Inference Theory for Volatility Functional Dependencies*

Jia Li† and Viktor Todorov‡ and George Tauchen§

January 23, 2016

Abstract

We develop inference theory for models involving possibly nonlinear transforms of the elements of the spot covariance matrix of a multivariate continuous-time process observed at high frequency. The framework can be used to study the relationship among the elements of the latent spot covariance matrix and processes defined on the basis of it such as systematic and idiosyncratic variances, factor betas and correlations on a fixed interval of time. The estimation is based on matching model-implied moment conditions under the occupation measure induced by the spot covariance process. We prove consistency and asymptotic mixed normality of our estimator of the (random) coefficients in the volatility model and further develop model specification tests. We apply our inference methods to study variance and correlation risks in nine sector portfolios comprising the S&P 500 index. We document sector-specific variance risks in addition to that of the market and time-varying heterogeneous correlation risk among the market-neutral components of the sector portfolio returns.

Keywords: high-frequency data, occupation measure, semimartingale, specification test, stochastic volatility.

JEL classification: C51, C52, G12.

*We would like to thank the editor, an associate editor and two anonymous referees for many useful comments and suggestions. We also thank Tim Bollerslev, Federico Bugni, Marine Carrasco, Silvia Goncalves, Wolfgang Hardle, Jean Jacod, Ilze Kalnina, Matthew Masten, Andrew Patton and Ya'acov Ritov for helpful discussions. Financial support from the NSF under grants SES-1227448 and SES-1326819 (Li) and SES-0957330 (Todorov) is gratefully acknowledged.

†Department of Economics, Duke University, Durham, NC 27708; e-mail: jl410@duke.edu.

‡Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@northwestern.edu.

§Department of Economics, Duke University, Durham, NC 27708; e-mail: george.tauchen@duke.edu.

1 Introduction

Economic models often place exact restrictions across the realizations of a set of random variables. One case in point is an affine term structure model for bond prices that constrains all bond yields to lie along a very low dimensional manifold; see, for example, Singleton (2006) and citations therein. More generally, factor models specify the pricing kernel as a function of a low-dimensional factor process. Combining this structure with a model for aggregate asset payoffs implies low-dimensional factor structure for the conditional distribution, and thereby the first and second conditional moments, of asset returns and derivative prices.

In a broad sense, dimension reduction is also often imposed for the purposes of parsimony in modeling high-dimensional objects so as to mitigate the statistical and/or computational complexity of econometric models. One example is the use of diffusion index in macroeconomic forecasting (see Stock and Watson (2002) and references therein). Other examples arise in models of the stochastic covariance matrix of a multivariate process, where dimension-reduction restrictions may take the form of time-invariant correlations (Bollerslev (1990)) or equal stochastic correlations among multiple time series (Engle and Kelly (2012)). Similar strategies have also been used in moderately high-dimensional models for nonlinear dependence, such as copula models (Oh and Patton (2013)).

The primary focus of this paper is the estimation and testing for such model restrictions among the elements of the spot covariance matrix of a multivariate process of asset returns. The estimation is based on high-frequency (intraday) observations of a multivariate Itô semimartingale process on a fixed time interval with mesh of the observation grid shrinking to zero. For this process, we are interested in estimating and testing pathwise models that impose time-invariant (over the observation interval) relations among possibly nonlinear transforms (e.g. beta, correlation and idiosyncratic variance) of the spot covariance matrix.

The statistical uncertainty in our setting arises from the fact that the spot covariance matrix is not directly observed and needs to be estimated, or "measured," from the discrete observations. The measurement error makes the functional relationship among latent risks hold only approximately for their estimated counterparts. Nevertheless, as the sampling frequency increases, the error vanishes asymptotically, so that we can uncover and rigorously test model restrictions for the latent risks. To the best of our knowledge, our test is the first general method for testing such model restrictions in the high-frequency setting, while allowing for general forms of nonstationarity and dependence in the data.

In the first part of our theoretical analysis, we consider estimation based on forming moment conditions under the covariance occupation measure which are implied by our pathwise volatility model. We then construct sample analogues of these moment conditions by plugging in local nonparametric estimators of volatility formed over blocks of high-frequency price increments with asymptotically decreasing length of each of the blocks. This is similar to Jacod and Rosenbaum (2013) who use block volatility estimates to construct estimators for integrated nonlinear functions of volatility. Finally, we weight the moment conditions using a feasible weight matrix and form a quadratic-form objective function that our estimator minimizes. We derive the limit behavior of our estimator not only in the case when the model is correctly specified but also in the case of model misspecification, and further provide feasible estimates for the standard errors of the parameters in the model.

Our estimator of the parameters of the volatility model can be viewed as an analogue to the classical minimum distance type estimators with several important differences. First, the moment conditions in our case are formed under the occupation measure and, hence, they hold for the observed path but not necessarily for the invariant distribution of the volatility process; indeed, the invariant distribution is not even required to exist. The strategy of framing inference procedures in terms of moments under the occupation measure opens the possibility of systematically reincarnating many classical moment-based econometric procedures (e.g. Hansen (1982)), which are framed under the probability measure, for conducting inference for multivariate volatility models. Second, the asymptotic behavior of the estimator is equivalent to that generated by observing the moment condition with the true value of the spot covariance matrix plus a Gaussian martingale defined on an extension of the original probability space. This Gaussian martingale has quadratic variation that is adapted to the original filtration and shrinks asymptotically at order Δ_n, where Δ_n is the length of the high-frequency interval. Our estimation problem is thus similar to the problem of estimating a signal with asymptotically shrinking Gaussian noise, see, for example, section VII.4 in Ibragimov and Has'minskii (1981). Third, the limit law of our estimator is mixed Gaussian which means that the precision of estimation will typically vary depending on the particular realization.

The second, and perhaps more important, part of our analysis is specification testing for the pathwise volatility model. Since the model holds almost everywhere in time over the fixed time interval, designing a test based on the distance from zero of the model-based moment conditions under the covariance occupation measure is not sufficient. The reason is fairly intuitive: the covariance occupation measure does not preserve the information about the value of the spot covariance matrix at a particular point in time. For this reason, we introduce the concept of the weighted covariance occupation measure which, unlike the original occupation measure, allows to weigh differently the values of the spot covariance matrix at different points in time. We derive an empirical-process-type theory for an estimator of the weighted occupation measure. We use the latter to design a specification test for our pathwise volatility model by comparing the distance from zero of a set of moment conditions under a family of weighted occupation measures. We show that if the family of weight functions is chosen appropriately, our test statistic produces an asymptotically valid test.

Finally, we apply our inference theory to study the stochastic covariance structure on the industry level using S&P 500 sector index exchange-traded funds (ETFs). We specify and test models for the spot covariance matrix of the components of the industry returns that are orthogonal to the market portfolio. Our results show that not all variations in the stochastic variances of the industry portfolios can be accounted for by their sensitivity to market returns and the market variance. Some sectors like the Financials and Energy have independent sources of variance shocks in addition to that of the market. We further document nontrivial temporal variation in correlation in the market-neutral industry portfolio returns with nontrivial cross-sectional differences. The temporal variation in the market-neutral industry portfolio correlations suggests the presence of additional factors in the industry portfolio returns to span their risks.

The inference methods developed in the current paper are related with several strands of literature. First, our work is closely related to the literature on volatility estimation using high-frequency data. Early work mainly focuses on the estimation of the integrated variance (Barndorff-Nielsen and Shephard (2002), Andersen et al. (2003)) and covariance (Barndorff-Nielsen and Shephard (2004b)), which can be considered as the mean of the covariance occupation measure. The estimation of nonlinear transforms of the volatility has been considered by Barndorff-Nielsen et al. (2005), Jacod (2008), Mykland and Zhang (2009), Todorov and Tauchen (2012), Jacod and Rosenbaum (2013), Li et al. (2013), Kalnina and Xiu (2014) and Aït-Sahalia and Xiu (2015) among others. Similar to Jacod and Rosenbaum (2013) our estimation is based on local estimates of volatility, which are local versions of the truncated variation of Mancini (2001), over blocks of decreasing length. Unlike the above cited literature on the estimation of volatility functionals, our focus here is on the estimation and specification of pathwise models for the spot covariance matrix. On the technical level, this requires a derivation of empirical-process-type limit results for a family of weighted covariance measures which is new.[1] Second, in Li et al. (2013) we advocate the volatility

[1]In a concurrent work, Li and Xiu (2015) study regression-type problems for noisy semimartingales using the stochastic variance as an explanatory variable in conditional moment equality models that arise from derivative

occupation measure as a unifying framework for high-frequency based volatility estimation, but in Li et al. (2013) we focus only on the estimation of the volatility occupation time in a univariate setting, without weighting the observations, and importantly without deriving feasible central limit theorems. Third, our inference can be compared with the literature on estimating parametric volatility models using realized measures under joint in-fill and long span asymptotics, see, for example, Bollerslev and Zhou (2002), Barndorff-Nielsen and Shephard (2002), Corradi and Distaso (2006), Todorov (2009) and Todorov and Tauchen (2012). Unlike our setup here, the estimation in these papers is always parametric (at least about the stochastic volatility part of the observed process) and relies crucially on the error due to the discrete sampling being dominated by the empirical process type error due to time aggregation. In contrast, our estimation here is performed on a fixed span and does not involve full parametric specification of the volatility process. The pathwise volatility models of interest here hold for whole families of parametric models. The fixed span setting also allows us to accommodate general forms of nonstationarity and dependence in the data.

The paper is organized as follows. Section 2 presents the setting, three motivating examples, and a heuristic overview of our inference methods. The theory is developed in Section 3. Section 4 presents our empirical application. Section 5 concludes. All proofs are in the appendix. The working paper version of this paper contains a comprehensive simulation study that supports our asymptotic theory; to save space, we do not include the simulation results here.

2 The setting

We start with some notation that we are going to use throughout. All limits in the paper are for $n \to \infty$. We use $\xrightarrow{\mathbb{P}}$ to denote convergence in probability and use $\xrightarrow{\mathcal{L}\text{-}s}$ to denote stable convergence in law. For any matrix A, we use A_{ij} to denote its (i,j) element and $A^\intercal$ to denote its transpose. We sometimes identify the matrix A with its elements by writing $A = [A_{ij}]$. For a matrix valued process A_t, the notations $A_{ij,t}$ and $A_t^\intercal$ are interpreted similarly. For a matrix A and a differentiable function g, we denote $\partial_{jk} g(A) = \partial g(A)/\partial A_{jk}$ and $\partial^2_{jk,lm} g(A) = \partial^2 g(A)/\partial A_{jk} \partial A_{lm}$. If A and B are matrices with the same number of rows, we use (A, B) to denote a matrix with columns being those of A and B. The column vectorization operator is denoted by vec. The Kronecker product for matrices is denoted by $\otimes$; this notation is also used for the product of σ-fields. For any $q \in \mathbb{N}$, we denote the q dimensional identity matrix by I_q. The symbol $\equiv$ indicates equality by definition. We

pricing and market microstructure models. In contrast, the current paper focuses on the stochastic covariance matrix itself and conducts inference for pathwise restrictions on its behavior.

use $\|\cdot\|$ to denote the Euclidean norm on any finite-dimensional linear space. Additional conventions used throughout the paper are: the symbols $x, y, z, z^{\star}, \tilde{x}, \tilde{y}, w, g, h$ and $h^{\star}$ are reserved to denote various deterministic functions and blackboard bold letters such as $\mathbb{F}$, $\mathbb{V}$ and $\mathbb{B}$ are functions that act on deterministic functions. Composite notations such as $\mathbb{F}\mathbb{B}g$ and $\mathbb{F}\mathbb{V}(g,h)$ are understood as $\mathbb{F}[\mathbb{B}(g)]$ and $\mathbb{F}[\mathbb{V}(g,h)]$, respectively.

2.1 The spot covariance and the covariance occupation measure

The discretely observed process X is a d dimensional Itô semimartingale, defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, with the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} \delta(s,u)\,\mu(ds,du), \tag{2.1}$$

where the variable X_0 is $\mathcal{F}_0$-measurable, the instantaneous drift b_t is d dimensional càdlàg (i.e. right continuous with left limit) adapted, W is a d' dimensional Brownian motion, σ_t is a $d \times d'$ dimensional càdlàg adapted, $\delta : \Omega \times \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}^d$ is a predictable function and μ is a Poisson random measure with compensator $\nu(dt,du) = dt \otimes \lambda(du)$ for some σ-finite measure λ. The *spot covariance matrix* of X is then given by $c_t \equiv \sigma_t \sigma_t^{\intercal}$, which takes value in the space $\mathcal{M}_d$ consisting of d dimensional positive definite matrices. In this paper we are interested in estimating and testing functional relationships among various components of c_t, while recognizing the presence of the drift and the jump components of X.

The occupation measure induced by the spot covariance process over $[0,T]$ is defined as $\mathbb{F}(B) = \int_0^T 1_{\{c_s \in B\}} ds$ for any Borel subset $B \subseteq \mathbb{R}^{d\times d}$. From basic integration theory, it is equivalent (and more convenient) to consider the occupation measure $\mathbb{F}$ as a linear functional that acts on measurable functions: for any measurable function $g : \mathcal{M}_d \mapsto \mathbb{R}^{\dim(g)}$, we denote

$$\mathbb{F}g \equiv \int g(c)\mathbb{F}(dc) = \int_0^T g(c_s)ds. \tag{2.2}$$

In other words, the integrated g-transform of the spot covariance process can be thought of as the "mean" of g under $\mathbb{F}$.[2]

We suppose that X is observed at discrete times $i\Delta_n$, $i = 0, 1, \ldots,$ over a fixed time interval $[0,T]$ with $\Delta_n \to 0$ asymptotically and we further assume the following for X.

Assumption HF: The process X is a $\mathbb{R}^d$-valued Itô semimartingale with the form (2.1) such that the following conditions hold for a sequence $(T_m)_{m\geq 1}$ of stopping times increasing to infinity.

[2] To make the analogy exact, one may normalize the expression in (2.2) by T^{-1} or simply normalize $T = 1$. Here, we follow the convention in the literature on occupation measures (see, e.g., Geman and Horowitz (1980)) without using this normalization.

(a) For some constant $r \in (0,1)$, there exists a sequence of λ-integrable nonnegative functions $(\Gamma_m)_{m\geq 1}$ such that $\|\delta(\omega,t,u)\|^r \wedge 1 \leq \Gamma_m(u)$ for all $(\omega,u) \in \Omega \times \mathbb{R}$ and $t \leq T_m$.

(b) The process σ is also an Itô semimartingale with the form

$$\begin{aligned} vec(\sigma_t) &= vec(\sigma_0) + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s \\ &\quad + \int_0^t \int_{\mathbb{R}} \tilde{\delta}\left(s,u\right) 1_{\left\{\|\tilde{\delta}(s,u)\| \leq 1\right\}} (\mu - \nu)(ds,du) \\ &\quad + \int_0^t \int_{\mathbb{R}} \tilde{\delta}\left(s,u\right) 1_{\left\{\|\tilde{\delta}(s,u)\| > 1\right\}} \mu(ds,du), \end{aligned}$$

where $\tilde{b}_t$ and $\tilde{\sigma}_t$ are respectively dd' and $dd' \times d'$ dimensional optional locally bounded processes, and $\tilde{\delta}$ is a dd' dimensional predictable function. Moreover, there exists a sequence of λ-integrable nonnegative functions $(\widetilde{\Gamma}_m)_{m\geq 1}$ such that $\|\tilde{\delta}(\omega,t,u)\|^2 \wedge 1 \leq \widetilde{\Gamma}_m(u)$ for all $(\omega,u) \in \Omega \times \mathbb{R}$ and $t \leq T_m$.

