09	2.23	2.64	2.01
Business Equipment	0.34	0.40	0.55	0.67	0.71	0.54
Telecommunications	1.36	1.53	1.62	1.73	2.11	1.65
Wholesale and Retail	0.51	0.60	0.89	1.02	1.09	0.83
Health Care, Medical Equipment, and Drugs	0.42	0.73	1.12	1.58	1.77	1.13
Utilities	0.67	0.74	0.85	1.58	2.06	1.14
Others	0.52	0.60	0.65	0.70	0.72	0.64

Table E.4: **Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (3-Factor VHAR).**

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0,25	0,5	0,75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	0.21	0.94	1.25	1.66	3.85	1.33
Consumer Durables	0.00	0.00	0.00	1.56	3.12	0.52
Manufacturing	0.33	0.83	1.02	1.23	3.36	1.08
Oil, Gas, and Coal Extraction	0.00	0.68	1.17	1.56	4.49	1.19
Business Equipment	0.40	0.94	1.18	1.40	3.60	1.21
Telecommunications	0.00	0.00	0.00	1.00	5.00	0.57
Wholesale and Retail	0.25	0.94	1.19	1.43	3.80	1.24
Health Care, Medical Equipment, and Drugs	0.00	0.74	1.04	1.48	4.88	1.16
Utilities	0.08	0.62	0.93	1.23	4.55	0.98
Others	0.31	0.67	0.80	0.94	1.63	0.82
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.10	0.25	0.33	0.40	0.64	0.33
Consumer Durables	0.00	0.00	0.45	0.89	3.12	0.50
Manufacturing	0.06	0.09	0.11	0.12	0.20	0.11
Oil, Gas, and Coal Extraction	0.15	0.45	0.56	0.69	1.16	0.58
Business Equipment	0.04	0.08	0.10	0.12	0.20	0.10
Telecommunications	0.00	0.44	0.67	1.11	3.33	0.78
Wholesale and Retail	0.07	0.15	0.18	0.22	0.42	0.19
Health Care, Medical Equipment, and Drugs	0.05	0.25	0.34	0.45	0.86	0.36
Utilities	0.08	0.21	0.31	0.42	0.78	0.33
Others	0.05	0.06	0.07	0.07	0.10	0.07

Table E.5: Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (5-Factor VHAR).

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0,25	0,5	0,75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	11.32	11.97	12.19	12.45	13.61	12.27
Consumer Durables	6.12	6.50	6.62	7.00	7.62	6.77
Manufacturing	15.91	16.46	16.72	16.97	17.57	16.72
Oil, Gas, and Coal Extraction	10.28	10.75	11.12	11.38	12.06	11.08
Business Equipment	15.00	15.34	15.52	15.72	16.39	15.55
Telecommunications	6.50	6.90	7.30	7.70	7.90	7.30
Wholesale and Retail	12.56	13.22	13.40	13.62	14.24	13.42
Health Care, Medical Equipment, and Drugs	11.69	12.15	12.54	12.96	13.50	12.55
Utilities	13.47	14.22	14.36	14.56	14.97	14.37
Others	15.35	15.94	16.60	17.22	18.12	16.61
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.59	0.83	1.33	1.43	1.52	1.17
Consumer Durables	0.21	0.39	0.46	0.57	0.79	0.48
Manufacturing	0.46	0.53	0.61	0.63	0.69	0.59
Oil, Gas, and Coal Extraction	1.38	1.67	1.73	1.95	2.33	1.80
Business Equipment	0.37	0.42	0.57	0.68	0.74	0.56
Telecommunications	1.13	1.44	1.56	1.69	1.91	1.55
Wholesale and Retail	0.59	0.70	0.97	1.06	1.13	0.90
Health Care, Medical Equipment, and Drugs	0.52	0.75	1.20	1.54	1.78	1.16
Utilities	0.64	0.74	0.95	1.15	1.39	0.95
Others	0.60	0.67	0.70	0.74	0.77	0.70

Table E.6: Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (5-Factor VHAR).

