Long-History PCA in a Dynamic Factor Model with Weak Loadings*

Robert M. Anderson† Baeho Kim‡ Dean Ryu§

First version: June 23, 2024
This version: May 15, 2025

Abstract

Estimated covariance and precision matrices of asset returns significantly influence the set of portfolios compliant with risk budgets and their potential losses. Statistical risk modeling approaches often assume temporal stability for consistency with a static factor structure, typically estimated within $T_M \sim 250$ days of data history, resulting in finite-sample estimation error when the dimension of the population exceeds the number of observations. Our study introduces Long-History PCA (LH-PCA), utilizing extended data histories (e.g., $T_L \sim 1500$ trading days) to forecast the daily risk profile based on dynamic factor structures with heterogeneous factor strengths. The use of a longer data history mitigates excess dispersion bias in the estimated factor loadings, particularly in the presence of weak factors. As shown in simulations and empirical data from the United States and European stock markets, our approach substantially mitigates second-order risk bias compared to traditional methods using medium horizons (T_M), both with and without the augmentation of Responsive Covariance Adjustment (RCA) using a short half-life (T_S) of 40 days.

Keywords— Factor Model, Principal Component Analysis (PCA), Long-History PCA (LH-PCA), Second Order Risk

*We thank Jeffrey Bohn, Young Ho Eom, Lisa Goldberg, Markus Pelger, Raman Uppal, and seminar participants at FMA 2023, Informs annual meeting 2024, Korea University Business School, SWUFE, UC Berkeley, University of Oxford Saïd Business School, and Yonsei University Business School for many valuable comments. Anderson gratefully acknowledges the financial support of Swiss Re through the Consortium for Data Analytics in Risk. This study is partially supported by Korea University Business School Research Grant.

†University of California, Berkeley, `robert.anderson@berkeley.edu`

‡Corresponding author. Korea University Business School, `baehokim@korea.ac.kr`

§Saïd Business School, University of Oxford, `dean.ryu@sbs.ox.ac.uk`

1 Introduction

In the classical mean-variance framework, second-order risk (SOR) arises from the uncertain nature of the joint distribution of asset returns, particularly due to inaccuracies in estimating the population covariance matrix (Σ) or its inverse, the precision matrix (Σ^{-1}). Small errors in these estimates can significantly distort efficient allocations in portfolio optimization,[1] potentially biasing the optimized portfolio toward the deficiencies of the estimated model (Shepard, 2009; Bernardi et al., 2019). Consequently, understanding and mitigating SOR bias is crucial for investors seeking admissible portfolio compositions that align with a predefined risk budget.

Nevertheless, quantifying and managing SOR remains challenging due to the unobservable and rapidly evolving nature of the true covariance structure of asset returns. Statistical factor-based risk modeling approaches, such as Principal Component Analysis (PCA), often rely on assumptions of temporal stability and uniform factor strengths[2] to ensure consistent factor identification. These assumptions frequently fail to capture the dynamic complexities of real-world financial data. As a result, practitioners tend to limit their analysis to medium-length data histories (e.g., $T_M \sim 250$ or 500 trading days)[3] to estimate factor exposures, while using short half-life techniques, such as the Exponentially Weighted Moving Average (EWMA) with a half-life of $T_S \sim 40$ days, to estimate current factor variances.[4] For portfolios of practical relevance, which often include $N \sim 2000$ or more stocks, this mismatch leads to an inaccurate representation of the underlying covariance structure and exacerbates the SOR bias in portfolio optimization.

To address this challenge, this paper introduces the Long-History PCA (LH-PCA), a method designed to accurately forecast the precision matrix on a daily basis using an extended historical window (e.g., $T_L \sim 1500$ trading days). Building on the seminal contributions of Bates et al. (2013) and Bai and Ng (2023), we propose a dynamic factor model that accounts for variable factor volatilities and temporal instability in factor loadings, allowing for variations in their strength. We prove that PCA consistently estimates the underlying factor structure in the 'large-N and large-T' framework, enabling precise risk forecasts for portfolios of practical and reasonable size.

Beyond theoretical advancements, our study makes practical contributions by developing a robust framework for detecting and quantifying second-order risk (SOR) bias in high-dimensional portfolios. As the true return covariance structure is unobservable in practice, we propose an observable proxy of the SOR bias by constructing the Global Minimum Variance Portfolio (GMVP) using the empirically estimated precision matrix and comparing its out-of-sample realized volatility to its predicted volatility. This approach enables the development of reliable bias statistics that capture discrepancies between estimated and actual portfolio risk, providing an effective tool for risk-aware investment decisions.

Empirically, we validate the effectiveness of LH-PCA through Monte Carlo simulations and real-

[1] Neglecting the misestimation of expected returns can have a significant impact, similar to the mismeasurement of the covariance matrix. However, this study focuses solely on covariance matrix estimation, leaving expected return estimation errors for future research. Notably, the significance of expected return estimation errors may diminish in short-horizon optimization scenarios.

[2] We use the terms *strong* and *weak* factor loadings within the framework of Bai and Ng (2023), which extends and generalizes the concepts introduced in Lettau and Pelger (2020), Freyaldenhoven (2022), Uematsu and Yamagata (2023) among others.

[3] Throughout the paper, we use T_L, T_M, and T_S to represent Long, Medium, and Short historical windows, respectively.

[4] An exception is Northfield, which employs a hybrid model using 60 months of exponentially weighted regressions on fundamental factors and PCA factors derived from residuals. However, it still contends with SOR bias due to high dimensionality, typically exceeding $T \sim 60$.

world stock market data, including U.S. and European equities. Compared to traditional benchmarks, such as variants of PCA with medium-length historical windows and shrinkage techniques (Ledoit and Wolf, 2004b), we demonstrate that LH-PCA significantly reduces second-order risk (SOR) bias by extending the data history from T_M to T_L. This improvement remains robust across diverse scenarios, including the use of responsive covariance adjustments (RCA) with short data histories (e.g., EWMA with a half-life of $T_S \sim 40$ days). Our findings confirm that LH-PCA enhances risk estimation accuracy in fast-evolving financial markets, establishing it as a valuable tool for high-dimensional portfolio construction and risk management.

Overall, the proposed LH-PCA approach provides a robust and intuitive framework for analyzing high-dimensional financial datasets, particularly for portfolio optimization in the context of modern *big data* analytics. By effectively reducing dimensionality while retaining key factors, it facilitates accurate estimation of covariance and precision matrices, which are critical for constructing resilient and well-diversified portfolios. The adaptability of the proposed approach to both strong and weak factors, as well as dynamic and static loadings, renders it well-suited for the modeling of pragmatic dynamic factor structures, thereby addressing the limitations of traditional approaches.

Finally, our work complements research on estimation error in covariance matrices, including model-free shrinkage methods (Ledoit and Wolf, 2003, 2004a,b), eigenvalue corrections (Ledoit and Péché, 2011; Wang and Fan, 2017), and eigenvector adjustments (Goldberg et al., 2022). By addressing the challenges of SOR in high-dimensional portfolios, our study provides a robust framework for improving precision matrix estimation and enhancing portfolio optimization.

2 Related literature

Our study contributes to the literature on portfolio risk optimization, particularly in improving covariance matrix estimation for large portfolios. Traditional research on estimation errors in the Markowitz framework (Markowitz, 1952) has largely focused on portfolio allocation strategies, such as constructing minimum-variance portfolios (GMVP). Among various asset allocation models (see Table 1 in DeMiguel et al., 2009b), GMVP has been widely studied due to its ability to minimize risk while avoiding reliance on noisy mean return estimates (Jagannathan and Ma, 2003). Moreover, GMVPs serve as foundational building blocks for other portfolio strategies (DeMiguel et al., 2009a) and are favored by investors for their attractive risk-adjusted returns. Recent refinements to GMVP have been proposed by Basak et al. (2009), Shi et al. (2020), and Ding et al. (2021).

Our approach diverges from this literature by focusing on daily forecasting of the precision matrix for large portfolios, a critical component in managing second-order risk (SOR). To achieve this, we emphasize identifying all relevant risk factors, regardless of their risk premiums. Including factors with negative or zero risk premiums can inadvertently increase portfolio volatility without enhancing expected returns. While we do not assess whether factors are priced—since a factor can influence volatility without carrying a risk premium—our findings on SOR contribute to recent asset pricing discussions (Aït-Sahalia et al., 2024; Anderson et al., 2009; Bekaert et al., 2009a; Brenner and Izhakian, 2018). Given the need for robust risk estimation, we leverage PCA-based methods, which are computationally efficient in capturing market dynamics.

We also build upon research improving PCA methodologies. RP-PCA (Lettau and Pelger, 2020) extracts weak factors with high Sharpe ratios, Projected-PCA (Kim et al., 2021) constructs portfolios that hedge systematic risks, and Instrumental-PCA (Kelly et al., 2019) identifies latent factors explaining cross-sectional stock returns. These methods estimate low-dimensional factor structures under Arbitrage

Pricing Theory (Ross, 1976), prioritizing priced risks but often overlooking non-benchmark assets' time-series information (Pastor, 2000), which is central to our study. While fundamental risk factors such as SMB and HML (Fama and French, 1993) are interpretable, they often exhibit strong correlations. In contrast, PCA extracts orthogonal *latent* factors that efficiently capture market trends despite limited interpretability.[5] Given that capturing short-term trends—whether or not they persist long enough to be integrated into a fundamental factor model—is crucial for dynamic investment strategies, PCA-based models play a key role in adapting to evolving market conditions (Fan et al., 2016; Lettau and Pelger, 2020).

The literature on Dynamic Factor Models (DFMs) has evolved to address challenges in modeling time-varying dynamics in high-dimensional data. Early contributions, such as Del Negro and Otrok (2008), introduced models with time-varying factor loadings and stochastic volatility, shedding light on international business cycle fluctuations. Su and Wang (2017) extended this work with a nonparametric approach to estimate DFMs with smoothly varying loadings, enabling tests for structural changes. More recently, Pelger and Xiong (2022b) developed an inferential framework for state-varying factor models, improving the capture of time-dependent factor structures, while Barigozzi et al. (2021) advanced estimation methods for locally stationary DFMs, emphasizing financial connectedness and heterogeneous shock responses. Complementing these studies, our work leverages DFMs to estimate covariance and precision matrices for asset returns, a key component of portfolio risk management. Our proposed LH-PCA utilizes extended data histories to mitigate estimation errors from dispersion bias. This approach enhances traditional medium-horizon approaches by reducing second-order risk bias, as we elaborate in subsequent sections.

3 Problem Formulation

Throughout the paper, we define the matrix norm of a matrix A as the Frobenius norm, represented by $\|A\|_F = \sqrt{\mathrm{tr}(A^\top A)}$. Given $b \in \mathbb{R}^m$, we define $\mathrm{diag}(b)$ to be the $m \times m$ matrix with the elements of b on the diagonal and zeros off the diagonal, while $\mathrm{diag}(B)$ for some $B \in \mathbb{R}^{m\times m}$ represents the vector in $\mathbb{R}^m$ whose elements are the diagonal elements of B. We define the volatility function $\sigma(\gamma, \mathbf{C}) = \sqrt{\gamma^\top \mathbf{C}\gamma} \geq 0$ as the volatility of a given portfolio $\gamma \in \mathbb{R}^{N\times 1}$ calculated under a covariance matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$. As usual, $\mathbf{I}_N \in \mathbb{R}^{N\times N}$ represents an identity matrix, and $\mathbf{1} \in \mathbb{R}^{N\times 1}$ denotes the vector of ones, where all N entries are identical and equal to one.

3.1 Preliminaries

Over a finite time window of length $T > 0$ days, we observe the dynamics of N tradable security returns in the market, where $X_t \in \mathbb{R}^{N\times 1}$ represents the column vector of daily security returns on each date $t \in \{1, \ldots, T\}$. The true population covariance matrix and its estimate of

$$\mathbf{X} = (X_1, \ldots, X_T)^\top \in \mathbb{R}^{T\times N}$$

are indicated by $\Sigma \in \mathbb{R}^{N\times N}$ and $\hat{\Sigma} \in \mathbb{R}^{N\times N}$, respectively.

Estimating the precision matrix (the inverse of the covariance matrix) of stock returns is crucial in optimizing asset allocation by minimizing portfolio risk and enhancing the power of diversification by identifying assets that reduce overall volatility and improving the effectiveness of hedging strategies.

[5] Pelger and Xiong (2022a) discusses interpretation of latent PCA factors.

While the precision matrix enables more informed decision-making by balancing risk and return in a portfolio, misestimating it can lead to flawed conclusions, suboptimal portfolio construction, and increased exposure to unintended risks. Recognizing the critical role of the precision matrix managing portfolio-level risk, we define the SOR bias matrix as $\epsilon = \hat{\Sigma}^{-1} - \Sigma^{-1} \in \mathbb{R}^{N \times N}$ and set the SOR bias measure as

$$\text{(SOR Bias)} = \left\| \frac{1}{N} \left(\hat{\Sigma}^{-1} - \Sigma^{-1} \right) \right\|_F := \|\epsilon\| \ , \tag{1}$$

where the average Frobenius norm serves as a measure of the average error within the entries of the estimated precision matrix. Our objective is to minimize the SOR bias measured by $\|\epsilon\|$ in (1).

A fundamental challenge in detecting and addressing SOR bias lies in its measurement, as the true precision matrix Σ^{-1} is unobservable in practice. We therefore need an observable proxy to quantify the extent of SOR bias resulting from the misrepresentation in $\hat{\Sigma}^{-1}$. A reasonable approach is to construct optimized portfolios (such as the GMVP) based on the estimated $\hat{\Sigma}^{-1}$ and evaluate their out-of-sample performance by comparing it to the in-sample prediction.[6] In the absence of the short-selling constraint within the classical mean-variance framework, the true GMVP (ω) and the estimated GMVP ($\hat{\omega}$) can be expressed as $\omega = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^\top \Sigma^{-1}\mathbf{1}}$ and $\hat{\omega} = \frac{\hat{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1}}$, which are fully characterized by Σ^{-1} and $\hat{\Sigma}^{-1}$, respectively.

Then, the true GMVP variance of ω is given by

$$\text{(True variance of } \omega) = \sigma^2(\omega, \Sigma) = \omega^\top \Sigma \omega = \frac{1}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}} \ .$$

In addition, the (in-sample) predicted variance of the estimated $\hat{\omega}$ takes the form of

$$\text{(Predicted variance of } \hat{\omega}) = \sigma^2(\hat{\omega}, \hat{\Sigma}) = \hat{\omega}^\top \hat{\Sigma} \hat{\omega} = \frac{1}{\mathbf{1}^\top \hat{\Sigma}^{-1} \mathbf{1}} \ ,$$

whereas the (out-of-sample) actual variance of $\hat{\omega}$ to be realized from Σ is derived as

$$\text{(Actual variance of } \hat{\omega}) = \sigma^2(\hat{\omega}, \Sigma) = \hat{\omega}^\top \Sigma \hat{\omega} = \frac{\mathbf{1}^\top \hat{\Sigma}^{-1} \Sigma \hat{\Sigma}^{-1} \mathbf{1}}{(\mathbf{1}^\top \hat{\Sigma}^{-1} \mathbf{1})^2} \ .$$

3.2 Excess Dispersion Bias from Finite Sample Error

Statistical risk modeling approaches typically assume temporal stability to ensure consistency with a static risk structure over the estimation horizon. However, when the number of observations (T) is significantly less than the dimension (N) of $\mathbf{X} \in \mathbb{R}^{T \times N}$, finite-sample estimation error arises and the estimated model is subject to the SOR bias. When $N \gg T$, for example, the sample covariance matrix $\frac{\mathbf{X}^\top \mathbf{X}}{T}$ is singular,[7] so its inverse matrix is not well-defined, the GMVP is not unique, and has estimated variance of zero. One of the most widely used methods to mitigate this bias is shrinkage estimation of the sample covariance matrix, as proposed by Ledoit and Wolf (2004b), which serves as one of our benchmark methods.

In practice, variants of PCA estimation for linear factor models are typically applied using a *medium* horizon of one or two years of historical data to ensure the temporal stability of the factor structure.

[6] As illustrated by Proposition 3, achieving the minimum realized variance of the constructed GMVP, which is a practically important goal, may be conceptually different objective from measuring SOR in our context.

[7] As mean daily returns of stocks are small, and the sample mean is a very noisy estimate of the true mean, it is customary not to demean, and therefore to divide by T rather than $T - 1$.

Assuming $r < \min\{N, T\}$ where r represents the true number of factors, the factor structure is often represented by a *static* linear factor model in the form of

$$\underbrace{\mathbf{X}}_{(T\times N)} = \underbrace{\mathbf{F}}_{(T\times r)}\underbrace{\mathbf{\Lambda}^\top}_{(r\times N)} + \underbrace{\mathbf{e}}_{(T\times N)} ,$$

where $\mathbf{\Lambda} \in \mathbb{R}^{N\times r}$ represents the static loadings of the factor returns $\mathbf{F} = [F_1, \ldots, F_T]^\top \in \mathbb{R}^{T\times r}$, and $\mathbf{e} \in \mathbb{R}^{T\times N}$ denotes the idiosyncratic return matrix.

