Observable versus Latent Risk Factors

Yuan Liao[*] Viktor Todorov[†]

August 21, 2024

Abstract

We test for temporal stability in local linear projection coefficients of observable risk factors on latent ones embedded in the cross-section of asset prices and extracted via Principal Component Analysis (PCA). The test can be used for deciding if and over what horizon conventional linear asset pricing techniques can be employed for studying the pricing of observable factors. The proposed test explores the fact that under the null hypothesis residuals from global linear projections of observable factors on latent ones, computed over a fixed time interval via PCA, should be also locally uncorrelated with the PCA factors. The test is fully nonparametric. Its asymptotic behavior is derived under a joint in-fill and large cross-section asymptotic setup. In an empirical application, we show that a linear relation between the market volatility factor and the latent systematic risk factors embedded in the cross-section of stock returns exists only over short periods of length of one trading day.

Keywords: asset pricing, high-frequency data, latent factor model, nonparametric test, PCA, systematic risk.

JEL classification: C51, C52, G12.

1 Introduction

Understanding what risk is priced by investors is at the core of asset pricing. In equilibrium asset pricing models, the priced risk is typically captured by a small set of risk factors, related to e.g., features of the consumption dynamics of a representative

[*]Department of Economics, Rutgers University; e-mail: yuan.liao@rutgers.edu.
[†]Kellogg School of Management, Northwestern University; e-mail: v-todorov@northwestern.edu.

agent in the economy. These systematic risk factors are usually latent and not directly observable by economists. One can often form observable proxies for the latent risk factors, e.g., by measuring features of the consumption dynamics or by constructing factor mimicking portfolios. These observable factors carry economic meaning related to the model that is estimated or tested. Nevertheless, there is typically a gap between the latent risk factors and the observable ones, as noted by Connor and Korajczyk (1991), Fan et al. (2021) and Giglio and Xiu (2021). This can be due to measurement error or more generally because the observable factors might contain risks that do not generate systematic movements in asset prices. Moreover, the observable factors' exposure towards the latent ones can change even when the assets' exposure towards them do not. This would invalidate the use of standard techniques for evaluating linear asset pricing models, see e.g., Cochrane (2005). The goal of this paper, therefore, is to design a test for deciding if this is the case over a fixed time interval.

The key underlying premise of our analysis is that, apart from a few possible discrete changes (at random times and of random size), there is temporal stability in the exposures of assets towards the latent systematic risk factors. However, time-varying measurement error in the observable factor proxies and heteroskedasitcity in the latent factor returns can make assets' exposures towards the observable proxies for the risk factors vary nontrivially even over very short time periods, see e.g., Andersen et al. (2021) and Andersen et al. (2022), who provide evidence for such variation within the course of a single trading day.[1] This variation in estimated exposures towards observable factors provides a challenge for the direct use of high-frequency data for asset pricing applications as the time-varying factor exposures need to be accounted for in some way.

If there is a linear relation between the observable and latent factors over short time periods during which the assets' exposures to the latent factors stay constant (our null hypothesis), however, then one can still harness the rich information in the high-frequency data for the assets' sensitivities towards systematic risk. This can be done using conventional inference techniques for linear asset pricing models, see e.g., Giglio and Xiu (2021) and Buchta (2024). Mainly, if our null hypothesis holds, then one can extract the latent risk factors via Principal Component Analysis (PCA) analysis. The potential changes in the assets' exposures to the latent risk factors will result in an augmented PCA factor space, with the extra factors being present only on parts of the time interval and accounting for these changes. One can then explore the pricing of the PCA factors and in turn relate their prices to those of the observable factors.

[1]This analysis has been further extended by Liao and Todorov (2024) to the latent factor case.

The key for applying the conventional methods for linear asset pricing is to know the time interval over which any changes in the linear relation between observable and latent factors occur only at the times when the assets' exposures to the latent risk factors change as well. This is the focus of the current paper.

Our test utilizes high-frequency observations of the observable factors and a large cross-section of asset prices over a time interval of fixed length. The null hypothesis can be characterised in terms of the residual of a global linear projection of the observable factors on the latent ones, extracted via PCA, over the interval of fixed length. Under the null hypothesis, these residuals should have zero quadratic covariation with the PCA factors over the entire time interval. That is, the orthogonality between the residuals and the latent factors should hold locally over any sub-interval. In technical terms, the residuals from the global linear projection should be a martingale that is orthogonal in a martingale sense to the latent risk factors. If this is not the case, these residuals from the global linear projection can have asset pricing implications that need to be accounted for.

To build a test on the basis of this characterization of the null hypothesis, we first estimate the latent risk factors via a standard PCA analysis. We then run a linear regression of the high-frequency returns of the observable factors on the PCA ones and collect the residuals from this regression. Our test statistic is a weighted sum of the product of the estimated residuals and the latent factors. By varying the weights over a class of weight functions, as done by Bierens and Ploberger (1997) in a different context, we can determine whether the orthogonality of the residuals and latent factors holds locally at any point in time. Our test, therefore, is based on the supremum of the test statistic over a suitable class of weight functions. The latter should converge to zero under the null hypothesis and diverge otherwise.

We derive the limiting distribution of the test statistic considered as a function of the weights used in its construction. The limiting distribution of the statistic is non-standard. To compute the critical values of the test, we develop a wild bootstrap approach based on multiplying each of the residuals from the linear regression of observed factors on latent ones by i.i.d. standard normal random variables and recomputing the test statistic in the bootstrapped data. We show that, with this way of computing the critical values, the test achieves asymptotically the desired size and has an asymptotic power of one.

We implement the developed test to study the temporal stability of a linear relation between the market volatility factor and the latent ones embedded in the cross-section of stock returns. Market volatility is well-known to be a source of systematic risk that investors care about. Typically, due to difficulties with measuring volatility, earlier work has assumed implicitly that there is a temporal stability in the linear relation between the volatility factor and the latent ones over periods of length at least one month. Our evidence, however, shows that such temporal stability exists on much shorter periods of length of one trading day only.

The current paper relates to several strands of work. First, it is connected to earlier work that considers estimation of assets' exposures towards observable factors from high-frequency data. Contributions in this area include Barndorff-Nielsen and Shephard (2004), Andersen et al. (2006), Mykland and Zhang (2006), Mykland and Zhang (2009), Todorov and Bollerslev (2010), Mancini and Gobbi (2012), Li et al. (2017), Aït-Sahalia et al. (2020) and Chen et al. (2024), among many others. Second, our paper is also related to earlier work on estimation of latent factors from high-frequency returns via PCA analysis, see e.g., Aït-Sahalia and Xiu (2017), Aït-Sahalia and Xiu (2019) and Pelger (2019, 2020). Third, similar to Zaffaroni (2019), our focus is on conditional asset pricing models without imposing structure on the time-variation in the factor loadings. Unlike Zaffaroni (2019), our focus is on the observable factors and their relation with the latent ones. Fourth, the current work, like Connor and Korajczyk (1991) and Giglio and Xiu (2021), considers projection of observable on PCA factors. The PCA setup in Giglio and Xiu (2021) is static, assuming that the linear relationship between observable and latent factors holds over the long time span and this has been recently extended to a dynamic setting by Buchta (2024). By contrast, the goal of this paper is to test whether such temporal stability of a linear relation between the factors exists. A related but different connection between observable factors is studied in Fan et al. (2021). Mainly, Fan et al. (2021) apply PCA on the expected returns of assets after projecting them on a set of observed factors. Finally, the relation between observable and latent risk factors in low-frequency setting is studied by Andreou et al. (2019) and Fortin et al. (2023). Unlike these studies, in our work, we allow for a gap between the risks spanned by the latent and the observable factors which is important as noted recently by Giglio and Xiu (2021).

The rest of the paper is organized as follows. Section 2 presents the setup and formulates the testing problem. We develop our test and characterize its asymptotic properties in Section 3. We evaluate the finite sample behavior of the test in Section 4. Our empirical application is in Section 5. Section 6 concludes. The assumptions and proofs are given in Section 7.

2 Setup and formulation of the testing problem

2.1 The model

On a standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the filtration $(\mathcal{F}_t)_{t\geq 0}$, we have the following continuous-time dynamic latent factor model for the $p \times 1$ vector of asset prices Y_t over the fixed time interval $[0, T]$:

$$dY_t = \alpha_t dt + \tilde{\beta}_t d\tilde{f}_t + dJ_t + d\epsilon_t, \tag{1}$$

where α_t is the drift term, $\tilde{\beta}_t$ is a matrix of factor loadings, $\tilde{f}_t$ is a vector of continuous (latent) systematic factors, J_t is the jump part of Y_t and ϵ_t captures the idiosyncratic risk. α_t, J_t and ϵ_t are $p \times 1$ vectors, $\tilde{f}_t$ is $\tilde{K} \times 1$ vector and $\tilde{\beta}_t$ is $p \times \tilde{K}$ matrix, for some integer $\tilde{K}$.

The above continuous-time factor model is rather general. Typically, the time variation in the factor loadings in existing work is modeled as a function of either economy-wide (macro) state variables or firm characteristics, see e.g., Connor et al. (2012), Fan et al. (2016) and Kelly et al. (2019).[2]

An alternative to this approach, which avoids modeling the time variation in $\tilde{\beta}_t$, is to assume that $\tilde{\beta}_t$ is piece-wise constant over the interval $[0, T]$. We will adopt this approach here and we will not need to make any assumption about the number of changes in $\tilde{\beta}_t$ on the interval or about their location. A stronger version of this assumption is either imposed explicitly or implicitly in existing work whenever the latent factors are extracted via PCA.

With a piece-wise assumption for $\tilde{\beta}_t$, there is an equivalent static factor model representation for Y_t as we show now. To see this, suppose that $\tilde{\beta}_t$ has the following representation

$$\tilde{\beta}_t = \tilde{\beta}_0 + \sum_{i=1}^{m} (\tilde{\beta}_{t_i} - \tilde{\beta}_{t_{i-1}}) 1(t < t_i), \tag{2}$$

where for some non-negative integer m, we denote $t_0 = 0 < t_1 < ... < t_{m+1} = T$. The case $m = 0$ corresponds to the classical static factor case commonly assumed in applications. For $m > 0$, we allow the factor loadings to change discretely over the interval $[0, T]$. The first component $\tilde{\beta}_0$ is the baseline beta. The second component $\sum_{i=1}^{m} (\tilde{\beta}_{t_i} - \tilde{\beta}_{t_{i-1}}) 1(t < t_i)$ identifies m number of "significant jumps" in the beta,

[2] This approach of modeling factor exposures is also commonly done when factors are observed, see e.g., Shanken (1990), Jagannathan and Wang (1996), Ferson and Harvey (1999), Gagliardini et al. (2016) and Gagliardini et al. (2020), among others.

occurring at times $t_1, ..., t_m$.[3] They lead to static factors and betas as elements in the following matrices:

$$\beta = \left(\ \tilde{\beta}_0, \tilde{\beta}_{t_1} - \tilde{\beta}_{t_0}, \cdots, \tilde{\beta}_{t_m} - \tilde{\beta}_{t_{m-1}} \ \right), \quad f_t = \begin{pmatrix} \tilde{f}_t \\ \tilde{f}_t 1(t \in (t_1, t_2]) \\ \vdots \\ \tilde{f}_t 1(t \in (t_{m-1}, t_m]) \end{pmatrix}, \tag{3}$$

with β and f_t having dimensions $p \times \tilde{K}(m+1)$ and $\tilde{K}(m+1) \times 1$, respectively. By expanding the factor space, we can represent the dynamic factor model via the following static model:

$$\tilde{\beta}_t d\tilde{f}_t = \beta df_t, \ t \in [0, T].$$

Note that f_t capture the static factor components, which have a special structure. Mainly, apart from the first one, they are active only on parts of the time interval $[0, T]$. This, and the fact that f_t exhibits jumps at the times $\{t_i\}_{i=1}^m$, does not matter for our inference procedures. For this reason, from now on, we will work with the static factor model representation, given by:

$$dY_t = \alpha_t dt + \beta df_t + dJ_t + d\epsilon_t, \tag{4}$$

where β is a $p \times K$ matrix and f_t is a $K \times 1$ vector.

Our interest in the paper is one-dimensional observable factor g_t and its relation with the latent factors f_t. Towards this end, we can decompose the continuous part of g_t, denoted by g_t^c, as

$$g_t^c = \gamma_{t,0} + \int_0^t \gamma_s^\top df_s + \epsilon_t^g, \tag{5}$$

for some continuous finite variation process $\gamma_{t,0}$, some process γ_t, and a continuous martingale ϵ_t^g with $\langle f, \epsilon^g \rangle_t = \mathbf{0}$, for $t \in [0, T]$, where we denote the quadratic covariation between two arbitrary stochastic processes x and y as

$$\langle x, y \rangle_t = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \sum_{i=1}^{\lfloor t/\Delta \rfloor} (x_{i\Delta} - x_{(i-1)\Delta})(y_{i\Delta} - y_{(i-1)\Delta}). \tag{6}$$

The finite variation process $\gamma_{t,0}$ is the continuous-time analogue of the conditional mean of g_t in discrete settings. Our interest in this paper is in the martingale component of g_t, which contains the risks in g_t.

[3]In addition to these significant jumps in the factor loadings, our setup can also accommodate "small" jumps which translate into weak factors in the static factor representation. For simplicity of exposition, we do not consider such generalization of our setup.

We can view $\int_0^t \gamma_s^\top df_s$ in the above decomposition as the local linear projection of the martingale component of g_t on f_t, see Mykland and Zhang (2006). We can give an explicit expression for γ_t. For this, we denote

$$c_t^{ff} = \frac{d\langle f, f\rangle_t}{dt} \quad \text{and} \quad c_t^{fg} = \frac{d\langle f, g\rangle_t}{dt}, \tag{7}$$

which are the spot variance of f and the spot covariance between f and g. Without loss of generality we can assume that c_t^{ff} is diagonal, i.e., that the factors in the vector f_t are orthogonal in a martingale sense. With this notation, γ_t is given by

$$\gamma_t = (c_t^{ff})^{-1} c_t^{fg}, \tag{8}$$

where in the above, the inverse should be interpreted as a generalized inverse. This can cover the situation in which at certain points in time some of the factors in f_t are dormant.

2.2 Formulation of the testing problem

Given the martingale orthogonality condition $\langle f, \epsilon^g\rangle_t = \mathbf{0}$, we have

$$\mathbb{E}_t(f_{t+\Delta}\epsilon^g_{t+\Delta}) = \mathbf{0}, \quad \text{for } \forall\Delta > 0,$$

where as usual $\mathbb{E}_t$ denotes the $\mathcal{F}_t$-conditional expectation. If the latent factor f_t captures the priced risks in the cross-section of stocks as in the classical Arbitrage Pricing Theory (APT), see e.g., Ross (1976) and Chamberlain and Rothschild (1983) as well as the conditional version of it (Hansen and Richard (1987)), this means that $\int_0^t \gamma_s^\top df_s$ captures the part of g_t that is relevant for asset pricing and ϵ_t^g the part which is not.

The nonparametric inference for the projection coefficients γ_t is very difficult in general and practically impossible. Without any assumptions for the time variation in γ_t, one needs to compute this quantity over a short time window by "pretending" that it stays constant over that window. The shorter the window, the noisier the estimation. Thus, to reduce the noise in the inference, one would like to use a time window which is as long as possible. In fact, in linear asset pricing models, which are commonly used in practice, see e.g., Cochrane (2005), the projection coefficients remain constant, at least over short time intervals whenever the factors are present on the interval (i.e., whenever their quadratic variation is different from zero). Motivated by this, our aim in this paper is to test this temporal stability of the projection coefficients γ_t on the interval $[0, T]$.

To this end, we propose a novel formulation for the constant-γ hypothesis, which provides an easy way of testing for it. Consider the global linear projection of g_t on f_t over the interval $[0,T]$:

$$g_t^c = \gamma_{t,0} + \gamma^\top f_t + u_t, \tag{9}$$

for the same $\gamma_{t,0}$ as above, some random variable γ, and a continuous process u_t satisfying $\langle f, u\rangle_T = \mathbf{0}$. The random variable γ (adapted to $\mathcal{F}_T$) is the counterpart of the (global) linear projection coefficient in classical regression settings and is given by

$$\gamma = \langle f, f\rangle_T^{-1}\langle f, g\rangle_T \equiv \left(\int_0^T c_t^{ff} dt\right)^{-1} \int_0^T c_t^{fg} dt. \tag{10}$$

Using the expression for the local projection coefficients, we can express equivalently γ as a weighted average of these local coefficients:

$$\gamma = \int_0^T \omega_t \gamma_t dt, \quad \text{where } \omega_t = \left(\int_0^T c_t^{ff} dt\right)^{-1} c_t^{ff}.$$

In general, the error from the global projection u_t does not need to be a martingale and $\langle f, u\rangle_t$ does not need to equal $\mathbf{0}$, for $\forall t < T$. On one hand, as we discussed above, from an asset pricing perspective, the part of f_t of interest is $\int_0^t \gamma_s^\top df_s$ and not $\gamma^\top f_t$. This is because u_t and f_t can be locally correlated, and this can matter for pricing, i.e., part of the risk in u_t can be priced in the cross-section of stock returns.