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0,25	0,5	0,75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	0.10	0.83	1.25	1.56	3.95	1.27
Consumer Durables	0.00	0.00	0.00	0.00	3.12	0.41
Manufacturing	0.31	0.88	1.07	1.30	3.98	1.12
Oil, Gas, and Coal Extraction	0.10	0.59	0.88	1.27	4.49	0.97
Business Equipment	0.27	0.97	1.16	1.37	3.55	1.19
Telecommunications	0.00	0.00	0.00	1.00	6.00	0.61
Wholesale and Retail	0.20	0.94	1.23	1.53	4.10	1.27
Health Care, Medical Equipment, and Drugs	0.00	0.59	1.04	1.63	4.44	1.17
Utilities	0.00	0.85	1.16	1.47	5.71	1.19
Others	0.42	0.66	0.77	0.91	1.66	0.81
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.10	0.26	0.33	0.41	0.68	0.33
Consumer Durables	0.00	0.00	0.00	0.45	3.57	0.41
Manufacturing	0.05	0.09	0.10	0.12	0.17	0.10
Oil, Gas, and Coal Extraction	0.14	0.42	0.49	0.58	0.95	0.50
Business Equipment	0.04	0.08	0.10	0.12	0.18	0.10
Telecommunications	0.00	0.22	0.67	0.89	3.11	0.68
Wholesale and Retail	0.08	0.16	0.20	0.24	0.39	0.20
Health Care, Medical Equipment, and Drugs	0.07	0.26	0.39	0.50	1.11	0.39
Utilities	0.08	0.20	0.28	0.36	0.59	0.28
Others	0.05	0.07	0.07	0.08	0.10	0.07

Table E.7: Average Equation Size Distribution for Blocks in the Residual Covariance Matrix (7-Factor HVAR).

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0,25	0,5	0,75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	11.29	11.94	12.13	12.48	13.74	12.24
Consumer Durables	6.12	6.38	6.50	6.88	7.62	6.65
Manufacturing	15.82	16.38	16.65	16.91	17.37	16.65
Oil, Gas, and Coal Extraction	10.12	10.81	11.03	11.34	11.97	11.09
Business Equipment	14.72	15.20	15.38	15.56	16.05	15.37
Telecommunications	6.40	6.80	7.30	7.50	8.00	7.23
Wholesale and Retail	12.73	13.31	13.47	13.69	14.33	13.49
Health Care, Medical Equipment, and Drugs	11.77	12.31	12.69	13.15	13.73	12.73
Utilities	13.69	14.22	14.36	14.53	14.92	14.35
Others	15.36	15.93	16.57	17.12	18.05	16.59
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.59	0.82	1.33	1.41	1.51	1.15
Consumer Durables	0.14	0.36	0.46	0.54	0.75	0.45
Manufacturing	0.50	0.56	0.63	0.65	0.72	0.61
Oil, Gas, and Coal Extraction	1.40	1.64	1.77	2.05	2.42	1.84
Business Equipment	0.54	0.62	0.77	0.90	0.95	0.76
Telecommunications	1.58	1.82	1.91	2.02	2.56	1.93
Wholesale and Retail	0.62	0.73	1.03	1.11	1.16	0.93
Health Care, Medical Equipment, and Drugs	0.54	0.77	1.02	1.41	1.65	1.06
Utilities	0.78	0.88	1.06	1.12	1.25	1.00
Others	0.63	0.71	0.75	0.78	0.81	0.74

Table E.8: **Average Parameter Change Distribution for Blocks in the Residual Covariance Matrix (7-Factor HVAR).**

The distribution represents the percentiles and mean value calculated over 473 estimated models. The 430 stocks are separated into 10 groups, according to SIC. Averages are calculated among stocks in the same group.