Let $\text{rank}(\mathbf{X})$ be the full-rank of $\mathbf{X}$ and the time-normalized data $\mathbf{Z} = \frac{\mathbf{X}}{\sqrt{T}}$ admit the singular value decomposition (SVD) expressed as

$$\mathbf{Z} = \mathbf{U}_r\mathbf{D}_r\mathbf{V}_r^\top + \sum_{j=r+1}^{\text{rank}(\mathbf{X})} \mathbf{d}_j\mathbf{u}_j\mathbf{v}_j^\top ,$$

where $\mathbf{D}_r = \text{diag}(\mathbf{d}_1, \ldots, \mathbf{d}_r) \in \mathbb{R}^{r\times r}$ contains the top r singular values of $\mathbf{Z}$ in descending order, and $\mathbf{U}_r = [\mathbf{u}_1, \ldots, \mathbf{u}_r] \in \mathbb{R}^{T\times r}$ and $\mathbf{V}_r = [\mathbf{v}_1, \ldots, \mathbf{v}_r] \in \mathbb{R}^{N\times r}$ form the corresponding left and right singular vectors, respectively.[8] The principal component (PC) estimator with K (that may be different from r) eigenfactors is obtained by solving the optimization problem given by

$$\begin{aligned} &\underset{\hat{\mathbf{F}}_K\in\mathbb{R}^{T\times K}, \hat{\mathbf{\Lambda}}_K\in\mathbb{R}^{N\times K}}{\text{minimize}} \left\| \mathbf{X} - \hat{\mathbf{F}}_K\hat{\mathbf{\Lambda}}_K^\top \right\|_F^2 \\ &\text{subject to} \quad \frac{\hat{\mathbf{F}}_K^\top\hat{\mathbf{F}}_K}{T} = \mathbf{I}_K \quad \text{and} \quad \hat{\mathbf{\Lambda}}_K^\top\hat{\mathbf{\Lambda}}_K \text{ being diagonal} , \end{aligned}$$

where the solution is given by $\left(\tilde{\mathbf{F}}_K, \tilde{\mathbf{\Lambda}}_K\right) = \left(\sqrt{T}\mathbf{U}_K, \mathbf{V}_K\mathbf{D}_K\right)$.

While the factor-based approach is effective in mitigating finite-sample error problems when the number of factors (r) is effectively smaller than the number of observations, it still yields inaccurate estimated factor loadings when this medium horizon (T_M) is smaller than the number of securities. This discrepancy leads to what is known as excess dispersion bias in the estimated factor loadings, which remains one of the primary sources of second-order risk in factor-based models. More specifically, estimated factor loadings are contaminated by idiosyncratic returns when using static PCA with a medium data history $T_M \ll N$, as we have with $\mathbf{Y} = \frac{1}{T}\left(\mathbf{D}_K^{-1}\mathbf{V}_K^\top\right)$

$$\begin{aligned} \tilde{\mathbf{\Lambda}}_K^\top = \mathbf{D}_K\mathbf{V}_K^\top &= \underbrace{\frac{1}{T}\left(\mathbf{D}_K^{-1}\mathbf{V}_K^\top\right)}_{:=\mathbf{Y}} \underbrace{\left(T\mathbf{V}_K\mathbf{D}_K^2\mathbf{V}_K^\top\right)}_{\approx\mathbf{X}^\top\mathbf{X}} \\ &\approx \mathbf{Y}\left(\mathbf{F}\mathbf{\Lambda}^\top + \mathbf{e}\right)^\top\left(\mathbf{F}\mathbf{\Lambda}^\top + \mathbf{e}\right) \\ &= \mathbf{Y}\left(\mathbf{\Lambda}\mathbf{F}^\top\mathbf{F}\mathbf{\Lambda}^\top + \underbrace{\mathbf{\Lambda}\mathbf{F}^\top\mathbf{e} + (\mathbf{\Lambda}\mathbf{F}^\top\mathbf{e})^\top}_{\text{primary source of SOR bias}} + \mathbf{e}^\top\mathbf{e}\right) . \end{aligned}$$

In finite sample, with $K \ll T\,(= T_M) \ll N$, the idiosyncratic returns of some of the stocks will appear to be very significantly correlated with the returns of some of the factors. The realized correlations between the factor returns and the idiosyncratic returns appear explicitly in the terms $\mathbf{\Lambda}\mathbf{F}^\top\mathbf{e} + (\mathbf{\Lambda}\mathbf{F}^\top\mathbf{e})^\top$, adding

[8] The Eckart and Young (1936) theorem implies that $\mathbf{U}_r\mathbf{D}_r\mathbf{V}_r^\top$ is the best rank-r approximation of $\mathbf{Z}$.

dispersion to the estimated factor loadings.

3.3 Biased Estimation of Weak Factor Loadings

We next turn to an examination of weak latent factors. Standard factor models typically assume that the factors are *strong*, which implies that $\frac{\mathbf{\Lambda}^\top \mathbf{\Lambda}}{N}$ is asymptotically positive definite as $N \to \infty$. Recent literature has made efforts to relax this condition, allowing $\frac{\mathbf{\Lambda}^\top \mathbf{\Lambda}}{N^\alpha}$, where $\alpha \in (0, 1]$, to be asymptotically positive definite in the limit; e.g., see Freyaldenhoven (2022), Bai and Ng (2023), and the references therein. Factors with $\alpha < 1$, in particular, are commonly referred to as *weak*.

As noted by Freyaldenhoven (2022), a risk factor can be weak if its influence is limited to a subset of stocks in the market. We classify such weak factors in this context as *narrow*, whereas Freyaldenhoven (2022) refers to them as *local*. The crucial point is that PCA cannot accurately capture narrow factors, as the extracted eigenfactors are supposed to be orthogonal to one another. That is, each eigenfactor may be an amalgam of several narrow factors with broad factors and idiosyncratic returns, while each true narrow factor may be partially captured in each of multiple (orthogonal) eigenfactors.

Although a narrow factor does not appear as an *explicit* factor in the output of PCA, it nonetheless is a source of risk. This risk will be captured in the estimated covariance matrix produced by PCA, as long as we extract enough eigenvectors to ensure that 'the best rank K approximation' in the form of $\mathbf{X}/\sqrt{T} \sim \mathbf{U}_K \mathbf{D}_K \mathbf{V}_K^\top$ is sufficiently accurate.[9] A portfolio that is concentrated in a single industry or country will, all else being equal, be riskier than a portfolio that is diversified over industries and countries. In order to properly evaluate the risk of a portfolio, and especially to produce optimized portfolios, we need our estimated covariance matrix to accurately reflect the narrow factors.

When $K \ll T \ (= T_M) \ll N$, the idiosyncratic returns of some stocks will appear to be very significantly correlated with the returns of some narrow factors, just as is the case with the broad factors. For example, all European stocks are exposed to the market factor, with true factor loadings that are roughly equal. By contrast, only some stocks are exposed to each narrow factor, with true factor loadings that are roughly equal and comparable in magnitude to the true loadings on the market portfolio. The idiosyncratic return contaminates the narrow and market factor loadings essentially equally, but the relative contamination of the narrow factor is much greater than that of the market factor. This contamination assigns nonzero estimated factor loadings on the narrow factors, when the stocks are in fact not exposed to them. One can think of this as excess dispersion of the estimated factor loadings that are, in fact, zero.[10]

3.4 The SOR Bias Mechanism for GMVP

Consider how the excess dispersion affects the construction of the GMVP. The factor loadings of some stocks appear to be smaller than they actually are, and the optimizer will choose to increase their weights in the portfolio. Conversely, the factor loadings of some stocks appear to be larger than they actually are, and the optimizer will choose to decrease their weights in the portfolio. As stated in Remark 1, the excess dispersion resulting from contamination by idiosyncratic volatility results in estimated GMVPs that have

[9] If the approximation is poor, then the risk of the narrow factor will not be captured by PCA. As a consequence, an optimizer will not attempt to diversify away the risk of that narrow factor, and potentially expose the investor to unnecessary and unexpected risk.

[10] Goldberg et al. (2022) (GPS) propose a correction to the estimated *eigenvector* corresponding to the largest eigenvalue (i.e., market factor). But they do not address this excess dispersion in the narrow (or weak) factors.

significantly higher variance than the true GMVP. Moreover, the estimated GMVP is substantially more volatile than predicted.

In the presence of narrow factors, such as those driven by industry or country-specific influences, an optimizer subject to excess dispersion may attempt to offset a narrow factor's exposure. It does so by taking long positions in stocks that appear to be negatively correlated with the factor and short positions in stocks that appear to be positively correlated. However, if these stocks have no true exposure to the narrow factor, the positions fail to provide any actual offset to the factor exposure.[11]

Remark 1 (A consequence of the SOR bias). *A stylized fact stemming from the excess dispersion bias, as detailed in Sections 3.2 & 3.3, is that as the error in the finite-sample estimate of $\hat{\Sigma}^{-1}$ increases, it tends to underestimate the volatility of the estimated GMVP. This indicates that finite-sample estimation errors, in the presence of weak factors, cause an apparent inflation of $\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1}$ relative to its true value, $\mathbf{1}^\top \Sigma^{-1}\mathbf{1}$, as illustrated by*

$$\sigma^2(\hat{\omega}, \hat{\Sigma}) = \frac{1}{\mathbf{1}^\top \Sigma^{-1}\mathbf{1} + \underbrace{\mathbf{1}^\top \epsilon \mathbf{1}}_{\uparrow \ as \ \|\epsilon\|\uparrow}} \downarrow \quad as \quad \|\epsilon\| \uparrow , \tag{2}$$

which shows the mechanism behind the underprediction of the estimated GMVP's (in-sample) predicted volatility.[12]

To tackle the issue of excess dispersion bias, for example, Goldberg et al. (2022) proposed the GPS correction method that corrects the leading eigenvector based on PCA using the medium horizon T_M. They show that the GPS correction can asymptotically mitigate this bias under the *'large-N and fixed-T'* framework. However, our preliminary simulation study indicates that the marginal benefit of using GPS correction is limited with a single layer of narrow factors (e.g., industries) and deteriorates further with multiple layers (e.g., countries and industries). Detailed results are available upon request.

Instead, our approach employs a longer history ($T_L > T_M$) of sample data to estimate the dynamic factor structure, with the goal of mitigating the SOR bias at the portfolio level under the *'large-N and large-T'* framework. This involves extending the history of the sample data to a longer horizon (T_L) in the hope that we can reduce the contamination of the estimated factor loadings. Of course, several challenges arise when extending covariance estimation into the temporal domain, presenting both theoretical and empirical hurdles. As the factor loadings of the securities change over longer horizons, extending the sample period creates challenges in ensuring consistent estimates of principal components in a dynamic factor model. In addition, detecting and extracting weaker (eigen-)factors, which typically contain multiple narrow factor components, poses a challenge due to inconsistencies in principal component estimates when weaker factors predominate. From this perspective, Bai and Ng (2023) argue that accurately estimating weaker loadings requires using a longer T.

[11] When over 90% of the stocks have zero true exposure to a given narrow factor, the number of false positives—stocks that falsely appear to have nonzero exposure at the 5% confidence level—will be about half the number of true positives. As a consequence, about one-third of the stocks in our estimate (embedded in the PCA output) of the narrow factor will be false positives, and we cannot capture the narrow factor as cleanly as we would wish. Because the longer history (T_L) reduces the correlation between factor and idiosyncratic return, it reduces the estimated loadings of the false positives, but does not reduce the number of false positives. This suggests that the factor return model is misspecified in the presence of two layers of narrow factors, most likely because it does not allow for interaction terms between narrow factors in the two layers.

[12] When $\hat{\Sigma}$ is estimated by the sample covariance matrix with $T > N$ and the true covariance matrix Σ is time-invariant, the estimated precision matrix $\hat{\Sigma}^{-1}$ is characterized by the *inverse-Wishart* distribution.

3.5 Observable measures of the SOR bias

The volatility ratio (VR) is defined as the actual volatility, $\sigma(\hat{\omega}, \Sigma)$, divided by the predicted volatility of the estimated GMVP, $\sigma(\hat{\omega}, \hat{\Sigma})$. Simply put, the VR of $\hat{\omega}$ captures the discrepancy between the in-sample $\sigma(\hat{\omega}, \hat{\Sigma})$ and the out-of-sample $\sigma(\hat{\omega}, \Sigma)$. The VR of $\hat{\omega}$ is an observable measure in a statistical sense, and its ideal value would be one when the estimation of the precision matrix is perfectly accurate. According to Proposition 2, the VR can be approximated by the square root of a ratio expression where the error term in the numerator is twice that in the denominator. This suggests that the VR is monotonically increasing in the magnitude of $\|\epsilon\|$, subject to the SOR bias, as noted in Remark 1.

Proposition 2 (A definitive SOR bias measure)**.** *Under the condition specified in (2) of Remark 1, the SOR bias measured by $\|\epsilon\|$ monotonically increases the volatility ratio (VR) of the estimated $\hat{\omega}$; i.e., VR can be approximated as*

$$\textit{(VR)} := \frac{\sigma(\hat{\omega}, \Sigma)}{\sigma(\hat{\omega}, \hat{\Sigma})} = \frac{\textit{(Actual volatility of } \hat{\omega})}{\textit{(Predicted volatility of } \hat{\omega})} \approx \sqrt{\frac{\mathbf{1}^\top \Sigma^{-1} \mathbf{1} + 2 \cdot \mathbf{1}^\top \epsilon \mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1} + \mathbf{1}^\top \epsilon \mathbf{1}}} \geq 1 \ ,$$

which is monotone increasing as $\|\epsilon\|$ gets larger.

Proof of Proposition 2. Refer to Section A in the Electronic Companion. □

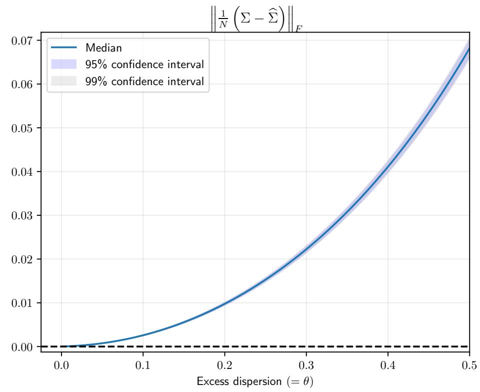
One might argue that comparing the realized risk of the estimated GMVP would be the primary criterion for assessing the degrees of SOR bias across different methodologies, rather than focusing on the 'risk-prediction error' that the VR attempts to capture. The intuition is that a more accurately estimated precision matrix would result in lower realized volatilities of the estimated GMVP. However, we have discovered that this intuition may not always hold true and can be misleading in many realistic cases. Specifically, recall that the actual volatility is the product of the predicted volatility and the volatility ratio by definition. As a result, the actual volatility of $\hat{\omega}$ $\big($i.e., $\sigma(\hat{\omega}, \Sigma)\big)$ may not have a monotone relationship with $\|\epsilon\|$ as

$$\sigma(\hat{\omega}, \Sigma) = \underbrace{\sigma(\hat{\omega}, \hat{\Sigma})}_{\downarrow \text{ as } \|\epsilon\| \uparrow} \cdot \underbrace{(\text{VR})}_{\uparrow \text{ as } \|\epsilon\| \uparrow} \ ,$$

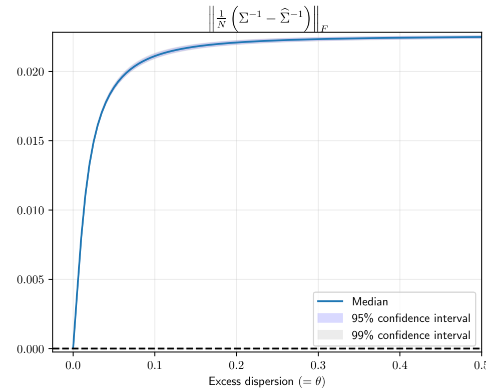
where $\sigma(\hat{\omega}, \hat{\Sigma})$ decreases and the VR increases with the estimation error under the condition in (2). As a result, the overall behavior may not be a monotonic function of the error, which suggests that VR is a more reliable proxy for SOR than $\sigma(\hat{\omega}, \Sigma)$ for measuring the estimation error in $\hat{\Sigma}^{-1}$ across different approaches.

We conduct a numerical experiment to quantify the relationship between different SOR bias metrics and the excess dispersion bias in the estimated factor exposures specific to the GMVP. The data generating process involves the contaminated mean-variance stock return model with $K = 5$ mutually uncorrelated factors, where the idiosyncratic returns of $N = 2000$ stocks are uncorrelated with each other as well as with the factor returns. For simplicity, we assume homogeneous factor volatilities of 10% and idiosyncratic volatilities of 5%. The true factor loadings are drawn from a normal distribution with mean one, with the true dispersion is represented by $\theta_0 = 1/200$. The estimated covariance matrix, $\hat{\Sigma}$, is subject to contamination by an excess dispersion parameter, $\theta \geq 0$, where the total dispersion is given by $\theta_0 + \theta$. Sample paths are generated from 1000 different simulation seeds.