On the other hand, γ is much easier to measure from the data than γ_t as the former can use the data from the whole interval $[0,T]$ while the latter can use only a small fraction of it, mainly the observations in a local neighborhood of t. To illustrate, an estimator of γ based on time series data in $[0,T]$ has an asymptotic variance that is proportional to T in a baseline homoskedastic setting.

Motivated by the above discussion, our null hypothesis is that u_t is a martingale that is locally uncorrelated with the latent factors f_t over the whole time interval $[0,T]$. That is,

$$\mathbb{H}_0 : \langle f, u\rangle_t = \mathbf{0}, \quad \text{for } \forall t \in [0,T]. \tag{11}$$

We will use this formulation of the null hypothesis to derive our test in the next section. To see the connection with the time variation in γ_t, note that we can write

$$\langle f, u\rangle_t = \int_0^t (\gamma_s - \gamma)^\top d\langle f, f\rangle_s.$$

Therefore, if γ_t varies over some part of $[0,T]$, then the process $\langle f, u\rangle_t$ will deviate from zero at some point in $[0,T]$.

2.3 An illustrative example

We now illustrate the various concepts introduced above and highlight the importance of the testing problem for asset pricing in a simplified setting. Suppose the only factor driving asset returns is the market portfolio. In our notation, see equation (1) above, that means that $\widetilde{f}_t$ is the value of the market portfolio at time t. Suppose further that the stochastic discount factor is driven by the market portfolio, i.e., that the conditional CAPM holds. We assume that at time $\tau_f \in [0, T]$, assets' exposures towards the market factor (denoted with $\widetilde{\beta}_t$) change:

$$\widetilde{\beta}_t = \begin{cases} \beta_b, & t \leq \tau_f \\ \beta_a, & t \in (\tau_f, T], \end{cases}$$

for some vectors β_b and β_a. In this case, the equivalent static factor model representation in equation (1) holds with

$$f_t = \begin{pmatrix} \widetilde{f}_t \\ \widetilde{f}_t 1(t \in (\tau_f, T]) \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_b \\ \beta_a \end{pmatrix},$$

and we can write

$$\widetilde{\beta}_t d\widetilde{f}_t = \beta df_t, \quad t \in [0, T].$$

That is, we have two static factors, over the interval $[0, T]$, which account for the time variation in the assets' exposures towards the market portfolio. Note that the two static factors are perfectly correlated on $[0, \tau_f]$.

We are interested in the pricing of an observable factor g_t. For example, in our empirical application g_t will be the level of market volatility at time t. Since the conditional CAPM holds in our example, g_t does not matter for asset pricing once the dynamic exposure towards market risk is properly accounted for.

Suppose the relation between g_t and $\widetilde{f}_t$ is given by

$$g_t = \widetilde{\gamma}_t \widetilde{f}_t + \epsilon_t^g, \ \widetilde{\gamma}_t = \begin{cases} \gamma_b, \ \text{if } t \leq \tau_g \\ \gamma_a, \ \text{if } t > \tau_g \end{cases}, \ \tau_g \in (0, T], \tag{12}$$

where ϵ_t^g is a martingale that is orthogonal in a martingale sense to $\widetilde{f}_t$. That is, at time τ_g, we can have a change in the linear relation between the volatility factor and the market factor. The times τ_f and τ_g do not need to coincide. We can equivalently represent g_t in terms of f_t:

$$\begin{aligned} g_t = & \gamma_b f_{t,1} + (\gamma_a - \gamma_b) f_{t,2} + \epsilon_t^g \\ & + (\gamma_a - \gamma_b) f_{t,1} 1_{\{t \in (\tau_g, \tau_f]\}} + (\gamma_b - \gamma_a) f_{t,1} 1_{\{t \in (\tau_f, \tau_g]\}}, \end{aligned} \tag{13}$$

with $f_{t,1}$ and $f_{t,2}$ being the first and second element, respectively, of f_t.

Given this representation of g_t, the null hypothesis that we test corresponds to the situation $\tau_f = \tau_g$ and the alternative hypothesis to $\tau_f \neq \tau_g$. If the null hypothesis is true, then one can study the priced risk in g_t using standard methods for static linear asset pricing models.[4] This is in spite of the fact that the linear relation between the market factor and the volatility factor changes at some point in the time interval. For this to happen, however, we need that the exposure towards the market factor to change at exactly the same time. If this is not the case (our alternative hypothesis), then static linear asset pricing methods will not lead to a valid inference for the priced risk in g_t. This is because the static risk factors driving the cross-section of stock returns are not sufficient to capture all the priced risk in g_t. Indeed, they will not account for the risk in g_t given in the second line of equation (13).

We can extend this example to the more general situation when there is more than one change in the assets' exposures to the latent risk factors. Our null hypothesis then does not test against any time-variation in the exposure of the observable factor towards the latent dynamic factors. Indeed, we allow under the null hypothesis for $\widetilde{\gamma}_t$ given in (12) to change. However, this should happen only at the times $t_1, t_2, ..., t_m$ when the assets' exposures towards the latent dynamic factors change. That is, we are testing for *economically significant* changes in the relation between observable and latent dynamic factors as far as the pricing of the observable factor is concerned.

3 The test

We now develop the test for the null hypothesis and analyze its asymptotic properties. We start with introducing the observation scheme and constructing the test statistic.

3.1 Construction of the test statistic

Asset prices and the risk factor g_t are observed on the equidistant grid $0, \frac{T}{n}, ..., T$, for T fixed and $n \to \infty$. We use $\Delta_n = \frac{T}{n}$ for the mesh of the observation grid. The

[4]Note, however, that if $\tau_f = \tau_g < T$, both $\widetilde{f}_t$ and g_t will matter for asset pricing when using techniques for static linear asset pricing models with *observable* factors. This is because we need two static factors to account properly for the priced risk in the cross-section of asset returns (f_t in this example is of dimension 2). Due to that, as a static factor, g_t will have additional pricing ability over what a static market factor can do while this is not the case in a dynamic setting.

increments of a generic process G are denoted with

$$\Delta_i^n G = G_{i\Delta_n} - G_{(i-1)\Delta_n}.$$

To remove the jumps from the asset prices, we will do trimming of the increments following Mancini (2001). More specifically, denote $h(x, a) = (x \vee -a) \wedge a$, for some $x \in \mathbb{R}$ and $a \in \mathbb{R}_+$. Then, the scaled and trimmed price increments are denoted with

$$\overline{Y}_{tj} = \frac{h(\Delta_t^n Y, \gamma_j \Delta_n^{\varpi})}{\sqrt{\Delta_n}}, \quad \overline{Y} = \left(\overline{Y}_{tj} : t = 1, ..., n, j = 1, ..., p\right), \tag{14}$$

for some $\gamma_j > 0$ that is uniformly bounded in j and a constant $\varpi \in (0, 1/2)$.

We introduce similar notation for the scaled increments of the factors and of the idiosyncratic risk in the asset prices:

$$\overline{f}_t = \frac{\Delta_t^n f}{\sqrt{\Delta_n}}, \quad \overline{g}_t = \frac{\Delta_t^n g}{\sqrt{\Delta_n}}, \quad \overline{\epsilon}_{tj} = \frac{\Delta_t^n \epsilon}{\sqrt{\Delta_n}}, \; t = 1, .., n, \; j = 1, ., p, \tag{15}$$

and

$$\overline{F} = \left(\overline{f}_t : t = 1, .., n\right), \overline{G} = \left(\overline{g}_t : t = 1, .., n\right), \overline{\mathcal{E}} = \left(\overline{\epsilon}_{tj} : t = 1, .., n, j = 1, .., p\right). \tag{16}$$

In matrix form, the observed data can be compactly written as

$$\overline{Y} = \beta \overline{F}^{\top} + \overline{\mathcal{E}} + R, \tag{17}$$

for some "residual" matrix R capturing the error from trimming and the drifts in the prices.

Using the discrete observations and PCA analysis, we form estimates for the latent factors up to a rotation matrix. We denote these estimates with $\hat{f}_t$. We then run a standard OLS regression of g_t on $\hat{f}_t$ and denote the estimated residuals with $\hat{u}_t$.

One can show (see Theorem 3.2 below) that, under our assumptions, we have the following convergence in probability

$$\frac{1}{n} \sum_{t=1}^{n} \hat{f}_t \hat{u}_t w(t/n) \xrightarrow{\mathbb{P}} \int_0^T w(t/T) d\langle f, u \rangle_t, \tag{18}$$

for some deterministic weight function $w : [0, 1] \to \mathbb{R}$. Obviously, if the null hypothesis is true and $\langle f, u \rangle_t$ is identically zero on $[0, T]$, the limit in the above will be zero as well. However, this will not be in general the case when the null does not hold.

Our test consists of checking whether the limit in the above is statistically different from zero for a whole class of weight functions. In particular, we introduce the following family of weight functions:

$$\mathcal{W}_\eta(t) = \cos(\eta t) + \sin(\eta t), \quad \text{for } \eta \in [0, 2\pi] \text{ and } t \in [0, 1]. \tag{19}$$

Using Theorem 1 of Bierens and Ploberger (1997), we have that for a cadlag function $f : [0, T] \to \mathbb{R}$, $f(t) = 0$ almost everywhere on $[0, T]$ if and only if $\int_0^T \mathcal{W}_\eta(t/T) f(t) dt = 0$ for all $\eta \in [0, 2\pi]$.

This suggests the following test statistic for our null hypothesis:

$$\widehat{\mathbb{S}} = \sum_{k=1}^{K} \sup_{\eta_k \in [0,2\pi]} \left| \widehat{\mathbb{S}}_k(\eta_k) \right|^2, \tag{20}$$

where

$$\widehat{\mathbb{S}}_k(\eta) = \frac{1}{n} \sum_{t=1}^{n} \hat{f}_{t,k} \hat{u}_t \mathcal{W}_\eta(t/n), \quad k = 1, ..., K. \tag{21}$$

Under the null hypothesis, $\widehat{\mathbb{S}}$ converges in probability to zero and to a strictly positive number under the alternative hypothesis.

3.2 Asymptotic properties of the test statistic

We start with deriving a Central Limit Theorem (CLT) for our statistic under the null hypothesis. In the following theorem, $\xrightarrow{\mathcal{L}-s}$ denotes stable convergence stable in law, see e.g., Section VIII.5.c of Jacod and Shiryaev (2003).

Our CLT is for the vector

$$\widehat{\mathbb{S}}(\bar{\eta}) = \left(\widehat{\mathbb{S}}_k(\eta_k) : k = 1, ..., K \right), \quad \bar{\eta} = (\eta_1, ..., \eta_K).$$

Theorem 3.1. *Suppose assumptions A1-A3 hold. Let $n, p \to \infty$ with*

$$n^{1/2}/p \to 0, \quad p/n^{2\widetilde{\varpi}+1} \to 0, \quad \textit{and} \quad \varpi > 1/4, \tag{22}$$

for some $\widetilde{\varpi} < \varpi$. Under $\mathbb{H}_0$ in (11):

$$\sqrt{n}\, \widehat{\mathbb{S}}(\bar{\eta}) \xrightarrow{\mathcal{L}-s} \mathbb{Z}_K(\bar{\eta}), \tag{23}$$

uniformly in $\bar{\eta} \in [0, 2\pi]^K$, and where $\mathbb{Z}_K(\bar{\eta})$ is a K-dimensional vector-valued process, defined on an extension of the original probability space, which is $\mathcal{F}$-conditionally a zero-mean Gaussian process, with $\mathcal{F}$-conditional covariance kernel given in the appendix.

We make several comments about the above limit result. First, we need a restriction on the growth of the size of the cross-section of assets relative to the sampling frequency. This condition is given in (22) and is relatively weak. In particular, the empirically relevant case in which p and n grow at the same rate is allowed for. Second, given the in-fill asymptotic setup, the limit is mixed-Gaussian process with $\mathcal{F}$-conditional covariance kernel that is in general random. Third, the rate of convergence depends only on the sampling frequency and is determined by the martingale component of the observable factor that is orthogonal to the latent factors (in a martingale sense). One can show that the error in recovering the latent factors from the cross-section of assets is of higher asymptotic order and does not impact the limiting distribution in (23).

We next present a convergence in probability result for the process $\widehat{\mathbb{S}}(\bar{\eta})$.

Theorem 3.2. *Suppose assumptions A1-A3 hold, let $n, p \to \infty$, and assume that the condition in (22) holds. We have*

$$\widehat{\mathbb{S}}(\bar{\eta}) \overset{\mathbb{P}}{\to} \frac{1}{T}\int_0^T diag\,(w_{\eta_1}(t/T), ..., w_{\eta_K}(t/T))\bar{H}'\,d\langle f, u\rangle_t, \tag{24}$$

uniformly in $\bar{\eta} \in [0, 2\pi]^K$, and where $\bar{H}$ is some random $K \times K$ invertible matrix.

The limit of $\widehat{\mathbb{S}}(\bar{\eta})$ depends on a random invertible matrix $\bar{H}$. This is because the latent factors can be recovered only up to a rotation matrix. This, however, will not impact our testing problem as our interest is only whether the limit in (24) is different from zero for any η.

3.3 Bootstrap critical values

For constructing a test on the basis of the limit result in Theorem 3.1, we need feasible estimates of the covariance kernel. Instead of constructing such estimates, we will develop a bootstrap approach which is straightforward to implement. In the high-frequency setting, the factors typically exhibit nontrivial variation in their volatility. To account for that, therefore, we employ a wild bootstrap as in Mammen (1993). Specifically,

Step 1. Generate η_t^*, $t = 1, ..., n$ as i.i.d. $N(0, 1)$ random variables, and define

$$g_t^* = \widehat{\gamma}_0 + \widehat{\gamma}^\top \widehat{f}_t + \eta_t^* \widehat{u}_t$$

Step 2. Let $\widehat{\gamma}_1^*, ..., \widehat{\gamma}_K^*$ denote the estimated slope coefficients when regressing g_t^* on $\widehat{f}_t$. Let $\widehat{u}_t^*$ denote the corresponding time series residuals.

Step 3. Let

$$\widehat{\mathbb{S}}^* = \sum_{k=1}^{K} \sup_{\eta_k \in [0,2\pi]} \left| \widehat{\mathbb{S}}_k^*(\eta_k) \right|^2, \tag{25}$$

where

$$\widehat{\mathbb{S}}_k^*(\eta) = \frac{1}{n} \sum_{t=1}^{n} \widehat{f}_{t,k} \widehat{u}_t^* \mathcal{W}_\eta(t/T), \quad k = 1, ..., K. \tag{26}$$

Let

$$\widehat{\mathbb{S}}^*(\eta) = \left(\widehat{\mathbb{S}}_k^*(\eta) : k = 1, ..., K \right).$$

The asymptotic behavior of the bootstrap statistic is given in the following theorem. In it, we denote $\mathcal{F}$-conditional convergence in law with $\xrightarrow{\mathcal{L}|\mathcal{F}}$, see Section VIII.5b of Jacod and Shiryaev (2003).

Theorem 3.3. *Suppose assumptions A1-A3 hold, let $n, p \to \infty$ and assume that the condition in (22) holds. We have*

$$\sqrt{n}\, \widehat{\mathbb{S}}^*(\eta) \ \xrightarrow{\mathcal{L}|\mathcal{F}} \ \mathbb{Z}_K(\eta), \tag{27}$$

uniformly in $\eta \in [0, 2\pi]^K$ and where $\mathbb{Z}_K(\eta)$ is a process defined on an extension of the original probability space and which has the same $\mathcal{F}$-conditional law as the limiting process in Theorem 3.1.

As seen from the result of the above theorem, the bootstrap statistic has the same $\mathcal{F}$-conditional limiting distribution as our test statistic. This allows for an easy way of constructing test of given asymptotic size as we show next.

3.4 Asymptotic size and power of the test

For constructing the critical region of the test with asymptotic size of $\alpha \in (0, 1)$, we repeat steps 1-3 of the previous section B times and obtain $\widehat{\mathbb{S}}_1^*, ..., \widehat{\mathbb{S}}_B^*$. We then compute the bootstrap p-value

$$p^* = \frac{1}{B} \sum_{b=1}^{B} 1\{\widehat{\mathbb{S}}_b^* + \epsilon\, \zeta_b^2/n > \widehat{\mathbb{S}} + \epsilon\, \zeta^2/n\}, \tag{28}$$

where $\{\zeta_b\}_{b=1,...,B}$ and ζ are standard normal variables, defined on an extension of the probability space, that are independent of $\mathcal{F}$ and of each other, and $\epsilon > 0$ is a constant. The reason for adding a scaled and squared standard normal variable to our statistic is to ensure that the limiting distribution of the statistic has a probability density. It appears non-trivial to show that this is the case for that of $\widehat{\mathbb{S}}$. By adding an independent random variable to the test statistic, with a smooth probability density, we can ensure that this is the case for the limiting distribution of $\widehat{\mathbb{S}} + \epsilon\,\zeta^2/n$ under the null hypothesis. In order to minimize the impact from this on the power of the test, one should pick ϵ very small as we do in our empirical application.