Distribution	Min	0.25	0.5	0.75	Max	Mean
	Variance Equations					
<i>Group</i>						
Consumer Non-Durables	0.00	0.94	1.25	1.66	3.64	1.32
Consumer Durables	0.00	0.00	0.00	0.00	3.12	0.40
Manufacturing	0.38	0.88	1.04	1.25	3.74	1.10
Oil, Gas, and Coal Extraction	0.00	0.59	0.88	1.17	3.71	0.92
Business Equipment	0.46	0.91	1.13	1.35	3.14	1.16
Telecommunications	0.00	0.00	0.00	1.00	8.00	0.67
Wholesale and Retail	0.25	0.89	1.19	1.53	3.90	1.24
Health Care, Medical Equipment, and Drugs	0.00	0.59	1.04	1.63	4.44	1.19
Utilities	0.23	0.85	1.16	1.47	6.10	1.22
Others	0.36	0.65	0.77	0.92	1.73	0.80
	Covariance Equations					
<i>Group</i>						
Consumer Non-Durables	0.05	0.25	0.33	0.40	0.65	0.33
Consumer Durables	0.00	0.00	0.00	0.89	2.23	0.43
Manufacturing	0.06	0.09	0.10	0.12	0.18	0.11
Oil, Gas, and Coal Extraction	0.17	0.43	0.51	0.61	0.92	0.52
Business Equipment	0.06	0.11	0.12	0.14	0.25	0.13
Telecommunications	0.00	0.44	0.89	1.33	4.00	0.93
Wholesale and Retail	0.08	0.17	0.20	0.24	0.40	0.20
Health Care, Medical Equipment, and Drugs	0.06	0.25	0.34	0.44	0.75	0.35
Utilities	0.10	0.22	0.29	0.36	0.63	0.30
Others	0.05	0.07	0.08	0.08	0.10	0.08

F

Minimum Variance Portfolio: Partial Rebalancing

In this section, we show the return statistics for cases in which we do not fully rebalance the portfolio on a daily basis. We assume that each day, the investor updates only $1/22$ of the portfolio based on the predicted covariance matrix for the following day, implying that every daily portfolio update has an effective one-month holding period. This change accounts for the possible effect of transaction costs, as only a fraction of the portfolio changes in any single day. Indirectly, we also test whether investors with a one-month horizon could also benefit from using covariance estimates from the proposed model, which are designed to forecast one-day ahead covariance matrices.

Tables F.1, F.2 and F.3 present results for the unrestricted, restricted, and long-only cases, respectively. Our proposed model continues to perform better than the alternatives, even when we focus on longer holding periods.

Table F.1: **Statistics for Daily Average Portfolios - Global Minimum Variance.**

The table reports the results for portfolios constructed according to the optimization problem as in (6-1) with partial rebalancing. Different models are used to provide the one-step-ahead forecasts for the realized covariance matrix: RW is the random walk model, and Blocks 1F, 3F, 5F, and 7F are random walks applied to the residual covariance matrix after the four different factor decompositions.