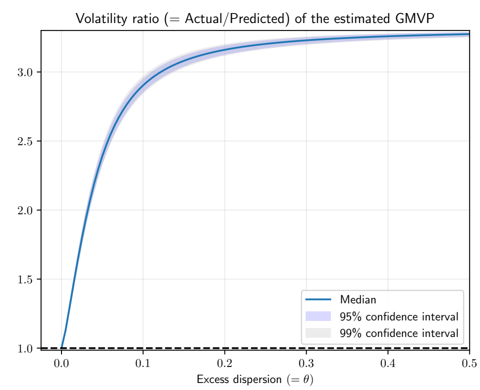
Figure 1 illustrates the results of our numerical experiment. As shown in panels (a) and (b), the ideal (or unobservable) SOR bias measure, computed using the average Frobenius norm of the estimation error in both the covariance and precision matrices, exhibits a monotonically increasing association with



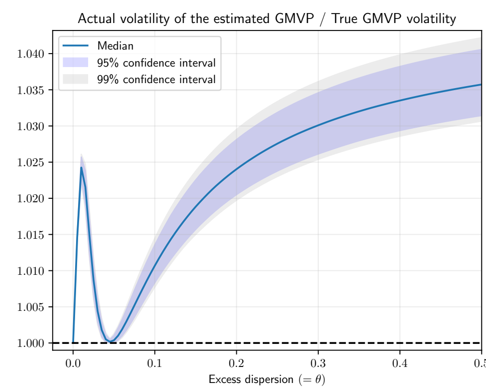
(a) Covariance matrix error



(b) Precision matrix error



(c) $\frac{(\text{Actual vol.})_{\text{GMVP}}}{(\text{Predicted vol.})_{\text{GMVP}}} = \frac{\sigma(\hat{\omega},\Sigma)}{\sigma(\hat{\omega},\hat{\Sigma})}$



(d) $\frac{(\text{Actual vol.})_{\text{GMVP}}}{(\text{True vol.})_{\text{GMVP}}} = \frac{\sigma(\hat{\omega},\Sigma)}{\sigma(\omega,\Sigma)}$

Figure 1: The metrics for SOR bias as functions of the excess dispersion imposed in the factor loadings

Note: This figure illustrates the relationship between various SOR bias metrics and the excess dispersion bias in estimated factor exposures for the GMVP, using a contaminated mean-variance stock return model with $K = 5$ mutually uncorrelated factors and $N = 2000$ stocks with uncorrelated idiosyncratic returns and factor returns. Assuming homogeneous factor volatilities of 10% and idiosyncratic volatilities of 5%, the true factor loadings are drawn from a normal distribution with a mean of one, with true dispersion represented by $\theta_0 = 1/200$. The estimated covariance matrix, $\hat{\Sigma}$, is contaminated by an excess dispersion parameter, $\theta \geq 0$, resulting in total dispersion $\theta_0 + \theta$. Sample paths are generated from 1000 different simulation seeds.

the excess dispersion parameter, θ. In particular, the estimation error in $\hat{\Sigma}^{-1}$ appears to be more sensitive to relatively small deviations from the true Σ^{-1}. Panels (c) and (d) confirm that VR serves as a more definitive measure of the accuracy of $\hat{\Sigma}^{-1}$ compared to the actual volatility of the estimated GMVP. The latter displays a non-monotonic pattern influenced by the excess dispersion of the estimated factor loadings, along with its noisier reflection of the SOR bias, as indicated by relatively wider confidence intervals. Related to this observation, Proposition 3 draws attention to a potential SOR measurement problem concerning the actual volatility of $\hat{\omega}$, revealing that the true precision matrix may not be the unique minimizer of the GMVP's actual volatility. This illustrates that a naive comparison of realized GMVP risks may result in an inaccurate assessment of the extent of the SOR bias across different methods employing the same dataset.

Proposition 3 (An illusive SOR bias measure)**.** *For $\hat{\Sigma} \neq \Sigma$, suppose that $\Sigma\hat{\Sigma}^{-1}$ has an eigenvector close to the vector of ones; i.e., $\Sigma\hat{\Sigma}^{-1}\mathbf{1} \approx \lambda\mathbf{1}$ for some $\lambda > 0$. Then, we observe*

$$\textit{(Actual variance of } \hat{\omega}\textit{)} \approx \textit{(True variance of } \omega\textit{)}$$

and the relationship between the SOR bias ($\|\epsilon\|$) and the actual (i.e., to-be-realized) volatility of the estimated GMVP may not be monotonic.

Proof of Proposition 3. Refer to Section A in the Electronic Companion. □

4 A Dynamic Factor Model with Weak Loadings

This section introduces our dynamic factor model (DFM) with weak loadings for stock returns and outlines the key assumptions. Furthermore, it explores the identification approaches and the theoretical justification of Long-History PCA (LH-PCA) in an asymptotic framework, where both N and T are large.

4.1 Econometric Framework and Assumptions

The DFM of stock returns considered in this paper generalizes the standard definition by incorporating time-varying security sensitivities to factors and heterogeneous strengths in factor loadings. This setup builds on the frameworks of Bates et al. (2013), who propose a decomposition of factor loadings into a static term and a purely innovative component, and Bai and Ng (2023), who extend the assumption of homogeneous factor loading strengths to allow heterogeneity across factors. It also aligns with the local factor setup of Freyaldenhoven (2022), making it a versatile framework for capturing the empirical dynamics of observables with heterogeneous factor strengths.

Definition 4. *A dynamic factor model*

The log returns of the securities are generated by the linear factor model specified as

$$\underbrace{X_t}_{(N\times 1)} = \underbrace{\Lambda_t}_{(N\times r)} \underbrace{F_t}_{(r\times 1)} + \underbrace{e_t}_{(N\times 1)} \qquad \textit{for } \; t = 1, \ldots, T \; , \tag{3}$$

where Λ_t, F_t and e_t represent the time-varying factor loadings, latent factor returns, and idiosyncratic returns at time t, respectively.

Definition 4 establishes the econometric framework for our DFM. We presume that the daily log returns of N stocks are generated by a fixed number ($r > 0$) of unobserved common factors ($F_t \in \mathbb{R}^{r\times 1}$)

with the possibly time-varying factor loadings ($\Lambda_t \in \mathbb{R}^{N\times r}$) and idiosyncratic errors $e_t \in \mathbb{R}^{N\times 1}$ on each date $t = 1,\ldots,T$. Assumptions 5 and 6, which are central to our framework, rely on the existence of a finite $M < \infty$ to ensure that the DFM remains well-posed and robust across varying empirical contexts.

Assumption 5. *Latent factor and idiosyncratic returns*

A. *(Latent factor returns):* $F_t = (F_{1t},\ldots,F_{rt})^\top \in \mathbb{R}^{r\times 1}$

(a) $\mathbf{E}(F_t) = \mathbf{0}$ *and* $\mathbf{E}\left(\|F_t\|_F^4\right) \le M$.

(b) *For* $k = 1,\ldots,r$, $\mathbf{E}\left(F_{kt}^2\right) = \mu_{kt}^2$, *where* $\mu_{kt} \in [0, M]$ *is the nonrandom (but possibly time-varying) factor volatility of* F_{kt}.

(c) *Define* $\mu_t = diag(\mu_{1t},\ldots,\mu_{rt}) \in \mathbb{R}^{r\times r}$ *and the diagonal elements of*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mu_t^2 := \Sigma_F$$

are strictly positive.

(d) *We have*

$$\frac{1}{T}\sum_{t=1}^{T} F_t F_t^\top \xrightarrow{p} \Sigma_F \quad as \quad T\to\infty \ .$$

B. *(Idiosyncratic returns):* $e_t = (e_{1t},\ldots,e_{Nt})^\top \in \mathbb{R}^{N\times 1}$

(a) *For* $n = 1,\ldots,N$, *we have* $\mathbf{E}(e_{nt}) = 0$, $\mathbf{E}\left(|e_{nt}|^8\right) \le M$ *and* $\mathbf{E}\left(e_{nt}^2\right) = \nu_{nt}^2$, *where* $\nu_{nt} \in [0, M]$ *is the nonrandom idiosyncratic volatility of* e_{nt}. *For notational consistency, define* $\nu_t = diag(\nu_{1t},\ldots,\nu_{Nt}) \in \mathbb{R}^{N\times N}$ *and*

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \nu_t^2 := \Sigma_\varepsilon \ .$$

(b) *Both* $\gamma_N(s,t) = \frac{1}{N}\mathbf{E}\left(e_s^\top e_t\right)$ *and* $\tau_{ij,st} = \mathbf{E}(e_{it}e_{js})$ *exist for all* $i,j \in \{1,\ldots,N\}$ *and* $s,t \in \{1,\ldots,T\}$.

(c) *We have* $\left|\gamma_N(s,s)\right| \le M$ *for all* $s \in \{1,\ldots,T\}$, *and*

$$\frac{1}{T}\sum_{s,t=1}^{T} \left|\gamma_N(s,t)\right| \le M \ .$$

(d) *For all* $t \in \{1,\ldots,T\}$, $|\tau_{tt,ij}| \le |\tau_{ij}|$ *holds for some* τ_{ij} *and we have*

$$\frac{1}{N}\sum_{i,j=1}^{N} |\tau_{ij}| \le M \quad and \quad \frac{1}{NT}\sum_{s,t=1}^{T}\sum_{i,j=1}^{N} |\tau_{ts,ij}| \le M \ .$$

(e) *For all* $s,t \in \{1,\ldots,T\}$, *we have*

$$\mathbf{E}\left(\left|\frac{1}{\sqrt{N}}\sum_{n=1}^{N} e_{sn}e_{nt} - \mathbf{E}(e_{sn}e_{nt})\right|^4\right) \le M \ .$$

(f) For all $i \in \{1,\ldots,N\}$ and $s,t \in \{1,\ldots,T\}$, e_{it} is independent of F_s.

The conditions outlined in Assumption 5 align closely with those in Bai and Ng (2002), accommodating weak dependence between factors and idiosyncratic returns, as well as limited cross-sectional and temporal dependence among the idiosyncratic errors. As such, the dynamic nature of the factor model arises from two sources: (i) variable volatilities of both factor and idiosyncratic returns, and (ii) the time-varying factor loadings, with their conditions detailed in Assumption 6.

Assumption 6. *Assumptions on the factor loading dynamics*

A. *(Factor loading dynamics)*

(a) *The dynamics of Λ_t for $t = 1,\ldots,T$ is expressed as*

$$\Lambda_t = \Lambda_0 + h_{NT}\xi_t \ , \tag{4}$$

where $\Lambda_0 \in \mathbb{R}^{N\times r}$ specifies the (static) baseline factor loadings, $h_{NT} \in \mathbb{R}$ is a deterministic scalar that may depend on the pair of (N,T) and $\xi_t = (\xi_{1t},\ldots,\xi_{Nt})^\top \in \mathbb{R}^{N\times r}$ is a (possibly degenerate) stochastic innovation process.

(b) *For all $i,j \in \{1,\ldots,N\}$ and $s,t \in \{1,\ldots,T\}$, e_{it} is independent of ξ_{js}.*

B. *(Static baseline component): Λ_0*

(a) *Λ_0 has rank r, the columns of Λ_0 are orthogonal, and $\|\Lambda_0\|_F \leq M$ holds.*

(b) *Allowing the strengths of baseline loadings to vary across factors, let*

$$1 \geq \alpha_1 \geq \cdots \geq \alpha_r > 0$$

so that the weakest baseline loading has strength $\alpha_r \in (0,1]$. Letting the $r\times r$ normalization matrix

$$B_N = diag\left(N^{\frac{\alpha_1}{2}},\ldots,N^{\frac{\alpha_r}{2}}\right) \ ,$$

there exists some diagonal matrix $D \in \mathbb{R}^{r\times r}$ such that

$$B_N^{-1}\Lambda_0^\top\Lambda_0 B_N^{-1} \xrightarrow{p} D \quad as \quad N \to \infty \ .$$

C. *(Stochastic innovation component): $h_{NT}\xi_t$*

(a) *There exist envelope functions $Q_1(N,T)$, $Q_2(N,T)$ and $Q_3(N,T)$ such that the following conditions hold for all N, T and factor indices $p,q,k,\ell = 1,\ldots,r$.*

$$\sup_{s,t\leq T}\sum_{i,j=1}^{N}\left|\mathbf{E}\Big((\xi_{is})_p(\xi_{jt})_q \ F_{ps}F_{qt}\Big)\right| \leq Q_1(N,T)$$

$$\sum_{s,t=1}^{T}\sum_{i,j=1}^{N}\left|\mathbf{E}\Big((\xi_{is})_p(\xi_{js})_q \ F_{ps}F_{qs}F_{kt}F_{\ell t}\Big)\right| \leq Q_2(N,T)$$

$$\sum_{s,t=1}^{T}\sum_{i,j=1}^{N}\left|\mathbf{E}\Big((\xi_{is})_p(\xi_{js})_q(\xi_{it})_k(\xi_{jt})_\ell \ F_{ps}F_{qs}F_{kt}F_{\ell t}\Big)\right| \leq Q_3(N,T)$$

(b) The following conditions hold:

$$h_{NT}^2\ Q_1(N,T) = \mathcal{O}(N)$$
$$h_{NT}^2\ Q_2(N,T) = \mathcal{O}(NT^2)$$
$$\min\{N,T\}\ h_{NT}^4\ Q_3(N,T) = \mathcal{O}(N^2T^2)\ .$$

Assumption 6 demonstrates that our approach significantly advances the conventional static framework by extending the setup of Bates et al. (2013) to account for time-varying security sensitivities to factors with heterogeneous strengths. The time-varying loadings (Λ_t) are decomposed into a static component (Λ_0) and a term driven by pure innovations ($h_{NT}\xi_t$). As illustrated by Bates et al. (2013), this structure facilitates the consistent estimation of factor models through PCA under mild regularity conditions, which we also adopt.[13] These assumptions include mean-reverting[14] stationary dynamics of factor exposures, occasionally interrupted by structural breaks, a feature commonly observed in empirical datasets. This enhancement builds on the contributions of Bai and Ng (2023), further expanding the local factor structure proposed by Freyaldenhoven (2022).[15]

The static component of the factor loadings, Λ_0, is the primary focus of our analysis, as it approximates the baseline factor sensitivities that cannot be diversified away. PCA-based estimation of such models typically assumes homogeneous factor loadings ($\alpha = 1$), a framework extended by Bai and Ng (2023) to allow heterogeneity using a parameter $\alpha \in (0, 1]$. We incorporate this extension in our normalization matrix B_N, enabling the model to accommodate variations in factor strength across different securities and factors. Accordingly, we adhere to the assumptions of Bai and Ng (2023) for addressing heterogeneity in the strengths of the static component. These assumptions ensure consistency in estimation while allowing for flexibility in modeling empirical data. For example, when Λ_t, μ_t and ν_t are independent of t, the model reduces to the standard definition of a static factor model with constant factor sensitivities, volatilities, and idiosyncratic terms over time. Overall, the proposed model combines the strengths of existing literature while providing greater flexibility to capture time-varying dynamics and heterogeneous strengths in factor loadings.

4.2 Identification approaches

Our DFM specification can also be reformulated as a static factor model under additional restrictive conditions, accommodating the stochastic variation in the loadings by attributing some of this variability to the factors. Specifically, assume there exists $\xi_0 \in \mathbb{R}^{N\times r}$ with full column rank r. Let $\xi_0^+ = \left(\xi_0^\top \xi_0\right)^{-1} \xi_0^\top \in \mathbb{R}^{r\times N}$ denote the Moore-Penrose pseudo-inverse of ξ_0. Under these conditions, our DFM model can be interpreted as a static factor model with time-invariant loadings in the form of

$$X_t = (\Lambda_0 + h_{NT}\xi_t)\, F_t + e_t = \Lambda_0 F_t + \tilde{\Lambda}_0 \tilde{F}_t + e_t\ ,$$

where $\tilde{\Lambda}_0 = h_{NT}\xi_0 \in \mathbb{R}^{N\times r}$ and $\tilde{F}_t = \xi_0^+(\xi_t F_t) \in \mathbb{R}^{r\times 1}$. Observe that this formulation highlights that our DFM with time-varying loadings is both more general and parsimonious, utilizing fewer factors and

[13] The envelope conditions outlined in Assumption 6-C. align with those in Corollary 1 of Bates et al. (2013), serving as sufficient criteria for the envelope functions specified in Assumption 4 of Bates et al. (2013).

[14] Empirical evidence suggests that the factor loadings of individual stocks exhibit mean-reverting behavior over time (Blume, 1971, Alexander and Chervany, 1980).

[15] More specifically, Freyaldenhoven (2022) states that a factor can be considered weak if it influences only a subset of the observables, where such factors are referred to as *local*.

avoiding reliance on the construction of ξ_0^+ that must be invertible.

Our approach also introduces a simplified representation of a linear factor model in the sense that the observed data process can be expressed as

$$X_t = \Lambda_0 F_t + \omega_t + e_t \ , \tag{5}$$

where $\omega_t = h_{NT}\xi_t F_t$ represents additional time-varying noise term. Such representation is comparable to Pelger and Xiong (2022b), where the time-varying factor loadings are assumed to be $\Lambda_t = \Lambda(S_t)$ as a deterministic function of the observable state process S_t. Then, the state-dependent factor structure is given by

$$X_t = \left(\Lambda(S_t) + \mathcal{E}_t\right) F_t + e_t \ ,$$

where the noise term $\mathcal{E}_t$ captures variations in loading coefficients that cannot be fully explained by the state process (S_t), potentially arising from measurement errors or omitted variables. Although Pelger and Xiong (2022b) assumes the existence of an observable state process that accounts for the factor loading dynamics in real time, if this is not feasible, the state-dependent dynamic factor structure can be reformulated as given in Equation (5) by imposing[16]

$$\Lambda(S_0) = \Lambda_0 \quad \text{and} \quad \Lambda(S_t) - \Lambda(S_0) + \mathcal{E}_t = h_{NT}\xi_t \ .$$

In short, the representation in Equation (5) simplifies our DFM by assuming time-invariant loadings, while treating the composite noise as a combination of idiosyncratic and measurement errors that can be diversified away under mild assumptions, providing a robust yet interpretable formulation for dynamic factor analysis.