(c) For a sequence of convex compact subsets $(\mathcal{K}_m)_{m\geq 1}$ of $\mathcal{M}_d$, $c_t \in \mathcal{K}_m$ for all $t \leq T_m$.

Assumption HF is very general and nests most of the multivariate continuous-time models used in economics and finance. The part of the assumption concerning the jump component J_t restricts the jump process to be of finite variation, but allowing it to have infinite activity. This assumption is perhaps unavoidable if one wants $\Delta_n^{-1/2}$ asymptotic mixed normality for our estimators introduced later on. Similar assumptions regarding the jumps have been employed in the context of estimating functionals of c_t from high-frequency data; see, for example, Jacod and Protter (2012) and many references therein. Assumption HF can be further relaxed to allow for infinite-variational jumps in X if one is only interested in consistency results. The Itô semimartingale assumption for σ is very general but it nevertheless rules out some stochastic volatility models of long-memory type that are driven by fractional Brownian motions (Comte and Renault (1998)). Finally, the localized support condition (c) facilitates the use of the spatial localization technique of Li et al. (2014); this allows us to avoid Jacod and Rosenbaum's (2013) polynomial growth condition on $g(\cdot)$, which is (prohibitively) restrictive for our applications.

2.2 Model restrictions on volatilities

Our main interest is to test model restrictions on the spot covariance matrix c_t, which we now formally describe. We consider a setting with $\bar{m}$ linear restrictions on possibly nonlinear transforms of c_t: for $m \in \{1,\ldots,\bar{m}\}$,

$$\tilde{y}_m(c_t) - \tilde{x}_m\left(c_t\right)^{\intercal} \theta_{m,0} = 0, \quad \text{for Lebesgue almost every } t \in [0,T], \tag{2.3}$$

where, for each m, $\tilde{y}_m : \mathcal{M}_d \mapsto \mathbb{R}$ and $\tilde{x}_m : \mathcal{M}_d \mapsto \mathbb{R}^{\dim(\theta_m)}$ are known deterministic transformations, and the vector $\theta_{m,0} : \Omega \mapsto \mathbb{R}^{\dim(\theta_m)}$ is an unknown parameter which is allowed to be random. We allow $\tilde{x}_m$ and $\theta_{m,0}$ to be empty, in which case the term $\tilde{x}_m(c_t)^\intercal \theta_{m,0}$ is considered absent from (2.3). In the general case with unknown $\theta_{m,0}$, we are also interested in its estimation and inference.

The key restriction imposed by (2.3) is that the transforms $\tilde{y}_m(\cdot)$ and $\tilde{x}_m(\cdot)$, as well as the parameter $\theta_{m,0}$, are all time-invariant. Hence, the model can be used to investigate whether the variation of $\tilde{y}_m(c_t)$ can be spanned by that of $\tilde{x}_m(c_t)$. Relationships between (transforms of) the elements of the covariance matrix c_t as in (2.3) arise in many applications in economics and finance, as we now illustrate with a few examples. Additional novel examples are studied in our empirical application; see Section 4.

Example (Factor Models): A continuous-time factor model for the d dimensional asset price process X can be written as $dX_t = Adf_t + d\widetilde{X}_t$, where A is a $d \times k$ constant factor loading matrix, f is a k dimensional factor process with $k < d$, and $\widetilde{X}$ is the residual component. The continuous martingale components of $\widetilde{X}$ are mutually orthogonal and are orthogonal to that of f.[3] The dynamics of the factors is given by $df_t = b_{f,t}dt + \Sigma_{f,t}^{1/2} dW_t + dJ_{f,t}$, where the factor spot covariance matrix process $\Sigma_{f,t}$ is normalized to be diagonal, W is a k dimensional standard Brownian motion and J_f is the jump part of f. The factors may be latent and we do not assume that they can be recovered from observing X. This setup covers many uses of factor models in asset pricing.[4] The factor structure implies linear restrictions on the off-diagonal elements of the spot covariance matrix c_t of the process X. Indeed, the time variation of all $d(d-1)/2$ off-diagonal elements are completely captured by the time variation in the k factor variances. To be explicit, we denote $\tilde{A} \equiv [A_{1j}A_{(i+1)j}]_{1 \le i,j \le k}$ and assume $\tilde{A}$ has full rank. We can then write $(\Sigma_{11,f,t}, \ldots, \Sigma_{kk,f,t})^\intercal = \tilde{A}^{-1}(c_{12,t}, \ldots, c_{1(k+1),t})^\intercal$, which implies the following time-invariant linear restrictions

$$c_{ij,t} = (A_{i1}A_{j1}, \ldots, A_{ik}A_{jk})\tilde{A}^{-1}(c_{12,t}, \ldots, c_{1(k+1),t})^\intercal, \qquad \text{for } i \neq j. \tag{2.4}$$

If the residual component, $\widetilde{X}$, is absent, the linear restrictions can be extended to also include the diagonal elements of c_t (i.e., spot variances): $c_{ij,t} = (A_{i1}A_{j1}, \ldots, A_{ik}A_{jk})\bar{A}^{-1}(c_{11,t}, \ldots, c_{kk,t})^\intercal$ for

[3]Two continuous (local) martingales are called orthogonal if their quadratic covariation process is identically zero, up to an evanescent set.

[4]Examples include models in the arbitrage pricing theory with exact factors for pricing individual stocks, as well as models of linkages between international stock markets (King et al. (1994)), although the exact factor structure is restrictive in high-dimensional settings (Fan et al. (2014)). Another important example is the term structure model where bond yields follow the above factor specification with the idiosyncratic component $\widetilde{X}$ absent (see, e.g., chapters 12 and 13 in Singleton (2006) and references therein). Since f is allowed to be latent, our setup is general enough to accommodate term structure models in which volatility is not spanned by the yield curve. We note finally that, factor models often impose structure on the drift and the jumps of X; estimating and testing model restrictions of these sorts, however, is out of the scope of the current paper.

all $1 \leq i, j \leq d$, where the matrix $\bar{A} \equiv [A_{ij}^2]_{1\leq i,j\leq k}$ is assumed to have full rank.

Example (Spot Correlation Models): Many multivariate models, including Bollerslev (1990) and Engle and Kelly (2012), impose restrictions on the spot correlation matrix. The simplest case is the continuous-time analogue of the constant conditional correlation model of Bollerslev (1990) which imposes that

$$\rho_t \equiv \begin{pmatrix} c_{11,t}^{-1/2} & & 0 \\ & \ddots & \\ 0 & & c_{dd,t}^{-1/2} \end{pmatrix} c_t \begin{pmatrix} c_{11,t}^{-1/2} & & 0 \\ & \ddots & \\ 0 & & c_{dd,t}^{-1/2} \end{pmatrix} = R, \tag{2.5}$$

where ρ_t denotes the spot correlation matrix and R is a time-invariant (positive semidefinite) correlation matrix. This model has the form of (2.3) with $\tilde{x}(c_t)$ containing only the constant term and $\tilde{y}(c_t)$ being the spot correlation matrix ρ_t. More generally, extensions of the above model, such as the dynamic equicorrelation model of Engle and Kelly (2012) and its generalization to block equicorrelation model, allow the spot correlation matrix ρ_t to vary over time but impose linear time-invariant restrictions between the elements of ρ_t. These restrictions can be casted in the model setting of (2.3) by redefining $\tilde{y}(c_t)$ as the proper linear transformation of ρ_t.

Example (Idiosyncratic Variance Models): In empirical finance, it is common to define the idiosyncratic variance of an asset as the variance of the residual of the stock return obtained from a linear projection on systematic risk factors, where the slope coefficient in the linear projection is called beta. Restricting attention to the one-factor market model for simplicity, the beta for the diffusive movement of the stock with respect to the market is given by $\beta_t \equiv c_{12,t}/c_{11,t}$, where the market and the stock are labelled by 1 and 2 respectively. The idiosyncratic spot variance of the stock is thus $c_{22,t} - \beta_t^2 c_{11,t} = c_{22,t} - c_{12,t}^2/c_{11,t}$. The idiosyncratic variance has received a lot of attention in the empirical finance literature (see, e.g., Ang et al. (2006, 2009)). A natural question concerning whether the idiosyncratic variance captures an independent source of risk is to examine whether it can be spanned by systematic factors such as the stochastic variance of the market. In the univariate setting, this can be conveniently casted in the form of the following parsimonious linear model $c_{22,t} - c_{12,t}^2/c_{11,t} = \theta_0 + \theta_1 c_{11,t}$, which corresponds to (2.3) with $\theta = (\theta_0, \theta_1)$, $\tilde{x}(c_t) = (1, c_{11,t})^\intercal$ and $\tilde{y}(c_t) = c_{22,t} - c_{12,t}^2/c_{11,t}$. Here, the parameter θ_1 may be referred to as the idiosyncratic variance beta of the stock with respect to the market. Extensions to the case with multiple risk factors is obvious.

In the above three examples, the model restrictions imply that the deterministic relationship between the elements of the spot variance matrix hold true for any time interval, and not just for the fixed one over which we test the model restriction. However, if one allows for presence of structural breaks in the above models (e.g., market beta that remains constant only over a week or a month) with known times of the structural breaks, then the testing of such deterministic connections between the elements of the spot variance matrix is only for a given fixed interval of time, exactly as in our asymptotic setup.

As indicated by the three preceding examples, model restrictions like (2.3) can be conveniently used to investigate the relationship between many latent risk measures that are of practical interest. We note that model restrictions like (2.3) are semiparametric in nature: they only impose parametric constraints on the spot covariance of the Itô semimartingale model (2.1) of the asset prices, while leaving other model components, such as drift, jumps and the marginal law of each component of c_t, completely nonparametric.

The current setting is nonstandard in several aspects. First, the identification of the spot covariance process is from observing the sample path of the studied process in continuous time, rather than from the invariant distribution of observed data. Indeed, we do not even require the invariant distribution of the price process or that of the volatility process to exist. Doing so allows us to accommodate essentially arbitrary forms of nonstationarity, dependence and heterogeneity in the data. As the identification is obtained in continuous time, we consider an in-fill asymptotic setting with data sampled at high frequency. Second, the fixed-span setting can be implemented over short samples (e.g., one quarter) but still provides statistically accurate inference. By implementing the method over relatively short subsamples using high-frequency data, one can readily accommodate slowly varying parameters like in structural break models.

Model (2.3) may be further generalized in several directions. First, one may extend the set of random variables entering the model (2.3) to include any directly observable processes. Another extension is to augment the model (2.3) by allowing for various latent quantities associated with the process X which can be recovered asymptotically as we sample more frequently. These include the jumps on a given interval and various measures associated with them. Given the highly nonstandard nature of the estimation problem concerning jumps and the associated different rates of convergence of the corresponding estimators, formulating such a problem in a general setting is rather nontrivial. Finally, our setup in (2.3) is linear in the parameter vector θ and a natural extension is to consider models that are nonlinear in parameters.

We conclude this subsection by formalizing the notion of correct specification. We consider two sets of sample paths

$$\Omega_{0,T} \equiv \Big\{\omega \in \Omega : \tilde{y}_m(c_t(\omega)) - \tilde{x}_m\,(c_t(\omega))^\intercal\,\theta_m(\omega) = 0 \text{ for some}$$
$$\text{real vector } \theta_m(\omega), \text{ all } 1 \le m \le \bar{m} \text{ and Lebesgue almost every } t \in [0,T]\Big\},$$

and $\Omega_{a,T} \equiv \Omega \setminus \Omega_{0,T}$. That is, $\Omega_{0,T}$ collects the sample paths on which the model restrictions in (2.3) hold for some vector $\theta(\omega)$, whereas its complement $\Omega_{a,T}$ is the event of misspecification. Model (2.3) is called correctly specified if the observed sample path falls in $\Omega_{0,T}$ and is called misspecified otherwise. The sets $\Omega_{0,T}$ and $\Omega_{a,T}$ play the role of the null and the alternative hypotheses, respectively, in a specification test. Specifying hypotheses in terms of random events is unlike the classical setting of hypothesis testing (e.g., Lehmann and Romano (2005)), but is standard in the study of high frequency data; see, e.g., Aït-Sahalia and Jacod (2012) and many references therein.

2.3 Heuristics for the inference procedures

We next describe the heuristics for our inference procedures with the formal theory presented in Section 3. For equation m, we consider a known measurable function $z_m : \mathcal{M}_d \mapsto \mathbb{R}^{q_m}$ for some $q_m \in \mathbb{N}$. Under correct specification, model (2.3) implies a set of moment conditions under the occupation measure

$$\mathbb{F}\left(y_m - x_m\theta_{m,0}\right) = 0, \quad \text{where} \quad y_m(\cdot) \equiv z_m(\cdot)\tilde{y}_m(\cdot), \quad x_m(\cdot) \equiv z_m(\cdot)\tilde{x}_m^\intercal\left(\cdot\right), \quad 1 \le m \le \bar{m}. \tag{2.6}$$

Analogous to standard econometric terminology, we refer to the function $z_m(\cdot)$ as an instrument. It is convenient to stack the equations in (2.6) by writing

$$\mathbb{F}\left(y - x\theta_0\right) = 0, \tag{2.7}$$

where

$$y\left(\cdot\right) = \left(y_1\left(\cdot\right)^\intercal, \ldots, y_{\bar{m}}\left(\cdot\right)^\intercal\right)^\intercal, \quad \theta_0 = \left(\theta_{1,0}^\intercal, \ldots, \theta_{\bar{m},0}^\intercal\right)^\intercal,$$
$$x\left(\cdot\right) = \begin{pmatrix} x_1(\cdot) & & 0 \\ & \ddots & \\ 0 & & x_{\bar{m}}\left(\cdot\right) \end{pmatrix}.$$

Below, we denote $q \equiv \sum_{m=1}^{\bar{m}} q_m$.

We conduct estimation via a classical minimum distance (CMD) procedure. Suppose that we can construct a sample analogue $\mathbb{F}^n$ (see Section 3.1) for the occupation measure $\mathbb{F}$, so that $\mathbb{F}y$ and $\mathbb{F}x$ can be estimated by $\mathbb{F}^n y$ and $\mathbb{F}^n x$ respectively. For any weighting matrix Ψ_n that satisfies $\Psi_n \xrightarrow{\mathbb{P}} \Psi$, where Ψ is a positive semidefinite matrix, the CMD estimator is given by

$$\theta_n \equiv \underset{\theta}{\operatorname{argmin}}\,(\bar{y}_n - \bar{x}_n\theta)^\intercal\Psi_n(\bar{y}_n - \bar{x}_n\theta),$$

where we use the shorthand

$$\bar{x}_n \equiv \mathbb{F}^n x, \quad \bar{y}_n \equiv \mathbb{F}^n y. \tag{2.8}$$

The CMD estimator has a simple closed-form solution

$$\theta_n = \left(\bar{x}_n^\intercal \Psi_n \bar{x}_n\right)^{-1} \bar{x}_n^\intercal \Psi_n \bar{y}_n, \tag{2.9}$$

provided that the matrix inversion is well-defined, at least asymptotically. Note that studying the asymptotic property of the CMD estimator amounts to studying the joint asymptotic behavior of $\bar{x}_n$ and $\bar{y}_n$.