The critical region $\mathcal{C}_{n,\alpha}$ of the test is then given by $\mathcal{C}_{n,\alpha} = \{p^* < \alpha\}$. We consider the following alternative:

$$\mathbb{H}_1 : \langle f, u \rangle_t = \mathbf{c}, \quad \text{for some } t \in [0,T] \text{ and } \mathbf{c} \neq 0.$$

Theorem 3.4. *Suppose assumptions A1-A4 hold, let $n, p, B \to \infty$ and assume that the condition in (22) holds. We have*

$$\mathbb{P}\left(\mathcal{C}_{n,\alpha} | \mathbb{H}_0\right) \to \alpha, \ \mathbb{P}\left(\mathcal{C}_{n,\alpha} | \mathbb{H}_1\right) \to 1. \tag{29}$$

4 Monte Carlo study

4.1 Setting

The baseline model for our Monte Carlo study is the one used in Liao and Todorov (2024):

$$dY_{t,j} = \sigma_t \left(\beta_j^\top \Sigma_f^{1/2} dW_t + \overline{\sigma}_j dW_{t,j} \right), \quad j = 1, ..., p, \tag{30}$$

where the univariate stochastic volatility process σ_t has the following square-root diffusion dynamics

$$d\sigma_t^2 = 8.3(1 - \sigma_t^2)dt + \sigma_t dB_t, \tag{31}$$

with B_t being a standard Brownian motion that is independent from W_t and $\{W_{t,j}\}_{j \geq 1}$.

We generate the *baseline* factor and betas from $\beta_j^0 \sim N(\mu_\beta, \Sigma_\beta)$ and $f_t^0 \sim N(0, \Sigma_f)$, for $j = 1, ..., p$ and $t \in [0,T]$, and where

$$\Sigma_f = \text{diag}(1.198, 0.377, 0.264) \times 10^{-6}, \tag{32}$$

$$\mu_\beta = \begin{pmatrix} 0.8763 \\ 0.2818 \\ -0.2965 \end{pmatrix}, \ \Sigma_\beta = \begin{pmatrix} 0.2326 & -0.2474 & 0.2603 \\ -0.2474 & 0.9224 & 0.0837 \\ 0.2603 & 0.0837 & 0.9139 \end{pmatrix}. \tag{33}$$

We consider three settings to generate the true factors and betas from the above baseline model:

Setting A: Set the true betas and factors as the baseline:

$$\beta_i = \beta_i^0, \quad f_t = f_t^0, \quad i = 1, ...p, \quad t \in [0, T].$$

Setting B: The third factor is not active on the interval $[0, T_0)$, for some $0 < T_0 < T$:

$$\beta_i = \beta_i^0, \quad f_{t,3} = \begin{cases} 0, & t \in [0, T_0) \\ f_{t,3}^0, & t \in [T_0, T]. \end{cases}$$

Setting C: The loadings for the third factor have a structural break:

$$\beta_{i,j} = \beta_{i,j}^0, \quad j = 1, 2, \quad \beta_{i,t,3} = \begin{cases} \beta_{i,3}^0, & t \in [0, T/2) \\ \beta_{i,3}^1 & t \in [T/2, T]. \end{cases}, \quad f_t = f_t^0,$$

where

$$\beta_{i,3}^1 = \beta_{i,3}^0 + \tau \times z_i, \quad z_i \sim N(0, 0.9139),$$

and τ is a parameter that determines the amount of change in the third beta.

Finally, as in Liao and Todorov (2024), the scale of the idiosyncrartic variance $\overline{\sigma}_j$ is cross-sectionally i.i.d. and is drawn according to

$$\overline{\sigma}_j \sim \text{Uniform}([0.5, 1.5]) \times 1.1 \times 10^{-3}, \quad j = 1, ..., p. \tag{34}$$

With such a choice for $\overline{\sigma}_j$, the share of idiosyncratic risk in total asset variance is around 40% for the median stock in the cross-section.

We turn next to the specification of the observable factor, which we model as

$$g_t = f_t' \gamma_t + u_t, \quad u_t \sim N(0, \sigma_g^2), \quad \sigma_g = 0.7 \times 10^{-5}.$$

The above value of σ_g is chosen so that the share of u_t in the variance of g_t is around 40%. γ_t is either constant (null hypothesis) or varies (alternative hypothesis). More specifically, we set

null: γ_t stays constant: $\gamma_t = \mu_\beta$.

alternative: γ_t shifts by one at some point in the observation interval:

$$\gamma_t = \begin{cases} \mu_\beta, & t \in [0, T - T_1), \\ \mu_\beta + 1, & t \in [T - T_1, T], \end{cases}$$

for some $0 < T_1 < T$ and where μ_β is the mean of the latent factor loadings.

Our simulation setup, presented above, is rather general. In particular, we consider cases in which some of the latent factors are "active" over parts of the estimation window only or when the exposure to them changes discretely. As explained in Section 2, these covered are all covered by our analysis.

Finally, for the observation scheme, we use a cross-section of $p = 500$ assets and assume that we sample asset prices either 80 or 40 times during a trading day. This corresponds approximately to sampling every five or ten minutes in a 6.5 hour trading day. In all considered cases in the Monte Carlo, the time window is fixed to ten trading days.

4.2 Results

The tuning parameters of the test are set as follows. First, the truncation parameter is set in the following data-driven way:

$$\varpi = 0.49 \quad \text{and} \quad \gamma_j = 4 \times \sqrt{RV_{d,j}}, \ j = 1, .., p, \tag{35}$$

where $RV_{d,j}$ is the realized variance of asset j on the (trading) day d the increment belongs to, given by

$$RV_{d,j} = \sum_{i=\lfloor (d-1)/(252\Delta_n) \rfloor + 1}^{\lfloor d/(252\Delta_n) \rfloor} (\Delta_i^n X_j)^2. \tag{36}$$

This way of setting truncation allows it to adapt to the level of volatility of the asset.

Next, for determining the number of latent factors, we follow Liao and Todorov (2024) and standardize the asset returns by estimates of their volatility in order to minimize the impact of idiosyncratic volatility. Then, the number of factors is estimated using the information criterion proposed by Bai and Ng (2002), where we set $g_{n,p} = \log\left(\frac{np}{n+p}\right)$ in the penalty term, see Liao and Todorov (2024).

Finally, the scaling constant ϵ (recall our definition of p^* in equation (28)) is set 1/1000 times the sample variance of the observable factor.

The Monte Carlo results are reported in Table 1. We can see from them that the test has overall good finite sample size properties. There is only a slight under-rejection in setting A for the coarser sampling frequency of ten minutes and window size of one day only. Note, in particular, that the size of the test is close to the nominal level in settings B and C when either a factor is present only for part of the time or when factor loadings change. This shows the robustness of the test to such features in the data.

Table 1 reveals also good power properties of the test in general. Not surprisingly, the power increases when either the length of the time window or the sampling frequency increases. Naturally, the harder to identify alternative from the data is the one in which the change in the relationship between observable and latent factors is present for a smaller fraction of the total sample.

Overall, Table 1 shows good finite sample properties of the developed test.

5 Market volatility risk and the cross-section of stock returns

We use the techniques developed in the paper to study the dynamic relation between latent risk factors embedded in the cross-section of stock returns and a market volatility factor. Time-varying volatility is an ubiquitous feature of asset prices, see e.g., the seminal work of Engle (1982) and Bollerslev (1986). Volatility risk can have effect on the real economy (Bloom (2009) and Fernández-Villaverde et al. (2011)) and it is a source of risk that investors care about (Bollerslev et al. (2009)).

The impact of volatility risk on the cross-section of stock returns has been studied by Ang et al. (2006) and Herskovic et al. (2016), among others. Due to the fact that volatility is latent, these studies use realized volatilities computed over months as nonparametric proxies for the unobserved volatility over these periods. Realized volatility, however, is a rather noisy measure of true volatility and importantly it is also affected by jumps, which generate risk with very different behavior from the one due to volatility. Indeed, jumps, unlike volatility, do not persist typically over time.

For this reason, we will use in our analysis a measure of market spot volatility that is constructed from options on the S&P 500 index by the CBOE options exchange that is less noisy than the realized volatility and is robust to jumps. The ticker for this volatility measure is SPOTVOL and its construction is based on the theoretical analysis of Todorov (2019). The series is available at intraday frequency and this will allow us to study the dynamic interaction between market volatility factor and the latent systematic risk in stock prices over relatively short intervals of time. We plot the spot volatility index over 2015-2021 (our sample period) in Figure 1. As seen from the figure, our sample covers both periods of very high and very low volatility. The figure also reveals both high and low frequency variation in volatility.

The cross-section of stocks we use in our analysis changes on a yearly basis and comprises of the 500 largest stocks by market capitalization determined at the end of the previous calendar year. To alleviate potential concerns about the impact of market microstructure noise, we sample asset prices at the relatively coarse frequency of ten minutes during each trading day. In addition to the cross-section of individual stocks, we will also use the SPY ETF on the S&P 500 index as our proxy for the

Table 1: Rejection Probabilities in Monte Carlo.

		Settings				
		A	B		C	
Frequency	Hypothesis		$T_0 = \frac{T}{4}$	$T_0 = \frac{T}{2}$	$\tau = 1$	$\tau = 0.1$
Panel A: T = 1 day						
10 minutes	Null	0.022	0.044	0.056	0.048	0.038
	Alter. $T_1 = T/10$	0.124	0.150	0.156	0.136	0.142
	Alter. $T_1 = 3T/10$	0.720	0.716	0.696	0.662	0.722
5 minutes	Null	0.054	0.040	0.038	0.056	0.054
	Alter. $T_1 = T/10$	0.164	0.196	0.226	0.176	0.196
	Alter. $T_1 = 3T/10$	0.844	0.880	0.844	0.864	0.876
Panel B: T = 5 days						
10 minutes	Null	0.050	0.054	0.072	0.052	0.046
	Alter. $T_1 = T/10$	0.696	0.684	0.686	0.720	0.712
	Alter. $T_1 = 3T/10$	1.000	1.000	1.000	1.000	1.000
5 minutes	Null	0.040	0.048	0.058	0.058	0.068
	Alter. $T_1 = T/10$	0.836	0.862	0.826	0.856	0.866
	Alter. $T_1 = 3T/10$	1.000	1.000	1.000	1.000	1.000
Panel C: T = 10 days						
10 minutes	Null	0.044	0.052	0.064	0.064	0.030
	Alter. $T_1 = T/10$	0.958	0.964	0.954	0.956	0.968
	Alter. $T_1 = 3T/10$	1.000	1.000	1.000	1.000	1.000
5 minutes	Null	0.038	0.044	0.050	0.046	0.048
	Alter. $T_1 = T/10$	0.994	0.986	0.986	0.990	0.996
	Alter. $T_1 = 3T/10$	1.000	1.000	1.000	1.000	1.000
Panel D: T = 20 days						
10 minutes	Null	0.046	0.044	0.054	0.066	0.040
	Alter. $T_1 = T/10$	1.000	0.998	1.000	1.000	1.000
	Alter. $T_1 = 3T/10$	1.000	1.000	1.000	1.000	1.000
5 minutes	Null	0.058	0.054	0.066	0.046	0.040
	Alter. $T_1 = T/10$	1.000	1.000	1.000	1.000	1.000
	Alter. $T_1 = 3T/10$	1.000	1.000	1.000	1.000	1.000

The results are based on 500 Monte Carlo replications and number of bootstrap replications of $\mathcal{B} = 1,000$. The time window is 10 days. 5 and 10 minute sampling frequencies correspond to 80 and 40 returns per day, respectively. The nominal size of the test is 5%.

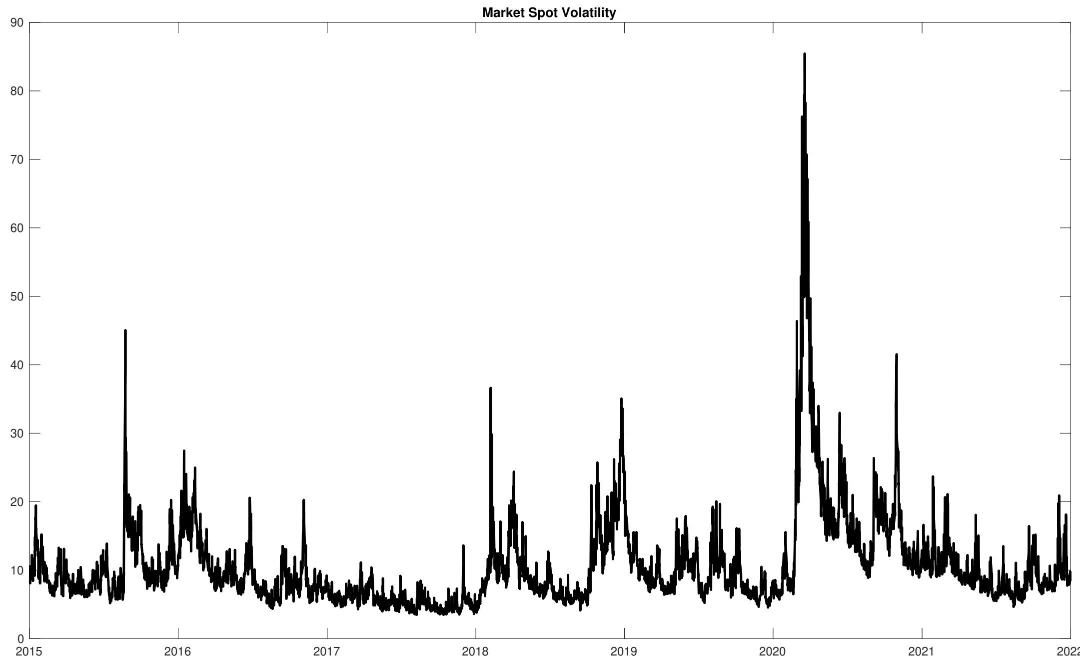

Figure 1: Market Spot Volatility.

market portfolio.

We implement our test for temporal stability of the relation between the volatility factor and the latent ones by setting the tuning parameters exactly as in the Monte Carlo study. In Table 2, we report the results from performing the test for different levels of time aggregation, ranging from 1 day to 40 trading days. As seen from the reported results, there is nontrivial rejection of the null hypothesis of a constant linear relation between the volatility factor and the latent risk factors for long periods of aggregation. Indeed, the empirical rejection rate gets close to the nominal size of the test only when the time window is of length one day. This means that for asset pricing applications, in order to capture fully the risks in the volatility factor, one has to estimate sensitivity towards volatility risk over much shorter periods than what has been done in earlier work, some of which was cited above. The average R^2 of the regressions, reported in Table 2, suggest that around 40% of the variation in the market spot volatility factor generates systematic moves in the cross-section of asset prices. The remaining part, which is rather nontrivial, is not relevant for the pricing of the assets in the cross section. Not accounting for this component of the volatility factor, as typically done in cross-sectional asset pricing, can lead to serious

Table 2: Empirical Results for Market Volatility

Time Aggregation	Empirical Rejection Rate of Test	Average R^2 of regression
1 day	0.0956	0.4483
5 days	0.2478	0.4040
10 days	0.3430	0.3994
20 days	0.4643	0.4036
40 days	0.6429	0.4100

The nominal size of the test is 5%. The selection of the latent factors is done as in the Monte Carlo.

distortions in the inference for the pricing of volatility risk. We leave the exploration of those for future work.

In Figure 2, we plot the daily realized volatility of market volatility and the component of it that generates systematic risk in stocks, i.e., the linear projection on the latent systematic risk factors. We use daily aggregation when computing the linear projections. Even though market volatility changes a lot over our sample, we can see that market volatility of volatility does not exhibit significant low frequency variation. There are a few episodes with spikes in the overall volatility of volatility. Interestingly, when this happens, most of this volatility of volatility generates systematic risk in asset prices. Overall, the figure reveals that there is nontrivial variation in the importance of market volatility risk for the stocks.

Market volatility is well-known to be correlated with market risk. More specifically, shocks to the market portfolio and its volatility are known to be highly negatively correlated. Following Black (1976), this negative dependence is referred to as leverage effect. Given this strong dependence and the fact that the market factor is well-known to generate systematic risk in asset prices, as in the classical CAPM model of Sharpe (1964) and Lintner (1965), a natural question is whether volatility generates systematic risk beyond the one due to the market risk factor. Put differently, can the component of volatility risk that generates systematic risk in the cross-section of asset prices be completely spanned by the market risk?