4.3 Theoretical Basis of Long-History PCA

To mitigate the realized correlation between idiosyncratic returns and factor returns, we consider a longer data history of length $T_L (> T_M)$, for example, daily returns over six years. As a result, we cannot presume that the factor loadings remain constant over the entire (extended) sample period T_L. Nevertheless, Theorem 7 and Corollary 8 confirm that, under the 'large-N and large-T' framework, PCA can consistently estimate the covariance and precision matrices in a dynamic factor model under mild regularity conditions outlined in Assumptions 5 and 6.

Theorem 7 (Convergence of a population covariance and precision matrices)**.** *Consider the population covariance matrix within the observation time window T given by*

$$\Sigma_N(T) = \frac{1}{T}\mathbf{X}^\top\mathbf{X} \in \mathbb{R}^{N\times N},$$

where $\mathbf{X} \in \mathbb{R}^{T\times N}$ is the observed daily return data of N securities. In addition, we introduce $\overline{\Sigma}_N(T)$ as a static version of $\Sigma_N(T)$ given by

$$\overline{\Sigma}_N(T) = \Lambda_0 \ \left(\frac{1}{T}\sum_{t=1}^{T}\mu_t^2\right)\Lambda_0^\top + \left(\frac{1}{T}\sum_{t=1}^{T}\nu_t^2\right) \ ,$$

[16] The positive recurrence of S_t and mild conditions of $\mathcal{E}_t$ ensure that the term $\Lambda(S_t) - \Lambda(S_0) + \mathcal{E}_t$ does not drift indefinitely, provided that the deterministic function $\Lambda(\cdot)$ is well-behaved.

where both $\Sigma_N(T)$ and $\overline{\Sigma}_N(T)$ are assumed to be invertible. By letting $N, T \to \infty$, we have

$$\left\| \frac{1}{N} \left(\Sigma_N(T) - \overline{\Sigma}_N(T) \right) \right\|_F^2 \asymp \left\| \frac{1}{N} \left(\Sigma_N^{-1}(T) - \overline{\Sigma}_N^{-1}(T) \right) \right\|_F^2 = O_p\left(\frac{1}{\min\{N, T\}} \right)$$

under the conditions stated in Assumptions 5 and 6, where $a_N(T) \asymp b_N(T)$ indicates $a_N(T) = O(b_N(T))$ and $b_N(T) = O(a_N(T))$ as $N, T \to \infty$.[17]

Proof of Theorem 7. Refer to Section A in the Electronic Companion. □

Theorem 7 asymptotically justifies that the variation introduced by $h_{NT}\xi_t$ remains sufficiently small, ensuring that it does not dominate the factor structure in the limiting representations of the covariance and precision matrices. Note that the speed of convergence is governed by the minimum of N and T, emphasizing the importance of extending T for large N, while ensuring that T is not excessively large relative to N.

Corollary 8 (Consistent estimation of LH-PCA)**.** *Let $(\tilde{F}_r, \tilde{\Lambda}_r)$ denote the principal component estimator of $\mathbf{X} \in \mathbb{R}^{N \times T}$, where r represents the true number of factors. Define $\tilde{e}_r := \mathbf{X} - \tilde{F}_r \tilde{\Lambda}_r^\top \in \mathbb{R}^{N \times T}$ and*

$$\widetilde{\Sigma}_N(T) := \tilde{\Lambda}_r \left(\frac{1}{T} \tilde{F}_r^\top \tilde{F}_r \right) \tilde{\Lambda}_r^\top + \frac{1}{T} \tilde{e}_r \tilde{e}_r^\top \ .$$

Then, as $N, T \to \infty$, we have

$$\left\| \frac{1}{N} \left(\Sigma_N(T) - \widetilde{\Sigma}_N(T) \right) \right\|_F^2 \asymp \left\| \frac{1}{N} \left(\Sigma_N^{-1}(T) - \widetilde{\Sigma}_N^{-1}(T) \right) \right\|_F^2 = O_p\left(\frac{1}{\min\{N, T\}} \right)$$

under the conditions stated in Assumptions 5 and 6.

Proof of Corollary 8. Refer to Section A in the Electronic Companion. □

In particular, Corollary 8 justifies the use of LH-PCA in the setting of variable factor and idiosyncratic volatility, along with time-varying factor loadings with heterogeneous strengths under our assumptions. This implies that, accounting for time variation in both factor and idiosyncratic volatilities, changes in the values of the factor-loading matrix under our assumption do not pose an additional obstacle to the consistent estimation of the factor-driven covariance matrix. Furthermore, the estimation remains consistent under the dynamic factor model specification, even with weaker loadings in their static components.

Empirically, factor volatility is highly variable, even over short horizons. Factor loadings are also variable over long, and possibly medium, horizons. The presence of weak factors complicates the process of portfolio volatility prediction. Note that even if factor loadings were constant, and there were no weak factors, variation in factor volatility would alter the return correlation matrix over short horizons. PCA estimation over medium horizons makes sense only if factor loadings can be consistently estimated in the presence of rapidly varying factor volatility. PCA estimation over long horizons requires, in addition, that factor loadings can be consistently estimated in the presence of variable factor loadings and weak factors. For this reason, commercial latent factor models have avoided using long histories of daily return data. Our theoretical results justify estimation over longer histories, even when factor volatilities and

[17] The average Frobenius norm serves as a measure of the average error within the entries of the approximated population covariance and precision matrices.

factor loadings are varying, and weak factors are present. Our empirical and simulation results show that the use of longer histories improves the prediction of portfolio volatility.[18]

4.4 Tuning Parameters

A key challenge in applying PCA is determining the appropriate number of eigenfactors to extract. In our main estimation, we adopt the Bai and Ng (2002) estimator with $PC_{p1}(k)$ as the penalty function, consistent with prior studies (see Electronic Companion D.3 for details). This data-driven approach dynamically selects the number of eigenfactors (K) within the estimation sample, allowing the model to adapt to the evolving nature of the stock market. By enhancing flexibility and reducing the risks of under- or overfitting, this method offers a more accurate representation of the underlying factor structure than approaches that fix the number of factors arbitrarily throughout the analysis.

There are several factor selection criteria that provide consistent estimates even in the presence of weak factors, as supported by recent empirical studies (e.g., Uematsu and Yamagata, 2023; Bai and Ng, 2023). Freyaldenhoven (2022) explores the selection of K in the presence of 'local' factors, which aligns with our concept of 'narrow' factors and the broader framework of 'weak' factors proposed by Bai and Ng (2023). As noted by Han and Caner (2017) and Freyaldenhoven (2022), the BN estimator (Bai and Ng, 2002) remains consistent in identifying both strong and weak factors when $\alpha > 0$, where $\alpha \in (0, 1]$ represents the strength of the factors and the associated eigenvalues scale as $\mathcal{O}_p(N^\alpha)$, with N denoting the population size.

Although Freyaldenhoven (2022) proposes filtering out factors with $\alpha \leq 1/2$ as *irrelevant* in a different context, this criterion may be too restrictive for our purposes, where even weak factors can provide valuable insights for modeling and estimating portfolio volatility. Transient shocks (such as flash crashes, liquidity disruptions, hedge fund deleveraging, and order flow imbalances) can significantly impact realized volatility. Even if weak factors do not meet conventional relevance criteria, they can still enhance volatility estimates by capturing short-term risk fluctuations. Allowing for a larger number of estimated factors thus helps identify volatility drivers that traditional methods might otherwise dismiss as noise. As a result, while the original BN estimator does not explicitly account for weak factors in its theoretical derivation, its estimation method remains well-suited to our framework for detecting latent factors with $\alpha > 0$.

To validate that the outperformance of LH-PCA stems from the use of a long history rather than the number of factors chosen, we conduct several robustness checks. Specifically, we test varying factor selections using both information criteria, adjusting the BN estimator by adding or subtracting certain number of factors, and considering fixed specifications (e.g., $K = 5, 10, \ldots, 100$). While a fixed choice is generally suboptimal compared to data-driven selection, it serves as a useful benchmark to isolate the impact of model selection on performance. Overall, our findings confirm that the outperformance of LH-PCA holds true across settings. The results are available in the Electronic Companion D.3.

5 Simulation Study

In this section, we employ Monte Carlo simulations to assess the performance of the SOR bias. Having access to the true risk factor structure in the simulations, we can directly compute the difference between

[18] In addition, Proposition 6 of Bai and Ng (2023) proves that the average error in estimating Λ_0 asymptotically vanishes provided that $\alpha_r > 0$ and $\frac{N^{1-\alpha_r}}{T} \to 0$, where α_r is the weakest factor strength out of the r factor loadings in Λ_0.

Broad Factors			Country Factors		
	Volatility	No. of Stocks		Volatility	No. of Stocks
Market	16.00%	2000	Country 1	13.00%	154
Style 1	8.00%	2000	Country 2	12.38%	69
Style 2	4.00%	2000	Country 3	18.53%	159
Style 3	4.00%	2000	Country 4	15.91%	76
Industry Factors			Country 5	24.93%	178
Industry 1	11.51%	143	Country 6	18.14%	99
Industry 2	12.74%	291	Country 7	11.86%	177
Industry 3	13.00%	131	Country 8	17.57%	19
Industry 4	15.08%	144	Country 9	16.43%	203
Industry 5	13.99%	115	Country 10	24.41%	73
Industry 6	17.80%	251	Country 11	14.31%	56
Industry 7	14.07%	258	Country 12	12.03%	247
Industry 8	12.22%	53	Country 13	13.33%	171
Industry 9	9.80%	109	Country 14	13.79%	75
Industry 10	13.11%	171	Country 15	18.29%	135
Industry 11	21.23%	334	Country 16	22.59%	109
Industry Sum		2000	Country Sum		2000

Table 1: True Factor Structure in Simulation

Note: This table reports the true structure of broad, country and industry factors in our simulation study. Factor volatilities are annualized, in percent. The true factors are assumed to be normally distributed with a mean of zero within the same regime, and are considered pairwise uncorrelated.

the true and estimated precision matrices. This not only supports the validation of our observable SOR bias metrics but also provides a solid basis for our subsequent empirical analysis, as detailed in the following section.

5.1 Simulation Setup

In our simulation, we assume the presence of four *broad* factors, consisting of one market factor and three style factors, along with 27 narrow factors, which include 16 country factors plus 11 industry factors, to mimic the narrow factor structure of the European stock market. We set the annualized daily volatility of the market factor at 16%, and the daily volatility of the three style factors at 8%, 8%, and 4%, respectively (Clarke et al., 2011; Morozov et al., 2012). The country and industry factor volatilities are calibrated from the European stock market data. We draw idiosyncratic volatilities uniformly from $[32\%, 64\%]$ per annum. We randomly assign $N = 2000$ stocks to countries and industries; see Table 1 for details. The true factors are normally distributed with mean zero, and they are assumed to be pairwise uncorrelated. The sample stock returns are generated across 1000 different seeds.

Variable volatility structure We assume the return-generating process follows a Markov regime-switching model (MRS). There is a latent Markov chain of the state process $s_t \in \{0, 1\}$, in which $s_t = 0$ denotes a normal factor-volatility structure, and $s_t = 1$ denotes a crisis structure. This Markov chain is described by its transition matrix $\mathbf{\Pi}$ with the probabilities $p_{i|j} = \mathbb{P}(s_t = i | s_{t-1} = j)$ of switching from regime j at time $t-1$ to regime i at time t. In particular, we have

$$\mathbf{\Pi} = \begin{bmatrix} p_{0|0} & p_{1|0} \\ p_{0|1} & p_{1|1} \end{bmatrix} = \begin{bmatrix} 0.998 & 0.002 \\ 0.008 & 0.992 \end{bmatrix} .$$

	Expected Performance Statistics (Bootstrapped)		
	Ideal SOR Bias	Volatility Ratio	Actual Volatility
	$\mathbb{E}\left\lVert \frac{1}{N}(\hat{\Sigma}^{-1}-\Sigma^{-1}) \right\rVert_F$	$\mathbb{E}\left(\frac{\sigma(\hat{\omega},\Sigma)}{\sigma(\hat{\omega},\hat{\Sigma})}\right)$	$\mathbb{E}\Big(\sigma(\hat{\omega},\Sigma)\Big)$ (%, p.a.)
LW	[51.5739, 56.4899]	[22.2182, 24.6615]	[9.5471, 9.8187]
PCA	[5.0011, 5.2837]	[3.1111, 3.1945]	[7.2323, 7.4567]
GPS	[4.9987, 5.2811]	[2.9618, 3.0429]	[7.2391, 7.4566]
LH-PCA	[4.7119, 5.0142]	[2.5992, 2.6837]	[7.3795, 7.6154]

Table 2: 99% confidence intervals for the expected performance metrics (w/o RCA)

Note: This table shows the 99% confidence intervals of the estimated GMVP's expected performance metrics obtained from 1000 different seeds under the simulation setup with the MRS volatility structure and the mean-reverting factor loading dynamics. The performance statistics are calculated from a sample of $N = 2000$ stocks. LW denotes the shrinkage estimation of the sample covariance matrix proposed by Ledoit and Wolf (2004b) with T_M. PCA stands for the plain PCA method with T_M, GPS indicates the methodology proposed by Goldberg et al. (2022) with T_M. We employ T_L for LH-PCA. These results are obtained from 1000 simulation runs with varying seeds, acquired through bootstrapping with replacements for cross-validation.

Furthermore, we assume that the initial distribution is given by $\mathbb{P}(s_t = 0) = 0.8$. During the normal period ($s_t = 0$), the factor volatilities are drawn from the same distribution as the baseline setup. For the crisis regime ($s_t = 1$), we assume that the volatilities of the (market, style, country, industry, idiosyncratic) factors are multiplied by (2.0, 1.5, 1.5, 1.5, 1.25), respectively. The true factors are normally distributed with mean zero within the same regime, and they are assumed to be pairwise uncorrelated.

Time-varying factor loadings We further extend the simulation setup by allowing variable factor loadings over time, following Adrian and Franzoni (2009) and the references therein. We first assume that the *broad* factor exposures (i.e., the market beta and the style factor loadings) are time-varying based on a discretized version of the mean-reverting Ornstein–Uhlenbeck process given by the following SDE:

$$d\Lambda_t^i = \kappa\left(\Lambda_0^i - \Lambda_t^i\right)dt + D_\sigma dW_t^i \ ,$$

where $\Lambda_t^i = \left[\lambda_t^{i0},\ \lambda_t^{i1},\ \lambda_t^{i2},\ \lambda_t^{i3}\right]^\top$ represents the *broad* factor loadings specific to the i-th stock, $\kappa = 1.0$ governs the (common) mean-reversion speed, $\Lambda_0^i = \left[\lambda_0^{i0},\ \lambda_0^{i1},\ \lambda_0^{i2},\ \lambda_0^{i3}\right]^\top$ specifies the long-run mean level of the factor loadings, D_σ is a 4×4 diagonal matrix with its diagonal entries equal to $[10\%, 5\%, 5\%, 5\%]$ per annum, and W_t^i is an independent 4-dimensional standard Brownian motion. We set the initial broad factor loadings as the long-run mean level of the factor loadings. For simplicity, we assume that the narrow factor loadings are binary constants.

5.2 Main Findings

Table 2 summarizes the simulation results within the setup characterized by the MRS volatility structure and the mean-reverting factor loading dynamics. We compare the performance of four distinct approaches; i.e., LW denotes the shrinkage estimation of the sample covariance matrix, proposed by Ledoit and Wolf (2004b) using one year of historical data.[19] PCA represents the naive PCA estimation based on one year ($T_M = 250$) of data history. GPS refers to the PCA method adjusted by the GPS correction in accordance with the guidelines of Goldberg et al. (2022) to correct the leading eigenvector. LH-PCA

[19]See Electronic Companion B for details.

represents the PCA estimation utilizing a six-year history ($T_L = 1500$) of sample data. The simulation results are obtained from 1000 simulation runs with varying seeds, acquired through bootstrapping with replacements for cross-validation.