A further important problem is specification testing. That is, we want to decide in which event, $\Omega_{0,T}$ or $\Omega_{a,T}$, the observed sample path falls. As is clear from the definition of these events, such a decision requires knowledge about c_t at almost every $t \in [0, T]$. Such temporal information is lost in the occupation measure $\mathbb{F}$ as a result of the temporal aggregation. To preserve the temporal information, it is useful to consider a generalized occupation measure as follows. Throughout the paper, we call a function $w : \mathbb{R} \mapsto \mathbb{R}$ a *weight function* if it is infinitely continuously differentiable. Consider a family of weight functions $w_\tau(\cdot)$ indexed by $\tau \in \mathcal{T}$ where $\mathcal{T} \subset \mathbb{R}$ is compact with positive Lebesgue measure. For each τ, the w_τ-weighted occupation measure, denoted by $\mathbb{F}_\tau$, is defined as a linear functional:

$$\mathbb{F}_\tau g \equiv \int_0^T g\left(c_s\right) w_\tau\left(s\right) ds. \tag{2.10}$$

If the model (2.3) is correctly specified, then

$$\mathbb{F}_\tau \left(\tilde{y}_m - \tilde{x}_m^\intercal \theta_{m,0}\right) = 0, \quad \text{for any } 1 \leq m \leq \bar{m} \text{ and } \tau \in \mathcal{T}. \tag{2.11}$$

Moreover, if the family of weight functions is properly chosen, the moment condition (2.11) holds if and only if (2.3) is correctly specified (see Proposition 3 below). A specification test then can be carried out by testing (2.11) via its empirical analogue. This idea motivates us to introduce the notion of weighted occupation measure $\mathbb{F}_\tau$ in the first place. It also motivates the study of the asymptotic theory concerning the estimation and inference for the τ-indexed process $\mathbb{F}_\tau g$ for some fixed test function g. This testing strategy is akin to the consistent model specification test of Bierens (1982), but applied in a nonstandard setting for investigating the pathwise properties of stochastic processes.

3 Theory

We now present the main theoretical results of the paper. Section 3.1 introduces the estimator of the weighted occupation measure $\mathbb{F}_\tau$ for a family of weight functions $\{w_\tau(\cdot) : \tau \in \mathcal{T}\}$ and discusses its asymptotic properties. We present asymptotic results for the CMD estimator θ_n in Section 3.2 and propose a specification test in Section 3.3.

3.1 The empirical covariance occupation measure

The empirical occupation measure, that is, the estimator of the covariance occupation measure from the discrete observations of X, is constructed in two steps. In the first step we recover nonparametrically the spot covariance process and in the second step we use the spot covariance estimates to construct a sample analogue of $\mathbb{F}_\tau$.

Let $\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}$ denote the ith increment of X at asymptotic stage n. Following Jacod and Protter (2012), we estimate the spot covariance at time $i\Delta_n$ via a local truncated variation estimator (Mancini (2001)). To define this local estimator, we consider a sequence of integers k_n that determines the number of increments in a local window for spot covariance estimation. The spot covariance estimator is then given by

$$\hat{c}_{i\Delta_n} \equiv \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n} \left(\Delta_{i+j}^n X\right) \left(\Delta_{i+j}^n X\right)^\intercal 1_{\{\|\Delta_{i+j}^n X\| \le \chi \Delta_n^\varpi\}}, \tag{3.1}$$

where $\chi > 0$ and $\varpi \in (0, 1/2)$ are constants that specify the truncation threshold. The truncation technique is used so that the local estimator $\hat{c}_{i\Delta_n}$ is robust to jumps in X. Along the same line of argument as in Li et al. (2013), we can allow the truncation threshold to have (a certain type of) data-dependence without affecting the results of this paper. In the setting without jumps, the estimation of spot variance can be dated back to Foster and Nelson (1996) and Comte and Renault (1998). We assume the following for the tuning parameters.

Assumption LW: $k_n \asymp \Delta_n^{-\varsigma}$ for some constant $\varsigma \in (\frac{r}{2} \vee \frac{1}{3}, \frac{1}{2})$ and $\varpi \in [\frac{1-\varsigma}{2-r}, \frac{1}{2})$.

Equipped with the local estimator $\hat{c}_{i\Delta_n}$, we set, for any measurable function $g : \mathcal{M}_d \mapsto \mathbb{R}^{\dim(g)}$ and $\tau, \eta \in \mathcal{T}$,

$$\begin{cases} \widehat{\mathbb{F}}_\tau^n g \equiv \Delta_n \displaystyle\sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} g(\hat{c}_{i\Delta_n}) w_\tau(i\Delta_n), \\ \widehat{\mathbb{F}}_{\tau,\eta}^n g \equiv \Delta_n \displaystyle\sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} g(\hat{c}_{i\Delta_n}) w_\tau(i\Delta_n) w_\eta(i\Delta_n), \end{cases}$$

where $\lfloor T/\Delta_n \rfloor$ denotes the integer part of T/Δ_n. Clearly, $\widehat{\mathbb{F}}_\tau^n$ is the sample analogue of $\mathbb{F}_\tau$. The double-indexed estimator $\widehat{\mathbb{F}}_{\tau,\eta}^n$ is the sample analogue of

$$\mathbb{F}_{\tau,\eta} g \equiv \int_0^T g(c_s) w_\tau(s) w_\eta(s) ds,$$

which is used below for denoting the asymptotic covariance function of our estimators. In the sequel, we refer to $\widehat{\mathbb{F}}_{\tau}^{n}$ as the *raw* empirical occupation measure, as it suffers from a high-order bias that needs to be corrected. Nevertheless, $\widehat{\mathbb{F}}_{\tau}^{n}$ is useful for constructing consistent estimators for various quantities, such as the asymptotic variances.

We impose a smoothness condition on the family $\{w_{\tau}(\cdot) : \tau \in \mathcal{T}\}$ of weight functions as follows, where WF stands for weight function.

Assumption WF: The index set $\mathcal{T}$ is a compact subset of $\mathbb{R}$. Moreover, for some constant $K > 0$, we have $|w_{\tau}(s) - w_{\eta}(s)| \leq K\,|\tau - \eta|$ and $|w_{\tau}(s) - w_{\tau}(t)| \leq K\,|s - t|$ for all $\tau, \eta \in \mathcal{T}$ and $s, t \in [0, T]$.

Theorem 1 below provides sufficient conditions and an exact sense for the uniform consistency of $\widehat{\mathbb{F}}_{\tau}^{n}$ and $\widehat{\mathbb{F}}_{\tau,\eta}^{n}$ towards $\mathbb{F}_{\tau}$ and $\mathbb{F}_{\tau,\eta}$, respectively.

Theorem 1. *Suppose (i) Assumptions HF, LW and WF and (ii) g is a continuous function on $\mathcal{M}_d$. Then, $\widehat{\mathbb{F}}_{\tau}^{n} g \xrightarrow{\mathbb{P}} \mathbb{F}_{\tau} g$ and $\widehat{\mathbb{F}}_{\tau,\eta}^{n} g \xrightarrow{\mathbb{P}} \mathbb{F}_{\tau,\eta} g$ uniformly in $\tau, \eta \in \mathcal{T}$.*

The convergence in Theorem 1 is not associated with a central limit theorem because of the presence of an asymptotic bias. To get a central limit theorem, we now consider a bias-corrected version of $\widehat{\mathbb{F}}_{\tau}^{n}$, denoted below by $\mathbb{F}_{\tau}^{n}$. With each g that is three times continuously differentiable, we associate a function $\mathbb{B}g$ given by

$$(\mathbb{B}g)(c) \equiv \frac{1}{2} \sum_{j,k,l,m=1}^{d} \partial_{jk,lm}^{2} g(c) \left(c_{jl} c_{km} + c_{jm} c_{kl}\right).$$

Now, we can define the *bias-corrected* empirical occupation measure $\mathbb{F}_{\tau}^{n}$ as

$$\mathbb{F}_{\tau}^{n} g \equiv \widehat{\mathbb{F}}_{\tau}^{n} g - k_n^{-1} \widehat{\mathbb{F}}_{\tau}^{n} \mathbb{B} g,$$

where we remind the reader that, by convention, $\widehat{\mathbb{F}}_{\tau}^{n} \mathbb{B} g$ is understood as $\widehat{\mathbb{F}}_{\tau}^{n}(\mathbb{B}g)$. To simplify notation, henceforth, when the weight function is identically one, we denote the raw and the bias-corrected empirical occupation measure respectively by $\widehat{\mathbb{F}}^{n}$ and $\mathbb{F}^{n}$, cf. (2.2).

In the next theorem we state the stable convergence in law[5] (denoted with $\xrightarrow{\mathcal{L}\text{-}s}$) of the sequence $\Delta_n^{-1/2}(\mathbb{F}_{\tau}^{n} g - \mathbb{F}_{\tau} g)$ of τ-indexed processes for some fixed test function g. To describe the asymptotic

[5]Stable convergence in law is stronger than the usual notion of weak convergence. It requires that the convergence holds jointly with any bounded random variable defined on the original probability space. Its importance for our problem stems from the fact that the limiting process of our estimator is an $\mathcal{F}$-conditionally Gaussian process and stable convergence allows for feasible inference using a consistent estimator for its $\mathcal{F}$-conditional variance. See Jacod and Shiryaev (2003) for further details on stable convergence on filtered probability spaces.

covariance function of $\Delta_n^{-1/2}(\mathbb{F}_\tau^n g - \mathbb{F}_\tau g)$, we define, for each pair g, h of vector-valued functions,

$$[\mathbb{V}(g,h)]\,(c) = \sum_{j,k,l,m=1}^{d} \partial_{jk} g\,(c)\, \partial_{lm} h\,(c)^\intercal \,(c_{jl}c_{km} + c_{jm}c_{kl})\,.$$

Theorem 2. *Suppose (i) Assumptions HF, LW and WF; (ii) g is three times continuously differentiable on $\mathcal{M}_d$. Then the sequence $\Delta_n^{-1/2}(\mathbb{F}_\tau^n g - \mathbb{F}_\tau g)$ of τ-indexed processes converges stably in law to a process under the uniform metric, which is defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$ and is, conditional on $\mathcal{F}$, a centered Gaussian process with covariance function $M(\tau,\eta) \equiv \mathbb{F}_{\tau,\eta}\mathbb{V}\,(g,g)$, for $\tau, \eta \in \mathcal{T}$.*

3.2 Asymptotic properties of the CMD estimator

We consider next the asymptotic behavior of the CMD estimator θ_n given by (2.9), with the (obvious) assumption that the term $\tilde{x}_m^\intercal \theta_{m,0}$ is present in (2.3). We complement the notation in (2.8) by setting

$$\bar{x} \equiv \mathbb{F}x, \quad \bar{y} \equiv \mathbb{F}y.$$

The following condition ensures that θ_n is well-defined, where ID stands for identification.

ASSUMPTION ID: We have $\Psi_n \xrightarrow{\mathbb{P}} \Psi$ for some $\mathcal{F}$-measurable positive semidefinite random matrix Ψ. Moreover, the random matrix $\bar{x}^\intercal \Psi \bar{x}$ is almost surely nonsingular.

The convergence in probability of θ_n is given in the following theorem.

Theorem 3. *Suppose (i) Assumptions HF, LW and ID; (ii) the functions $x\,(\cdot)$ and $y\,(\cdot)$ are three times continuously differentiable on $\mathcal{M}_d$. Then $\theta_n \xrightarrow{\mathbb{P}} \theta_0^* \equiv (\bar{x}^\intercal \Psi \bar{x})^{-1} \bar{x}^\intercal \Psi \bar{y}$.*

Remark 1. *The convergence result in Theorem 3 does not depend on (2.3) being correctly specified. Hence, in general, θ_0^* is interpreted as the pseudo-true parameter in the current estimation setting. In the basic case with $\tilde{x}(c_t) = z\,(c_t) \equiv 1$, the pseudo-true parameter θ_0^* reduces to the integrated volatility functional $\int_0^T \tilde{y}\,(c_s)\,ds$ for a general smooth function $\tilde{y}\,(\cdot)$.*

We now turn to the second-order asymptotic behavior, namely the stable convergence in law, of the CMD estimator θ_n. We start with a general result (Theorem 4 below) without assuming correct specification. We state this result under a high-level, but easily verifiable, condition that is given as follows, where SC stands for stable convergence.

ASSUMPTION SC: $\Delta_n^{-1/2}(\mathbb{F}^n x - \mathbb{F}x, \mathbb{F}^n y - \mathbb{F}y, \Psi_n - \Psi) \xrightarrow{\mathcal{L}\text{-}s} (\xi_x, \xi_y, \xi_\Psi)$, where the limit variables are defined on an extension of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The joint convergence of $\Delta_n^{-1/2}(\mathbb{F}^n x - \mathbb{F}x, \mathbb{F}^n y - \mathbb{F}y)$ in Assumption SC can be derived using Theorem 2 applied to $g(\cdot) = (vec(x(\cdot))^\intercal, y(\cdot)^\intercal)^\intercal$ and a constant weight function. In the special case with $\Psi_n \equiv \Psi$, Assumption SC is verified with (ξ_x, ξ_y) being an $\mathcal{F}$-conditionally centered Gaussian variable and $\xi_\Psi \equiv 0$. More generally, if Ψ_n has the form $\mathbb{F}^n\psi$ for some three times continuously differentiable function ψ, then Assumption SC can again be verified using Theorem 2 applied to $g(\cdot) = (vec(x(\cdot))^\intercal, y(\cdot)^\intercal, vec(\psi(\cdot))^\intercal)^\intercal$.

Theorem 4. *Under Assumptions ID and SC, we have* $\Delta_n^{-1/2}(\theta_n - \theta_0^*) \xrightarrow{\mathcal{L}\text{-}s} (\bar{x}^\intercal \Psi \bar{x})^{-1}(\xi_x^\intercal \Psi (\bar{y} - \bar{x}\theta_0^*) + \bar{x}^\intercal \xi_\Psi (\bar{y} - \bar{x}\theta_0^*) + \bar{x}^\intercal \Psi (\xi_y - \xi_x \theta_0^*))$.

Theorem 4 shows the stable convergence of θ_n in a nonparametric setting that allows for misspecification. In particular, θ_n is centered at the pseudo-true parameter θ_0^* which, as mentioned in Remark 1, includes integrated volatility functionals as special cases. We note that the first two components of the limit variable in Theorem 4 are zero if $\bar{y} = \bar{x}\theta_0^*$. This condition holds when model (2.3) is correctly specified. This condition also holds under misspecification provided that θ_0^* is exactly identified. The latter case is of particular practical interest since empirical workers typically use an exactly identified system of moment conditions.