To answer this question we repeat the above analysis but for volatility risk that is locally orthogonal to market risk. We do this by estimating the linear projection of volatility returns on the market returns, and looking at the residuals. To alleviate the impact due to the error from estimating the linear projection coefficients on the

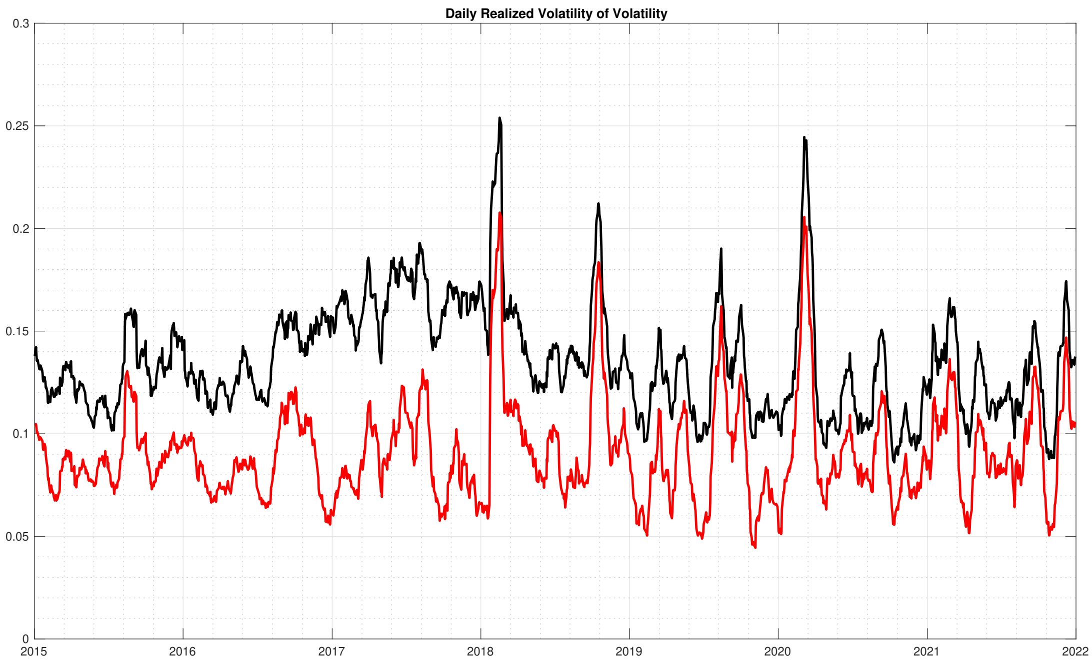

Figure 2: Market Volatility of Volatility. Black line corresponds to 20-day moving average of daily realized volatility of market volatility and red line is its counterpart for the daily linear projection of market volatility on the PCA factors.

test, we use projection coefficients computed from the previous period when forming the estimates for volatility residuals over the current period. The test results are reported in Table 3.

These results mirror those for the volatility factor reported in Table 2. Indeed, the empirical rejection rate of the test is significantly above the nominal one for time windows above one trading day. This, of course, implies that the residual market volatility risk does generate systematic risk in the cross-section of assets. That said, the R^2 of the regressions for the residual market volatility are significantly smaller than those for the market volatility itself. This means that, although present, the component of market volatility that is not spanned linearly by market risk and generates systematic risk in assets is relatively small.

Overall, our empirical analysis shows that market volatility generates systematic risk in asset prices which cannot be fully spanned by the market risk. To account for it, one should consider linear projections on the latent systematic risk factors embedded in the cross-section of stocks on relatively short time windows of around

Table 3: Empirical Results for Residual Market Volatility

Time Aggregation	Empirical Rejection Rate of Test	Average R^2 of regression
1 day	0.0845	0.2417
5 days	0.2588	0.1122
10 days	0.3455	0.0841
20 days	0.4286	0.0890
40 days	0.6571	0.0957

Residual market volatility is the residual from linear projection of market volatility returns on market returns using linear projection coefficients estimated from the preceding period. The nominal size of the test is 5%. The selection of the latent factors is done as in the Monte Carlo.

one day. More generally, we expect that the observable risk factors will differ in terms of their relation with the latent ones embedded in the cross-section of stock returns. For some, which are more dynamic in nature, this relationship will hold only over very short periods while for others, e.g., the market factor, this relationship will hold for much longer periods. These differences need to be accounted for when studying the cross-sectional pricing of the observable factors.

6 Conclusion

In this paper we develop a nonparametric test for the temporal stability of a linear relation between observable risk factors and latent ones embedded in the cross-section of asset returns. The test uses high-frequency data of the observable factors and of a large cross-section of asset prices. It evaluates the local behavior of residuals from global linear projections of the observed factors on the estimates of the latent ones. The test is used to study the dynamic relation between the market volatility factor and the systematic risk in the cross-section of asset prices.

7 Assumptions and proofs

7.1 Assumptions

The latent factors obey the following dynamics

$$df_t = \Lambda_t dW_t, \tag{37}$$

where W_t is $K \times 1$ standard Brownian motion while Λ_t is a $K \times K$ matrix, for some positive integer K.

The observable factor g, on the other hand, is given by the following Ito semimartingale

$$dg_t = \alpha_t^g dt + \lambda_t^g dW_t + \sigma_t^g dW_t^g + dJ_t^g, \tag{38}$$

where W_t^g is a univariate Brownian motion independent of W_t, α_t^g and σ_t^g are univariate and λ_t^g is a $1 \times K$ vector, J_t^g is a pure-jump vector. Under the null hypothesis, $\lambda_t^g = \lambda \times \Lambda_t$, for some constant $1 \times K$ vector λ.

The jump parts of Y and g are of the form, $J_t = \sum_{s \leq t} \Delta J_s$ and $J_t^g = \sum_{s \leq t} \Delta J_s^g$, where $\Delta J_t = J_t - J_{t-}$ and $\Delta J_t^g = J_t^g - J_{t-}^g$.

Finally, the idiosyncratic diffusive components of the asset prices are given by

$$d\epsilon_{t,j} = \sigma_{tj} dW_{t,j}, \tag{39}$$

where $(W_{t,j})_{j \geq 1}$ is a sequence of independent univariate Brownian motions that are also independent of W_t and W_t^g.

We have the following assumptions for the various processes that enter the dynamics of the asset prices and the risk factors.

A1. *There exists a sequence of stopping times $T_1, T_2, \ldots$ increasing to infinity such that for $s, t \leq T_m$, we have*

$$\sup_{j \geq 1} \mathbb{E}|\chi_{t,j}|^q + \sup_{j \geq 1} \mathbb{E}|J_{t,j}|^q + \mathbb{E}|\alpha_t^g|^q + \mathbb{E}|J_t^g|^q < \infty, \quad \text{for any } q > 0, \tag{40}$$

$$\sup_{j \geq 1} \mathbb{E}|\chi_{t,j} - \chi_{s,j}|^2 + |\sup_{j \geq 1} \mathbb{E}(\chi_{t,j} - \chi_{s,j})| \leq C_m |t - s|, \tag{41}$$

$$\sup_{j \geq 1} \mathbb{P}\left(J_{t,j} - J_{s,j} \neq 0\right) + \mathbb{P}\left(J_t^g - J_s^g \neq 0\right) \leq C_m |t - s|, \tag{42}$$

$$\mathbb{E}\|\Lambda_t - \Lambda_s\|_F^2 + \|\mathbb{E}(\Lambda_t - \Lambda_s)\|_F + \mathbb{E}|\lambda_t^g - \lambda_s^g|^2 + |\mathbb{E}(\lambda_t^g - \lambda_s^g)| \leq C_m |t - s|, \tag{43}$$

for some sequence of positive constants C_m and where $\chi_{t,j}$ is either $\alpha_{t,j}$ or $\sigma_{t,j}$.

To state the next assumption, we introduce the common set σ-algebra $\mathcal{C}$ that contains the information about the systematic risks. The factors and the processes used in defining their dynamics are all adapted to $\mathcal{C}$.

A2. *Conditional on* $\mathcal{C}$*:*

(i) $W_{t,j}$*,* $\sigma_{t,j}$ *and* $J_{t,j}$ *are independent across* j*.*

(ii) σ_{tj} *is independent from* W_{tj} *and* $\sup_{t\in[0,T],\, j\geq 1} |\sigma_{tj}|$ *is bounded.*

A3. *Let* $\Sigma_\beta = plim \frac{1}{p}\beta'\beta$ *and* $\Sigma_f = \int_0^T \Lambda_t \Lambda_t^\top dt$ *as well as* $M = \Sigma_\beta^{1/2}\Sigma_f\Sigma_\beta^{1/2}$*. Then,* M *is of full rank and have distinct non-zero eigenvalues. That is, if the eigenvalues of any of these matrices are denoted with* $v_1, ..., v_k$*, then there is* $c_0 > 0$ *so that* $|v_i - v_j| > c_0 > 0$ *for all* $i \neq j$*.*

A4. $\int_0^T (\sigma_t^g)^2 dt > 0$ *a.s.*

Throughout the proofs we will work with the following strengthened version of assumption A1:

SA1. *We have assumption A1 for* $s, t \in [0, T]$.

Using a standard localization argument, see e.g., Section 4.4.1 in Jacod and Protter (2011), one can extend the proof to the general (weaker) assumption A1.

7.2 Preliminaries

We start with setting some notation that will be used throughout the proof. Let

$$\mathcal{W}_\eta = \mathrm{diag}(\mathcal{W}_\eta(1/n), ..., \mathcal{W}_\eta(1)).$$

Next, we set

$$\overline{u}_t = \frac{1}{\sqrt{\Delta_n}} \int_{(t-1)\Delta_n}^{t\Delta_n} \sigma_s^g dW_s^g + \frac{1}{\sqrt{\Delta_n}} \left(\int_{(t-1)\Delta_n}^{t\Delta_n} \lambda_s^g dW_s - \gamma^\top \int_{(t-1)\Delta_n}^{t\Delta_n} \Lambda_s dW_s \right), \quad (44)$$

and from here the $n \times 1$ vector $\overline{U}$. We note that under the null hypothesis, the second term of $\overline{u}_t$ is zero.

Recall the matrix form of returns:

$$\overline{Y} = \beta \overline{F}^\top + \overline{\mathcal{E}} + R + \Upsilon.$$

The entries of the residual matrix R are given by

$$
\begin{aligned}
r_{tj} = & \sqrt{\Delta_n}\alpha_{(t-1)\Delta_n,j} - \sqrt{\Delta_n}\alpha_{(t-1)\Delta_n,j}1\left(|\Delta_t^n Y_j| \geq \gamma_j \Delta_n^{\varpi}\right) \\
& + \frac{1}{\sqrt{\Delta_n}}\left(\int_{(t-1\Delta_n}^{t\Delta_n}(\alpha_{s,j} - \alpha_{(t-1)\Delta_n,j})ds + \Delta_t^n J_j\right)1\left(|\Delta_t^n Y_j| < \gamma_j \Delta_n^{\varpi}\right) \\
& + \left(\gamma_j \Delta_n^{\varpi-1/2} - \iota_j'\beta\overline{f}_t - \overline{\epsilon}_{tj}\right)1\left(\Delta_t^n Y_j > \gamma_j \Delta_n^{\varpi}\right) \\
& - \left(\gamma_j \Delta_n^{\varpi-1/2} + \iota_j'\beta\overline{f}_t + \overline{\epsilon}_{tj}\right)1\left(\Delta_t^n Y_j < -\gamma_j \Delta_n^{\varpi}\right) \equiv \sum_{k=1}^{5} r_{tj}^{(k)}.
\end{aligned} \tag{45}
$$

We can define similarly the matrrix of residuals for the observed factor. This matrix is denoted as

$$
R_g = (r_t^{(g)} : t = 1, ..., n).
$$

We proceed with establishing several bounds that will be used in the proofs First, using Theorem 4.6.1 of Vershynin (2018) and assumption A2, we have

$$
\|\overline{\mathcal{E}}\| = O_p\left(\sqrt{n+p}\right). \tag{46}
$$

Second, Lemma A.1 of Liao and Todorov (2024), we have

$$
\|R\| = O_P(\sqrt{pn}\Delta_n^{\widetilde{\varpi}}), \tag{47}
$$

for any $\widetilde{\varpi} < \varpi$.

The (k, l) element of the $K \times K$ matrix $\overline{F}'\mathcal{W}_\eta R'\beta$ is given by $\sum_{j=1}^{p}\sum_{t=1}^{n}\overline{f}_{tk}r_{tj}\beta_{jl}\mathcal{W}_\eta(t/n)$. Using Theorem 12.3 of Billingsley (2013), we can establish that

$$
\sup_{\eta\in[0,2\pi]} \|\overline{F}'\mathcal{W}_\eta(R^{(1)})'\beta\| = O_P(p),
$$

where $R^{(1)}$ is the counterpart of R with r_{tj} replaced by $r_{tj}^{(1)}$.

Next, using the Holder inequality as well as bounds for powers of increments of Ito semimartingales in Section 2.1.5 in Jacod and Protter (2011), we have

$$
\mathbb{E}\left(\sum_{m=2}^{5}|\overline{f}_{tk}r_{tj}^{(m)}\beta_{jl}|\right) + \mathbb{E}|\overline{f}_{tk}r_t^{(g)}| \leq C\Delta_n^{1/2+\widetilde{\varpi}},
$$

for any $\widetilde{\varpi} < \varpi$. Altogether, using the algebraic inequality $\sqrt{\sum_{i=1}^{k}a_i^2} \leq \sum_{i=1}^{k}|a_i|$ for some integer k and a sequence of real numbers $(a_i)_{i=1,\dots,k}$, we can conclude

$$
\sup_{\eta\in[0,2\pi]} \|\overline{F}'\mathcal{W}_\eta R'\beta\| + \|\overline{F}'R'\beta\| = O_P\left(p\Delta_n^{\widetilde{\varpi}-1/2}\right), \tag{48}
$$

$$\|\overline{F}' R_g\| = O_P(\Delta_n^{\widetilde{\varpi}-1/2}). \tag{49}$$

Exactly in the same way, we can establish

$$\sup_{\eta\in[0,2\pi]} \|\overline{U}'\mathcal{W}_\eta R'\beta\| + \|\overline{U}' R'\beta\| = O_P\left(p\Delta_n^{\widetilde{\varpi}-1/2}\right). \tag{50}$$

7.3 Proof of Theorem 3.1

We start with introducing some notation. Let $H = \beta'\widehat{\beta}/p$, $\widehat{F}_k$ and H_k denote the k-th columns of $\widehat{F}$ and H, respectively. Let e_k denote the k-th natural canonical basis and set $H_k = He_k$. We can decompose

$$\widehat{\gamma} - H^{-1}\gamma = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\overline{F}H - \widehat{F})H^{-1}\gamma + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\overline{U} + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'R_g. \tag{51}$$

Denote

$$\begin{aligned} \Delta_1 &= (\widehat{F} - \overline{F}H)\widehat{\gamma}, \quad \Delta_2 = \overline{F}H(\widehat{\gamma} - H^{-1}\gamma) \\ \widehat{U} - \overline{U} &= -\Delta_1 - \Delta_2. \end{aligned} \tag{52}$$

Then

$$\frac{1}{n}\widehat{F}_k'\mathcal{W}_\eta\widehat{U} = \frac{1}{n}(\widehat{F}_k - \overline{F}H_k)'\mathcal{W}_\eta\widehat{U} + \underbrace{\frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\overline{U}}_{leading} - \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_1 - \underbrace{\frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_2}_{leading}.$$

We start the proof with establishing several auxiliary results. In what follows, we will denote

$$\begin{aligned} \Delta_4 :&= \frac{1}{pn}\|R\|^2 + \|\frac{1}{pn}\overline{U}'R'\widehat{\beta}\| + \|\frac{1}{pn}\overline{F}'R'\widehat{\beta}\| + \|\frac{1}{n}\overline{F}'R_g\| \\ &+ \sup_{\eta\in[0,2\pi]}\|\frac{1}{pn}\overline{F}'\mathcal{W}_\eta R'\beta\| + \sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\overline{U}'\mathcal{W}_\eta R'\beta\|. \end{aligned} \tag{53}$$

Lemma 7.1. *Suppose assumptions SA1 and A2-A3 hold. Then*

(i) There is $K\times K$ matrix H so that

$$\begin{aligned} \frac{1}{n}\|\widehat{F} - \overline{F}H\|^2 &= O_P(\frac{1}{p} + \frac{1}{n} + \frac{1}{pn}\|R\|^2) \\ \frac{1}{\sqrt{p}}\|\widehat{\beta} - \beta H^{-1}\| &= O_P(\frac{1}{\sqrt{n}} + \frac{1}{p} + \frac{1}{\sqrt{np}}\|R\|). \end{aligned}$$

(ii)

$$\widehat{\gamma} - H^{-1}\gamma = (\widehat{F}'\widehat{F})^{-1}H'\overline{F}'\overline{U} + O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4),$$

where Δ_4 is defined in (53).