	RW	Block 1F	Block 3F	Block 5F	Block 7F	EWMA (Returns)	BEKK-NL	DCC - NL	
Standard Deviation (%)	11.08	9.49	9.57	9.40	9.43	14.00	9.58	10.64	
Lower Partial Standard Deviation (%)	12.57	10.69	10.75	10.48	10.50	14.16	9.65	11.09	
Kurtosis	4.86	5.11	4.76	4.66	4.94	2.09	1.57	4.08	
Skewness	-0.72	-0.88	-0.85	-0.81	-0.84	-0.04	-0.18	-0.09	
Average Diversification Ratio	7.09	7.19	7.29	7.30	7.35	1.08	3.07	3.90	
Average Max. Weight	0.03	0.07	0.07	0.07	0.07	0.19	0.06	0.10	
Average Min. Weight	-0.02	-0.01	-0.01	-0.01	-0.01	-0.14	-0.04	-0.03	
Average Gross Leverage	2.57	2.14	2.15	2.15	2.16	11.89	5.01	3.74	
Proportion of Leverage (%)	37.74	42.09	41.75	41.36	41.54	49.07	45.15	50.94	
Average Turnover (%)	0.09	0.04	0.04	0.04	0.04	0.08	0.03	0.04	
Average Excess Return (%)	17.06	14.41	15.00	15.61	15.36	3.73	17.52	15.50	
Cumulative Return (%)	36.44	30.21	31.65	33.21	32.56	5.52	38.03	32.61	
Sharpe Ratio	1.54	1.52	1.57	1.66	1.63	0.27	1.83	1.46	
		1 Factor VHAR (Log matrix)		3 Factors VHAR (Log matrix)		5 Factors VHAR (Log matrix)		7 Factors VHAR (Log matrix)	
		LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO
Standard Deviation (%)	9.13	9.11	9.09	9.06	8.98	8.96	8.86	8.84	
Lower Partial Standard Deviation (%)	9.84	9.81	9.79	9.77	9.72	9.72	9.72	9.76	
Kurtosis	2.76	2.83	2.37	2.40	2.18	2.21	2.72	2.74	
Skewness	-0.53	-0.53	-0.50	-0.50	-0.47	-0.47	-0.58	-0.58	
Average Diversification Ratio	4.84	4.88	5.05	5.08	4.85	4.88	4.96	4.98	
Average Max. Weight	0.06	0.07	0.07	0.07	0.07	0.08	0.07	0.08	
Average Min. Weight	-0.01	-0.01	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	
Average Gross Leverage	2.56	2.56	2.70	2.70	2.74	2.74	2.84	2.84	
Proportion of Leverage (%)	45.61	45.70	44.64	44.78	44.65	44.81	44.95	45.14	
Average Turnover (%)	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	
Average Excess Return (%)	13.75	13.56	15.95	15.66	16.70	16.30	15.60	15.33	
Cumulative Return (%)	28.69	28.23	34.12	33.40	36.05	35.05	33.30	32.63	
Sharpe Ratio	1.51	1.49	1.75	1.73	1.86	1.82	1.76	1.73	

Table F.2: Statistics for Daily Average Portfolios - Restricted Minimum Variance.

The table reports the results for portfolios constructed according to the optimization problem as in (6-2) with partial rebalancing. Different models are used to provide the one-step-ahead forecasts for the realized covariance matrix: RW is the random walk model, and Blocks 1F, 3F, 5F, and 7F are random walks applied to the residual covariance matrix after the four different factor decompositions.

	RW	Block 1F	Block 3F	Block 5F	Block 7F	EWMA (Returns)	BEKK-NL	DCC - NL	
Standard Deviation (%)	13.10	12.79	12.62	12.56	12.63	14.80	15.01	14.55	
Lower Partial Standard Deviation (%)	13.85	13.34	13.19	13.03	13.16	15.53	15.58	15.40	
Kurtosis	3.33	3.10	3.27	3.29	3.40	3.35	3.26	3.14	
Skewness	-0.37	-0.31	-0.35	-0.34	-0.36	-0.35	-0.28	-0.29	
Average Diversification Ratio	4.61	4.53	4.78	4.78	4.78	2.38	2.44	2.35	
Average Max. Weight	0.12	0.11	0.11	0.11	0.11	0.18	0.19	0.19	
Average Min. Weight	-0.05	-0.05	-0.05	-0.05	-0.05	-0.15	-0.14	-0.10	
Average Gross Leverage	1.43	1.48	1.48	1.48	1.48	1.57	1.56	1.55	
Proportion of Leverage (%)	7.63	14.89	14.51	13.71	13.26	1.16	1.64	2.75	
Average Turnover (%)	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02	
Average Excess Return (%)	16.75	18.48	18.64	18.36	17.74	15.38	14.05	13.75	
Cumulative Return (%)	35.03	39.58	40.07	39.34	37.72	31.01	27.71	27.17	
Sharpe Ratio	1.28	1.44	1.48	1.46	1.41	1.04	0.94	0.95	
		1 Factor VHAR (Log matrix)		3 Factors VHAR (Log matrix)		5 Factors VHAR (Log matrix)		7 Factors VHAR (Log matrix)	
		LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO
Standard Deviation (%)	12.68	12.74	12.21	12.35	12.07	12.17	12.19	12.24	
Lower Partial Standard Deviation (%)	12.94	12.95	12.40	12.57	12.41	12.43	12.67	12.72	
Kurtosis	2.59	2.73	2.96	2.99	3.05	3.04	3.19	3.03	
Skewness	-0.23	-0.24	-0.24	-0.26	-0.27	-0.25	-0.30	-0.26	
Average Diversification Ratio	3.62	3.61	3.87	3.86	3.91	3.91	3.87	3.85	
Average Max. Weight	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	
Average Min. Weight	-0.07	-0.07	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	
Average Gross Leverage	1.58	1.58	1.59	1.59	1.59	1.59	1.59	1.59	
Proportion of Leverage (%)	8.52	8.71	7.72	7.97	7.68	7.85	7.12	7.34	
Average Turnover (%)	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	
Average Excess Return (%)	18.12	17.87	17.94	18.10	17.65	17.97	17.20	17.42	
Cumulative Return (%)	38.69	38.02	38.37	38.74	37.65	38.45	36.47	37.00	
Sharpe Ratio	1.43	1.40	1.47	1.47	1.46	1.48	1.41	1.42	