With this in mind, we specialize the result of Theorem 4 in the exact identification setting under primitive conditions. To facilitate application, we also provide the asymptotic variance and its estimator in explicit form, for which we need some additional notation. We set

$$\begin{cases} M_{yy} \equiv \mathbb{F}\mathbb{V}(y, y), \quad M_{xx} \equiv \mathbb{F}\mathbb{V}(vec(x), vec(x)), \\ M_{yx} \equiv \mathbb{F}\mathbb{V}(y, vec(x)), \quad M_{xy} \equiv \mathbb{F}\mathbb{V}(vec(x), y), \end{cases} \tag{3.2}$$

with sample analogue estimators

$$\begin{cases} M_{yy,n} \equiv \widehat{\mathbb{F}}^n\mathbb{V}(y, y), \quad M_{xx,n} \equiv \widehat{\mathbb{F}}^n\mathbb{V}(vec(x), vec(x)), \\ M_{yx,n} \equiv \widehat{\mathbb{F}}^n\mathbb{V}(y, vec(x)), \quad M_{xy,n} \equiv \widehat{\mathbb{F}}^n\mathbb{V}(vec(x), y). \end{cases} \tag{3.3}$$

Below, for a generic sequence Z_n of random variables, we write $Z_n \xrightarrow{\mathcal{L}\text{-}s} \mathcal{MN}(0, \Sigma_Z)$ if Z_n converges stably in law to a variable that is defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$ and is, conditional on $\mathcal{F}$, centered Gaussian with covariance matrix Σ_Z.

Proposition 1. *Let Ψ_n and Ψ be the identity matrix. Suppose (i) Assumptions HF and LW; (ii) $\bar{x} \equiv \mathbb{F}x$ is a square random matrix and is nonsingular almost surely; (iii) the functions $x(\cdot)$ and $y(\cdot)$ are three times continuously differentiable on $\mathcal{M}_d$. Then the following statements hold.*

(a) $\Delta_n^{-1/2}(\theta_n - \theta_0^*) \xrightarrow{\mathcal{L}\text{-}s} \mathcal{MN}(0, \Sigma^*)$, *where* $\Sigma^* \equiv \bar{x}^{-1} A^* (\bar{x}^\intercal)^{-1}$ *and*

$$A^* \equiv M_{yy} - M_{yx}\Theta_0^{*\intercal} - \Theta_0^* M_{xy} + \Theta_0^* M_{xx} \Theta_0^{*\intercal}, \quad \Theta_0^* \equiv \theta_0^{*\intercal} \otimes I_q.$$

(b) The asymptotic variance Σ^* can be consistently estimated by $\Sigma_n^* \equiv (\widehat{\mathbb{F}}^n x)^{-1} A_n (\widehat{\mathbb{F}}^n x^\intercal)^{-1}$, where A_n is given by

$$A_n \equiv M_{yy,n} - M_{yx,n}\Theta_n^\intercal - \Theta_n M_{xy,n} + \Theta_n M_{xx,n}\Theta_n^\intercal, \quad \textit{for } \Theta_n \equiv \theta_n^\intercal \otimes I_q. \tag{3.4}$$

3.3 Specification testing

We next derive a test for the pathwise relation in (2.3), that is, we test in which of the two complementary events, $\Omega_{0,T}$ or $\Omega_{a,T}$, the observed sample path falls. As hinted in Section 2.3, model restrictions in (2.3) can be equivalently represented by moment equalities under weighted occupation measures, provided that the family of weight functions is properly chosen. This argument is formalized in Proposition 2. Below, we consider a family of weight functions $\{w_\tau(\cdot) : \tau \in \mathcal{T}\}$ where the index set $\mathcal{T} \subset \mathbb{R}$ is compact with positive Lebesgue measure. This family is said to be *complete* if it satisfies the following property: for any càdlàg function $f : [0,T] \mapsto \mathbb{R}$, $f(t) = 0$ for Lebesgue almost every $t \in [0,T]$ if and only if $\int_0^T f(s) w_\tau(s) ds = 0$ for all $\tau \in \mathcal{T}$.

Proposition 2. *Suppose Assumption ID. For each $m \in \{1, \ldots, \bar{m}\}$, let $z_m^\star : \mathcal{M}_d \mapsto (0,\infty)$ be a (strictly positive) measurable function. If $\{w_\tau(\cdot) : \tau \in \mathcal{T}\}$ is complete, then $\Omega_{0,T} = \{\mathbb{F}_\tau(z_m^\star \tilde{y}_m) - \mathbb{F}_\tau(z_m^\star \tilde{x}_m^\intercal)\theta_{m,0}^* = 0$ for all $\tau \in \mathcal{T}$ and $1 \le m \le \bar{m}\}$.*

Proposition 2 shows that $\Omega_{0,T}$ can be equivalently specified as $\mathbb{F}_\tau(z_m^\star \tilde{y}_m) - \mathbb{F}_\tau(z_m^\star \tilde{x}_m^\intercal)\theta_{m,0}^* = 0$ for all $\tau \in \mathcal{T}$ and $m \in \{1, \ldots, \bar{m}\}$. A natural choice of the instrument $z_m^\star$ is $z_m^\star(\cdot) \equiv 1$, so that the moment condition to be tested corresponds to (2.11). We consider $z^\star$ with general functional forms with no additional cost in our derivations. In practice, the user may choose $z^\star$ to assign more weight on some regions of the state space. Below, for the sake of notational simplicity, we set $y_m^\star = z_m^\star \tilde{y}_m$, $x_m^\star = z_m^\star \tilde{x}_m^\intercal$ and

$$y^\star \equiv \begin{pmatrix} y_1^\star \\ \vdots \\ y_{\bar{m}}^\star \end{pmatrix}, \quad x^\star \equiv \begin{pmatrix} x_1^\star & & 0 \\ & \ddots & \\ 0 & & x_{\bar{m}}^\star \end{pmatrix}.$$

With this notation, the assertion of Proposition 2 can be written as

$$\Omega_{0,T} = \{\mathbb{F}_\tau(y^\star - x^\star\theta_0^*) = 0 \text{ for all } \tau \in \mathcal{T}\}. \tag{3.5}$$

Proposition 3 below provides a general way of constructing a complete class of weight functions.

Proposition 3. *Let $\mathcal{T}$ be a compact subset of $\mathbb{R}$ with strictly positive Lebesgue measure. Let $w : \mathbb{R} \mapsto \mathbb{R}$ be a power series on $\mathbb{R}$ such that the set $\left\{k \in \mathbb{N} : (d/du)^k w(u)\,|_{u=0} = 0\right\}$ is finite. Then the family of weight functions $w_\tau(s) \equiv w(\tau s)$, $\tau \in \mathcal{T}$, is complete.*

Proposition 3 is a special case of Theorem 1 of Bierens and Ploberger (1997). A concrete example of the function $w(\cdot)$ is $w(s) \equiv \cos(s) + \sin(s)$. The corresponding weight functions can be taken as $w_\tau(s) = w(\tau s)$ with $\mathcal{T}$ being a compact interval with positive length.

We consider the scaled sample analogue of the moment conditions in (3.5) given by

$$\zeta_n(\tau) \equiv \Delta_n^{-1/2}(\mathbb{F}_\tau^n y^\star - (\mathbb{F}_\tau^n x^\star)\, \theta_n).$$

The asymptotic behavior of $\zeta_n(\cdot)$ is described in Theorem 5 below. Under correct specification, that is, in restriction to $\Omega_{0,T}$, we show that the process $\zeta_n(\cdot)$ converges stably in law to a mixture Gaussian process. We need some notation to describe the conditional asymptotic covariance function of the limiting process and the consistent estimator for it. We set

$$\begin{cases} h^\star = (y^{\star\intercal}, vec\,(x^\star)^\intercal)^\intercal\,, \quad h = (y^\intercal, vec\,(x)^\intercal)^\intercal\,, \\ Q(\tau,\eta) \equiv \begin{pmatrix} \mathbb{F}_{\tau,\eta}\mathbb{V}\,(h^\star, h^\star) & \mathbb{F}_\tau\mathbb{V}\,(h^\star, h) \\ \mathbb{F}_\eta\mathbb{V}\,(h, h^\star) & \mathbb{F}\mathbb{V}\,(h,h) \end{pmatrix}, \\ Q_n(\tau,\eta) \equiv \begin{pmatrix} \widehat{\mathbb{F}}_{\tau,\eta}^n\mathbb{V}\,(h^\star, h^\star) & \widehat{\mathbb{F}}_\tau^n\mathbb{V}\,(h^\star, h) \\ \widehat{\mathbb{F}}_\eta^n\mathbb{V}\,(h, h^\star) & \widehat{\mathbb{F}}^n\mathbb{V}\,(h,h) \end{pmatrix}. \end{cases} \tag{3.6}$$

We further set

$$\begin{cases} \Xi \equiv (\bar{x}^\intercal\Psi\bar{x})^{-1}\bar{x}^\intercal\Psi, \\ \Xi_n \equiv (\widehat{\mathbb{F}}^n x^\intercal\Psi_n\widehat{\mathbb{F}}^n x)^{-1}\widehat{\mathbb{F}}^n x^\intercal\Psi_n, \\ \kappa(\tau;\theta) \equiv (I_{\bar{m}}, -\theta^\intercal \otimes I_{\bar{m}}, -(\mathbb{F}_\tau x^\star)\,\Xi, (\mathbb{F}_\tau x^\star)\,\Xi\,(\theta^\intercal \otimes I_q))\,, \\ \kappa_n(\tau;\theta) \equiv \Big(I_{\bar{m}}, -\theta^\intercal \otimes I_{\bar{m}}, -(\widehat{\mathbb{F}}_\tau^n x^\star)\Xi_n, (\widehat{\mathbb{F}}_\tau^n x^\star)\Xi_n\,(\theta^\intercal \otimes I_q)\Big)\,. \end{cases}$$

Finally, we set

$$\begin{cases} N(\tau,\eta;\theta) \equiv \kappa\,(\tau;\theta)\,Q(\tau,\eta)\kappa\,(\eta;\theta)^\intercal\,, \\ N_n(\tau,\eta;\theta) \equiv \kappa_n\,(\tau;\theta)\,Q_n(\tau,\eta)\kappa_n\,(\eta;\theta)^\intercal\,. \end{cases}$$

Theorem 5. *Suppose (i) Assumptions HF, LW, WF and ID; (ii) the functions $x^\star$, $y^\star$, x and y are three times continuously differentiable on $\mathcal{M}_d$. Then we have the following.*

(a) $\Delta_n^{1/2}\zeta_n(\tau) \xrightarrow{\mathbb{P}} \zeta^\star(\tau) \equiv \mathbb{F}_\tau y^\star - (\mathbb{F}_\tau x^\star)\,\theta_0^*$ uniformly in $\tau \in \mathcal{T}$.

(b) In restriction to $\Omega_{0,T}$, the sequence $\zeta_n(\tau)$ of τ-indexed processes converges stably in law under the uniform metric to a process $\zeta(\tau)$ which is defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$ and is, conditional on $\mathcal{F}$, centered Gaussian with covariance function $N(\cdot,\cdot;\theta_0)$.

(c) $N_n(\tau,\eta;\theta_n) \xrightarrow{\mathbb{P}} N(\tau,\eta;\theta_0^*)$ uniformly in $\tau,\eta \in \mathcal{T}$. In particular, $N_n(\tau,\eta;\theta_n) \xrightarrow{\mathbb{P}} N(\tau,\eta;\theta_0)$ uniformly in $\tau,\eta \in \mathcal{T}$ in restriction to $\Omega_{0,T}$.

We are now ready to describe the specification test. Among many possible choices, we consider a Kolmogorov-Smirnov type test statistic given by

$$S_n \equiv \varphi\left(\zeta_n(\cdot), N_n(\cdot,\cdot;\theta_n)\right) \equiv \sup_{\tau\in\mathcal{T}} \sup_{1\leq m\leq \bar{m}} \frac{|\zeta_{m,n}(\tau)|}{\sqrt{N_{mm,n}\left(\tau,\tau;\theta_n\right)}}, \tag{3.7}$$

where $\zeta_{m,n}\left(\cdot\right)$ is the mth component of $\zeta_n\left(\cdot\right)$, $N_{mm,n}\left(\cdot\right)$ is the mth diagonal element of $N_n(\cdot)$, and the function $\varphi\left(\cdot\right)$ is implicitly defined by the second equality. The normalization by the standard error makes the test statistic scale-invariant, which is desirable in practice.

The asymptotic properties of S_n follow directly from Theorem 5 and the continuous mapping theorem. More specifically, Theorem 5(a,c) shows that $\Delta_n^{1/2}S_n \xrightarrow{\mathbb{P}} \varphi\left(\zeta^\star\left(\cdot\right), N\left(\cdot,\cdot;\theta_0^*\right)\right)$. Under misspecification, $\zeta^\star(\cdot)$ is not identically zero by Proposition 2 and, hence, S_n diverges to infinity in probability. On the other hand, if model (2.3) is correctly specified, then Theorem 5(b,c) implies that $S_n \xrightarrow{\mathcal{L}\text{-}s} \varphi(\zeta(\cdot), N\left(\cdot,\cdot;\theta_0\right))$. To conduct the specification test at nominal level α, it remains to select a tight sequence $cv_{n,\alpha}$ of critical values, which, in restriction to $\Omega_{0,T}$, consistently estimates the $1-\alpha$ $\mathcal{F}$-conditional quantile of $\varphi(\zeta(\cdot), N\left(\cdot,\cdot;\theta_0\right))$. Since the limiting distribution is nonstandard, the quantile does not have a closed-form expression in general. Nevertheless, as is standard in this type of problems, the critical value can be obtained via simulation by following three steps. Step 1: estimate the conditional covariance function $N(\cdot,\cdot;\theta_0^*)$ using $N_n(\cdot,\cdot;\theta_n)$. Step 2: simulate a large number of centered Gaussian processes with covariance function $N_n(\cdot,\cdot;\theta_n)$; Step 3: set $cv_{n,\alpha}$ to be the $1-\alpha$ quantile of $\varphi(\zeta_n^{MC}\left(\cdot\right), N_n\left(\cdot,\cdot;\theta_n\right))$, where $\zeta_n^{MC}(\cdot)$ denotes a Monte Carlo realization of the Gaussian process simulated in Step 2. Corollary 1 below summarizes the testing result.

Corollary 1. *Let $\alpha \in (0, 1/2)$ be a constant. Consider a sequence $C_{n,\alpha}$ of critical regions given by $C_{n,\alpha} \equiv \{S_n > cv_{n,\alpha}\}$. Suppose (i) the conditions in Theorem 5; (ii) the family $\{w_\tau(\cdot) : \tau \in \mathcal{T}\}$ is complete and $z_m^\star$ takes values in $(0,\infty)$ for each $m \in \{1,\ldots,\bar{m}\}$; (iii) for each $m \in \{1,\ldots\bar{m}\}$, $\inf_{\tau\in\mathcal{T}} N_{mm}\left(\tau,\tau;\theta_0^*\right) > 0$ almost surely. Then the following statements hold.*

(a) The test associated with the critical region $C_{n,\alpha}$ has asymptotic size α under the null hypothesis that (2.3) is correctly specified: $\mathbb{P}\left(C_{n,\alpha}|\Omega_{0,T}\right) \longrightarrow \alpha$.

(b) The test associated with the critical region $C_{n,\alpha}$ has asymptotic power one under the alternative hypothesis that (2.3) is misspecified: $\mathbb{P}\left(C_{n,\alpha}|\Omega_{a,T}\right) \longrightarrow 1$.

4 Empirical applications

We illustrate the use of our inference techniques in an analysis of the dynamic and cross-sectional properties of the stochastic variance in the economy. In particular, we use data on the SPDR ETF tracking the S&P 500 index and the nine S&P 500 sector index ETFs, all traded on the New

York Stock Exchange. The nine sector ETFs conveniently separate the stocks of the S&P 500 index into the following sector categories: Consumer Discretionary, Consumer Staples, Energy, Financial, Health Care, Industrial, Materials, Technology and Utilities. Our data covers the period January 1, 2006 till December 31, 2012 for a total of 1,746 days, or 28 quarters. In order to mitigate the effect of microstructure noise, we sample the data sparsely at every 10 minutes during the trading hours, which results in 38 high-frequency returns per day for each of the studied ETFs. Below, we label the market to be 1 and the sector ETFs by $2, \ldots, 10$.