Proof. First, as in the usual PCA setting, $\widehat{F} = \overline{F}H + \overline{\mathcal{E}}'\widehat{\beta}/p + R'\widehat{\beta}/p$ for some H. We state, therefore, the following results without proof (their proof can be adopted from Liao and Todorov (2024)):

$$\begin{aligned} \frac{1}{n}\|\widehat{F} - \overline{F}H\|^2 &= O_P(\frac{1}{p} + \frac{1}{n} + \frac{1}{pn}\|R\|^2), \\ \frac{1}{n}\overline{F}'(\widehat{F} - \overline{F}H) &= O_P(\frac{1}{n} + \frac{1}{\sqrt{np}}) + O_P(\frac{1}{n\sqrt{p}}\|R\| + \|\frac{1}{pn}\overline{F}'R'\widehat{\beta}\|), \\ \frac{1}{n}\widehat{F}'(\widehat{F} - \overline{F}H) &= O_P(\frac{1}{p} + \frac{1}{n} + \frac{1}{pn}\|R\|^2 + \|\frac{1}{pn}\overline{F}'R'\widehat{\beta}\|). \end{aligned}$$

Also as in the proof of Lemma C.2 in Liao and Todorov (2024), we have

$$\begin{aligned} \frac{1}{\sqrt{p}}\|\widehat{\beta} - \beta H^{-1}\| &\leq O_P(\frac{1}{\sqrt{n}} + \frac{1}{p}) + O_P(\frac{1}{\sqrt{p}})\|\frac{1}{pn}\overline{\mathcal{E}}R'\widehat{\beta} + \frac{1}{pn}R\overline{Y}'\widehat{\beta}\| \\ &\leq O_P(\frac{1}{\sqrt{n}} + \frac{1}{p} + \frac{1}{\sqrt{np}}\|R\|). \end{aligned}$$

In addition, using assumption A2, we have

$$\|\overline{\mathcal{E}}\,\overline{U}\| + \|\beta'\overline{\mathcal{E}}\,\overline{U}\| = O_P(\sqrt{np}).$$

Therefore,

$$\begin{aligned} \frac{1}{n}\widehat{F}'\overline{U} &= \frac{1}{n}H'\overline{F}'\overline{U} + \frac{1}{pn}(\widehat{\beta} - \beta H^{-1})'\overline{\mathcal{E}}\,\overline{U} + \frac{1}{pn}H^{-1}\beta'\overline{\mathcal{E}}\,\overline{U} + \frac{1}{pn}\widehat{\beta}'R\overline{U} \\ &= \frac{1}{n}H'\overline{F}'\overline{U} + O_P(\frac{1}{n} + \frac{1}{\sqrt{np}} + \frac{1}{n\sqrt{p}}\|R\| + \frac{1}{pn}\widehat{\beta}'R\overline{U}). \end{aligned} \tag{54}$$

Thus, taking into account the decomposition of $\widehat{\gamma} - H^{-1}\gamma$ in (51) and the above bounds, we get the result (ii) of the lemma.

□

Lemma 7.2. *Suppose assumptions SA1 and A2-A3 hold. Then*

(i)

$$\sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{np}}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\| = O_P(1),$$

(ii)

$$\sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{np}}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta\| + \sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{np}}\overline{U}'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta\| = O_P(1),$$

(iii)

$$\sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{n}}\overline{U}'\mathcal{W}_\eta\overline{F}\| + \sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{n}}(\overline{F}'\mathcal{W}_\eta\overline{F} - \sum_{t=1}^{n}\int_{(t-1)\Delta_n}^{t\Delta_n}\Lambda_s\Lambda_s^\top ds\mathcal{W}_\eta(t/n))\| = O_P(1).$$

Proof. Let $\overline{\mathcal{E}}_i$ be the n-dim vector with entries of $\bar{\epsilon}_{ti}$, for $t = 1, ..., n$. Similarly, let $\overline{F}_k$ be the n-dim vector with entries $\overline{f}_{tk}$, for $t = 1, ..., n$. Denote with $\beta_{i,l}$ the (i, l) element of β. Fix any $k, l \leq K$ and denote

$$X_i(\eta) := (\frac{1}{\sqrt{n}}\overline{F}'_k\mathcal{W}_\eta\overline{\mathcal{E}}_i)^2 - \frac{1}{np}\sum_{t=1}^n\sum_{i=1}^n\int_{(t-1)\Delta_n}^{t\Delta_n}\sigma_{si}^2 ds\overline{f}_{tk}^2\mathcal{W}_\eta^2(t/n),\ Y_i(\eta) := \frac{1}{\sqrt{n}}G'\mathcal{W}_\eta\overline{\mathcal{E}}_i\beta_{i,l},$$

where $G \in \{\overline{F}_k, \overline{U}\}$. Then

$$\begin{aligned}
\|\frac{1}{\sqrt{np}}\overline{F}'_k\mathcal{W}_\eta\overline{\mathcal{E}}'\|^2 &= \frac{1}{p}\sum_{i=1}^p X_i(\eta) + \frac{1}{np}\sum_{t=1}^n\sum_{i=1}^n\int_{(t-1)\Delta_n}^{t\Delta_n}\sigma_{si}^2 ds\overline{f}_{tk}^2\mathcal{W}_\eta^2(t/n)\overline{\mathcal{E}}_i)^2 \\
\left|\frac{1}{\sqrt{np}}G'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta_l\right| &= \frac{1}{\sqrt{p}}|\sum_i Y_i(\eta)| \\
\left|\frac{1}{\sqrt{n}}\overline{U}'\mathcal{W}_\eta\overline{F}_k\right| &= \frac{1}{\sqrt{n}}|\sum_t u_t\overline{f}_{tk}\mathcal{W}_\eta(t/n)|.
\end{aligned}$$

Next, let

$$\mathcal{X}_n(\eta) := \frac{1}{\sqrt{p}}\sum_{i=1}^p X_i(\eta), \quad \mathcal{Y}_n(\eta) := \frac{1}{\sqrt{p}}\sum_{i=1}^p Y_i(\eta), \quad \mathcal{Z}_n(\eta) := \frac{1}{\sqrt{n}}\sum_t u_t\overline{f}_{tk}\mathcal{W}_\eta(t/n).$$

We now proceed in two steps.

Step 1: $\mathbb{E}\left|\frac{1}{\sqrt{n}}\overline{F}'_k\mathcal{W}_\eta\overline{\mathcal{E}}_i\right|^4 < C$, $\mathbb{E}|Y_i(\eta)|^4 < C$ and $\mathbb{E}|\mathcal{Z}_n(\eta)|^2 < C$, for some constant $C > 0$ that does not depend on η. By the Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
\mathbb{E}\left|\frac{1}{\sqrt{n}}\overline{F}'_k\mathcal{W}_\eta\overline{\mathcal{E}}_i\right|^4 &= \mathbb{E}\left[\frac{1}{\sqrt{n}}\sum_t\bar{\epsilon}_{ti}\overline{f}_{tk}\mathcal{W}_\eta(t/n)\right]^4 \leq C\mathbb{E}\left(\frac{1}{n}\sum_t\bar{\epsilon}_{ti}^2\overline{f}_{tk}^2\mathcal{W}_\eta(t/n)^2\right)^2 \\
&\leq C\frac{1}{n}\sum_t\mathbb{E}\bar{\epsilon}_{ti}^4\overline{f}_{tk}^4 < C,
\end{aligned}$$

where the constant C changes from line to line and importantly it does not depend on η. This also implies $\mathbb{E}|Y_i(\eta)|^4 < C$. Similar argument implies $\mathbb{E}|\mathcal{Z}_n(\eta)|^2 < C$.

Step 2: $\mathcal{X}_n(\eta)$, $\mathcal{Y}_n(\eta)$ and $\mathcal{Z}_n(\eta)$ are tight. To prove the tightness, we employ Theorem 12.3 of Billingsley (2013) by verifying its conditions. The first condition that $\mathcal{X}_n(0) = O_P(1)$ follows from step 1. As for the second condition, for any η_1, η_2,

$$
\begin{aligned}
& \mathbb{E}|\mathcal{X}_n(\eta_1) - \mathcal{X}_n(\eta_2)|^2 \\
= & \frac{1}{p}\sum_{i=1}^{p} \mathrm{Var}(X_i(\eta_1) - X_i(\eta_2)) = \frac{1}{p}\sum_{i=1}^{p} \mathrm{Var}[(\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_1}\overline{\mathcal{E}}_i)^2 - (\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_2}\overline{\mathcal{E}}_i)^2] \\
\leq & 4\frac{1}{p}\sum_{i=1}^{p} \mathbb{E}\left[(\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_1}\overline{\mathcal{E}}_i)^2 - (\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_2}\overline{\mathcal{E}}_i)^2\right]^2 \\
\leq & \frac{4}{p}\sum_{i=1}^{p} \mathbb{E}\left[\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_1}\overline{\mathcal{E}}_i + \frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_2}\overline{\mathcal{E}}_i\right]^2 \left[\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_1}\overline{\mathcal{E}}_i - \frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_2}\overline{\mathcal{E}}_i\right]^2 \\
\leq & 16\sqrt{\sup_{\eta\in[0,2\pi]} \mathbb{E}\left|\frac{1}{\sqrt{n}}\overline{F}'_k \mathcal{W}_\eta \overline{\mathcal{E}}_i\right|^4} \sqrt{\frac{1}{p}\sum_{i=1}^{p} \mathbb{E}\left[\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_1}\overline{\mathcal{E}}_i - \frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_2}\overline{\mathcal{E}}_i\right]^4}.
\end{aligned}
$$

By step 1, $\sup_{\eta\in[0,2\pi]} \mathbb{E}\left|\frac{1}{\sqrt{n}}\overline{F}'_k \mathcal{W}_\eta \overline{\mathcal{E}}_i\right|^4 < C$. Also, by the Burkholder-Davis-Gundy inequality,

$$
\begin{aligned}
& \mathbb{E}\left[\frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_1}\overline{\mathcal{E}}_i - \frac{1}{\sqrt{n}}\overline{F}'_k W_{\eta_2}\overline{\mathcal{E}}_i\right]^4 = \mathbb{E}\left[\frac{1}{\sqrt{n}}\sum_t \overline{\epsilon}_{ti}\overline{f}_{tk}[w(\eta_1,t) - w(\eta_2,t)]\right]^4 \\
\leq & C\mathbb{E}\left[\frac{1}{n}\sum_t \overline{\epsilon}^2_{ti}\overline{f}^2_{tk}[w(\eta_1,t) - w(\eta_2,t)]^2\right]^2 \leq C\frac{1}{n}\sum_t \mathbb{E}\overline{\epsilon}^4_{ti}\overline{f}^4_{tk}[w(\eta_1,t) - w(\eta_2,t)]^4 \\
\leq & C\max_{it} \mathbb{E}(\overline{\epsilon}^4_{ti}\overline{f}^4_{tk})[\eta_1 - \eta_2]^4 \leq C|\eta_1 - \eta_2|^4,
\end{aligned}
$$

where the last inequality follows from SA1. Together

$$
\mathbb{E}|\mathcal{X}_n(\eta_1) - \mathcal{X}_n(\eta_2)|^2 \leq C|\eta_1 - \eta_2|^2,
$$

which verifies the second condition of Theorem 12.3 in Billingsley (2013). Hence, $\mathcal{X}_n(\eta)$ is tight.

To prove $\mathcal{Y}_n(\eta)$ is tight, first note that $\mathcal{Y}_n(0) = O_P(1)$ following from step 1. Next, for g_t being the t th element of G,

$$
\begin{aligned}
& \mathbb{E}|\mathcal{Y}_n(\eta_1) - \mathcal{Y}_n(\eta_2)|^2 = \frac{1}{p}\sum_{i=1}^{p} \mathrm{Var}(\frac{1}{\sqrt{n}}G' W_{\eta_1}\overline{\mathcal{E}}_i\beta_{i,l} - \frac{1}{\sqrt{n}}G' W_{\eta_2}\overline{\mathcal{E}}_i\beta_{i,l}) \\
= & \frac{1}{pn}\sum_{i=1}^{p}\sum_t \mathrm{Var}(\overline{\epsilon}_{ti} g_t \beta_{i,l})[w(\eta_1,t) - w(\eta_2,t)]^2 \leq C|\eta_1 - \eta_2|^2.
\end{aligned}
$$

Hence, $\mathcal{Y}_n(\eta)$ is also tight. For $\mathcal{Z}_n(\eta)$, we have

$$\mathbb{E}|\mathcal{Z}_n(\eta_1) - \mathcal{Z}_n(\eta_2)|^2 = \frac{1}{n}\sum_{t=1}^{n} \mathrm{Var}(u_t \overline{f}_{tk})[w(\eta_1, t) - w(\eta_2, t)]^2 \leq C|\eta_1 - \eta_2|^2.$$

Hence $\mathcal{Z}_n(\eta)$ is also tight.

The tightness of the sequences implies:

$$\sup_{\eta\in[0,2\pi]} |\mathcal{X}_n(\eta)| = O_P(1), \quad \sup_{\eta\in[0,2\pi]} |\mathcal{Y}_n(\eta)| = O_P(1), \text{ and } \sup_{\eta\in[0,2\pi]} |\mathcal{Z}_n(\eta)| = O_P(1).$$

Hence

$$\sup_{\eta\in[0,2\pi]} \|\frac{1}{\sqrt{np}}\overline{F}'_k \mathcal{W}_\eta \overline{\mathcal{E}}'\|^2 \leq \frac{1}{\sqrt{p}} \sup_{\eta\in[0,2\pi]} |\mathcal{X}_n(\eta)| + \sup_{\eta\in[0,2\pi]} \frac{1}{p}\sum_{i=1}^{p} \mathbb{E}(\frac{1}{\sqrt{n}}\overline{F}'_k \mathcal{W}_\eta \overline{\mathcal{E}}_i)^2 = O_P(1).$$

Finally, the proof that $\sup_{\eta\in[0,2\pi]} \|\frac{1}{\sqrt{n}}(\overline{F}'\mathcal{W}_\eta\overline{F} - \sum_{t=1}^{n}\int_{(t-1)\Delta_n}^{t\Delta_n} \Lambda_s\Lambda_s^\top ds \mathcal{W}_\eta(t/n))\| = O_P(1)$ follows in the same manner.

□

Lemma 7.3. *Suppose assumptions SA1 and A2-A3 hold. Then, with the definitions of* Δ_1 *and* Δ_2 *in (52), we have*

(i)

$$\sup_{\eta\in[0,2\pi]} \frac{1}{n}|\overline{F}'\mathcal{W}_\eta\Delta_1| \leq O_P(\frac{1}{\sqrt{np}} + \frac{1}{n} + \Delta_4),$$

(ii)

$$\sup_{\eta\in[0,2\pi]} \frac{1}{n}|\overline{F}'\mathcal{W}_\eta\Delta_2 - \overline{F}'\mathcal{W}_\eta\overline{F}H(\widehat{F}'\widehat{F})^{-1}H'\overline{F}'\overline{U}| = O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4),$$

(iii)

$$\sup_{\eta\in[0,2\pi]} \|\frac{1}{n}\widehat{U}'\mathcal{W}_\eta(\widehat{F}_k - \overline{F}H_k)\| = O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4),$$

where Δ_4 *is given in (53).*

Proof. In the proof below, we shall repeatedly use Lemmas 7.1 and 7.2. First, $\|\overline{\mathcal{E}}\| = O_P(\sqrt{p+n})$ from (46). Next, Lemma 7.1 shows

$$\frac{1}{\sqrt{p}}\|\widehat{\beta} - \beta H^{-1}\| = O_P(\frac{1}{\sqrt{n}} + \frac{1}{p} + \frac{1}{\sqrt{np}}\|R\|).$$

Thus

$$
\begin{aligned}
\frac{1}{n}\|\Delta_1\|^2+\|\Delta_2\|^2 &\leq O_P(\frac{1}{p}+\frac{1}{n}+\Delta_4), \qquad (55)\\
\frac{1}{pn}\overline{U}'\mathcal{W}_\eta\overline{\mathcal{E}}'(\widehat{\beta}-\beta H_\beta) &\leq \frac{1}{pn}\|\overline{U}\|\|\overline{\mathcal{E}}\|\|\widehat{\beta}-\beta H_\beta\|\\
&\leq O_P(\sqrt{n+p})(\frac{\sqrt{np}}{pn})(\frac{1}{\sqrt{n}}+\frac{1}{p}+\frac{1}{\sqrt{np}}\|R\|)\\
&\leq O_P(\frac{1}{n}+\frac{1}{p}+\frac{1}{pn}\|R\|^2)\\
\frac{1}{pn}\overline{U}'\mathcal{W}_\eta R'(\widehat{\beta}-\beta H_\beta) &\leq \frac{1}{pn}\|\overline{U}\|\|R\|\|\widehat{\beta}-\beta H_\beta\|\leq O_P(\frac{1}{\sqrt{n}}+\frac{1}{p}+\frac{1}{\sqrt{np}}\|R\|)\frac{1}{\sqrt{pn}}\|R\|\\
&\leq O_P(\frac{1}{n}+\frac{1}{p^2}+\frac{1}{pn}\|R\|^2)
\end{aligned}
$$

where we used the inequality $|ab|\leq a^2+b^2$. Also, Lemma 7.2 shows

$$
\sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{np}}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta\|+\sup_{\eta\in[0,2\pi]}\|\frac{1}{\sqrt{np}}\overline{U}'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta\|=O_P(1).
$$