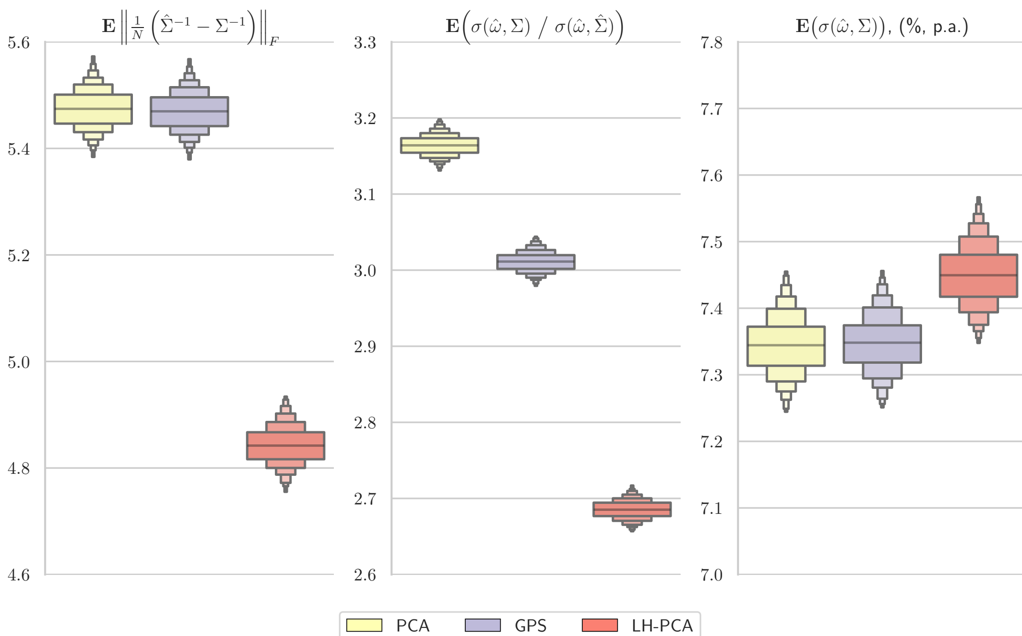


Figure 2: Expected performance statistics of the factor-based approaches (w/ RCA)

Note: This figure shows the distributions of the estimated GMVP's expected metrics obtained from 1000 different seeds under the extended setup with the MRS volatility structure and the mean-reverting factor loading dynamics. The performance statistics are calculated from a sample of $N = 2000$ stocks. PCA stands for the plain PCA method with T_M, GPS indicates the methodology proposed by Goldberg et al. (2022) with T_M. We employ T_L for LH-PCA. The estimated covariance matrices are adjusted by RCA with $T_S = 40$ days. These results are obtained from 1000 simulation runs with varying seeds, acquired through bootstrapping with replacements for cross-validation.

As indicated by the ideal SOR bias defined in Equation (1), which is quantified by the average Frobenius norm of the estimation error in the precision matrix and is only available in simulation, the factor-based approach demonstrates approximately 10 times superior performance to the use of the shrunk sample covariance matrix. This observation is corroborated by both metrics for expected volatility ratio and actual volatility, though the performance differentials in terms of actual volatility become much less pronounced. Moreover, the actual volatility measure exhibits a contrasting direction compared to the observations from the other two metrics within the factor-based approaches. This is consistent with our observations in Proposition 3 and Figure 1 in that the actual volatility may lose its monotone relationship with the estimation error when the estimated $\hat{\Sigma}^{-1}$ is reasonably close to the true Σ^{-1}.

In a dynamic factor model, the covariance matrix estimated by PCA (and its variant) is distinct from the covariance matrix encountered on the following day of estimation. Such discrepancies have the potential to result in empirically inaccurate out-of-sample predictions of the true factor structure on a daily basis. Understanding these discrepancies underscores the potential necessity for responsive covariance adjustments (RCA) in empirical analysis; see Electronic Companion C for more details of its implementation under the factor model approaches. Nevertheless, similar patterns can be observed in the enhanced box plots in Figure 2, where the estimated covariance matrices are adjusted by RCA with $T_S = 40$ days for PCA, GPS and LH-PCA. Specifically, the enhanced box plots consistently illustrate the distributions of the expected metrics within each panel, specific to the estimated GMVPs across different estimation schemes.

The fact that expected VR statistics are not adversely affected by the introduction of variable factor volatility and mean-reverting stock factor exposures is a direct consequence of Corollary 8, which asserts that (LH-)PCA consistently estimates the baseline factor exposures in the presence of (potentially) unstable factor structure dynamics under mild regularity conditions. In addition, the outcomes demonstrate that the use of a longer data history leads to a significant improvement in the estimation of factor loadings. This is evidenced by the fact that LH-PCA outperforms one-year PCA and GPS in the presence of narrow factors under both the ideal SOR bias and the volatility ratio metrics.[20] Overall, our numerical results confirm that the use of long histories such as $T_L = 1500$ trading days, or six years, substantially improves the performance of the unconstrained GMVPs, compared to using a medium history of $T_M = 250$ days, indicating superior SOR bias mitigation.

6 Empirical Analysis

6.1 Data and Samples

In our empirical study, we investigate two financial markets: United States and Europe. Daily stock return data in the United States are obtained from the Center for Research in Security Prices (CRSP), and European stock return data from 16 countries[21] are obtained from Compustat-Capital IQ's Global Daily (henceforth EURO). Our database spans from 2001 to 2021. We selected this period for several reasons: (i) it provides a substantial timeframe for estimating the plane spanned by factor returns over a long history (T_L); (ii) it encompasses major market shocks, including the global financial crisis and COVID-19; and (iii) most European countries transitioned from their national currencies to the Euro in the late 1990s and early 2000s. One particular feature of our data is that we consider the entire universe of stocks at a given point in time. This is crucial because small stocks, which tend to be highly volatile, can make the estimation problem more challenging than focusing on more established samples such as the S&P 500 constituents.

As in the previous section, we fix $T_L = 250 \times 6 = 1500$ and $T_M = 250$ trading days. Our decision to use $T_L = 1500$ trading days (approximately six trading years) is based on observed stock market patterns in the U.S. and Europe. For instance, the U.S. stock market experienced a bull market from late 2002 to 2007, followed by the Global Financial Crisis in 2008–2009 and a recovery beginning in 2009. Similarly, from 2013 to 2019, markets underwent a prolonged period of sustained growth, which was disrupted by the COVID-19-induced crash in 2020 and followed by a rapid recovery driven by significant stimulus measures and post-pandemic growth. In Europe, the Eurozone debt crisis (2010–2012) was preceded by a recovery from the financial crisis and followed by a stabilization phase from 2013 to 2018. Taken together, our choice of six trading years as the long-history window is grounded in the stock market cycles observed during our sample period, 2001–2021. In Section D.3 of the Electronic Companion, we cross-check our

[20] It is noteworthy that the performance of one-year GPS is fairly decent when only broad factors are present. However, the performance of the method deteriorates with the inclusion of a single layer of narrow factors, and deteriorates further with the addition of a second layer of narrow factors. Details are available upon request.

[21] These are Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Italy, Netherland, Norway, Portugal, and Sweden. These 16 countries are commonly studied in analyses of European stock markets (see Fama and French, 1998 and Bekaert et al., 2009b).

	United States			Europe		
Data Sources:	CRSP			Compustat-Capital IQ Global		
Data Dimensions:	$T = 5,284,\ N = 6,516$			$T = 5,359,\ N = 7,350$		
	(1)	(2)	(3)	(1)	(2)	(3)
	1 - Consumer Nondurables	346	19.32%	1 - Consumer Nondurables	679	12.08%
	2 - Consumer Durables	155	26.65%	2 - Consumer Durables	216	13.12%
	3 - Manufacturing	614	26.30%	3 - Manufacturing	1082	12.21%
	4 - Energy	302	39.02%	4 - Energy	344	13.00%
	5 - Chemicals	134	23.67%	5 - Chemicals	218	13.94%
Industry	6 - Business Equipment	1400	21.82%	6 - Business Equipment	1764	12.37%
	7 - Telecom	278	23.75%	7 - Telecom	250	12.92%
	8 - Utilities	157	18.94%	8 - Utilities	151	13.94%
	9 - Shops	719	23.28%	9 - Shops	767	11.83%
	10 - Health	812	20.80%	10 - Health	746	12.30%
	11 - Money (Finance)	1599	23.91%	11 - Money (Finance)	1347	11.98%
				1 - Austria (AUS)	98	14.46%
				2 - Belgium (BEL)	151	13.01%
				3 - Switzerland (CHE)	337	12.16%
				4 - Germany (DEU)	997	11.14%
				5 - Denmark (DNK)	231	13.76%
				6 - Spain (ESP)	201	15.25%
				7 - Finland (FIN)	203	11.92%
Country	United States	6,516		8 - France (FRA)	909	12.70%
				9 - United Kingdom (GBR)	2148	11.70%
				10 - Greece (GRC)	239	14.75%
				11 - Ireland (IRL)	78	15.87%
				12 - Italy (ITA)	487	16.47%
				13 - Netherland (NLD)	226	16.04%
				14 - Norway (NOR)	356	15.33%
				15 - Portugal (PRT)	48	17.88%
				16 - Sweden (SWE)	855	15.35%

Table 3: Data Description

Note: In each dataset, (1) denotes the narrow factor classification (industry factor classification or country factor classification), (2) indicates the number of stocks within each classification, and (3) reports the annualized realized volatility of each narrow factor, expressed in percentage terms. The industry classifications are based on the definitions provided on Kenneth French's website.

empirical results using alternative lengths for the T_L moving window to ensure that our main findings are not sensitive to this parameter choice.

Table 3 provides a comprehensive overview of our sample, presenting the number of stocks and the yearly realized volatility across countries and industries. On average, a moving window in the CRSP dataset contains approximately 2,260 stocks, with the total number of firms in the sample being $N = 6,516$. At the start of our moving window analysis, around 2,800 securities are available, but this number gradually declines to approximately 1,800 by the end of the sample period, consistent with the *listing gap* argument in Doidge et al. (2017). In the EURO dataset, a total of $N = 7,350$ firms are included in our sample, with an average of about 2,500 stocks available in each moving window. Unlike in CRSP, the number of available securities in EURO has increased over the sample period, likely due to the broader coverage of the database and the expanding stock market.

6.2 Empirical Assessment of SOR Bias

It is essential to employ a statistical proxy for VR in empirical analysis. As proposed by Shepard (2009), the building block of our performance statistics is *Z-score* of $\hat{\omega}$, defined as

$$z_t = \frac{(\text{Realized GMVP return})_{t+1}}{(\text{Predicted GMVP volatility})_t} = \frac{R_{t+1}\ \hat{\omega}_t}{\sigma(\hat{\omega}_t, \hat{\Sigma}_t)} \ , \tag{6}$$

where the variance of the Z-scores should be equal to one, provided that the predicted volatilities are perfectly precise. We consider *Bias Statistic* (BS) introduced in Menchero et al. (2013) as our main performance measure to evaluate the SOR bias for the GMVP, expressed as

$$\mathrm{BS}_t(\tau) = \sqrt{\frac{1}{\tau} \sum_{k=t-\tau+1}^{t} z_k^2} \ , \tag{7}$$

where τ denotes the number of trading days in the testing (out-of-sample) period. The BS can be interpreted as the sample standard deviation of the realized Z-scores, irrespective of their distribution. Ideally, in the absence of SOR bias, the BS value should be equal to one. The BS captures whether the risk forecasts were accurate, on average, for the testing portfolio over time. To assess the statistical reliability of BS, confidence intervals can be approximated by bootstrapping the realized Z-scores with replacement for cross-validation purposes.

A potential pitfall of the BS measure lies in the time-resolution problem, which arises because the statistic may appear overly optimistic, nearing one, due to cyclical prediction errors. These errors may alternatively over- or underestimate the predicted risk, but offset each other over the long term. The *Mean Rolling Absolute Deviation* (MRAD) is an alternative way to mitigate this problem, as it computes the bias statistic over the u-day horizon sub-periods and then calculates the average of those rolling windows until we exhaust the sample. The MRAD metric is defined as

$$\mathrm{MRAD}_t(u;\tau) = \frac{1}{\tau - u + 1} \sum_{k=t-\tau+u}^{t} \Big| \mathrm{BS}_k(u) - 1 \Big| \ , \tag{8}$$

which is the mean absolute deviation of $\mathrm{BS}_k(u)$ from the overall BS, defined in Equation (7). We set $u = 12$ following the approach of Menchero et al. (2013).[22]

We further employ Q-statistics, which is a QLIKE loss function proposed in Patton (2011). As mentioned in Menchero et al. (2013), the Q-statistics address both time-resolution and portfolio-resolution issues by penalizing both underprediction and overprediction of risks, defined as

$$\text{Q-statistic} = z_t^2 - \log(z_t^2) \ , \tag{9}$$

which is minimized in expectation when the estimation is perfectly accurate.[23]

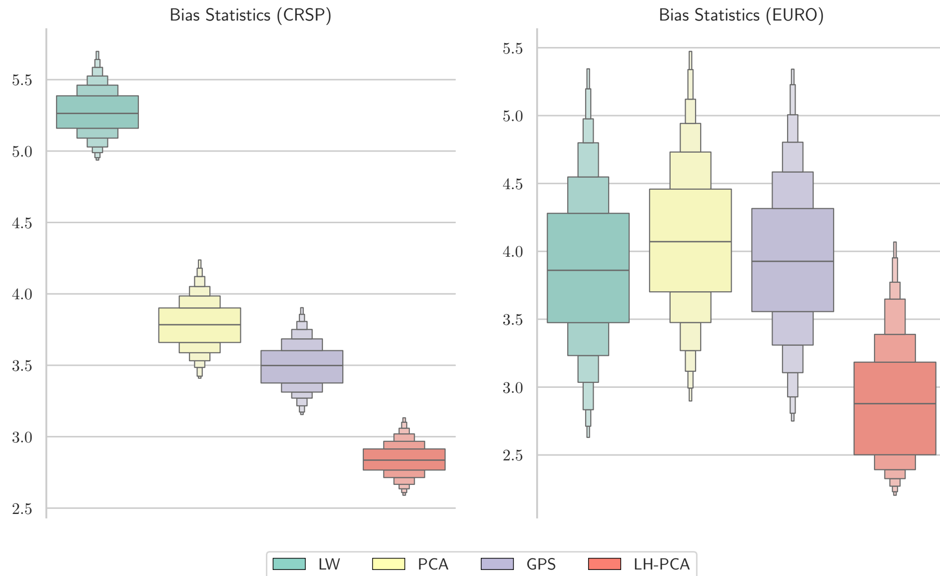


Figure 3: Bias Statistic (BS) Results

Note: This figure shows the distributions of the estimated GMVP's Bias Statistics obtained from daily CRSP (left panel) and EURO (right panel) stock returns data. LW stands for Ledoit-Wolf shrinkage method, PCA stands for the plain PCA method with T_M, GPS indicates the methodology proposed by Goldberg et al. (2022) with T_M. We employ T_L for LH-PCA. These results are obtained from 1000 simulation runs with varying seeds, acquired through bootstrapping with replacements for cross-validation. For the European data, we cap Z-scores at 100 and -100 to mitigate the impact of a small number of extreme days that cause factor models based on medium-length windows (such as LW, PCA, and GPS) to become overly sensitive to outliers. This conservative treatment favors competing methods over LH-PCA.

6.3 Main Findings

Similar to the simulation section, we assess the predictive performance of the four estimation methods (refer to Section 5.2). As introduced in Section 6.2, our main empirical measure is the Bias Statistic, defined in Equation (7). Across all factor-based methods, we use Bai and Ng (2002) estimator to determine the number of PCA eigenfactors to extract.[24] Figure 3 presents the enhanced box plots[25] using 1,000 times bootstrapped Z-score samples with replacement. In the CRSP sample, the bias statistic for the Ledoit-Wolf method are 5.13, which is significantly higher than the standard PCA estimation of 3.76. We see that eigenvector correction of GPS (purple) shows a modest improvement to 3.49. Our proposed method, LH-PCA (red), shows the best performance at the bias statistic of 2.84.

Our analysis now turns to the empirical results obtained from the EURO dataset.[26] The right panel

[22] Under the assumption of normally distributed daily returns, the ideal value of MRAD is approximately 0.17. In the presence of leptokurtic return distributions, this value tends to increase, implying that our evaluation based on the normal benchmark adopts a conservative perspective.

[23] Patton (2011) shows that the Q-statistic is the unique loss function that solely relies on the Z-scores. The ideal value of the expected Q-statistic is approximately 2.27 assuming the normal distribution of daily returns.

[24] The number of factors extracted varies depending on the moving window; see Electronic Companion D.3 for details.

[25] We use the Python library Seaborn to produce enhanced box plots. This enhanced version plots more quantiles than the standard, providing more information about the shape of the distribution, particularly in the tails.