We set tuning parameters as follows. In (3.1), the truncation threshold in day t is given by $3\bar{\sigma}_t \Delta^{0.49}$, where $\bar{\sigma}_t$ is a preliminary estimate of the day-t average volatility, here computed as the squared root of the annualized bipower variation (Barndorff-Nielsen and Shephard (2004a)). The specification test is conducted with $z^\star(\cdot) \equiv 1$ and the weight function $w_\tau(s) = \cos(\tau s) + \sin(\tau s)$ for $\tau \in \mathcal{T} \equiv [5, 100]$. To keep the computation manageable, $\mathcal{T}$ is discretized as $\{5, 10, \ldots, 100\}$. We set the local window $k_n = 14$.[6]

We now introduce some notation for our empirical analysis. The dynamics of the market portfolio are given by

$$dX_{1,t} = b_{1,t}dt + dX^c_{1,t} + dJ_{1,t}, \tag{4.1}$$

where X^c_1 and J_1 are respectively the diffusive and the jump parts of X_1 and we further assume the following dynamics for the nine sector portfolios

$$dX_{j,t} = b_{j,t}dt + \beta_{j,t}dX^c_{1,t} + d\widetilde{X}^c_{j,t} + dJ_{j,t}, \quad j = 2, ..., 10, \tag{4.2}$$

where J_j is the jump component of portfolio j, $\beta_{j,t} \equiv c_{1j,t}/c_{11,t}$ is the spot beta of portfolio j with respect to the market, and $\widetilde{X}^c_j$ is the residual diffusive component of X_j, which we henceforth refer to as the *market-neutral* component of X_j. We stress that $\beta_{j,t}$ is defined locally and nonparametrically. Also, the model (4.2) entails no *a priori* conditions on the cross-sectional dependencies among the market-neutral returns. The specification in (4.2) is a decomposition that formalizes the sense of market neutrality. We finally note that we do not separate the jumps into market and market-neutral components since the jumps are filtered out in the estimation.

In the empirical application, we are interested in estimating and testing model restrictions for the spot covariance and spot correlation matrices of $\widetilde{X}^c_t$, which are respectively denoted by $\tilde{c}_t$ and $\tilde{\rho}_t$ and defined as

[6] We use the same setup in a simulation setting with realistic features such as leverage effect, stochastic volatility and volatility jumps. The simulation results suggest that the inference procedure perform well and is robust to nontrivial perturbations in k_n. For brevity, we refer the reader to the working paper version of this paper for details regarding the simulation study.

$$\tilde{c}_{jk,t} \equiv \left\langle d\widetilde{X}^c_{j,t}, d\widetilde{X}^c_{k,t} \right\rangle = c_{jk,t} - c_{1j,t}c_{1k,t}/c_{11,t}, \quad \tilde{\rho}_{jk,t} = \frac{\tilde{c}_{jk,t}}{\sqrt{\tilde{c}_{jj,t}\tilde{c}_{kk,t}}}.$$

These quantities are smooth nonlinear transforms of c_t. We conduct our analysis on a quarterly basis (i.e. T =1 quarter), so that low-frequency structural changes across quarters are automatically accommodated. The short sample span is closely aligned with our "fixed-T" asymptotic setting, where we allow for essentially arbitrary forms of nonstationarity and data heterogeneity.

4.1 Variance spanning of market-neutral returns of sector portfolios

We start with the diagonal elements of $\tilde{c}_t$. Our initial working hypothesis is that all of the variation over time in the variances of market-neutral returns is attributable to the market variance itself, i.e. $c_{11,t}$. Formally, we examine the following set of linear specifications, both jointly and individually, for the spot variances

$$\tilde{c}_{jj,t} = \theta_{j,0} + \theta_{j,1}c_{11,t}, \quad j \in \{2, \ldots, 10\}. \tag{4.3}$$

We refer to the coefficients $\theta_{j,1}$, $2 \leq j \leq 10$, as variance betas. The set of restrictions with a single spanning factor embodied in (4.3) appears strong, but they appear in many equilibrium (and reduced form) models for the variance risk premium.[7] Also, a single macro volatility factor is the key feature of aggregate time change models.

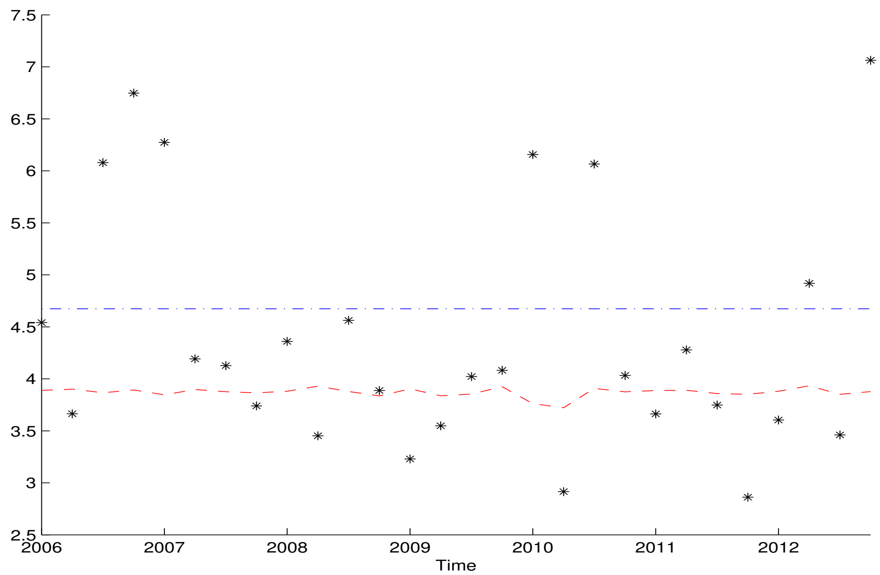
We first test whether this set of linear equations holds jointly for all nine sector portfolios over each of the 28 quarters in our sample. The results of these tests are shown in Figure 1, where we plot the test statistic and the (pointwise) 1% level critical value for each quarter. In order to draw inference across all quarters, we also plot the uniform 1% level critical value that is uniform with respect to all 28 quarters.[8] As seen from the figure, there is strong evidence that the market-neutral variance risk in the sector portfolios cannot be linearly spanned by the market variance risk. This is in spite of the fact that the parameters in (4.3) are kept fixed only during a relatively short period of time, that is, a quarter.

We next study whether the spanning restriction (4.3) holds for some sectors. To address this question, we perform tests for each of the sector portfolios individually. The results are shown on Figure 2. Some interesting findings emerge. First, there are a few sectors for which the linear

[7] See Drechsler and Yaron (2011), Bollerslev et al. (2012), and references therein.

[8] The uniform critical value is constructed as the 99% quantile of the supremum of the 28 quarterly test statistics under the null distribution. The null distribution of the quarterly statistics are asymptotically $\mathcal{F}$-conditionally independent since the limiting variable in the CLT is a process with $\mathcal{F}$-conditionally independent increments. Similar to the computation of the quarterly critical values, the uniform critical value can be estimated via simulation. To be specific, we independently simulate Monte Carlo samples of the null distribution of each quarterly test statistics, based on which we construct a Monte Carlo sample for the supremum of all quarterly statistics, and then report the 99% quantile of the latter Monte Carlo sample as the uniform critical value at the 1% significance level.

Figure 1: Joint Test for Variance Spanning.



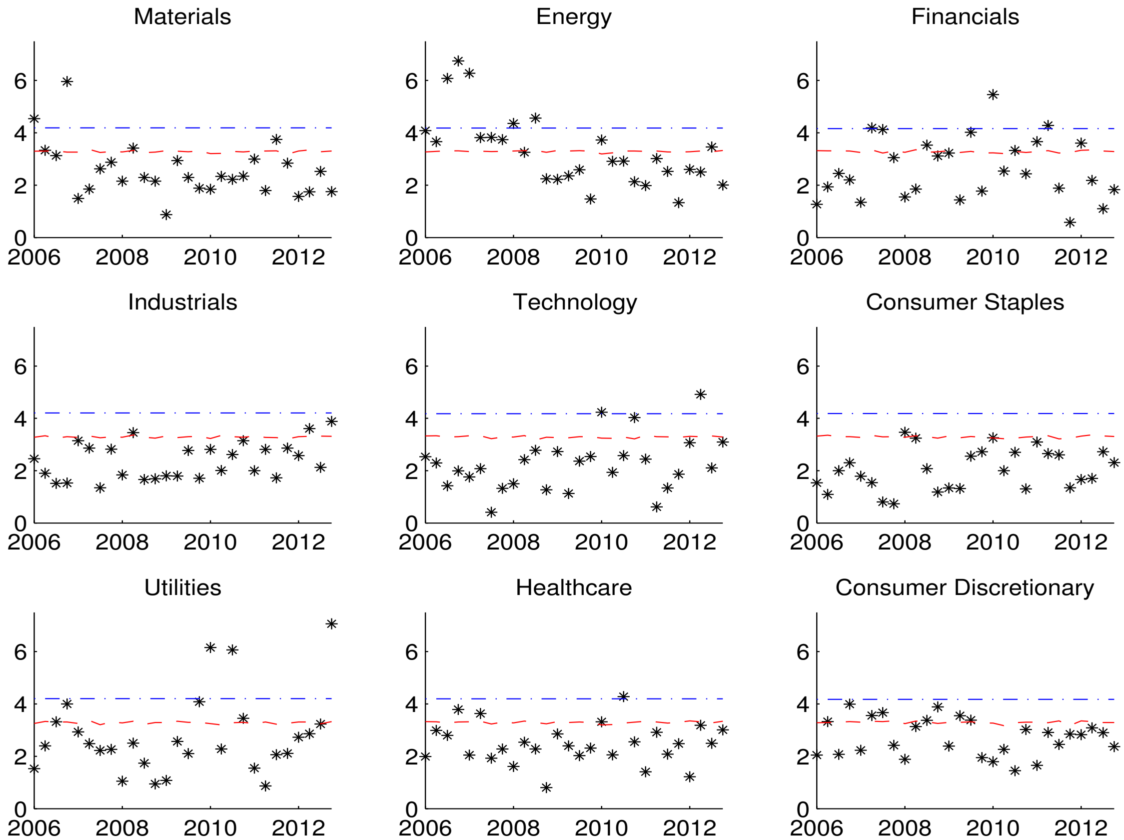
Asterisk: test statistic. Dashed line: 1% level pointwise critical value. Dash-dotted line: 1% level uniform critical value.

restriction (4.3) appears reasonable, in the sense that the specification (4.3) is not rejected jointly over all quarters (i.e., all quarterly test statistics are below the uniform critical value). These sectors include Consumer Staples, Consumer Discretionary and Industrials. On the other hand, for sectors such as Energy and Financials, we see that (4.3) is not only strongly rejected jointly for all quarters, but also rejected for many individual quarters over the entire sample.

While the tests on Figures 1 and 2 impose constancy of the parameters in (4.3) within each quarter, it is interesting to investigate whether the variance beta estimates are positive and whether they exhibit large variations across quarters in our sample.[9] On Figure 3 we plot the quarterly variance beta estimates, along with a two-sided 95% confidence band and a 99% lower confidence bound. We find strong evidence that the residual variances of sector portfolios, $(\widetilde{c}_{jj,t})_{j=2,\ldots,10}$, comove with the market variance, $c_{11,t}$, that is, variance betas are always positive with nontrivial statistical significance. Moreover, we can see from the plot that sectors like Consumer Staples, Consumer Discretionary and Industrials exhibit relatively small variation in their variance beta estimates. In sharp contrast, the variance betas of Financials and Energy vary substantially across quarters in our sample.

[9] In view of the evidence of misspecification mentioned above, we interpret the variance betas as pseudo-true parameters, or linear projection coefficients under the occupation measure, from model (4.3).

Figure 2: Industry-specific Tests for Variance Spanning.



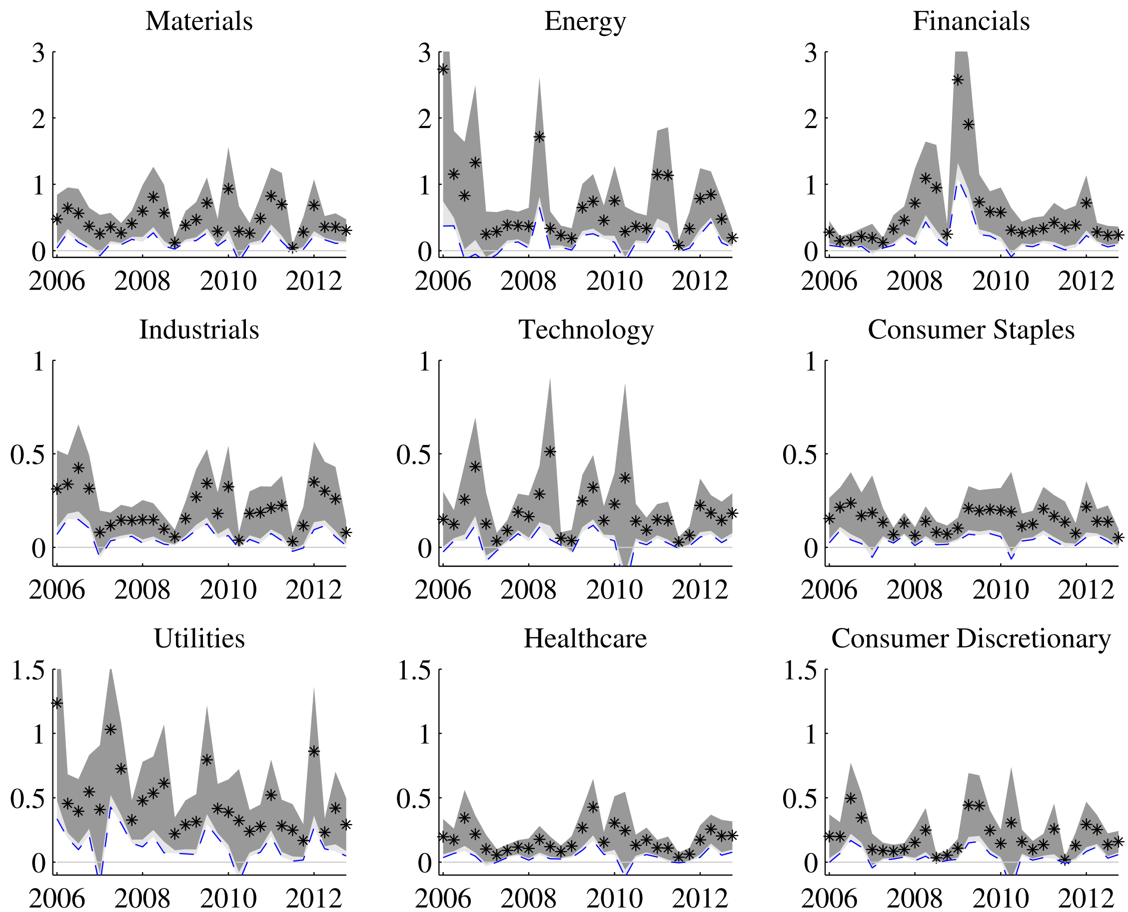
Asterisk: test statistic. Dashed line: 1% level pointwise critical value. Dash-dotted line: 1% level uniform critical value.