Therefore,

$$
\begin{aligned}
\frac{1}{n}\widehat{U}'\mathcal{W}_\eta(\widehat{F}_k-\overline{F}H_k) &= \frac{1}{n}\overline{U}'\mathcal{W}_\eta(\widehat{F}_k-\overline{F}H_k)-\frac{1}{n}(\Delta_1+\Delta_2)'\mathcal{W}_\eta(\widehat{F}_k-\overline{F}H_k)\\
&= \frac{1}{np}\overline{U}'\mathcal{W}_\eta(\overline{\mathcal{E}}+R)'\beta H^{-1}+\frac{1}{np}\overline{U}'\mathcal{W}_\eta(\overline{\mathcal{E}}+R)'(\widehat{\beta}-\beta H_\beta) \qquad (56)\\
&\quad +O_P(\|\Delta_1\|+\|\Delta_2\|)\frac{1}{n}\|\widehat{F}_k-FH_k\|\\
&\leq O_P(\frac{1}{\sqrt{np}}+\frac{1}{n}+\frac{1}{p}+\frac{1}{pn}\|R\|^2+\sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\overline{U}'\mathcal{W}_\eta R'\beta\|+\Delta_4)\\
&= O_P(\frac{1}{n}+\frac{1}{p}+\Delta_4). \qquad (57)
\end{aligned}
$$

In addition,

$$
\begin{aligned}
\frac{1}{n}\overline{F}'\mathcal{W}_\eta\Delta_1 &= \frac{1}{pn}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\widehat{\beta}\widehat{\gamma}+\frac{1}{pn}\overline{F}'\mathcal{W}_\eta R'\widehat{\beta}\widehat{\gamma}\\
&\leq \frac{1}{pn}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta H^{-1}\widehat{\gamma}+\frac{1}{pn}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'(\widehat{\beta}-\beta H_\beta)\widehat{\gamma}+\frac{1}{pn}\overline{F}'\mathcal{W}_\eta R'\widehat{\beta}\widehat{\gamma}\\
&\leq \|\frac{1}{pn}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\beta\|O_P(1)+\|\frac{1}{pn}\overline{F}'\mathcal{W}_\eta\overline{\mathcal{E}}'\|\|(\widehat{\beta}-\beta H_\beta)\widehat{\gamma}\|\\
&\quad +\frac{1}{pn}\overline{F}'\mathcal{W}_\eta R'\widehat{\beta}\widehat{\gamma}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \quad O_P(\frac{1}{\sqrt{np}} + \frac{1}{n} + \frac{1}{pn}\|R\|^2 + \|\frac{1}{pn}\overline{F}'\mathcal{W}_\eta R'\beta\|). && (58)\\
\frac{1}{n}\overline{F}'\mathcal{W}_\eta\Delta_2 &= \frac{1}{n}\overline{F}'\mathcal{W}_\eta\overline{F}H(\widehat{\gamma} - H^{-1}\gamma)\\
&= \frac{1}{n}\overline{F}'\mathcal{W}_\eta\overline{F}H(\widehat{F}'\widehat{F})^{-1}H'\overline{F}'\overline{U} + O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4). && (59)
\end{aligned}
$$

□

Lemma 7.4. *Suppose assumptions SA1 and A2-A3 hold. Then there is a limiting matrix $\bar{H}$ so that $H \xrightarrow{\mathbb{P}} \bar{H}$, and define*

$$
\bar{\lambda}_k(\eta)' = -e_k'\bar{H}'\Sigma_{F,\eta}\Sigma_F^{-1}\bar{H}', \tag{60}
$$

$$
\Sigma_F = \frac{1}{T}\int_0^T \Lambda_t\Lambda_t' dt, \quad \Sigma_{F,\eta} = \frac{1}{T}\int_0^T \Lambda_t\Lambda_t'\mathcal{W}_\eta(t/T)dt. \tag{61}
$$

Then, uniformly in $\eta \in [0, 2\pi]$,

$$
\sqrt{n}\,\widehat{\mathbb{S}}_k(\eta) = \frac{1}{\sqrt{n}}\sum_t Z_t(\eta, k) + O_P(n^{1/2}p^{-1}) + o_P(1) + O_P(\sqrt{n}\Delta_4),
$$

where $\widehat{\mathbb{S}}_k(\eta) = \frac{1}{n}\widehat{F}_k'\mathcal{W}_\eta\widehat{U}$, and

$$
\Gamma_t^n(\eta, k) := e_k'\bar{H}'w_\eta(t/n) + \bar{\lambda}_k(\eta), \quad Z_t(\eta, k) \;:=\; \Gamma_t^n(\eta, k)'\overline{f}_t\overline{u}_t.
$$

Proof. Define the following $1 \times K$ vector:

$$
\lambda_k(\eta)' = -e_k'H'\overline{F}'\mathcal{W}_\eta\overline{F}H(\widehat{F}'\widehat{F})^{-1}H'.
$$

We then have

$$
\frac{1}{n}\widehat{F}_k'\mathcal{W}_\eta\widehat{U} = \frac{1}{n}(\widehat{F}_k - \overline{F}H_k)'\mathcal{W}_\eta\widehat{U} + \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta U - \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_1 - \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_2. \tag{62}
$$

By Lemma 7.3,

$$
\begin{aligned}
\frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_1 &= O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4),\\
-\frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_2 &= \frac{1}{n}\lambda_k(\eta)'\overline{F}'\overline{U} + O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4),\\
\frac{1}{n}(\widehat{F}_k - \overline{F}H_k)'\mathcal{W}_\eta\widehat{U} &= O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4).
\end{aligned}
$$

Hence

$$\frac{1}{n}\widehat{F}_k'\mathcal{W}_\eta\widehat{U} = \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\overline{U} + \frac{1}{n}\lambda_k(\eta)'\overline{F}'\overline{U} + O_P(\frac{1}{n}+\frac{1}{p}+\Delta_4),$$

uniformly in η. We proceed next with showing $\lambda_k(\eta) \xrightarrow{\mathbb{P}} \bar{\lambda}_k(\eta)$. First, using the fact that Λ_t is matrix-valued process with cadlag paths, we can apply Theorem 3.4.1 of Jacod and Protter (2011), and get

$$\frac{1}{n}\sum_{t=1}^{n}\overline{f}_t\overline{f}_t' \;\xrightarrow{\mathbb{P}}\; \frac{1}{T}\int_0^T \Lambda_t\Lambda_t' dt, \tag{63}$$

as well as

$$\frac{1}{n}\sum_{t=1}^{n}\overline{f}_t\overline{f}_t'\mathcal{W}_\eta(t/n) \;\xrightarrow{\mathbb{P}}\; \frac{1}{T}\int_0^T \Lambda_t\Lambda_t'\mathcal{W}_\eta(t/T) dt, \tag{64}$$

for any fixed η. Showing that the latter holds uniformly in η can be done by establishing tightness of the sequence, which in turn can be done using Theorem 12.3 of Billingsley (2013).

In addition, it follows from Lemma C.11 (display C.38) of Liao and Todorov (2024) that $\text{plim}H \to \bar{H}$ for some nonrandom matrix $\bar{H}$. From here, the convergence $\lambda_k(\eta) \xrightarrow{\mathbb{P}} \bar{\lambda}_k(\eta)$ follows by application of Lemma 7.1, and hence the result of the lemma follows. $\square$

Lemma 7.5. *Suppose assumptions SA1 and A2-A3 hold. With the notation of Lemma 7.4 and for $Z_t(\eta,k) := \Gamma_t^n(\eta,k)'\overline{f}_t\overline{u}_t$, we have*

$$\widehat{\mathbb{Z}}_K(\eta) := \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix} Z_t(\eta_1,1) \\ \vdots \\ Z_t(\eta_K,K)\end{pmatrix} \xrightarrow{\mathcal{L}-s} \mathbb{Z}_K(\bar{\eta})$$

where $\bar{\eta} = (\eta_1,...,\eta_K)$ and $\mathbb{Z}_K(\bar{\eta})$ is a K-dimensional vector-valued process, defined on an extension of the original probability space, which is $\mathcal{F}$-conditionally a zero-mean Gaussian process, with $\mathcal{F}$-conditional covariance kernel being the following $K \times K$ matrix:

$$\mathbb{E}(\mathbb{Z}_K(\eta_1)\mathbb{Z}_K(\eta_2)'|\mathcal{F}) = \frac{1}{T}\int_0^T \Gamma_t(\eta_1)'V_t\Gamma_t(\eta_2)dt, \quad V_t := (\sigma_t^g)^2\Lambda_t\Lambda_t',$$

$$\Gamma_t(\eta_i)' = \begin{pmatrix}\Gamma_t(\eta_{i1},1)' \\ \vdots \\ \Gamma_t(\eta_{iK},K)'\end{pmatrix}_{K\times K}, \quad \Gamma_t(\eta,k) := e_k'\bar{H}'w_\eta(t/T), \quad i=1,2,$$

and $\eta_i = (\eta_{i1},...,\eta_{iK})$, for $i=1,2$.

Proof. We will use the following shorthand notation in the proof:

$$\mathcal{W}_{n,K}(\eta) = (\mathcal{W}_n(\eta_1, 1), \cdots \mathcal{W}_n(\eta_K, K))', \quad \mathcal{W}_n(\eta_k, k) := \frac{1}{\sqrt{n}} \sum_t Z_t(\eta_k, k).$$

Using Proposition VIII.5.33 (iv) of Jacod and Shiryaev (2003), we need to show: (1) finite-dimensional stable convergence of $\mathcal{W}_{n,K}(\eta)$ and (2) tightness of $\mathcal{W}_{n,K}(\eta)$.

Step 1 (fidi). Let d be any finite positive integer, and $\eta_1, ..., \eta_d$ be K-dimensional vectors with elements in $[0, 2\pi]$. Then, we need to show

$$(\widehat{\mathbb{Z}}_K(\eta_1)', \cdots, \widehat{\mathbb{Z}}_K(\eta_d)')' \xrightarrow{\mathbb{L}-s} (\mathbb{Z}_K(\eta_1)', \cdots, \mathbb{Z}_K(\eta_d)')'. \tag{65}$$

By application of Theorem 2.2.15 of Jacod and Protter (2011), the result will follow from the following:

$$\frac{1}{n} \sum_{t=1}^{n} \mathbb{E}_{(t-1)\Delta_n} (Z_t(\eta_1, k) Z_t(\eta_2, l)) \xrightarrow{\mathbb{P}} \frac{1}{T} \int_0^T \Gamma_t(\eta_1, k)' V_t \Gamma_t(\eta_2, l) dt, \tag{66}$$

$$\frac{1}{n^{1+\iota/2}} \sum_{t=1}^{n} \mathbb{E}_{(t-1)\Delta_n} |Z_t(\eta_1, k)|^{2+\iota} \xrightarrow{\mathbb{P}} 0, \text{for some } \iota > 0, \tag{67}$$

$$\frac{1}{n} \sum_{t=1}^{n} \mathbb{E}_{(t-1)\Delta_n} \left(Z_t(\eta_1, k)(M_{t\Delta_n} - M_{(t-1)\Delta_n}) \right) \xrightarrow{\mathbb{P}} 0, \tag{68}$$

for $\eta_1, \eta_2 \in [0, 2\pi]$, $k, l = 1, ..., K$ and a univariate bounded martingale M which either equals the elements of W_t, W_d, or is orthogonal to them in a martingale sense.

The first of the above results follows by application of Theorem 3.4.1 of Jacod and Protter (2011). The second one follows by an application of Burkholder-Davis-Gundy inequality. The third result in the case when M equals an element of W_t or W_t^d follows using the independence of these univariate Brownian motions. Finally, the third result in the case when M is a bounded martingale orthogonal to W_t and W_t^d follows by an application of the martingale representation theorem.

Step 2 (tightness). We now prove that $\mathcal{W}_{n,K}(\cdot)$ is tight by verifying the conditions of Theorem 20 in Ibragimov and Has'minskii (2013). First $W_{n,K}(0) = O_P(1)$ follows from the fidi step above. Next, for any $\eta_1, \eta_2 \in [0, 2\pi]^K$ and $p \geq 2$,

$$\mathbb{E}\|\mathcal{W}_{n,K}(\eta_1) - \mathcal{W}_{n,K}(\eta_2)\|^p \quad \leq \quad C_p \sum_{k=1}^{K} \mathbb{E}|\mathcal{W}_n(\eta_{1k}, k) - \mathcal{W}_n(\eta_{2k}, k)|^p$$

$$\leq \quad C_p \sum_{k=1}^{K} |\eta_{1k} - \eta_{2k}|^p,$$

for some constant C_p that depends on p and where $\eta_i = (\eta_{i1}, ..., \eta_{iK})$, for $i = 1, 2$. Hence $\mathcal{W}_{n,K}(\eta)$ is tight. □

We are now ready to prove the result of the theorem. First, using the preliminary bounds in (47)-(50) as well as Lemma 7.1, we have that

$$\sqrt{n}\Delta_4 = O_P(\sqrt{n}\Delta_n^{2\tilde{\varpi}} + \sqrt{n}\Delta_n^{\tilde{\varpi}+1}\sqrt{p}). \tag{69}$$

Then, given the condition in (22) in the statement of the theorem, we have from Lemma 7.4 that uniformly in $\eta \in [0, 2\pi]$,

$$\sqrt{n}\widehat{\mathbb{S}}_k(\eta) = \frac{1}{\sqrt{n}} \sum_t Z_t(\eta, k) + o_P(1).$$

Hence, applying Lemma 7.5, we have uniformly for $\bar{\eta} = (\eta_1 ... \eta_K) \in [0, 2\pi]^K$,

$$\sqrt{n} \begin{pmatrix} \widehat{\mathbb{S}}_1(\eta_1) \\ \vdots \\ \widehat{\mathbb{S}}_K(\eta_K) \end{pmatrix} = \frac{1}{\sqrt{n}} \sum_t \begin{pmatrix} Z_t(\eta_1, 1) \\ \vdots \\ Z_t(\eta_K, K) \end{pmatrix} + o_P(1) \xrightarrow{\mathcal{L}-s} \mathbb{Z}_K(\bar{\eta}),$$

which is the result of the theorem to be proved.

7.4 Proof of Theorem 3.2

First, it follows from Lemma C.11 (display C.38) of Liao and Todorov (2024) that plim$H \to \bar{H}$ for some nonrandom matrix $\bar{H}$.

Next, we note that the expression for γ given the assumption for the dynamics of the observable and latent factors is given by

$$\gamma \equiv \left(\int_0^T \Lambda_s \Lambda_s^\top ds \right)^{-1} \int_0^T \Lambda_s (\lambda_s^g)^\top ds.$$

As in the proof of Lemma 7.4 and with the notation of this lemma, we have

$$\frac{1}{n} \widehat{F}_k' \mathcal{W}_\eta \widehat{U} = \frac{1}{n} e_k' \bar{H}' \overline{F}' \mathcal{W}_\eta \overline{U} + \frac{1}{n} \lambda_k(\eta)' \overline{F}' \overline{U} + o_P(1),$$

uniformly in $\eta \in [0, 2\pi]$. Now, given the definition of the residual matrix $\overline{U}$, we have $\frac{1}{n}\overline{F}'\overline{U} = o_P(1)$. On the other hand, we have

$$\frac{1}{n}\overline{F}'\mathcal{W}_\eta\overline{U} = \frac{1}{n}\sum_{t=1}^{n}\overline{f}_t\overline{u}_t\mathcal{W}_\eta(t/n) \overset{\mathbb{P}}{\to} \mathcal{S}_1(\eta) := \frac{1}{T}\int_0^T \mathcal{W}_\eta(t/T)d\langle f, u\rangle_t. \tag{70}$$

This result for fixed η follows by using the fact that Λ_t and λ_t^g are processes with cadlag paths and applying Theorem 3.4.1 of Jacod and Protter (2011). Showing that the result holds uniformly over η can be established exactly as in the proof of Lemma 7.2.