[26] When analyzing the European results, we found that using a short moving window for risk forecasts produced

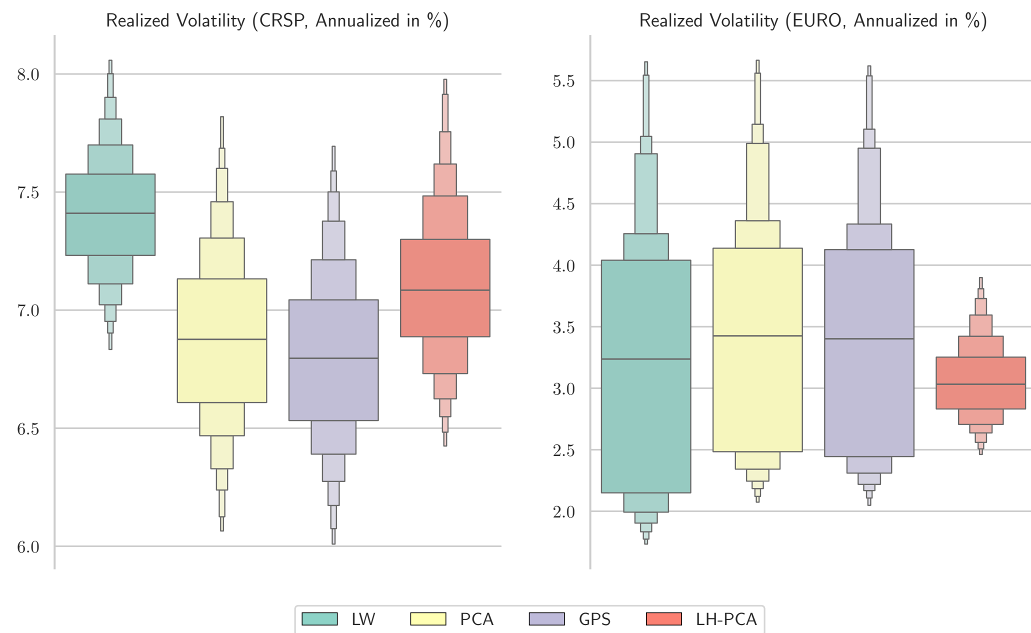


Figure 4: Realized Volatility Results

Note: This figure shows the distributions of the estimated GMVP's Realized Volatility obtained from daily CRSP (left panel) and EURO (right panel) stock returns data. LW stands for Ledoit-Wolf shrinkage method, PCA stands for the plain PCA method with T_M, GPS indicates the methodology proposed by Goldberg et al. (2022) with T_M. We employ T_L for LH-PCA. These results are obtained from 1000 simulation runs with varying seeds, acquired through bootstrapping with replacements for cross-validation. For the European data, we cap Z-scores at 100 and -100 to mitigate the impact of a small number of extreme days that cause factor models based on medium-length windows (such as LW, PCA, and GPS) to become overly sensitive to outliers. This conservative treatment favors competing methods over LH-PCA.

in the same figure paints the same pictures as in CRSP - contrary to the practitioners' unwillingness to use PCA on data histories longer than one or two years, using a long-history approach (LH-PCA) benefits portfolio volatility prediction. The improvements in the EURO case (right panel of Figure 3) appear less pronounced than in the CRSP case (left panel of Figure 3), but this is largely due to our treatment of outliers, as illustrated in Sections D.1 and D.2 of the Electronic Companion. Unlike the CRSP data, the European stock market dataset is less comprehensive and requires additional filtering, which likely excludes small and volatile stocks. Even after cleaning, we observed that certain methods—namely Ledoit-Wolf (LW), standard PCA, and GPS—occasionally produce extreme Z-scores. To ensure meaningful comparisons in the boxplot, we exclude these outliers when constructing Figure 3. In doing so, we are bending the rules in favor of these methods; even so, LH-PCA still outperforms them. If we had included those extreme values, the differences would have been stark: the mean Bias Statistic for LW, PCA, and GPS exceeds 15. In short, the small difference in the EURO boxplot is an artifact of conservative trimming—without it, LH-PCA's advantage would appear even more dramatic.

inaccurate estimates for a few days, leading to an unexpectedly high Z-score (for example, an estimated Z-score of 400 under the standard PCA factor model). Such bias greatly undermines our ability to effectively compare the performance of risk forecasts. To get around this problem, we cap Z-scores at 100 and -100 and bend the rule to favor short history algorithms. Nevertheless, we consistently observe that LH-PCA works better.

The realized (or actual) volatility $\sigma(\hat{\omega}, \Sigma)$ results are depicted in Figure 4. Here, we observe a slight deterioration in LH-PCA's realized volatility performance in both CRSP and EURO. The average realized volatility of LH-PCA in CRSP is 7.43% per annum, which is comparable to the 7.41% per annum performance of LW's realized volatility, but higher than that of PCA and GPS. In the case of EURO, LH-PCA achieves lower mean realized volatility (3.09%). However, other methods have a wider distribution of realized volatility, which sometimes leads to lower realized volatility. This mixed result is exactly what we documented in Section 3.5, where the level of actual volatility $\sigma(\hat{\omega}, \Sigma)$ is not a monotonic function of SOR bias $\|\epsilon\|$. Putting together, LH-PCA demonstrates powerful performance in mitigating SOR bias in both the United States and European stock markets, as indicated by Figure 3. However, due to the non-monotonic nature between the degree of SOR bias and realized volatility, Figure 4 produces less clear-cut results.

We recast the empirical comparison using an alternative performance metric, MRAD, to demonstrate that LH-PCA consistently outperforms competing methods in estimating factor models (again, outliers are removed in the EURO sample). Overall, the findings in this section provide strong evidence that employing a long-history approach such as LH-PCA yields more accurate GMVP risk forecasts and substantially mitigates second-order risk.

6.4 Robustness Checks

We conduct a series of robustness checks to ensure that our main results are not sensitive to specification choices. In Section D.3 of the Electronic Companion, we demonstrate that LH-PCA provides superior volatility prediction remains robust under various conditions. Specifically, our results hold when varying our choice of T_L, and when adjusting the number of PCA eigenfactors by either adding or subtracting factors, or by using a fixed number of factors.

In addition, we confirm that our findings remain consistent when using the extended CRSP sample (1961–2021), as well as its subsamples (1961–1980, 1981–2000). Table 4 presents subsample results using extended data from CRSP World. We extend the main CRSP sample (2001–2021) back to 1961, providing 61 years of daily stock return data. The dataset is divided into three periods (1961–1980; 1981–2000; and our main CRSP sample of 2001–2021), creating a framework that, while not strictly a three-fold validation, offers a similar structured approach that researchers can leverage for robustness checks. This finding is particularly valuable, as both earlier periods experienced major stock market shocks, including the 1973 oil crisis and the stagflation of the 1970s, as well as the Black Monday crash in 1987.

Each panel employs a different test statistic to ensure that our findings are not driven by a specific statistical measure. Panel A uses Bias Statistics, which has an ideal value of one under the assumption of normally distributed daily returns and when volatility predictions are perfect. Panel B employs MRAD, while Panel C utilizes Q-statistics. Across all three subsamples, as well as the full 61-year dataset, we consistently find that LH-PCA outperforms standard PCA, as its volatility predictions are consistently closer to the ideal values suggested by each test statistic. Therefore, we conclude that the performance advantage of LH-PCA is neither sample-specific nor method-specific.

7 Conclusion

This paper introduces Long-History Principal Component Analysis (LH-PCA), a robust framework for estimating precision matrices in high-dimensional portfolios under dynamic factor structures with weak loadings. By extending the data history used for PCA estimation from traditional short-to-medium

Panel A. Excess Bias Statistics	PCA	LH-PCA	Difference
CRSP 1961 - 1980	1.40	0.90	0.50*** (6.93)
CRSP 1981 - 2000	2.81	1.81	1.00*** (9.04)
CRSP 2001 - 2021	2.73	2.11	0.62*** (5.55)
CRSP 1961 - 2021	2.25	1.55	0.69*** (14.20)

Panel B. Excess MRAD	PCA	LH-PCA	Difference
CRSP 1961 - 1980	0.98	0.55	0.43*** (6.02)
CRSP 1981 - 2000	2.00	1.27	0.73*** (6.58)
CRSP 2001 - 2021	1.87	1.36	0.52*** (4.69)
CRSP 1961 - 2021	1.60	1.00	0.60*** (12.24)

Panel C. Excess Q-statistics	PCA	LH-PCA	Difference
CRSP 1961 - 1980	3.39	1.73	1.66*** (27.72)
CRSP 1981 - 2000	11.52	5.34	6.19*** (3.23)
CRSP 2001 - 2021	10.87	7.12	3.76*** (15.27)
CRSP 1961 - 2021	7.70	4.19	3.51*** (28.97)

Table 4: CRSP Longer Sample Results

Note: This table presents robustness checks using an extended CRSP sample, covering 61 years of daily data (1961–2021). The dataset is divided into three periods: 1961–1980, 1981–2000, and 2001–2021, with the 2001–2021 results corresponding to our main findings in the text. Additionally, we analyze results for the entire 1961–2021 period. PCA refers to the standard PCA method using T_M, while LH-PCA employs T_L. For each panel, we report the deviation of the corresponding test statistics from their ideal values under the assumption of normally distributed daily returns, adopting a conservative perspective. Specifically, the reported figures include Excess Bias Statistics (Panel A), Excess MRAD (Mean Rolling Absolute Deviation) (Panel B), and the Excess mean of Q-statistics (Panel C), all computed using Global Minimum Variance Portfolios (GMVP) across different subsamples and estimation methods. The last column (Diff) reports the difference in the test statistic along with the associated t-statistics.

horizons (e.g., $T_M = 250$ days) to a long history ($T_L = 1500$ days), LH-PCA mitigates second-order risk (SOR) bias more effectively than competing methodologies, including Ledoit-Wolf shrinkage, standard PCA, and the GPS eigenvector correction (Goldberg et al., 2022). Theoretically, we show that LH-PCA maintains consistency in estimating the factor-driven covariance and precision matrices even under variable factor volatilities and time-varying loadings. Empirically, both simulation results and real-world stock return data from the U.S. (CRSP) and European markets demonstrate that LH-PCA produces more accurate forecasts of GMVP volatility, as measured by bias statistics, volatility ratios, MRAD, and Q-statistics.

Our findings have important implications for portfolio risk management in the presence of high-dimensional and dynamically evolving return structures. The superior performance of LH-PCA is particularly evident in environments where narrow factors are prominent and traditional PCA-based estimators are vulnerable to dispersion bias. Moreover, the method remains effective even when responsive covariance adjustments (RCA) are incorporated, further validating its applicability in fast-changing financial markets. In light of increasing investor reliance on data-driven risk models, LH-PCA offers a scalable, theoretically grounded, and empirically validated approach to improving portfolio construction and volatility forecasting under realistic and noisy conditions.

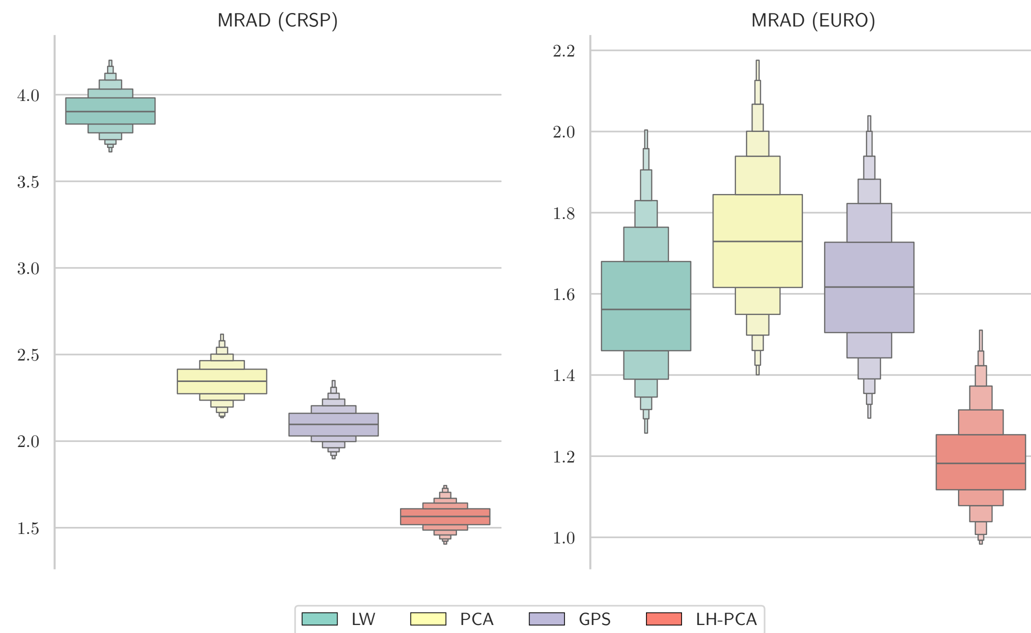


Figure 5: MRAD Results

Note: This figure shows the distributions of the estimated GMVP's MRAD obtained from daily CRSP (left panel) and EURO (right panel) stock returns data. LW stands for Ledoit-Wolf shrinkage method, PCA stands for the plain PCA method with T_M, GPS indicates the methodology proposed by Goldberg et al. (2022) with T_M. We employ T_L for LH-PCA. These results are obtained from 1000 simulation runs with varying seeds, acquired through bootstrapping with replacements for cross-validation. For the European data, we cap Z-scores at 100 and -100 to mitigate the impact of a small number of extreme days that cause factor models based on medium-length windows (such as LW, PCA, and GPS) to become overly sensitive to outliers. This conservative treatment favors competing methods over LH-PCA.

References

Adrian, T. and Franzoni, F. (2009). Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM. *Journal of Empirical Finance*, 16(4):537–556.

Aït-Sahalia, Y., Matthys, F., Osambela, E., and Sircar, R. (2024). When uncertainty and volatility are disconnected: Implications for asset pricing and portfolio performance. *Journal of Econometrics,*, page In Press.

Alexander, G. J. and Chervany, N. L. (1980). On the Estimation and Stability of Beta. *Journal of Financial and Quantitative Analysis*, 15(1):123–137.

Anderson, E. W., Ghysels, E., and Juergens, J. L. (2009). The impace of risk and uncertainty on expected returns. *Journal of Financial Economics*, 94(2):233–263.

Bai, J. and Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70(1):191–221.

Bai, J. and Ng, S. (2023). Approximate factor models with weaker loadings. *Journal of Econometrics*, 235:1893–1916.

Barigozzi, M., Hallin, M., Soccorsi, S., , and von Sachs, R. (2021). Time-varying general dynamic factor models and the measurement of financial connectedness. *Journal of Econometrics*, 222(1):324–343.

Basak, G. K., Jagannathan, R., and Ma, T. (2009). Jackknife Estimator for Tracking Error Variance of Optimal Portfolios. *Management Science*, 55(6):990–1002.

Bates, B. J., Plagborg-Møller, M., Stock, J. H., and Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, 177:289–304.

Bekaert, G., Engstrom, E., and Xing, Y. (2009a). Risk, uncertainty, and asset prices. *Journal of Financial Economics*, 91(1):59–81.

Bekaert, G., Hodrick, R. J., and Zhang, X. (2009b). International Stock Return Comovements. *Journal of Finance*, 64(6):2591–2626.

Bernardi, S., Leippold, M., and Lohre, H. (2019). Second-order risk of alternative risk parity strategies. *Risk*, 21(3):1–25.

Blume, M. E. (1971). On the Assessment of Risk. *Journal of Finance*, 26(1):1–10.

Bollerslev, T., Hood, B., Huss, J., and Pedersen, L. H. (2018). Risk Everywhere: Modeling and Managing Volatility. *Review of Financial Studies*, 31(7):2729–2773.

Brenner, M. and Izhakian, Y. (2018). Asset pricing and ambiguity: Empirical evidence. *Journal of Financial Economics*, 130(3):503–531.

Clarke, R., Silva, R. D., and Thorley, S. (2011). Minimum-variance portfolio composition. *Journal of Portfolio Management*, 2(37):31–45.

Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *FRB of New York Staff Report*, 326:1–44.

DeMiguel, V., Garlappi, L., Nogales, F. J., and Uppal, R. (2009a). A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. *Management Science*, 55(5):798–812.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009b). Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies*, 22(5):1915–1953.

Ding, Y., Li, Y., and Zheng, X. (2021). High Dimensional Minimum Variance Portfolio Estimation under Statistical Factor Models. *Journal of Econometrics*, 222(1):502–515.

Doidge, C., Karolyi, A. G., and Stulz, R. M. (2017). The U.S. listing gap. *Journal of Financial Economics*, 123(3):464–487.

Eckart, C. and Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, 123(3):441–463.

Fama, Eugene, F. and French, K. R. (1998). Value versus Growth: The International Evidence. *Journal of Finance*, 53(6):1975–1999.

Fan, J., Liao, Y., and Wang, W. (2016). Projected Principal Component Analysis in Factor Models. *Annals of Statistics*, 44(1):219–254.

Freyaldenhoven, S. (2022). Factor models with local factors: Determining the number of relevant factors. *Journal of Econometrics*, 229(1):80–102.

Goldberg, L., Papanicolaou, A., and Shkolnik, A. (2022). The Dispersion Bias. *SIAM Journal of Financial Mathematics*, 13(2):521–550.

Han, X. and Caner, M. (2017). Determining the number of factors with potentially strong within-block correlation in error terms. *Econometric Reviews*, 36(6-9):946–969.

Harvey, C. R., Hoyle, E., Korgaonkar, R., Rattray, S., Sargaison, M., and Hemert, O. V. (2018). The Impact of Volatility Targeting. *Journal of Portfolio Management*, 45(1):14–33.

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58(4):1651–1684.

Kelly, B., Pruitt, S., and Su, Y. (2019). Characteristics Are Covariances: A Unified Model of Risk and Return. *Journal of Financial Economics*, 134(3):501–524.

Kim, S., Korajczyk, R. A., and Neuhierl, A. (2021). Arbitrage Portfolios. *Review of Financial Studies*, 34(6):2813–2856.

Ledoit, O. and Péché, S. (2011). Eigenvectors of Some Large Sample Covariance Matrix Ensembles. *Probability Theory and Related Fields*, 151:233–264.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5):603–621.

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

Ledoit, O. and Wolf, M. (2004b). Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management*, 30(4):110–119.

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

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

Menchero, J., Morozov, A., and Pasqua, A. (2013). Predicting Risk at Short Horizons: A Case Study for the USE4D Model. MSCI Model Insight, MSCI Barra documentation.

Minka, T. P. (2000). Automatic choice of dimensionality for PCA. *Advances in Neural Information Processing Systems*, 13:598–604.

Morozov, A., Wang, J., Borda, L., and Mencehro, J. (2012). The barra global equity model (gem3). *MSCI Barra model documentation*.

Onatski, A. (2010). Determining the Number of Factors from Empirical Distribution of Eigenvalues. *Review of Economics & Statistics*, 92(4):1004–1016.