To sum up, we find that, for sector portfolios, the temporal variation of the market-neutral stochastic variances cannot be explained (i.e., linearly spanned) completely by the variation in the market variance. This finding is most pronounced for Financials and Energy sectors, suggesting that they load on important systematic variance risk factors that are not captured by the market variance risk.

4.2 Correlation of market-neutral returns of the sector portfolios

We next study the market-neutral correlation matrix $\tilde{\rho}_t$. Our goal is to specify and test model restrictions for the dynamic behavior of $\tilde{\rho}_t$ within each of the 28 quarters in the sample. Before diving into the within-quarter analysis, we briefly describe the between-quarter variations of correlations. In Figure 4, we plot the quarterly averages of the market-neutral correlation $\tilde{\rho}_{jk,t}$ and, for

Figure 3: Variance Betas.

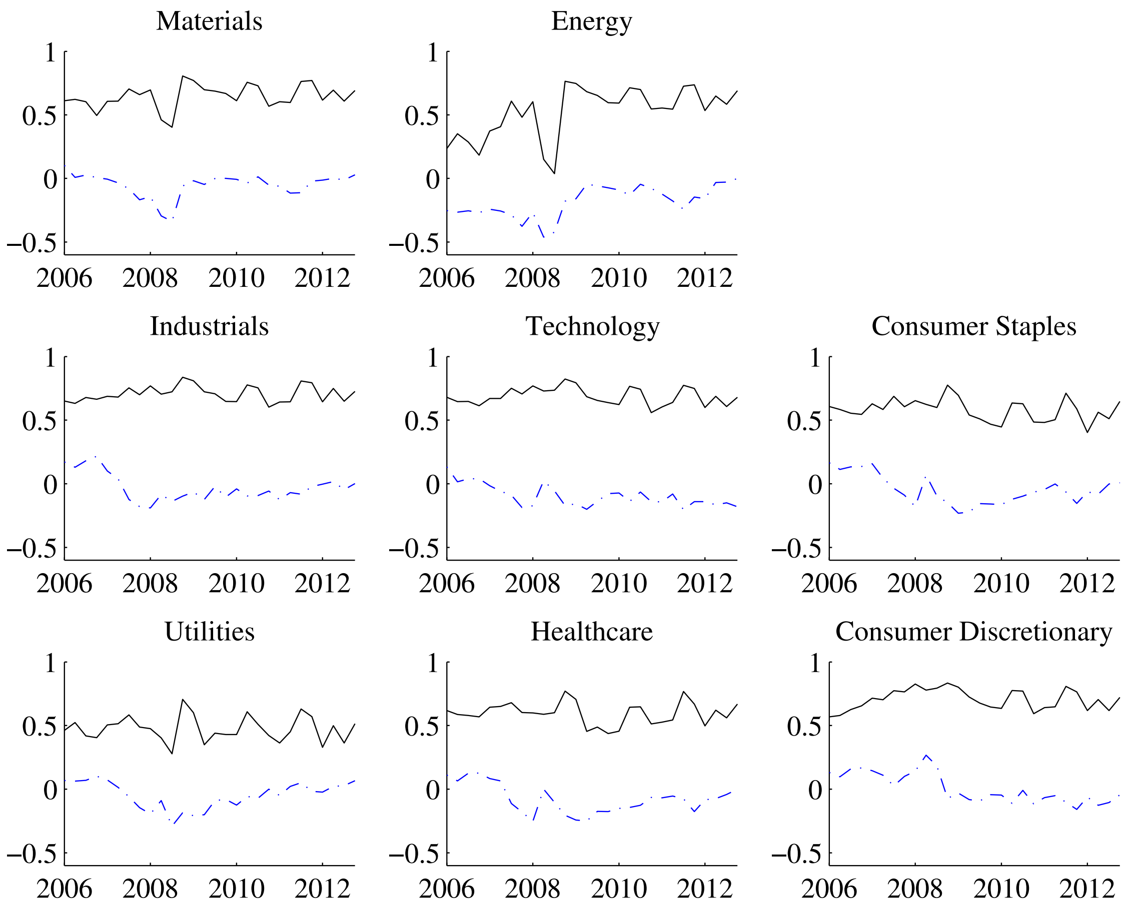


Asterisk: variance beta estimate. Shaded area: 95% pointwise two-sided confidence band. Dashed: 99% pointwise lower confidence bound.

comparison, the quarterly averages of the raw correlations $\rho_{jk,t} \equiv c_{jk,t}/\sqrt{c_{jj,t}c_{kk,t}}$. To save space, we only show correlations of the Financials sector with the other eight sectors, while noting that the empirical regularities discussed below are not limited to this choice.

Some interesting patterns emerge from Figure 4. We first note that the raw correlations are typically positive and high in magnitude, but the market-neutral correlations are much lower. This evidence suggests, perhaps not surprisingly, that a significant part of the positive correlations between the raw returns of the sector portfolios is driven by their (time-varying) loadings on the market returns. Second, the market-neutral correlations appear to be somewhat smoother than the corresponding raw correlations. This is in spite of the fact that $\widetilde{\rho}_t$ is harder to estimate than ρ_t, as the former involves more "layers of latency." The occasional spikes in ρ_t evident from Figure

Figure 4: Quarterly Average Correlations of Financials Sector with other Industry Portfolios.

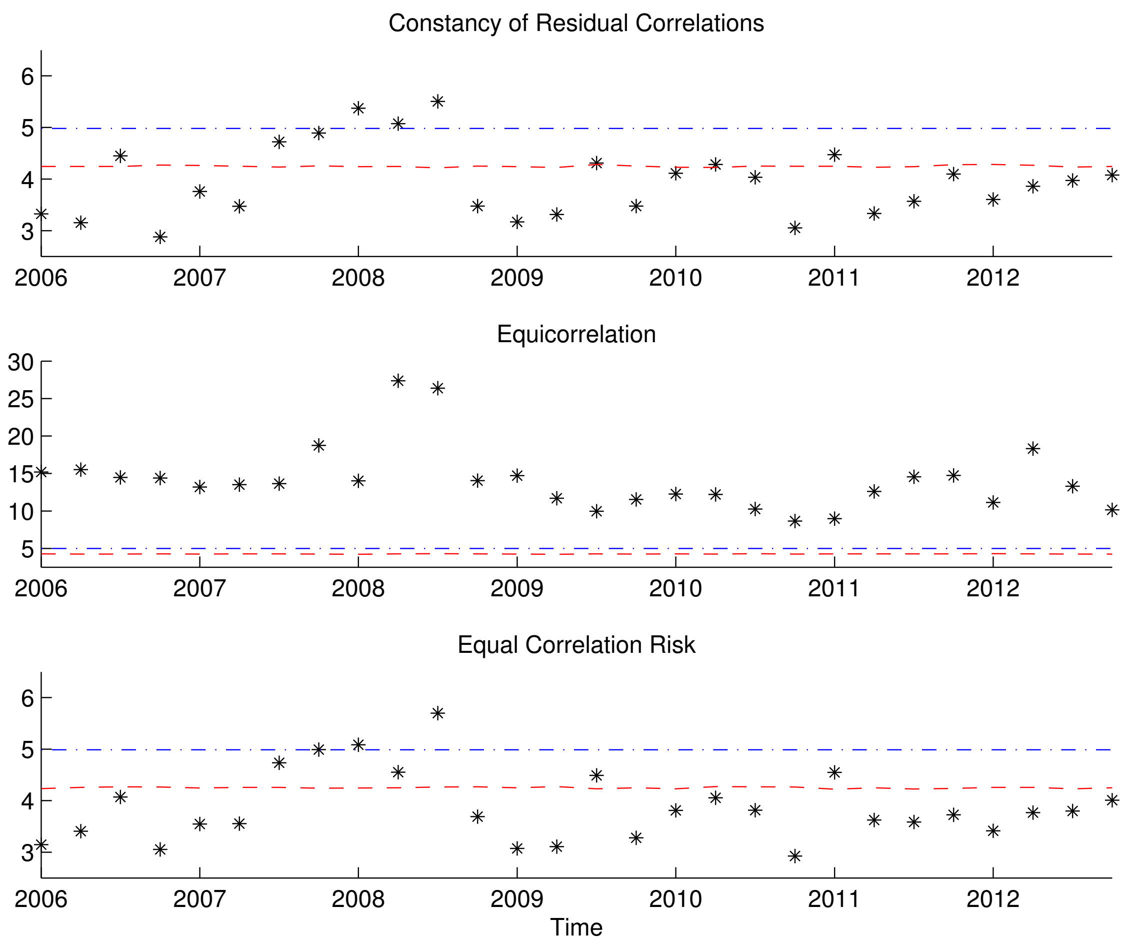


Solid line: raw correlations $\rho_{jk,t} \equiv c_{jk,t}/\sqrt{c_{jj,t}c_{kk,t}}$. Dash-dotted line: correlations of market-neutral returns $\tilde{\rho}_{jk,t} \equiv \tilde{c}_{jk,t}/\sqrt{\tilde{c}_{jj,t}\tilde{c}_{kk,t}}$.

4 can be explained with the factor structure of the industry portfolio returns and the (random) shocks in the market portfolio. Indeed, once we account for the exposure of the industry portfolios to the market risk, the correlations in the residual returns, $\tilde{\rho}_t$, are less erratic. Remarkably, the market-neutral correlations are fairly stable after 2009.[10] Nevertheless, we still see some variation in $\widetilde{\rho}_t$, particularly during the turbulent period of 2007-2008. This evidence suggests that we need more factors, than the market portfolio, to span the risks in the industry portfolios, a possibility that we shall explore later in this subsection.

[10]The post-2009 stability in the market-neutral correlation is shared by all pairs of sector portfolios. The market-neutral correlation between the Financials and other sectors in the post-crisis period is, interestingly, close to zero. The latter pattern is not shared by all sector pairs.

Figure 5: Correlation Risk Tests.



Top panel: constancy of partial correlation. Middle panel: equal partial correlations. Bottom panel: equal partial correlation risk. Asterisk: test statistic. Dashed line : 1% level pointwise critical value. Dash-dotted line: 1% level uniform critical value.

We now turn to the formal analysis of the market-neutral correlations within quarters. We start with testing two hypotheses. The first regards the presence of correlation risks in the market-neutral sector portfolio returns, that is, we study whether $\tilde{\rho}_t$ actually varies within each quarter; this amounts to a specification test of the model restriction

$$\tilde{\rho}_{jk,t} = \widetilde{R}_{jk}, \text{ for all } 2 \leq j, k, l, m \leq 10, \tag{4.4}$$

for some parameters $\widetilde{R}_{jk}$, which are time-invariant within each quarter but are allowed to differ across quarters. The second hypothesis is that market-neutral correlations can be restricted to be the same over the cross section, that is,

$$\tilde{\rho}_{jk,t} = \tilde{\rho}_{lm,t}, \text{ for all } 2 \leq j, k, l, m \leq 10. \tag{4.5}$$

This is essentially a test for an equicorrelation model, like that proposed in Engle and Kelly (2012).[11] We note that the equicorrelation restriction concerns the cross section, while allowing the spot correlation process to be time-varying and stochastic. The two restrictions, (4.4) and (4.5), respectively impose homogeneity on the temporal and the cross-sectional dimensions.

To cast the testing problem of the equicorrelation model (4.5) in the setting of Section 3.3, we define, for $2 \leq j < k \leq 10$,

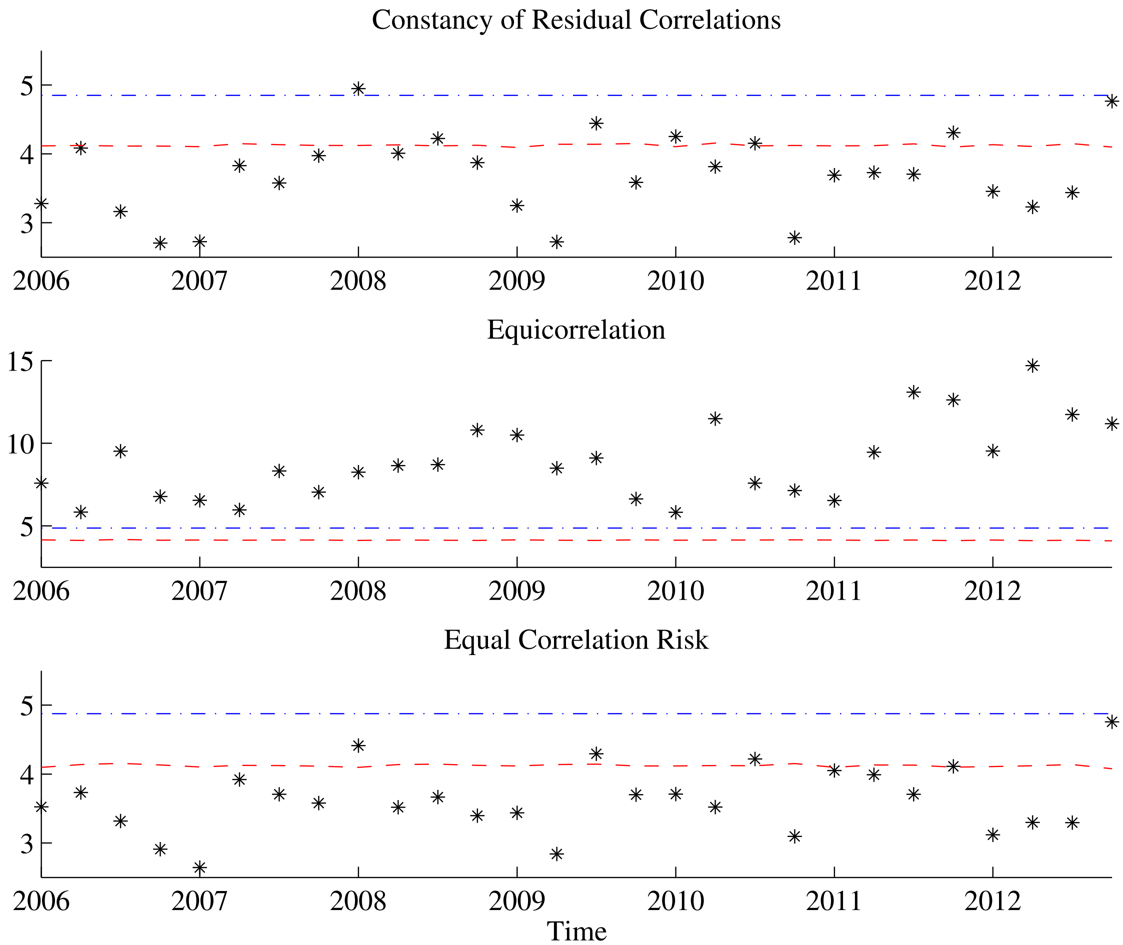
$$\tilde{y}_{jk}(c_t) = \tilde{\rho}_{jk,t} - \frac{1}{36} \sum_{2 \leq l < m \leq 10} \tilde{\rho}_{lm,t},$$

and let $\tilde{y}(c_t)$ be a vectorization of the double-indexed array $\tilde{y}_{jk}(c_t)$. By construction, $\tilde{y}(c_t)$ collects the deviation of the spot market-neutral correlation of each pair of sector ETFs from the average of the market-neutral correlations of all 36 such pairs. The equicorrelation hypothesis (4.5) can be equivalently written as $\tilde{y}(c_t) = 0$, which corresponds to a simple special case of (2.3) without the regressor $\tilde{x}(c_t)$ or the parameter θ.