Hence, $\frac{1}{n}\widehat{F}_k'W_{\eta_k}\widehat{U} \overset{\mathbb{P}}{\to} e_k'\bar{H}'\mathcal{S}_1(\eta_k)$, uniformly in $\eta_k \in [0, 2\pi]$, and from here:

$$\bar{\eta} = (\eta_1, ..., \eta_K)', \quad \widehat{\mathbb{S}}(\bar{\eta}) = \begin{pmatrix} \frac{1}{n}\widehat{F}_1'W_{\eta_1}\widehat{U} \\ \vdots \\ \frac{1}{n}\widehat{F}_K'W_{\eta_K}\widehat{U} \end{pmatrix} \overset{\mathbb{P}}{\to} \begin{pmatrix} e_1'\bar{H}'\mathcal{S}_1(\eta_1) \\ \vdots \\ e_K'\bar{H}'\mathcal{S}_1(\eta_K) \end{pmatrix}$$
$$= \frac{1}{T}\int_0^T \operatorname{diag}\left(w_{\eta_1}(t/T), ..., w_{\eta_K}(t/T)\right)\bar{H}'\, d\langle f, u\rangle_t.$$

7.5 Proof of Theorem 3.3

In the proof of the theorem, we denote with $\mathbb{P}^*$ and $\mathbb{E}^*$, the probability and expectation, respectively, on the extended probability space on which the variables $\{\eta_t^*\}_{t\geq 1}$ are defined, conditional on $\mathcal{F}$. Let G^*, U^* and $\widehat{U}^*$ denote the $n \times 1$ vectors of g_t^*, $\eta_t^*\widehat{u}_t$ and $\widehat{u}_t^*$, and let $\bar{U}^*$ denote the vector of $u_t\eta_t^*$. We then have

$$\widehat{\gamma}^* - \widehat{\gamma} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'U^*.$$

Differently from (52), we do not reestimate factors in the bootstrap, so we have:

$$\widehat{U}^* - U^* \;=\; -\Delta_2^*, \quad \Delta_2^* = \widehat{F}(\widehat{\gamma}^* - \widehat{\gamma}).$$

and hence

$$\begin{aligned} \frac{1}{n}\widehat{F}_k'\mathcal{W}_\eta\widehat{U}^* &= \Delta_3^* + \frac{1}{n}H_k'F'\mathcal{W}_\eta U^* - \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta\Delta_2^*, \\ \Delta_3^* &= \frac{1}{n}(\widehat{F}_k - \overline{F}H_k)'\mathcal{W}_\eta\widehat{U}^*. \end{aligned}$$

The proof of the theorem starts with establishing several lemmas.

Lemma 7.6. *Suppose assumptions SA1 and A2-A3 hold. Then,*

$$\sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\beta'\overline{\mathcal{E}}\mathcal{W}_\eta U^*\| = O_{P^*}(\frac{1}{p}+\frac{1}{\sqrt{np}}+\frac{1}{\sqrt{p}}\sqrt{\Delta_4}),$$

$$\sup_{\eta\in[0,2\pi]}\|\frac{1}{n}\overline{F}'\mathcal{W}_\eta U^*\| = O_{P^*}(\frac{1}{\sqrt{n}}+\frac{1}{\sqrt{p}}+\sqrt{\Delta_4}),$$

$$\sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\beta' R\mathcal{W}_\eta U^*\| = O_{P^*}(\Delta_4+\frac{1}{n}+\frac{1}{p}).$$

Proof. Using the same proof as that of Lemma 7.2,

$$\|\frac{1}{np}\beta'\overline{\mathcal{E}}\mathcal{W}_\eta \bar{U}^*\| = O_{P^*}(\frac{1}{\sqrt{np}}), \quad \sup_\eta\|\frac{1}{n}\overline{F}'\mathcal{W}_\eta \bar{U}^*\| = O_{P^*}(\frac{1}{\sqrt{n}}).$$

In addition, for any $n\times 1$ vector $A=(a_1...a_n)'$, we have

$$|\frac{1}{n}A'\mathcal{W}_\eta(\bar{U}^*-U^*)| = |\frac{1}{n}\sum_t a_t(u_t-\widehat{u}_t)\eta_t^*\mathcal{W}_\eta(t/n)| \leq \frac{2}{\sqrt{n}}\|\widehat{U}-\overline{U}\|M$$

where $M^2=\frac{1}{n}\sum_t a_t^2\eta_t^{*2}$, and $\mathbb{E}^*M^2=\frac{1}{n}\sum_t a_t^2\mathbb{E}^*\eta_t^{*2}=\frac{1}{n}\|A\|^2$. Hence

$$\frac{1}{n}|A'\mathcal{W}_\eta(\bar{U}^*-U^*)| \leq O_{P^*}\left(\frac{1}{n}\|A\|\|\widehat{U}-\overline{U}\|\right).$$

Using this and (55), we have

$$\begin{aligned}
\|\frac{1}{np}\beta'\overline{\mathcal{E}}\mathcal{W}_\eta U^*\| &\leq O_{P^*}\left(\|\frac{1}{np}\beta'\overline{\mathcal{E}}\|\|\widehat{U}-\overline{U}\|\right)+\|\frac{1}{np}\beta'\overline{\mathcal{E}}\mathcal{W}_\eta\bar{U}^*\| \\
&\leq O_{P^*}(\frac{1}{\sqrt{p}})\frac{1}{n}\|\Delta_1+\Delta_2\|+\|\frac{1}{np}\beta'\overline{\mathcal{E}}\mathcal{W}_\eta\bar{U}^*\| \\
&\leq O_{P^*}(\frac{1}{\sqrt{p}})(\frac{1}{\sqrt{n}}+\frac{1}{\sqrt{p}}+\sqrt{\Delta_4})+\|\frac{1}{np}\beta'\overline{\mathcal{E}}\mathcal{W}_\eta\bar{U}^*\| \\
&= O_{P^*}(\frac{1}{p}+\frac{1}{\sqrt{np}}+\frac{1}{\sqrt{p}}\sqrt{\Delta_4}). \\
\|\frac{1}{n}\overline{F}'\mathcal{W}_\eta U^*\| &\leq \|\frac{1}{n}\overline{F}\|\|\widehat{U}-\overline{U}\|+\|\frac{1}{n}\overline{F}'\mathcal{W}_\eta\bar{U}^*\| \\
&\leq O_{P^*}(\frac{1}{\sqrt{n}}+\frac{1}{\sqrt{p}}+\sqrt{\Delta_4}).
\end{aligned}$$

Also, $\sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\beta' R\mathcal{W}_\eta \bar{U}^*\| = O_{P^*}(\Delta_4)$ so

$$\begin{aligned}\sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\beta' R\mathcal{W}_\eta \bar{U}^*\| &\leq \sup_{\eta\in[0,2\pi]}\|\frac{1}{np}\beta' R\mathcal{W}_\eta (\bar{U}^*-U^*)\| + O_{P^*}(\Delta_4)\\ &\leq O_{P^*}(\Delta_4+\frac{1}{n}\|\widehat{U}-\overline{U}\|^2) = O_{P^*}(\Delta_4+\frac{1}{n}+\frac{1}{p}).\end{aligned}$$

□

Lemma 7.7. *Suppose assumptions SA1 and A2-A3 hold. Then,*
(i)

$$\begin{aligned}\|\frac{1}{n}\widehat{F}'U^* - \frac{1}{n}H'\overline{F}'\bar{U}^*\| &= O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5),\\ \|\frac{1}{n}\widehat{F}'U^*\| &= O_{P^*}(\frac{1}{\sqrt{n}}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5),\end{aligned}$$

(ii)

$$\begin{aligned}&\sup_{\eta\in[0,2\pi]}\frac{1}{n}\|\overline{F}'\mathcal{W}_\eta U^* - \overline{F}'\mathcal{W}_\eta \bar{U}^*\| = O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4),\\ &\frac{1}{n}\overline{F}'\mathcal{W}_\eta\Delta_2^* = \frac{1}{n}\overline{F}'\mathcal{W}_\eta\widehat{F}(\widehat{F}'\widehat{F})^{-1}H'\overline{F}'\bar{U}^* + O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5),\end{aligned}$$

where $\Delta_5 := \sqrt{\frac{1}{pn}\sum_t\sum_j \|\overline{f}_t\|^2 r_{t,j}^2}$.

Proof. (i) We have $\frac{1}{n}\widehat{F}'U^* = \frac{1}{n}\widehat{F}'(U^*-\bar{U}^*) + \frac{1}{n}\widehat{F}'\bar{U}^*$. Similar to the proof of (54),

$$\frac{1}{n}\widehat{F}'\bar{U}^* = \frac{1}{n}H'\overline{F}'\bar{U}^* + O_{P^*}(\frac{1}{n}+\frac{1}{\sqrt{np}}+\frac{1}{n\sqrt{p}}\|R\|+\|\frac{1}{pn}\widehat{\beta}'RU\|).$$

In addition,

$$\frac{1}{n}\widehat{F}'(U^*-\bar{U}^*) \leq \|\frac{1}{n}(\widehat{F}-\overline{F}H)\|\|\widehat{U}-\overline{U}\| + \frac{1}{n}H'\overline{F}'(U^*-\bar{U}^*).$$

To bound $\frac{1}{n}H'\overline{F}'(U^*-\bar{U}^*)$, we have

$$\begin{aligned}\frac{1}{n}\overline{F}'(U^*-\bar{U}^*) &= \frac{1}{n}\sum_t \overline{f}_t\eta_t^*(\widehat{u}_t-u_t)\\ &= -\frac{1}{n}\sum_t \overline{f}_t\eta_t^*(\widehat{f}_t-H'\overline{f}_t)'\widehat{\gamma} - \frac{1}{n}\sum_t \overline{f}_t\eta_t^*\overline{f}_t' H'(\widehat{\gamma}-H^{-1}\gamma).\end{aligned}$$

We focus on the first term on the right hand side.

$$\|\frac{1}{n}\sum_t \overline{f}_t\eta_t^*(\widehat{f}_t - H'\overline{f}_t)'\| \leq \|\frac{1}{pn}\sum_t \overline{f}_t\eta_t^*\overline{\epsilon}_t'\widehat{\beta}\| + \|\frac{1}{pn}\sum_t \overline{f}_t\eta_t^* R_t'\widehat{\beta}\| := a+b,$$

$$\begin{aligned}
\mathbb{E}^*a^2 &= \sum_{j,k}\frac{1}{p^2n^2}\sum_t \mathrm{Var}^*(\overline{f}_{t,j}\eta_t^*\overline{\epsilon}_t'\widehat{\beta}_k) \\
&\leq \frac{2}{p^2n^2}\sum_t \|\overline{f}_t\|^2\|\overline{\epsilon}_t\|^2\|\widehat{\beta}-\beta H_\beta\|^2 + \frac{1}{p^2n^2}\sum_t \|\overline{f}_t\|^2\|\overline{\epsilon}_t'\beta H_\beta\|^2 \\
a &\leq O_{P^*}(\frac{1}{n}+\frac{1}{p\sqrt{n}}+\frac{1}{\sqrt{np}}+\frac{1}{n\sqrt{p}}\|R\|), \\
\mathbb{E}^*b^2 &\leq \frac{1}{pn^2}\sum_t\sum_j \|\overline{f}_t\|^2 r_{t,j}^2 \\
b &\leq O_{P^*}(\frac{1}{\sqrt{n}}\Delta_5).
\end{aligned} \tag{71}$$

Together:

$$\|\frac{1}{n}\sum_t \overline{f}_t\eta_t^*(\widehat{f}_t - H'\overline{f}_t)'\| \leq O_{P^*}(\frac{1}{n}+\frac{1}{p\sqrt{n}}+\frac{1}{\sqrt{np}}+\frac{1}{n\sqrt{p}}\|R\|+\frac{1}{\sqrt{n}}\Delta_5).$$

Also $\|\frac{1}{n}\sum_t \overline{f}_t\eta_t^*\overline{f}_t'\| = O_{P^*}(n^{-1/2})$. Thus, by Lemma 7.1,

$$\begin{aligned}
\frac{1}{n}\overline{F}'(U^*-\bar{U}^*) &= -\frac{1}{n}\sum_t \overline{f}_t\eta_t^*(\widehat{f}_t - H'\overline{f}_t)'\widehat{\gamma} - \frac{1}{n}\sum_t \overline{f}_t\eta_t^*\overline{f}_t'H'(\widehat{\gamma}-H^{-1}\gamma) \\
&\leq O_{P^*}(\frac{1}{n}+\frac{1}{p\sqrt{n}}+\frac{1}{\sqrt{np}}+\frac{1}{n\sqrt{p}}\|R\|+\frac{1}{\sqrt{n}}\Delta_5+\frac{1}{\sqrt{n}}\Delta_4) \\
\frac{1}{n}\widehat{F}'(U^*-\bar{U}^*) &\leq \|\frac{1}{n}(\widehat{F}-\overline{F}H)\|\|\widehat{U}-U\| + \frac{1}{n}H'\overline{F}'(U^*-\bar{U}^*) \\
&\leq O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5).
\end{aligned} \tag{72}$$

Hence

$$\begin{aligned}
\frac{1}{n}\widehat{F}'U^* &= \frac{1}{n}\widehat{F}'(U^*-\bar{U}^*) + \frac{1}{n}\widehat{F}'\bar{U}^* \\
&= \frac{1}{n}H'\overline{F}'\bar{U}^* + O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5).
\end{aligned}$$

This also shows $\|\frac{1}{n}\widehat{F}'U^*\| = O_{P^*}(\frac{1}{\sqrt{n}}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5)$, because $\frac{1}{n}H'\overline{F}'\bar{U}^* = O_{P^*}(\frac{1}{\sqrt{n}})$.

(ii) Let $\overline{F}^*$ be $n\times K$ matrix of $\overline{f}_t\eta_t^*$. Then, $\frac{1}{n}\overline{F}'\mathcal{W}_\eta U^* = \frac{1}{n}\overline{F}'\mathcal{W}_\eta \bar{U}^* + \Delta_6^*$, where

$$\begin{aligned}\Delta_6^* &= \frac{1}{n}\sum_t \overline{f}_t\mathcal{W}_\eta(t/n)\eta_t^*(\widehat{u}_t-u_t) = \frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta(\widehat{U}-U)\\ &= -\frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\Delta_1 - \frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\Delta_2.\end{aligned}$$

For $\frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\Delta_1$, by the same proof as that of Lemma 7.2,

$$\sup_\eta\|\frac{1}{pn}\overline{F}^{*'}\mathcal{W}_\eta\overline{\mathcal{E}}'\beta\| = O_{P^*}(\frac{1}{\sqrt{pn}}),\quad \sup_\eta\|\frac{1}{pn}\overline{F}^{*'}\mathcal{W}_\eta\overline{\mathcal{E}}'\| = O_{P^*}(\frac{1}{\sqrt{pn}}).$$

Hence, by the same proof of (58),

$$\frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\Delta_1 \le O_{P^*}(\frac{1}{\sqrt{np}}+\frac{1}{n}+\frac{1}{pn}\|R\|^2).$$

For $\frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\Delta_2$, still by the same proof as that of (58), $\mathbb{E}^*\left(\frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\overline{F}\right)=0$ and it is tight, so $\sup_\eta \frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\overline{F} = O_{P^*}(n^{-1/2})$. Hence

$$\begin{aligned}\frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\Delta_2 &= \frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\overline{F}H(\widehat{\gamma}-H^{-1}\gamma)\\ &= \frac{1}{n}\overline{F}^{*'}\mathcal{W}_\eta\overline{F}H(\widehat{F}'\widehat{F})^{-1}H'\overline{F}'U + O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4)\\ &= O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4)\\ \frac{1}{n}\overline{F}'\mathcal{W}_\eta\Delta_2^* &= \frac{1}{n}\overline{F}'\mathcal{W}_\eta\widehat{F}(\widehat{\gamma}^*-\widehat{\gamma}) = \frac{1}{n}\overline{F}'\mathcal{W}_\eta\widehat{F}(\widehat{F}'\widehat{F})^{-1}\widehat{F}'U^*\\ &= \frac{1}{n}\overline{F}'\mathcal{W}_\eta\widehat{F}(\widehat{F}'\widehat{F})^{-1}H'\overline{F}'\bar{U}^* + O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5).\end{aligned}$$

$\square$

Lemma 7.8. *Suppose assumptions SA1 and A2-A3 hold. Uniformly in η:*

$$\frac{1}{\sqrt{n}}\widehat{F}_k'\mathcal{W}_\eta\widehat{U}^* = \frac{1}{\sqrt{n}}\sum_t Z_t(\eta,k)\eta_t^* + O_{P^*}(\frac{1}{n}+\frac{1}{p}+\Delta_4+\frac{1}{\sqrt{n}}\Delta_5).$$

where we recall

$$\Gamma_t^n(\eta,k) := e_k'\bar{H}'w_\eta(t/n)+\bar{\lambda}_k(\eta),\quad Z_t(\eta,k) := \Gamma_t^n(\eta,k)'\overline{f}_t\overline{u}_t.$$

Proof. We start the proof with showing that

$$\Delta_3^* := \frac{1}{n}(\widehat{F}_k - \overline{F}H_k)'\mathcal{W}_\eta \widehat{U}^* = O_{P^*}(\frac{1}{n} + \frac{1}{p} + \Delta_4).$$

Because $\widehat{F} - \overline{F}H = \overline{\mathcal{E}}'\widehat{\beta}/p + R'\widehat{\beta}/p$ and $\widehat{U}^* - U^* = -\Delta_2^*$, we have

$$\Delta_3^* = -\frac{1}{n}(\widehat{F}_k - \overline{F}H_k)'\mathcal{W}_\eta \Delta_2^* + \frac{1}{np}e_k'\widehat{\beta}'(\overline{\mathcal{E}} + R)\mathcal{W}_\eta U^*.$$