Pastor, L. (2000). Portfolio Selection and Asset Pricing Models. *Journal of Finance*, 55(1):179–223.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.

Pelger, M. and Xiong, R. (2022a). Interpretable Sparse Proximate Factors for Large Dimensions. *Journal of Business & Economic Statistics*, 40(4):1642–1664.

Pelger, M. and Xiong, R. (2022b). State-varying factor models of large dimensions. *Journal of Business and Economic Statistics*, 40(3):1315–1333.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, 13(3):341–360.

Schwert, G. W. (1989). Why Does Stock Market Volatility Change Over Time? *Journal of Finance*, 44(5):1115–1153.

Shepard, P. (2009). Second Order Risk. *Working Paper, MSCI Barra*.

Shi, F., Shu, L., Yang, A., and He, F. (2020). Improving Minimum-Variance Portfolios by Alleviating Overdispersion of Eigenvalues. *Journal of Financial and Quantitative Analysis*, 55(8):2700–2731.

Su, L. and Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics*, 198(1):84–101.

Uematsu, Y. and Yamagata, T. (2023). Estimation of sparsity induced weak factor models. *Journal of Business & Economic Statistics*, 41(1):213–227.

Wang, W. and Fan, J. (2017). Asymptotics of Empirical Eigenstructure for High Dimensional Spiked Covariance. *Annals of Statistics*, 45(3):1342–1374.

Long-History PCA in a Dynamic Factor Model with Weak Loadings

Electronic Companions

A Proofs

Proof of Proposition 2. Define $\tilde{\Sigma} := \hat{\Sigma}\Sigma^{-1}\hat{\Sigma}$ as the *sandwich* covariance matrix. Then, the actual variance of the estimated GMVP is given by

$$(\text{Actual variance of } \hat{\omega}) = \sigma^2(\hat{\omega}, \Sigma) = \frac{\mathbf{1}^\top \tilde{\Sigma}^{-1}\mathbf{1}}{(\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1})^2} = \frac{\sigma^4(\hat{\omega}, \hat{\Sigma})}{\sigma^2(\tilde{\omega}, \tilde{\Sigma})} = \frac{(\text{Predicted variance of } \hat{\omega})^2}{(\text{Predicted variance of } \tilde{\omega})} ,$$

where $\tilde{\omega}$ is the weight vector of the estimated GMVP based on $\tilde{\Sigma}$. This implies that $\sigma(\hat{\omega}, \hat{\Sigma})$ is the geometric mean of $\sigma(\hat{\omega}, \Sigma)$ and $\sigma(\tilde{\omega}, \tilde{\Sigma})$; i.e.,

$$(\text{VR}) = \frac{\sigma(\hat{\omega}, \Sigma)}{\sigma(\hat{\omega}, \hat{\Sigma})} = \frac{\sigma(\hat{\omega}, \hat{\Sigma})}{\sigma(\tilde{\omega}, \tilde{\Sigma})} . \tag{10}$$

Next, by the definitions of ϵ and $\tilde{\Sigma}$, we have

$$\tilde{\Sigma}^{-1} = \left(\Sigma^{-1} + \epsilon\right) \Sigma \left(\Sigma^{-1} + \epsilon\right) = \Sigma^{-1} + 2\epsilon + \epsilon\Sigma\epsilon \approx \Sigma^{-1} + 2\epsilon ,$$

where the approximation is justified by small $\|\epsilon\|$. It follows that

$$\sigma^2(\tilde{\omega}, \tilde{\Sigma}) = \frac{1}{\mathbf{1}^\top \tilde{\Sigma}^{-1}\mathbf{1}} \approx \frac{1}{\mathbf{1}^\top \Sigma^{-1}\mathbf{1} + 2 \cdot \mathbf{1}^\top \epsilon \mathbf{1}}$$

and combining (2) and (10) completes the proof. $\square$

Proof of Proposition 3. Under the hypothesis of $\Sigma\hat{\Sigma}^{-1}1 \approx \lambda 1$ with some $\lambda > 0$, we have

$$\mathbf{1}^\top \hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{1} \approx \frac{\left(\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1}\right)^2}{\mathbf{1}^\top \Sigma^{-1}\mathbf{1}}$$

This implies that

$$(\text{Actual variance of } \hat{\omega}) = \frac{\mathbf{1}^\top \hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\mathbf{1}}{\left(\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1}\right)^2} \approx \frac{\left(\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1}\right)^2}{\mathbf{1}^\top \Sigma^{-1}\mathbf{1}} \frac{1}{\left(\mathbf{1}^\top \hat{\Sigma}^{-1}\mathbf{1}\right)^2} = (\text{True variance of } \omega) ,$$

where the approximation becomes exact when $\Sigma\hat{\Sigma}^{-1}\mathbf{1} = \lambda\mathbf{1}$. $\square$

Proof of Theorem 7. Recall $\Lambda_t = \Lambda_0 + h_{NT}\xi_t \in \mathbb{R}^{N\times r}$. For a fixed pair of (N,T), we have

$$\begin{aligned}
\Sigma_N(T) &= \frac{1}{T}\sum_{t=1}^{T}\left(\Lambda_t F_t + e_t\right)\left(\Lambda_t F_t + e_t\right)^\top \\
&= \frac{1}{T}\sum_{t=1}^{T}\left((\Lambda_0 + h_{NT}\xi_t)F_t + e_t\right)\left((\Lambda_0 + h_{NT}\xi_t)F_t + e_t\right)^\top \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(a_t + b_t + b_t^\top + c_t\right) \ ,
\end{aligned}$$

where a_t, b_t and c_t are given by

$$\begin{aligned}
a_t &= \Lambda_0 F_t F_t^\top \Lambda_0^\top \\
b_t &= \Lambda_0 F_t \left(h_{NT}\xi_t F_t + e_t\right)^\top \\
c_t &= \left(h_{NT}\xi_t F_t + e_t\right)\left(h_{NT}\xi_t F_t + e_t\right)^\top \ .
\end{aligned}$$

As $\left(\sum_{j=1}^{m} x_j\right)^2 \leq m\sum_{j=1}^{m} x_j^2$ holds for all integers $m \geq 1$, we have

$$\frac{1}{N^2}\left\|\Sigma_N(T) - \overline{\Sigma}_N(T)\right\|_F^2 \leq 6\sum_{j=1}^{6}\Psi_j \ ,$$

where letting $N, T \to \infty$ justifies with our assumptions that

$$\begin{aligned}
\Psi_1 &= \frac{1}{N^2}\Big\|\Lambda_0 \underbrace{\frac{1}{T}\sum_{t=1}^{T}\left(F_t F_t^\top - \mu_t^2\right)}_{=o_p(1)} \Lambda_0^\top\Big\|_F^2 \leq \frac{1}{N^2}\underbrace{\left\|\Lambda_0\Lambda_0^\top\right\|_F^2}_{=O_p(N)} O_p(1) = O_p\left(\frac{1}{N}\right) \\
\Psi_2 &= \frac{1}{N^2}\Big\|\frac{1}{T}\sum_{t=1}^{T}\left(e_t e_t^\top - \nu_t^2\right)\Big\|_F^2 \leq \frac{1}{N^2}\left(O_p(N) + O_p\Big(\frac{N(N-1)}{T}\Big)\right) = O_p\left(\frac{1}{\min\{N,T\}}\right) \\
\Psi_3 &= \frac{1}{N^2}\left\|\frac{2}{T}\sum_{t=1}^{T}\Lambda_0 F_t e_t^\top\right\|_F^2 \leq \frac{1}{N^2}O_p\left(N\right) \leq O_p\left(\frac{1}{N}\right) \\
\Psi_4 &= \frac{h_{NT}^2}{N^2}\left\|\frac{2}{T}\sum_{t=1}^{T}\Lambda_0 F_t F_t^\top \xi_t^\top\right\|_F^2 \leq O_p\left(\frac{h_{NT}^2}{N^2T^2}Q_2(N,T)\right) = O_p\left(\frac{1}{N}\right) \\
\Psi_5 &= \frac{h_{NT}^2}{N^2}\left\|\frac{2}{T}\sum_{t=1}^{T}e_t F_t^\top \xi_t^\top\right\|_F^2 \leq O_p\left(\frac{h_{NT}^2}{N^2}Q_1(N,T)\right) = O_p\left(\frac{1}{N}\right) \\
\Psi_6 &= \frac{h_{NT}^4}{N^2}\left\|\frac{1}{T}\sum_{t=1}^{T}\xi_t F_t F_t^\top \xi_t^\top\right\|_F^2 = O_p\left(\frac{h_{NT}^4}{N^2T^2}Q_3(N,T)\right) = O_p\left(\frac{1}{\min\{N,T\}}\right) \ ,
\end{aligned}$$

which completes the proof

$$\left\|\frac{1}{N}\left(\Sigma_N(T) - \overline{\Sigma}_N(T)\right)\right\|_F^2 = O_p\left(\frac{1}{\min\{N,T\}}\right) \ .$$

In addition, our assumptions yield that there exists some $c > 0$ such that

$$\min\left\{\lambda_{\min}\left(\Sigma_N(T)\right),\ \lambda_{\min}\left(\overline{\Sigma}_N(T)\right)\right\} > c\ ,$$

where $\lambda_{\min}(A)$ represents the smallest eigenvalue of A. Note that this condition regarding the uniform boundedness (away from zero) of the minimum eigenvalue is consistent with a fixed $r < \infty$. This implies that we have

$$\begin{aligned}\left\|\frac{1}{N}\left(\overline{\Sigma}_N^{-1}(T) - \Sigma_N^{-1}(T)\right)\right\|_F &= \left\|\frac{1}{N}\left(\Sigma_N^{-1}(T)\ \left(\Sigma_N(T) - \overline{\Sigma}_N(T)\right)\ \overline{\Sigma}_N^{-1}(T)\right)\right\|_F \\ &\leq \underbrace{\frac{1}{\underline{\lambda}(\Sigma_N(T))}}_{<1/c} \cdot \left\|\frac{1}{N}\left(\Sigma_N(T) - \overline{\Sigma}_N(T)\right)\right\|_F \cdot \underbrace{\frac{1}{\underline{\lambda}(\overline{\Sigma}_N(T))}}_{<1/c} \\ &= \mathcal{O}\left(\left\|\frac{1}{N}\left(\Sigma_N(T) - \overline{\Sigma}_N(T)\right)\right\|_F\right)\ ,\end{aligned}$$

which directly completes the proof. □

Proof of Corollary 8. Let $\overline{X}_t = \Lambda_0 F_t + e_t \in \mathbb{R}^{N\times 1}$ and $\overline{\mathbf{X}} = [\overline{X}_1, \ldots, \overline{X}_T]^\top \in \mathbb{R}^{T\times N}$, where we have

$$\frac{1}{T}\overline{\mathbf{X}}^\top\overline{\mathbf{X}} = \Lambda_0\left(\frac{1}{T}\mathbf{F}^\top\mathbf{F}\right)\Lambda_0^\top + \frac{1}{T}\mathbf{e}^\top\mathbf{e}\ .$$

Using the same notations as in the proof of Theorem 7, we obtain that the perturbation matrix defined as

$$\Delta_N(T) := \frac{1}{N}\left(\frac{1}{T}\mathbf{X}^\top\mathbf{X} - \frac{1}{T}\overline{\mathbf{X}}^\top\overline{\mathbf{X}}\right)$$

satisfies the condition expressed as

$$\|\Delta_N(T)\|_F^2 \leq 2\left(\Psi_4 + \Psi_6\right) = \mathcal{O}_p\left(\frac{1}{\min\{N,T\}}\right)\ .$$

Since the leading eigenvalues of $\frac{1}{T}\overline{\mathbf{X}}^\top\overline{\mathbf{X}}$ are of order $\mathcal{O}_p(N^{\alpha_k})$ for $k = 1, \ldots, r$ and of order $\mathcal{O}_p(1)$ for the remaining components, the Davis-Kahan theorem ensures that the PC estimators based on $\mathbf{X}$ and $\overline{\mathbf{X}}$ must be asymptotically equivalent. As Λ_0 is time-invariant, the singular value decomposition of $\frac{\overline{\mathbf{X}}}{\sqrt{T}}$ yields $\overline{\mathbf{X}} = \sqrt{T}\mathbf{U}_r\mathbf{D}_r\mathbf{V}_r^\top + \tilde{e}_r$, which leads to

$$\frac{1}{T}\overline{\mathbf{X}}^\top\overline{\mathbf{X}} = \tilde{\Lambda}_r\left(\frac{1}{T}\tilde{F}_r^\top\tilde{F}_r\right)\tilde{\Lambda}_r^\top + \frac{1}{T}\tilde{e}_r\tilde{e}_r^\top := \widetilde{\Sigma}_N(T)$$

within the PCA framework specified in Section 3.2, where the PC estimators of $\overline{\mathbf{X}}$ are expressed as $\left(\tilde{F}_r, \tilde{\Lambda}_r\right) = \left(\sqrt{T}\mathbf{U}_r, \mathbf{V}_r\mathbf{D}_r\right)$. It follows that we have

$$\left\|\frac{1}{N}\left(\overline{\Sigma}_N(T) - \widetilde{\Sigma}_N(T)\right)\right\|_F^2 \leq 2\left(\Psi_1 + \Psi_2\right) = O_p\left(\frac{1}{\min\{N,T\}}\right)\ .$$

Since our assumptions ensure that the smallest eigenvalue of $\overline{\mathbf{X}}$ is uniformly bounded away from zero, the convergence in probability holds for their inverses as well. Applying Theorem 7 along with the triangle inequality for the Frobenius norm completes the proof. □

B Ledoit and Wolf model constructions

One of our benchmark methods in assessing the risk forecasting performance is Ledoit and Wolf (2004b) shrinkage approach. This shrinkage estimator is computed as

$$\widehat{\Sigma}_{Shrink} = \widehat{\delta}^{*}\mathbf{F} + (1 - \widehat{\delta}^{*})\mathbf{S} \ , \tag{11}$$

where $\mathbf{F}$ is the shrinkage target, $\mathbf{S}$ is the sample covariance matrix. The shrinkage constant, $\widehat{\delta}^{*} \in [0, 1]$, allows users to form a convex linear combination of the two matrices. We follow Appendix B in the Ledoit and Wolf (2004b) in constructing the optimal shrinkage parameter, and use constant-correlation model as suggested in the same paper in constructing the shrinkage target matrix $\mathbf{F}$. Consistent with our simulation and empirical approach, we compute the shrinkage estimator at each moving window in obtaining the risk forecasts.

C Responsive Covariance Adjustment

In real-world financial markets, volatility is constantly changing and the distribution of stock returns is not stationary over time. Accordingly, practitioners often assign more weight to recently realized returns than to distant observations to capture the contemporaneous covariance structure with a *responsive covariance adjustment* (RCA) based on the Exponentially Weighted Moving Averages (EWMA) model; see Harvey et al. (2018) and Bollerslev et al. (2018), for instance. The generalization to variable volatility is implicitly used in commercially available fundamental and latent factor models, which estimate factor and idiosyncratic volatilities using a short history, such as the EWMA with a half-life T_S of around 40 days (Menchero et al., 2013).

In what follows, we provide a detailed description of our approach to the implementation of the RCA framework. Let $\Sigma_N^{\psi}(t)$ denote the $N \times N$ covariance matrix of the $N \times 1$ factor-based return vector $\psi_t := \Lambda_t F_t$. Within each trailing window of T days, a decay factor $\eta \in (0, 1)$ yields the adjusted factor-based return covariance as a weighted average of $\psi_{t-m}\psi_{t-m}^{\top}$ with $m \in \{0, \ldots, T-1\}$. Simply put, the daily forecast of factor-based return covariance matrix estimated on day t (for day $t+1$) is given by

$$\widehat{\Sigma}_t^{\psi}(T) = \sum_{m=0}^{T-1} \omega_m \psi_{t-m}\psi_{t-m}^{\top} \ , \tag{12}$$

where the weights satisfy $\sum_{m=0}^{T-1} \omega_m = 1$ and decrease by fixed proportion as $\omega_{m+1} = \eta\omega_m$ for $m \geq 0$.[27] We set $T_S = 40$ days, which yields a decay factor of $\eta = \exp(\frac{log0.5}{40}) \approx 0.9828$.[28]

We apply the RCA schemes independently to both the factor returns and the idiosyncratic factor returns to correct for changing factor volatilities. Note that we treat the idiosyncratic returns as uncor-

[27] It is worth noting that our RCA formulation in Formula (12) does not rely on the estimated factor model. This approach is motivated by its ability to provide superior risk predictions across various methods of estimating the factor model.

[28] We are motivated by Menchero et al. (2013) that use 42-day EWMA volatility half-life. Harvey et al. (2018) do volatility scaling using a half-life of 20 days. In the empirical section, we conduct a sensitivity analysis using different T_S and provide further discussion of this issue. Notice that EWMA has been applied in similar contexts; e.g., see Harvey et al. (2018) and Bollerslev et al. (2018).

related across stocks by zeroing out the non-diagonal elements of the idiosyncratic part of the covariance matrix to estimate the $N \times N$ idiosyncratic return covariance matrix on day t with window width T, which is denoted by $\widehat{\Sigma}_t^{\varepsilon}(T)$. This is consistent with the purpose of dimensional reduction by using a factor model for reducing the tendency of the minimum variance optimizer to overfit the data. We finally obtain the daily forecast of covariance matrix estimated on day t as $\widehat{\Sigma}_t(T) = \widehat{\Sigma}_t^{\psi}(T) + \widehat{\Sigma}_t^{\varepsilon}(T)$ by assuming that the idiosyncratic return is uncorrelated with the factor return in the population.