The results of the two tests for the hypotheses in (4.4) and (4.5) are shown on the top two panels of Figure 5 respectively. From the top panel, we find evidence for the presence of correlation risk even within the quarterly horizon. The highest values of the test statistic for the hypothesis (4.4) (i.e., the null of no market-neutral correlation risk within the quarter) are during the last two quarters of 2007 and the first three quarters of 2008 which cover a relatively turbulent period in the market. The results from the equicorrelation test plotted on the middle panel of Figure 5 indicate a strong rejection, in every quarter, of this cross-sectional restriction among the (stochastic) market-neutral correlations. The equicorrelation test statistic also peaks during the 2008 crisis period.

[11] We stress however that we conduct the test for the market-neutral correlations and not for the raw correlations among sector portfolio returns. In particular, due to the presence of the market factor with stochastic variance (and possibly dynamic loadings on it), the stochastic correlation matrix of the raw returns in general does not have an equicorrelation structure even if (4.5) holds.

Figure 6: Correlation Risk Tests without Financials and Energy.



Top panel: constancy of partial correlation. Middle panel: equal partial correlations. Bottom panel: equal partial correlation risk. Asterisk: test statistic. Dashed line : 1% level pointwise critical value. Dash-dotted line: 1% level uniform critical value.

With the heteroskedasticity/heterogeneity on both time-series and cross-sectional dimensions formally documented above, we further investigate the plausibility of model restrictions that are less restrictive. We start with a hypothesis that the cross-sectional heterogeneity in the stochastic correlations are only in (time-invariant) levels, with their intraquarter time variation driven by a common factor. That is, we test whether the following is true

$$\tilde{\rho}_{jk,t} - \tilde{\rho}_{lm,t} = \widetilde{\delta}^{\rho}_{jk,lm}, \text{ for all } 2 \leq j, k, l, m \leq 10, \tag{4.6}$$

for some time-invariant parameters $\widetilde{\delta}^{\rho}_{jk,lm}$. We refer to this hypothesis as a hypothesis for *equal correlation risk (ECR).* This ECR restriction is perhaps the most parsimonious one that allows the market-neutral correlation processes to be time-varying and cross-sectionally heterogenous. The testing results for ECR are reported on the bottom panel of Figure 5. As seen from the figure, the hypothesis is rejected jointly for all quarters and 7 individual quarters, with the test statistics peaked around the 2008 financial crisis. Nevertheless, the ECR hypothesis is not rejected for 21 out of 28 quarters, which is in sharp contrast to the strong rejection of the equicorrelation test reported above.

Finally, we explore whether additional factors can account for the presence of correlation risk in the market-neutral components of the industry portfolio returns. Given the special role of the financial sector during the 2008 crisis and the distinct role of oil shocks in general, we use them as additional factors to span dependencies between the industry portfolios. We thus replace (4.2) with the following decomposition of the sector portfolio returns:

$$dX_{j,t} = b_{j,t}dt + \beta_{j,t}dX^c_{1,t} + \beta'_{j,t}df_{1,t} + \beta''_{j,t}df_{2,t} + d\widetilde{X}^c_{j,t} + dJ_{j,t}, \quad j = 2, ..., 10, \tag{4.7}$$

where the additional systematic risk factors f_1 and f_2 are taken as the diffusive components of returns of Financials and Energy sector portfolios and $\widetilde{X}^c_j$ is, by definition, orthogonal to X^c_1, f_1 and f_2. Clearly, the residual diffusive return $d\widetilde{X}^c_{j,t}$ defined by (4.7) is different from that defined by (4.2). In particular, the residual returns of Financials and Energy sectors are identically zero by construction. The other seven sectors give 21 pairs of residual correlation processes, for which we apply tests similar to those reported in Figure 5. Figure 6 shows the testing results. We see that the constancy of the residual correlations is rejected jointly over all quarters, but now only marginally so.[12] The equicorrelation hypothesis is again strongly rejected. The equal correlation risk hypothesis is rejected in a few quarters, but the test does not reject jointly for all quarters.

[12] As further evidence for the decreased correlation risk, the quarterly averages of the residual returns, after controlling for the sensitivity towards market returns and Financial and Energy sector returns, become closer to zero and somewhat smoother compared with the corresponding quarterly averages of the market-neutral return correlations.

This finding suggests that the equal correlation risk is a reasonable restriction for the residual correlation structure after controlling for the exposure to the Financial and Energy sectors returns as well as the market returns.[13]

To conclude, we find that the market-neutral correlation structure is time-varying and cross-sectionally heterogeneous. In particular, the equicorrelation hypothesis and the ECR hypothesis, especially the former, are both rejected. When we further control for the sensitivity to the Financial and Energy sectors returns, the residual correlation within quarters decreases and the ECR hypothesis is not rejected for the residual correlation dynamics jointly for all quarters in our sample. These findings suggest that a three-factor model with ECR may capture the dynamic correlation structure of the sector portfolios.

5 Conclusion

This paper develops inference theory for models involving the covariance occupation measure of a discretely-observed multivariate Itô semimartingale. Time-invariant (but possibly random) relations between nonlinear transforms of the elements of the stochastic spot covariance matrix of a multivariate stochastic process arise in many applications in economics and finance, such as factor models. We propose minimum distance type estimators for the random parameters of the above relations. We prove consistency and asymptotic mixed normality of our estimators. We further derive specification tests for the path-wise models concerning the covariance occupation measure of the discretely-observed multivariate process. We use the developed inference techniques to study the variance risk of a set of well-diversified industry portfolios comprising the S&P 500 index market portfolio.

Our empirical results indicate the presence of sector-specific variance risks in addition to that of the market. We further document time variation in cross-sectional correlations of market-neutral industry returns. The magnitude and cross-sectional heterogeneity of the residual industry return correlations decreases when, in addition to the market, we control for exposure to shocks in the financial and energy sectors.

[13] A by-product of using the Financials and the Energy sector portfolio returns to span the systematic returns is that they are effectively excluded in the testing exercise, in the sense that their residual correlations with other sectors are, by construction, zero. It is conceivable that, even in the one-factor model (4.2), if we exclude the Financials and the Energy sector, the ECR hypothesis may not be rejected, because the joint test would involve less restrictions. We implement our tests in this setting and find that the ECR hypothesis is still rejected at the 1% level jointly for all quarters.

6 Appendix: Proofs

Throughout the appendix, we use K to denote a generic positive constant which may change from line to line. For a generic sequence $Z_n(\tau)$ of τ-indexed processes, we write $Z_n = o_{pu}(1)$ if $Z_n(\tau) \xrightarrow{\mathbb{P}} 0$ uniformly in $\tau \in \mathcal{T}$. Similarly, for a double-indexed process $Z_n(\tau, \eta)$, $\tau, \eta \in \mathcal{T}$, we write $Z_n = o_{pu}(1)$ if $Z_n(\tau, \eta) \xrightarrow{\mathbb{P}} 0$ uniformly in τ, η.

By a standard localization argument (see Section 4.4.1 of Jacod and Protter (2012)), we can strengthen Assumption HF to Assumption SHF below, without loss of generality.

Assumption SHF: We have Assumption HF. The processes b_t, σ_t, $\tilde{b}_t$ and $\tilde{\sigma}_t$ are uniformly bounded. There exists a bounded λ-integrable function Γ on $\mathbb{R}$, such that $\|\delta(\omega, t, u)\|^r \leq \Gamma(u)$ and $\|\tilde{\delta}(\omega, t, u)\|^2 \leq \Gamma(u)$ for all $(\omega, t, u) \in \Omega \times [0, T] \times \mathbb{R}$. Finally, c_t takes value in a convex compact set $\mathcal{K}$ for all $t \leq T$.

Proof of Theorem 1. *Step 1.* By arguing component by component, we can assume that g is scalar-valued without loss of generality. By Lemma 1 of Li et al. (2014), the variables $(\hat{c}_{i\Delta_n})_{i\geq 0}$ are uniformly bounded with probability approaching one. By using the spatial location argument as in Theorem 2 of Li et al. (2014), we can assume that g is compactly supported without loss of generality. In this step, we fix some $\tau \in \mathcal{T}$, and show that

$$\widehat{\mathbb{F}}_\tau^n g \xrightarrow{\mathbb{P}} \mathbb{F}_\tau g. \tag{6.1}$$

We set $\hat{c}_t^+ = \hat{c}_{i\Delta_n}$ and $w_{n,\tau}(t) = w_\tau(i\Delta_n)$ for $t \in [(i-1)\Delta_n, i\Delta_n)$. Then

$$\begin{aligned} \widehat{\mathbb{F}}_\tau^n g &= \Delta_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} g(\hat{c}_{i\Delta_n}) w_\tau(i\Delta_n) \\ &= \Delta_n g(\hat{c}_0) w_\tau(0) + \int_0^{(\lfloor T/\Delta_n \rfloor - k_n)\Delta_n} g(\hat{c}_s^+) w_{n,\tau}(s) ds. \end{aligned}$$

Since g and $w_\tau(\cdot)$ are bound, we further deduce

$$\left|\widehat{\mathbb{F}}_\tau^n g - \mathbb{F}_\tau g\right| \leq K k_n \Delta_n + \int_0^{(\lfloor T/\Delta_n \rfloor - k_n)\Delta_n} \left| g(\hat{c}_s^+) w_{n,\tau}(s) - g(c_s) w_\tau(s) \right| ds.$$

By Theorem 9.3.2 of Jacod and Protter (2012), $\hat{c}_s^+ \xrightarrow{\mathbb{P}} c_s$ for each $s \in [0, T]$. Since g and w_τ are continuous, $g(\hat{c}_s^+) w_{n,\tau}(s) \xrightarrow{\mathbb{P}} g(c_s) w_\tau(s)$ for each $s \in [0, T]$. By the bounded convergence theorem, we derive $\widehat{\mathbb{F}}_\tau^n g \xrightarrow{\mathbb{P}} \mathbb{F}_\tau g$.

Step 2. By Assumption WF, $\sup_{s\in[0,T]} |w_\tau(s) - w_\eta(s)| \leq K|\tau - \eta|$. Note that for any $\tau, \eta \in \mathcal{T}$,

$$\left|\widehat{\mathbb{F}}_\tau^n g - \widehat{\mathbb{F}}_\eta^n g\right| \leq K \widehat{\mathbb{F}}^n (|g|) |\tau - \eta|.$$

By Step 1, $\widehat{\mathbb{F}}^n(|g|) = O_p(1)$. Hence, the sequence of τ-indexed processes $\widehat{\mathbb{F}}^n_\tau g$ is stochastically equicontinuous. Combining this with the pointwise convergence in (6.1), we have $\widehat{\mathbb{F}}^n_\tau g \xrightarrow{\mathbb{P}} \mathbb{F}_\tau g$ uniformly.

Step 3. Finally, we note that by essentially the same argument as above, we can also show $\widehat{\mathbb{F}}^n_{\tau,\eta} g \xrightarrow{\mathbb{P}} \mathbb{F}_{\tau,\eta} g$ uniformly. The details are omitted. *Q.E.D.*

Proof of Theorem 2. *Step 1.* By Lemma 1 and the spatial localization argument underlying Theorem 2 of Li et al. (2014), we can again assume that the variables $(\hat{c}_{i\Delta_n})_{i\geq 0}$ are uniformly bounded and that g is compactly supported without loss of generality. In this step, we outline the proof of Theorem 2. We denote the continuous component of X by

$$X'_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s.$$

The spot covariance estimator for X' is given by, for $i \in \{0, \ldots, \lfloor T/\Delta_n \rfloor - k_n\}$,

$$\hat{c}'_{i\Delta_n} \equiv \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n} (\Delta^n_{i+j} X')(\Delta^n_{i+j} X')^\intercal.$$

We then set

$$\alpha^n_i = \left(\Delta^n_{i+1} X'\right) \left(\Delta^n_{i+1} X'\right)^\intercal - c_{i\Delta_n}\Delta_n, \quad \beta^n_i = \hat{c}'_{i\Delta_n} - c_{i\Delta_n},$$

and denote the (j,k) components of α^n_i and β^n_i by $\alpha^n_{jk,i}$ and $\beta^n_{jk,i}$ respectively.

We consider the decomposition

$$\Delta_n^{-1/2} \left(\mathbb{F}^n_\tau g - \mathbb{F}_\tau g\right) = \sum_{j=1}^{5} V_{j,n}(\tau),$$

where

$$
\begin{aligned}
V_{1,n}(\tau) \quad &\equiv \quad \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} (g(\hat{c}_{i\Delta_n}) - g(\hat{c}'_{i\Delta_n}))w_\tau(i\Delta_n) \\
&\quad -\Delta_n^{1/2} k_n^{-1} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} (\mathbb{B}g(\hat{c}_{i\Delta_n}) - \mathbb{B}g(\hat{c}'_{i\Delta_n}))w_\tau(i\Delta_n), \\
V_{2,n}(\tau) \quad &\equiv \quad \Delta_n^{-1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \int_{i\Delta_n}^{(i+1)\Delta_n} (g(c_{i\Delta_n})w_\tau(i\Delta_n) - g(c_s)w_\tau(s))ds \\
&\quad -\Delta_n^{-1/2} \int_{(\lfloor T/\Delta_n \rfloor - k_n + 1)\Delta_n}^{T} g(c_s)w_\tau(s)ds, \\
V_{3,n}(\tau) \quad &\equiv \quad \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n})w_\tau(i\Delta_n) \frac{1}{k_n} \sum_{u=1}^{k_n} (c_{lm,(i+u-1)\Delta_n} - c_{lm,i\Delta_n}), \\
V_{4,n}(\tau) \quad &\equiv \quad \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \Big(g(c_{i\Delta_n} + \beta_i^n) - g(c_{i\Delta_n}) \\
&\quad - \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n})\beta_{lm,i}^n - k_n^{-1}\mathbb{B}g(\hat{c}'_{i\Delta_n}) \Big) w_\tau(i\Delta_n), \\
V_{5,n}(\tau) \quad &\equiv \quad \Delta_n^{-1/2} k_n^{-1} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n}) \sum_{u=1}^{k_n} \alpha_{lm,i+u-1}^n \right) w_\tau(i\Delta_n).
\end{aligned}
$$

We prove the assertion of the theorem by showing

$$
\sup_{\tau \in \mathcal{T}} |V_{j,n}(\tau)| = o_p(1), \quad \text{for } j = 1, 2, 3, 4, \tag{6.2}
$$

and

$$
V_{5,n}(\cdot) \xrightarrow{\mathcal{L}\text{-}s} \xi(\cdot), \tag{6.3}
$$

where $\xi(\cdot)$ denotes a process that is defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$ and, conditional on $\mathcal{F}$, is centered Gaussian with covariance function $M(\cdot,\cdot)$.

Step 2. This step contains four substeps which respectively show (6.2) for $j = 1, 2, 3, 4$. By using a componentwise argument, we can assume that g is $\mathbb{R}$-valued without loss of generality. Below, the notations ζ_i^n, $\zeta_i'^n$ and $\zeta_i''^n$ are defined differently across substeps 2(i)–2(iv).