Uniformly in η,

$$\begin{aligned}
\frac{1}{n}\|\Delta_2^*\|^2 &\leq O_P(\|\widehat{\gamma}^* - \widehat{\gamma}\|^2) \leq O_P(\|\frac{1}{n}\widehat{F}'U^*\|^2) = O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4) \\
\frac{1}{np}e_k'\widehat{\beta}'(\overline{\mathcal{E}} + R)\mathcal{W}_\eta U^* &\leq \frac{1}{np}O_P(\|\widehat{\beta} - \beta H_\beta\|\|\overline{\mathcal{E}}\|\|U^*\|) + \|\frac{1}{np}H_\beta'\beta'\overline{\mathcal{E}}\mathcal{W}_\eta U^*\| \\
&\quad + O_P(\frac{1}{p}\|\widehat{\beta} - \beta H_\beta\|^2 + \frac{1}{np}\|R\|^2) + \frac{1}{np}e_k'H_\beta'\beta' R\mathcal{W}_\eta U^* \\
&\leq O_P(\frac{1}{n} + \frac{1}{p} + \Delta_4) \\
\|\Delta_3^*\| &\leq O_{P^*}(\frac{1}{n}\|\widehat{F} - \overline{F}H\|^2 + \frac{1}{n}\|\Delta_2^*\|^2 + \frac{1}{n} + \frac{1}{p} + \Delta_4) \\
&= O_{P^*}(\frac{1}{n} + \frac{1}{p} + \Delta_4).
\end{aligned}$$

We are now ready to prove the result of the lemma. Recall the notation from the proof of Theorem 3.1:

$$\lambda_k(\eta)' = -e_k'H'\overline{F}'\mathcal{W}_\eta \overline{F}H(\widehat{F}'\widehat{F})^{-1}H'.$$

We have $\widehat{\gamma}^* - \widehat{\gamma} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'U^*$. By Lemmas 7.7 and 7.8 as well as the result for Δ_3^* above, we have that

$$\begin{aligned}
\frac{1}{n}\widehat{F}_k'\mathcal{W}_\eta \widehat{U}^* &= \Delta_3^* + \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta U^* - \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta \Delta_2^* \\
&= \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta \bar{U}^* - H_k'\frac{1}{n}\overline{F}'\mathcal{W}_\eta \widehat{F}(\widehat{F}'\widehat{F})^{-1}H'\overline{F}'\bar{U}^* \\
&\quad + O_{P^*}(\frac{1}{n} + \frac{1}{p} + \Delta_4 + \frac{1}{\sqrt{n}}\Delta_5) \\
&= \frac{1}{n}H_k'\overline{F}'\mathcal{W}_\eta \bar{U}^* + \frac{1}{n}\lambda_k(\eta)'\overline{F}'\bar{U}^* + O_{P^*}(\frac{1}{n} + \frac{1}{p} + \Delta_4 + \frac{1}{\sqrt{n}}\Delta_5) \\
&= \frac{1}{n}\sum_t Z_t(\eta, k)\eta_t^* + O_{P^*}(\frac{1}{n} + \frac{1}{p} + \Delta_4 + \frac{1}{\sqrt{n}}\Delta_5).
\end{aligned}$$

□

Lemma 7.9. *Suppose assumptions SA1 and A2-A3 hold. Then,*

$$\frac{1}{\sqrt{n}}\sum_t Z_t(\eta,k)\eta_t^* \xrightarrow{\mathcal{L}|\mathcal{F}} \mathbb{Z}_K(\eta),$$

where $\mathbb{Z}_K(\eta)$ is a K-dimensional vector-valued process, defined on an extension of the original probability space, which is $\mathcal{F}$-conditionally a zero-mean Gaussian process, with $\mathcal{F}$-conditional covariance kernel being the following $K \times K$ matrix

$$\mathbb{E}(\mathbb{Z}_K(\eta_1)\mathbb{Z}_K(\eta_2)'|\mathcal{F}) = \frac{1}{T}\int_0^T \Gamma_t(\eta_1)'V_t\Gamma_t(\eta_2)dt, \quad V_t := (\sigma_t^g)^2\Lambda_t\Lambda_t'.$$

Proof. The proof consists of establishing finite dimensional convergence and tightness of the sequence. For the former, we can apply Theorem VIII.5.25 of Jacod and Shiryaev (2003). It involves showing the convergences in (66)-(67). They can be established in exactly the same way as in the proof of Lemma 7.5.

To prove the tightness, note that by the definition of the variables η_t^*, we have

$$\mathbb{E}^*\left(|\frac{1}{\sqrt{n}}\sum_t (Z_t(\eta_1,k) - Z_t(\eta_2,k))\eta_t^*|^2\right) = \frac{1}{n}\sum_t |Z_t(\eta_1,k) - Z_t(\eta_2,k)|^2. \qquad (73)$$

From here the proof of tightness proceeds in a similar way to the one in the proof of Lemma 7.5. □

We are now ready to complete the proof of the theorem. Given the bounds in (47)-(50) as well as Lemma 7.1, we have that

$$\sqrt{n}\Delta_4 + \Delta_5 = o_{P^*}(1).$$

due to the condition in (22). Applying Lemma 7.8, we therefore have that

$$\frac{1}{\sqrt{n}}\widehat{F}_k'\mathcal{W}_\eta\widehat{U}^* = \frac{1}{\sqrt{n}}\sum_t Z_t(\eta,k)\eta_t^* + o_{P^*}(1).$$

From here, the result of the theorem follows by application of Lemma 7.9.

7.6 Proof of Theorem 3.4

From Theorems 3.1 and 3.3, we have

$$\widetilde{\mathcal{S}} := \sum_k n \sup_{\eta_k\in[0,2\pi]} \|\widehat{\mathbb{S}}_k(\eta_k)\|^2 + \epsilon\,\zeta^2 \quad \xrightarrow{\mathcal{L}-s} \quad Z := \sum_k \sup_{\eta_k}\|\mathbb{Z}_k(\eta_k)\|^2 + \epsilon\,\zeta^2, \text{ (null)}$$

$$\widetilde{\mathcal{S}}^* := \sum_k n \sup_{\eta_k \in [0,2\pi]} \|\widehat{\mathbb{S}}_k^*(\eta_k)\|^2 + \epsilon\,(\zeta^*)^2 \quad \xrightarrow{\mathcal{L}|\mathcal{F}} \quad Z, \quad \text{(both null and alternative)},$$

where $(\mathbb{Z}_1(\eta_k), ..., \mathbb{Z}_K(\eta_K))$ is a $\mathcal{F}$-conditionally multivariate Gaussian process and ζ^* is a standard normal variable defined on an extension of the original probability space, which is independent of ζ and $\mathcal{F}$. Let q^* be the τ-th upper quantile of $\widetilde{\mathcal{S}}^*$ so that

$$\mathbb{P}\left(\widetilde{\mathcal{S}}^* > q^*\right) = \tau.$$

We now show that the test has the correct asymptotic size, i.e., that $\mathbb{P}(\widetilde{\mathcal{S}} > q^*) \to \tau$ under the null. To this end, first note that $\widetilde{\mathcal{S}}^* \xrightarrow{\mathcal{L}|\mathcal{F}} Z$ implies $q^* \xrightarrow{\mathbb{P}} \widetilde{q}$, for $\widetilde{q}$ being the τ-th upper quantile of Z, by e.g., Lemma 21.2 of Van der Vaart (2000), since the $\mathcal{F}$-conditional law of Z has no atom. The latter is due to the fact that Z is a convolution of two independent random variables, one of which has a probability density. Next, for any $\delta > 0$,

$$\begin{aligned}
\mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} + \delta) &\leq \mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} + \delta, |q^* - \widetilde{q}| < \delta) + \mathbb{P}(|q^* - \widetilde{q}| > \delta) \leq \mathbb{P}(\widetilde{\mathcal{S}} > q^*) + o(1), \\
\mathbb{P}(\widetilde{\mathcal{S}} > q^*) &\leq \mathbb{P}(\widetilde{\mathcal{S}} > q^*, |q^* - \widetilde{q}| < \delta) + o(1) \leq \mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} - \delta) + o(1).
\end{aligned}$$

Therefore, $\mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} + \delta) + o(1) \leq \mathbb{P}(\widetilde{\mathcal{S}} > q^*) \leq \mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} - \delta) + o(1)$, which implies

$$\begin{aligned}
\left|\mathbb{P}(\widetilde{\mathcal{S}} > q^*) - \tau\right| &\leq \left|\mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} + \delta) - \tau\right| + \left|\mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} - \delta) - \tau\right| + o(1) \\
&\leq \left|\mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} + \delta) - \mathbb{P}(Z > \widetilde{q} + \delta)\right| + \left|\mathbb{P}(\widetilde{\mathcal{S}} > \widetilde{q} - \delta) - \mathbb{P}(Z > \widetilde{q} - \delta)\right| \\
&\quad + |\mathbb{P}(Z > \widetilde{q} + \delta) - \mathbb{P}(Z > \widetilde{q})| + |\mathbb{P}(Z > \widetilde{q} - \delta) - \mathbb{P}(Z > \widetilde{q})| + o(1) \\
&\leq o(1) + C\delta,
\end{aligned}$$

for some $C > 0$ that depends on the $\mathcal{F}$-conditional density of Z. The last inequality follows from the stable convergence result for $\widetilde{\mathcal{S}}$. Because $\delta > 0$ is arbitrarily small, $\mathbb{P}(\widetilde{\mathcal{S}} > q^*) \to \tau$. Hence, under the null,

$$\mathbb{P}(p^* < \tau | \mathbb{H}_0) = \mathbb{P}(\widetilde{\mathcal{S}} > q^* | \mathbb{H}_0) \to \tau.$$

Under the alternative,

$$\widehat{\mathbb{S}}(\bar{\eta}) \xrightarrow{\mathbb{P}} \bar{\mathbb{S}}(\bar{\eta}) := \bar{H}'\mathcal{S}_1(\bar{\eta}), \quad \mathcal{S}_1(\bar{\eta}) := \frac{1}{T}\int_0^T \operatorname{diag}(w_{\eta_1}(t/T), ..., w_{\eta_K}(t/T))d\langle f, u\rangle_t.$$

uniformly in $\bar{\eta} \in [0, 2\pi]^K$. Since $\langle f, u\rangle_t$ is not identically zero on $[0, T]$ and since $\bar{H}$ is full rank, the same holds true for $\bar{H}'\langle f, u\rangle_t$ as well. That means that at least one of its elements is not identically zero on $[0,T]$. This, however, implies that for that element, say k, $\sup_{\eta\in[0,2\pi]}\left|e_k'\bar{H}'\int_0^T \mathcal{W}_\eta(t/T)d\langle f,u\rangle_t\right|$ is strictly bigger than zero by an application of Theorem 1 of Bierens and Ploberger (1997). Thus, $\widetilde{\mathcal{S}} > C$ with probability approaching one, for some $C > 0$. Meanwhile, $q^* = O_P(1)$ under the alternative, hence

$$\mathbb{P}(p^* < \tau|\mathbb{H}_1) = \mathbb{P}(\widetilde{\mathcal{S}} > q^*|\mathbb{H}_1) \to 1.$$

References

Aït-Sahalia, Y., I. Kalnina, and D. Xiu (2020). High-frequency factor models and regressions. *Journal of Econometrics 216*(1), 86–105.

Aït-Sahalia, Y. and D. Xiu (2017). Using principal component analysis to estimate a high dimensional factor model with high-frequency data. *Journal of Econometrics 201*(2), 384–399.

Aït-Sahalia, Y. and D. Xiu (2019). Principal component analysis of high-frequency data. *Journal of the American Statistical Association 114*(525), 287–303.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and G. Wu (2006). Realized beta: Persistence and predictability. In *Econometric Analysis of Financial and Economic Time Series*, pp. 1–39. Emerald Group Publishing Limited.

Andersen, T. G., R. Riva, M. Thyrsgaard, and V. Todorov (2022). Intraday cross-sectional distributions of systematic risk. *Journal of Econometrics*.

Andersen, T. G., M. Thyrsgaard, and V. Todorov (2021). Recalcitrant betas: Intraday variation in the cross-sectional dispersion of systematic risk. *Quantitative Economics 12*(2), 647–682.

Andreou, E., P. Gagliardini, E. Ghysels, and M. Rubin (2019). Inference in group factor models with an application to mixed-frequency data. *Econometrica 87*(4), 1267–1305.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *The Journal of Finance 61*(1), 259–299.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica 72*(3), 885–925.

Bierens, H. J. and W. Ploberger (1997). Asymptotic theory of integrated conditional moment tests. *Econometrica*, 1129–1151.

Billingsley, P. (2013). *Convergence of probability measures.* John Wiley & Sons.

Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business & Economic Statistics Section, 1976*.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*(3), 623–685.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies 22*(11), 4463–4492.

Buchta, M. (2024). Roughing up risk premia: A robust strategy for the estimation of continuous and discontinuous risk prices. *Available at SSRN 4739685*.

Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure and mean-variance analysis in large asset markets. *Econometrica 51*, 1305–1324.

Chen, D., P. A. Mykland, and L. Zhang (2024). Realized regression with asynchronous and noisy high frequency and high dimensional data. *Journal of Econometrics 239*(2), 105446.

Cochrane, J. H. (2005). *Asest Pricing.* Princeton university press.

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

Connor, G. and R. A. Korajczyk (1991). The attributes, behavior, and performance of us mutual funds. *Review of Quantitative Finance and Accounting 1*, 5–26.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 987–1007.

Fan, J., Y. Ke, and Y. Liao (2021). Augmented factor models with applications to validating market risk factors and forecasting bond risk premia. *Journal of Econometrics 222*(1), 269–294.

Fan, J., Y. Liao, and W. Wang (2016). Projected principal component analysis in factor models. *Annals of Statistics 44*(1), 219–254.

Fernández-Villaverde, J., P. Guerrón-Quintana, J. F. Rubio-Ramirez, and M. Uribe (2011). Risk matters: The real effects of volatility shocks. *American Economic Review 101*(6), 2530–2561.

Ferson, W. E. and C. R. Harvey (1999). Conditioning variables and the cross section of stock returns. *The Journal of Finance 54*(4), 1325–1360.

Fortin, A.-P., P. Gagliardini, and O. Scaillet (2023). Latent factor analysis in short panels. *arXiv preprint arXiv:2306.14004*.

Gagliardini, P., E. Ossola, and O. Scaillet (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica 84*(3), 985–1046.

Gagliardini, P., E. Ossola, and O. Scailletd (2020). Estimation of large dimensional conditional factor models in finance. *Handbook of Econometrics*, 219.

Giglio, S. and D. Xiu (2021). Asset pricing with omitted factors. *Journal of Political Economy 129*(7), 1947–1990.

Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica 55*(3), 587–613.

Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics 119*(2), 249–283.

Ibragimov, I. A. and R. Z. Has'minskii (2013). *Statistical estimation: asymptotic theory*, Volume 16. Springer Science & Business Media.

Jacod, J. and P. E. Protter (2011). *Discretization of processes*, Volume 67. Springer Science & Business Media.

Jacod, J. and A. Shiryaev (2003). *Limit Theorems For Stochastic Processes* (2nd ed.). Berlin: Springer-Verlag.

Jagannathan, R. and Z. Wang (1996). The conditional capm and the cross-section of expected returns. *The Journal of Finance 51*(1), 3–53.

Kelly, B. T., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics 134*(3), 501–524.

Li, J., V. Todorov, and G. Tauchen (2017). Jump regressions. *Econometrica 85*(1), 173–195.

Liao, Y. and V. Todorov (2024). Changes in the span of systematic risk exposures. *Quantitative Economics 15*(3), 817–847.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics 47*, 13—-37.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *The Annals of Statistics 21*(1), 255–285.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari 64*(19-47), 44.

Mancini, C. and F. Gobbi (2012). Identifying the brownian covariation from the co-jumps given discrete observations. *Econometric Theory 28*(2), 249–273.

Mykland, P. A. and L. Zhang (2006). Anova for diffusions and ito processes. *The Annals of Statistics 34*(4), 1931–1963.

Mykland, P. A. and L. Zhang (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica 77*(5), 1403–1445.

Pelger, M. (2019). Large-dimensional factor modeling based on high-frequency observations. *Journal of Econometrics 208*(1), 23–42.

Pelger, M. (2020). Understanding systematic risk: A high-frequency approach. *The Journal of Finance 75*(4), 2179–2220.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory 13*, 341–360.

Shanken, J. (1990). Intertemporal asset pricing: An empirical investigation. *Journal of Econometrics 45*(1-2), 99–120.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.

Todorov, V. (2019). Nonparametric spot volatility from options. *The Annals of Applied Probability 29*(6), 3590–3636.

Todorov, V. and T. Bollerslev (2010). Jumps and betas: A new framework for disentangling and estimating systematic risks. *Journal of Econometrics 157*(2), 220–235.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, Volume 3. Cambridge university press.

Vershynin, R. (2018). *High-dimensional probability: An introduction with applications in data science*, Volume 47. Cambridge university press.

Zaffaroni, P. (2019). Factor models for conditional asset pricing. *Available at SSRN 3398169*.