Table F.3: **Statistics for Daily Average Portfolios - Restricted Minimum Variance (Long Only).**

The table reports the results for portfolios constructed according to the optimization problem as in (6-3) with partial rebalancing. Different models are used to provide the one-step-ahead forecasts for the realized covariance matrix: RW is the random walk model, and Blocks 1F, 3F, 5F, and 7F are random walks applied to the residual covariance matrix after the four different factor decompositions.

	RW	Block 1F	Block 3F	Block 5F	Block 7F	EWMA (Returns)	BEKK-NL	DCC - NL
Standard Deviation (%)	16.66	16.73	16.60	16.49	16.51	17.53	17.73	17.58
Lower Partial Standard Deviation (%)	17.85	17.69	17.45	17.40	17.49	18.56	18.78	18.62
Kurtosis	2.70	2.76	2.79	2.78	2.83	2.78	2.53	2.63
Skewness	-0.33	-0.32	-0.32	-0.32	-0.33	-0.32	-0.31	-0.29
Average Diversification Ratio	3.37	3.27	3.30	3.32	3.32	2.72	2.69	2.61
Average Max. Weight	0.15	0.17	0.17	0.17	0.17	0.17	0.17	0.17
Average Min. Weight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Gross Leverage	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Proportion of Leverage (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Turnover (%)	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Average Excess Return (%)	17.41	16.96	17.35	17.36	17.38	17.74	16.90	16.28
Cumulative Return (%)	35.36	34.18	35.23	35.30	35.35	35.81	33.61	32.12
Sharpe Ratio	1.05	1.01	1.05	1.05	1.05	1.01	0.95	0.93

	1 Factor		3 Factors		5 Factors		7 Factors	
	VHAR		VHAR		VHAR		VHAR	
	(Log matrix)		(Log matrix)		(Log matrix)		(Log matrix)	
	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO	LASSO	adaLASSO
Standard Deviation (%)	16.66	16.73	16.60	16.49	16.51	17.53	17.73	17.58
Lower Partial Standard Deviation (%)	17.85	17.69	17.45	17.40	17.49	18.56	18.78	18.62
Kurtosis	2.70	2.76	2.79	2.78	2.83	2.78	2.53	2.63
Skewness	-0.33	-0.32	-0.32	-0.32	-0.33	-0.32	-0.31	-0.29
Average Diversification Ratio	3.37	3.27	3.30	3.32	3.32	2.72	2.69	2.61
Average Max. Weight	0.15	0.17	0.17	0.17	0.17	0.17	0.17	0.17
Average Min. Weight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Gross Leverage	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Proportion of Leverage (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Turnover (%)	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Average Excess Return (%)	17.41	16.96	17.35	17.36	17.38	17.74	16.90	16.28
Cumulative Return (%)	35.36	34.18	35.23	35.30	35.35	35.81	33.61	32.12
Sharpe Ratio	1.05	1.01	1.05	1.05	1.05	1.01	0.95	0.93