D Additional Information on the Empirical Analyses

This section details the empirical analyses discussed in the main body of the paper.

D.1 Data Construction

In constructing the dataset, we apply a set of filters consistent with standard practices in the empirical finance literature. Specifically, we focus on common stocks listed on major exchanges and retain only those with valid country and industry classifications, ensuring exposure to both country- and industry-specific risk factors. For the European sample, we identify 16 countries as European (in alphabetical order): Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. This is consistent with the existing finance literature that studies international stock returns, such as Fama and French (2017). As noted in Section 6, both the CRSP and EURO datasets begin in 2001. This is because we want to make a direct comparison between the United States financial market and the European stock market. Even within European countries, some countries adopted Euros later than others. For instance, Greece adopted the euro at the start of 2001. We can thus extend the data up to 1999 by applying a fixed conversion rate (1€ = 340.75 GRD) to the closing price on Greece stocks recorded prior to 2000. But we choose to focus on daily price recorded from 2001 to ensure the reliability of the data. For the Greece conversion issue, see https://ec.europa.eu/info/business-economy-euro/euro-area/euro/eu-countries-and-euro/greece-and-euro_en.

D.2 Moving window analysis

We provide more details on our moving window approach that calculates the risk forecasts at a daily frequency. To maintain uniformity of the stock universe across all methods, we generate a single-day forecast on day $t+1$ only for the stocks that are present throughout the longer T_L moving window spanning from day $t - T_L + 1$ to t, as well as on the forecast day $t+1$. If a stock is delisted on the prediction day $t+1$, we impute a loss on the prediction day as described in the previous section. Moreover, if a stock declines in value over several days leading up its delisting, then each of those days will have been counted as a prediction day in a previous rolling window. Thus, all of the losses leading up to the demise of a given stock are included in our analysis for the appropriate day and contribute to the bias statistics we compute; thus, we argue that our analysis is free of survivorship bias. Note, however, that our analysis in any given window excludes stocks that were first listed after the first day of the window. Thus, our analysis omits newly created firms, until there is a full T_L window's worth of data from which to make predictions.

Moreover, using the same period, one day, for both data and predictions obviates the need for serial correlation adjustments that would arise if the horizons were different; refer to the related discussion in Menchero et al. (2013). Schwert (1989) show that the estimates of volatility from daily data have much less error than the estimates from low frequency data. In this study, we argue that the benefits of the additional observations from a six-year history of daily returns outweigh the cost of changing factor loadings over that long history; note that six years of weekly data contain only 20% more observations than one year of daily data. We could increase the number of observations by using high-frequency intraday data, but this would require us to extrapolate overnight risk from trade during the day, when the market is open.

D.3 Tuning Parameters

In the main analysis, we used Bai and Ng (2002) (BN) estimator at each moving window to extract the number of PCA eigenfactors. Under BN environment, factors' cumulative effect on the n cross-sectional units grows proportionally to n. Hence, the K eigenvalues (that corresponds to the K PCA eigenfactors) of the data's covariance matrix grow proportionally to n while the rest of the eigenvalues stay bounded (Onatski, 2010). In this case, the number of factors can be estimated by separating relatively large eigenvalues from the rest using threshold functions. We denote $m = \min(N, T)$ and $M = \max(N, T)$, where N is the number of stocks and T is the number of days. Note that only X is observed.

Bai and Ng (2002) solve this issue through the lens of model selection problem,

$$\underset{1\leq k\leq kmax}{\operatorname{argmin}} \underbrace{f(V(k))}_{\text{signal}} + k \cdot \underbrace{g(N,T)}_{\text{penalty}} ,$$

where $V(k)$ is the minimal possible sum of squared errors of a rank k approximation of data matrix divided by NT expressed as

$$V(k) = \min_{\Lambda, F^k} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^k F_t^k)^2.$$

The proposed estimators vary on their penalty terms:

$$PC_{p1}(k) = V(k) + k \cdot \hat{\sigma}^2 \cdot \Big(\frac{N+T}{NT}\Big) \cdot \ln\Big(\frac{NT}{N+T}\Big) ,$$

$$IC_{p1}(k) = \ln(V(k)) + k \cdot \Big(\frac{N+T}{NT}\Big) \cdot \ln\Big(\frac{NT}{N+T}\Big) ,$$

$$PC_{p2}(k) = V(k) + k \cdot \hat{\sigma}^2 \cdot \Big(\frac{N+T}{NT}\Big) \cdot \ln(m) ,$$

$$IC_{p2}(k) = \ln(V(k)) + k \cdot \Big(\frac{N+T}{NT}\Big) \cdot \ln(m) ,$$

where $\hat{\sigma}^2$ is a consistent estimate of $\hat{\sigma}^2 = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (e_{it})^2$. Consistent with prior studies, we adopt $PC_{p1}(k)$ as the penalty function in our principal component estimation.

While the BN estimator was not originally derived to handle weak factors in their original study (Bai and Ng, 2002), several studies, including Freyaldenhoven (2022) and Han and Caner (2017), state that the BN estimator remains consistent in detecting both strong and weak factors when $\alpha > 0$, where $\alpha \in (0, 1]$ represents the factor strength such that the corresponding eigenvalue of the population covariance matrix is of order $\mathcal{O}_p(N^\alpha)$, with N denoting the population size. These studies also emphasize that the BN criteria tend to overestimate (rather than underestimate) the number of relevant factors when $\alpha > 1/2$, which mitigates concerns of excluding meaningful weak factors. Furthermore, as noted in the conclusion of Bai and Ng (2023), the BN estimator does not inherently lead to underestimation, even in the presence of weak or "local" factors.

In our setting, this potential overestimation is not a drawback but rather an advantage. Unlike macroeconomic or asset pricing contexts where only pervasive factors are typically deemed relevant, volatility modeling benefits from capturing transient, high-impact events such as flash crashes or liquidity disruptions. These may correspond to weak factors that do not meet traditional thresholds (e.g., $\alpha > 1/2$) but still enhance short-term volatility forecasts. Thus, using the BN estimator—even if it selects a higher number of factors—remains appropriate for our objective of estimating precision matrices.

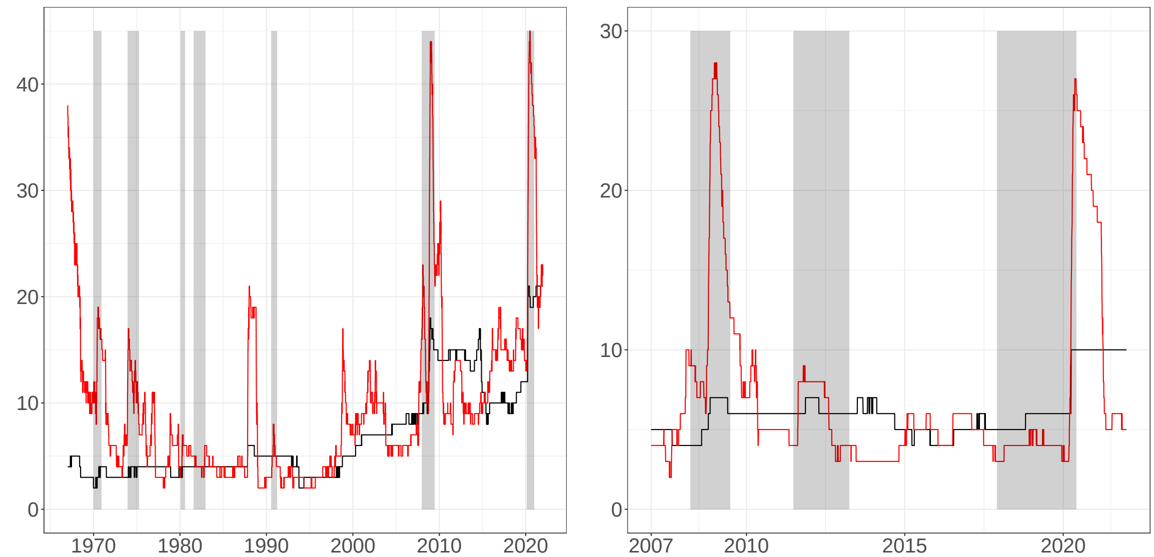


Figure 6: Number of PCA Eigenfactors using Bai and Ng (2002) Estimator

Note: The left (right) panel shows the number of PCA eigenfactors extracted at each moving window in the CRSP (EURO) daily stock return sample from 1961–2021 (2001–2021). The red solid line represents factors extracted using the Bai and Ng (2002) estimator with $T_M = 250$, while the black solid line corresponds to $T_L = 1500$. The longer U.S. market history allows estimation back to 1961. The x-axis represents the prediction date, and the y-axis denotes the number of extracted eigenfactors. Shaded areas indicate NBER (OECD-based) recessions for the U.S. (EURO area).

The plot presented in Figure 6 shows the number of PCA eigenfactors extracted using BN estimator where the estimator applied to the 250 (1500) trading moving windows is depicted by the red (black) line. Following other literature, we use $PC_{p1}(k)$ at each moving window to determine the K number of PCA eigenfactors to extract. We observe that the BN estimator applied to the medium-history window (red solid line) reveals a notable variation. In particular, it is suggested that additional factors are needed to explain the recession period, such as the Global Financial Crisis in 2008 and the COVID-19 Crisis at the beginning of 2020. This indirectly hints that disaster-related factors that are transient in nature are present during the economic downturn. The CRSP requires a higher number of factors than EURO, indicating the complexity of the United States market compared to its EURO counterpart. Finally, we confirm that this result is not method-specific as we also observe a similar pattern using Minka (2000)'s PCA (results are unreported).

For all three methodologies (PCA, GPS, and LH-PCA) that require a designated number of eigenfactors, we apply the BN estimator to a medium-history moving window of 250 trading days. In other

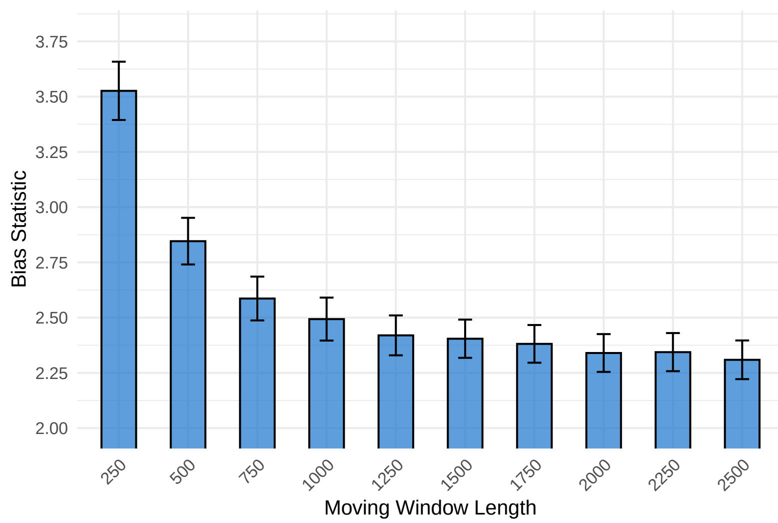


Figure 7: Bias Statistic Across Different Moving Window Lengths

Note: The figure reports the Bias Statistic (BS) across varying moving window lengths, using CRSP data from 2001 to 2021. A window length of 250 (far left) corresponds to the standard PCA method with T_M, while a window length of 1500 (center) represents the LH-PCA approach with T_L. To ensure an apples-to-apples comparison, we fix the starting point across all specifications so that any differences in BS can be attributed solely to the length of the moving window. Confidence intervals for each BS estimate are derived from 1000 simulation runs with varying random seeds, using bootstrapping with replacement for cross-validation.

words, the number of PCA eigenfactors used to construct the LHPCA risk forecasts is obtained from the medium-history moving window rather than the long-history moving window. This is to provide an apple-to-apple comparison where the differing performance comes solely from the factor model approaches.

We conclude this section with two additional figures. First, we show that selecting a moving window length different from our benchmark choice of $T_L = 1500$ trading days does not materially affect the results. As illustrated in Figure 7, the difference in Bias Statistic performance between plain PCA with $T_M = 250$ days and LH-PCA with $T_L = 1000$ days or more remains substantial. At the same time, the differences between $T_L = 1500$ and the longest window considered in this exercise, $T_L = 2500$, are small and statistically insignificant, as indicated by the overlapping confidence intervals. Next, Figure 8 confirms that either adding or subtracting a certain number (α) of factors in addition to the Bai and Ng (2002) estimator does not change the main implication that LH-PCA works better than the standard PCA. We do observe that when the number of factors is set to a very small number of fixed value, such as $K = 3$ or $K = 5$ (which likely underrepresents the true number needed to capture portfolio volatility), the performance of LH-PCA and PCA becomes more similar. However, as we increase the number of fixed factors K to extract, the performance gap widens, reassuring the superior performance of LH-PCA. In both figures, similar patterns are observed when using MRAD and Q-statistics (results unreported).

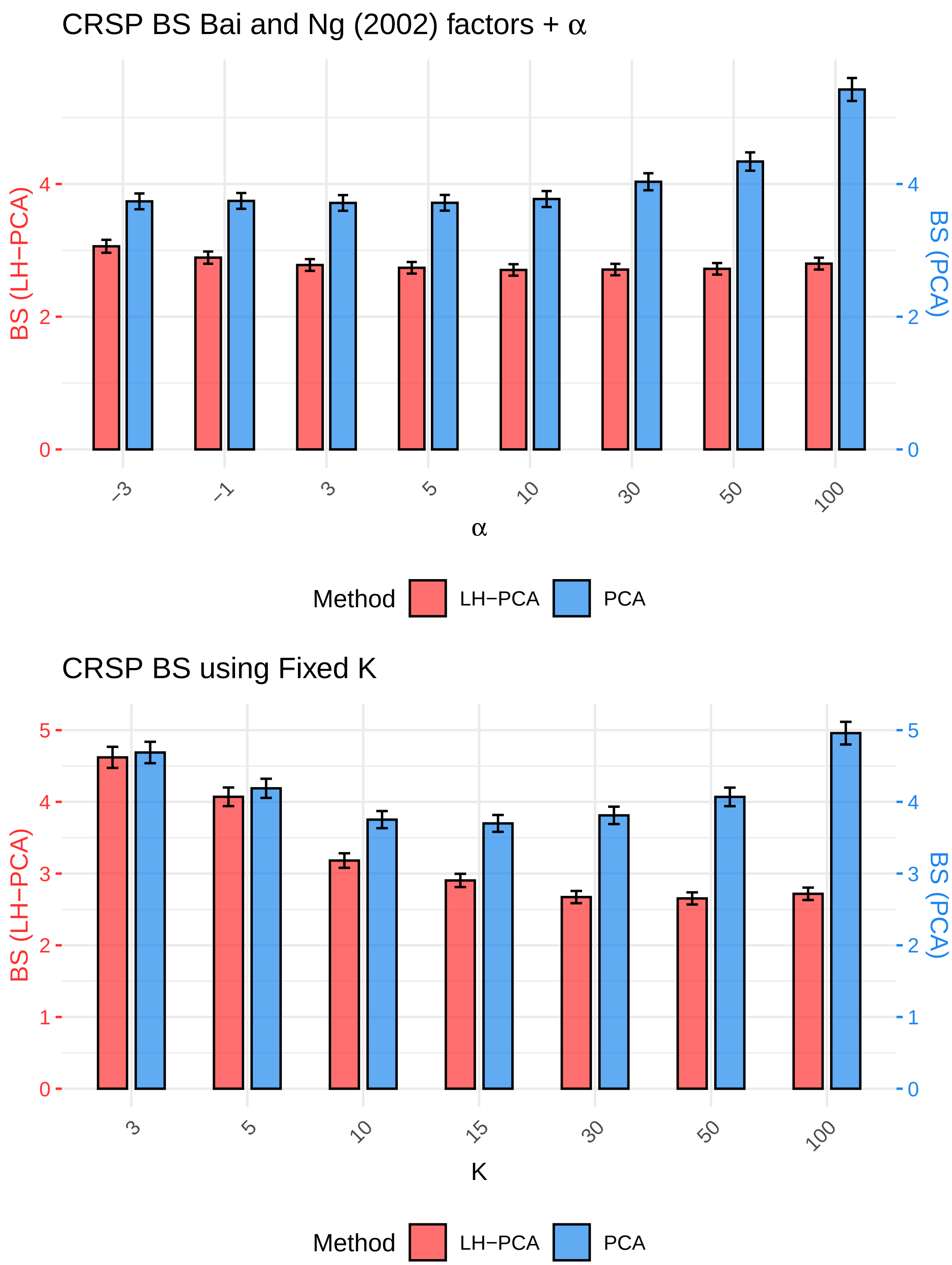


Figure 8: Robustness Checks on the Number of PCA eigenfactors

Note: The above (below) panel displays the Bias Statistic (BS) performance, where the number of PCA factors is determined using the Bai and Ng (2002) estimator with $T_M = 250$, adjusted by adding or subtracting a fixed number of factors, α. The data used is CRSP 2001 - 2021 daily. The red (blue) bar plots represent the BS values for LH-PCA (PCA). Confidence intervals for each BS estimate are derived from 1,000 simulation runs with varying random seeds, using bootstrapping with replacement for cross-validation.