Step 2(i). Since g is compactly supported, its derivatives are bounded. Then, by a mean-value expansion and (4.8) in Jacod and Rosenbaum (2013),

$$
\mathbb{E}\left| g(\hat{c}_{i\Delta_n}) - g(\hat{c}'_{i\Delta_n}) \right| \leq K a_n \Delta_n^{(2-r)\varpi} \tag{6.4}
$$

for some deterministic sequence a_n that satisfies $a_n \to 0$. Under Assumption WF, it is easy to see that

$$\sup_{\tau \in \mathcal{T}, s \in [0,T]} |w_\tau(s)| < \infty. \tag{6.5}$$

By (6.4), (6.5) and the triangle inequality,

$$\mathbb{E}\left[\sup_{\tau\in\mathcal{T}}\left|\Delta_n^{1/2}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}(g(\hat{c}_{i\Delta_n})-g(\hat{c}'_{i\Delta_n}))w_\tau(i\Delta_n)\right|\right]\leq Ka_n\Delta_n^{(2-r)\varpi-1/2}\to 0, \tag{6.6}$$

where the convergence follows from $\varpi \geq 1/2(2-r)$, which is implied by Assumption LW. Similarly,

$$\mathbb{E}\left[\sup_{\tau\in\mathcal{T}}\left|\Delta_n^{1/2}k_n^{-1}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}(\mathbb{B}g(\hat{c}_{i\Delta_n})-\mathbb{B}g(\hat{c}'_{i\Delta_n}))w_\tau(i\Delta_n)\right|\right]\to 0. \tag{6.7}$$

Combining (6.6) and (6.7), we derive (6.2) for $j = 1$.

Step 2(ii). By (6.5) and $k_n^2\Delta_n \to 0$ (Assumption LW), it is easy to see that

$$\sup_{\tau\in\mathcal{T}}\left|\Delta_n^{-1/2}\int_{(\lfloor T/\Delta_n\rfloor-k_n+1)\Delta_n}^{T}g(c_s)w_\tau(s)ds\right|=O_p(k_n\Delta_n^{1/2})=o_p(1).$$

To simplify notation, we set

$$R_{2,n}(\tau)\equiv\Delta_n^{-1/2}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\int_{i\Delta_n}^{(i+1)\Delta_n}(g(c_{i\Delta_n})w_\tau(i\Delta_n)-g(c_s)w_\tau(s))ds.$$

We decompose

$$R_{2,n}(\tau)=R_{2,1,n}(\tau)+R_{2,2,n}(\tau)+R_{2,3,n}(\tau),$$

where

$$\begin{aligned}
R_{2,1,n}(\tau) &\equiv \Delta_n^{-1/2}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\int_{i\Delta_n}^{(i+1)\Delta_n}g(c_{i\Delta_n})\left(w_\tau(i\Delta_n)-w_\tau(s)\right)ds,\\
R_{2,2,n}(\tau) &\equiv \Delta_n^{-1/2}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\int_{i\Delta_n}^{(i+1)\Delta_n}(g(c_{i\Delta_n})-g(c_s)-\sum_{l,m=1}^{d}\partial_{lm}g(c_{i\Delta_n})(c_{lm,i\Delta_n}-c_{lm,s}))w_\tau(s)ds,\\
R_{2,3,n}(\tau) &\equiv \Delta_n^{-1/2}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\int_{i\Delta_n}^{(i+1)\Delta_n}\sum_{l,m=1}^{d}\partial_{lm}g(c_{i\Delta_n})(c_{lm,i\Delta_n}-c_{lm,s})w_\tau(s)ds.
\end{aligned}$$

By Assumption WF, $\sup_{\tau\in\mathcal{T}}|w_\tau(i\Delta_n)-w_\tau(s)|\leq K\Delta_n$ for any $s\in[i\Delta_n,(i+1)\Delta_n]$. Hence,

$$\sup_{\tau\in\mathcal{T}}|R_{2,1,n}(\tau)|\leq K\Delta_n^{1/2}\to 0. \tag{6.8}$$

Next, consider $R_{2,2,n}(\tau)$. By a mean-value expansion and the boundedness of $\|c_s\|$, we have

$$\left| g(c_{i\Delta_n}) - g(c_s) - \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n})(c_{lm,i\Delta_n} - c_{lm,s}) \right| \le K \left\| c_s - c_{i\Delta_n} \right\|^2 .$$

By a standard estimate, $\mathbb{E}\left\| c_s - c_{i\Delta_n} \right\|^2 \le K \left| s - i\Delta_n \right|$, where K does not depend on s and i. Since $w_\tau(s)$ is uniformly bounded in τ and s, we readily derive

$$\mathbb{E}\left[\sup_{\tau \in \mathcal{T}} \left| R_{2,2,n}(\tau) \right| \right] \le K \Delta_n^{1/2}. \tag{6.9}$$

Now turn to $R_{2,3,n}(\tau)$. We first show that $R_{2,3,n}(\tau)$ is stochastically equicontinuous. Let $\tau, \eta \in \mathcal{T}$. We observe

$$\begin{aligned}
& \left| R_{2,3,n}(\tau) - R_{2,3,n}(\eta) \right| \\
\le \; & \Delta_n^{-1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \int_{i\Delta_n}^{(i+1)\Delta_n} \left| \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n})(c_{lm,i\Delta_n} - c_{lm,s}) \left(w_\tau(s) - w_\eta(s) \right) \right| ds \\
\le \; & K \left| \tau - \eta \right| \Delta_n^{-1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \int_{i\Delta_n}^{(i+1)\Delta_n} \left| \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n})(c_{lm,i\Delta_n} - c_{lm,s}) \right| ds \\
= \; & \left| \tau - \eta \right| O_p(1),
\end{aligned}$$

where the first inequality follows from the triangle inequality, the second equality is by Assumption WF, and the last line is derived by observing that $\left| \partial_{lm} g(c_{i\Delta_n}) \right|$ is bounded and $\mathbb{E}\left\| c_s - c_{i\Delta_n} \right\| \le K\Delta_n^{1/2}$. The stochastic equicontinuity of $R_{2,3,n}(\tau)$ readily follows.

In order to show

$$\sup_{\tau \in \mathcal{T}} \left| R_{2,3,n}(\tau) \right| = o_p(1), \tag{6.10}$$

it remains to show that $R_{2,3,n}(\tau) = o_p(1)$ for each fixed $\tau \in \mathcal{T}$. To simplify notation, let

$$\zeta_i^n(\tau) \equiv \int_{i\Delta_n}^{(i+1)\Delta_n} \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n})(c_{lm,i\Delta_n} - c_{lm,s}) w_\tau(s) ds.$$

We denote $\zeta_i'^n(\tau) \equiv \mathbb{E}\left[\zeta_i^n(\tau) | \mathcal{F}_{i\Delta_n} \right]$ and $\zeta_i''^n(\tau) \equiv \zeta_i^n(\tau) - \zeta_i'^n(\tau)$. Then we can decompose

$$R_{2,3,n}(\tau) = R'_{2,3,n}(\tau) + R''_{2,3,n}(\tau), \tag{6.11}$$

where

$$R'_{2,3,n}(\tau) = \Delta_n^{-1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i'^n(\tau), \quad R''_{2,3,n}(\tau) = \Delta_n^{-1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i''^n(\tau).$$

Note that for $s \geq i\Delta_n$, $\|\mathbb{E}[c_s - c_{i\Delta_n}|\mathcal{F}_{i\Delta_n}]\| \leq K|s - i\Delta_n|$. Hence, $\mathbb{E}|\zeta_i'^n(\tau)| \leq K\Delta_n^2$ and

$$\mathbb{E}\left|R_{2,3,n}'(\tau)\right| \leq K\Delta_n^{1/2}. \tag{6.12}$$

Note that $\zeta_i''^n(\tau)$ forms an array of martingale differences by construction. Moreover, $\mathbb{E}|\zeta_i''^n(\tau)|^2 \leq K\Delta_n^3$. Hence, for each $\tau \in \mathcal{T}$, $\mathbb{E}\left|R_{2,3,n}''(\tau)\right|^2 \leq K\Delta_n$, which further implies

$$R_{2,3,n}''(\tau) = o_p(1). \tag{6.13}$$

Combining (6.11)–(6.13), we derive (6.10). We then derive (6.2) for $j = 2$ by combining (6.8)–(6.10).

Step 2(iii). To simplify notation, we set (by overloading the notation ζ_i^n)

$$\zeta_i^n \equiv \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n}) \frac{1}{k_n} \sum_{u=1}^{k_n} (c_{lm,(i+u-1)\Delta_n} - c_{lm,i\Delta_n}),$$

so that $V_{3,n}(\tau)$ can be written as $V_{3,n}(\tau) = \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i^n w_\tau(i\Delta_n)$. We then set $\zeta_i'^n = \mathbb{E}[\zeta_i^n|\mathcal{F}_{i\Delta_n}]$ and $\zeta_i''^n = \zeta_i^n - \zeta_i'^n$. It is easy to see that

$$\left|\zeta_i'^n\right| \leq K k_n \Delta_n, \quad \mathbb{E}\left|\zeta_i''^n\right|^2 \leq K k_n \Delta_n. \tag{6.14}$$

Now, consider the decomposition

$$V_{3,n}(\tau) = V_{3,n}'(\tau) + V_{3,n}''(\tau),$$

where

$$\begin{aligned} V_{3,n}'(\tau) &\equiv \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i'^n w_\tau(i\Delta_n), \\ V_{3,n}''(\tau) &\equiv \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i''^n w_\tau(i\Delta_n). \end{aligned}$$

By (6.14) and $k_n^2 \Delta_n \to 0$ (Assumption LW), we derive

$$\sup_{\tau \in \mathcal{T}} \left|V_{3,n}'(\tau)\right| \leq K\Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left|\zeta_i'^n\right| = O_p(k_n \Delta_n^{1/2}) = o_p(1). \tag{6.15}$$

Now, note that $\zeta_i''^n$ and $\zeta_l''^n$ are uncorrelated whenever $|i - l| \geq k_n$. Hence, by the Cauchy–Schwarz inequality, (6.5), (6.14) and Assumption LW,

$$\mathbb{E}\left|V_{3,n}''(\tau)\right|^2 \leq K\Delta_n k_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \mathbb{E}\left|\zeta_i''^n\right|^2 \leq K k_n^2 \Delta_n \to 0.$$

In particular, $V''_{3,n}(\tau) = o_p(1)$ for each $\tau \in \mathcal{T}$. We now show that $V''_{3,n}(\tau)$ is stochastically equicontinuous. Note that for $\tau, \eta \in \mathcal{T}$, by the Cauchy–Schwarz inequality, (6.14) and Assumption WF,

$$\begin{aligned}
& \mathbb{E}\left|V''_{3,n}(\tau) - V''_{3,n}(\eta)\right|^2 \\
= \; & \mathbb{E}\left|\Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i''^n \left(w_\tau(i\Delta_n) - w_\eta(i\Delta_n)\right)\right|^2 \\
\leq \; & K\left(\Delta_n k_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \mathbb{E}\left|\zeta_i''^n\right|^2\right) |\tau - \eta|^2 \\
\leq \; & K\left|\tau - \eta\right|^2.
\end{aligned}$$

Therefore, the L_2 norm of $V''_{3,n}(\tau) - V''_{3,n}(\eta)$ is bounded by $K\left|\tau - \eta\right|$. Since $\mathcal{T}$ is a one-dimensional space, $V''_{3,n}(\tau)$ is stochastic equicontinuous. Combining this with the pointwise convergence, we have

$$V''_{3,n}(\tau) = o_{pu}(1). \tag{6.16}$$

Combining (6.15) and (6.16), we derive (6.2) for $j = 3$.

Step 2(iv). In this substep, we set

$$\begin{aligned}
\zeta_i^n \; & \equiv \; g(c_{i\Delta_n} + \beta_i^n) - g(c_{i\Delta_n}) - \sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n}) \beta_{lm,i}^n - k_n^{-1} \mathbb{B}g(\hat{c}'_{i\Delta_n}), \\
\zeta_i'^n \; & \equiv \; \frac{1}{2} \sum_{j,k,l,m=1}^{d} \partial_{jk,lm}^2 g(c_{i\Delta_n}) \left(\beta_{jk,i}^n \beta_{lm,i}^n - \frac{1}{k_n}\left(c_{jl,i\Delta_n} c_{km,i\Delta_n} + c_{jm,i\Delta_n} c_{kl,i\Delta_n}\right)\right) \\
\zeta_i''^n \; & \equiv \; \zeta_i^n - \zeta_i'^n.
\end{aligned}$$

Note that $V_{4,n}(\tau) = \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \zeta_i^n w_\tau(i\Delta_n)$. In the proof of Lemma 4.4 in Jacod and Rosenbaum (2013), it was shown that

$$\mathbb{E}\left[\Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\left|\zeta_i''^n\right| + \left|\mathbb{E}\left[\zeta_i'^n | \mathcal{F}_{i\Delta_n}\right]\right|\right)\right] \to 0.$$

Hence, by (6.5), we have

$$\sup_{\tau \in \mathcal{T}} \left|\Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\zeta_i''^n + \mathbb{E}\left[\zeta_i'^n | \mathcal{F}_{i\Delta_n}\right]\right) w_\tau(i\Delta_n)\right| = o_p(1).$$

To show (6.2) for $j = 4$, it remains to show that,

$$V'_{4,n}(\tau) \equiv \Delta_n^{1/2} \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\zeta_i'^n - \mathbb{E}\left[\zeta_i'^n | \mathcal{F}_{i\Delta_n}\right]\right) w_\tau(i\Delta_n) = o_{pu}(1). \tag{6.17}$$

As shown in Jacod and Rosenbaum (2013) (see (4.11)), $\mathbb{E}\,\|\beta_i^n\|^4 \le K(k_n^{-2}+k_n\Delta_n)$. By Assumption LW, we have

$$\mathbb{E}\left|\zeta_i'^n\right|^2 \le K(k_n^{-2}+k_n\Delta_n) \le Kk_n\Delta_n.$$

Further note that $\zeta_i'^n - \mathbb{E}\left[\zeta_i'^n|\mathcal{F}_{i\Delta_n}\right]$ and $\zeta_l'^n - \mathbb{E}\left[\zeta_l'^n|\mathcal{F}_{l\Delta_n}\right]$ are uncorrelated whenever $|i-l|\ge k_n$. Hence, for $\tau,\eta\in\mathcal{T}$,

$$\begin{aligned}
\mathbb{E}|V_{4,n}'(\tau)-V_{4,n}'(\eta)|^2 &= \Delta_n\mathbb{E}\left[\left(\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\left(\zeta_i'^n-\mathbb{E}\left[\zeta_i'^n|\mathcal{F}_{i\Delta_n}\right]\right)\left(w_\tau(i\Delta_n)-w_\eta(i\Delta_n)\right)\right)^2\right]\\
&\le K\Delta_n k_n\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\mathbb{E}\left|\zeta_i'^n\right|^2\left(w_\tau(i\Delta_n)-w_\eta(i\Delta_n)\right)^2\\
&\le K\left|\tau-\eta\right|^2.
\end{aligned}$$

Since $\mathcal{T}\subseteq\mathbb{R}$, the above estimate implies that $V_{4,n}'(\tau)$ is stochastically equicontinuous. The proof of (6.17) is now reduced to showing the pointwise convergence, that is, $V_{4,n}'(\tau)=o_p(1)$ for fixed $\tau\in\mathcal{T}$. To see this, we observe that

$$\mathbb{E}\left|V_{4,n}'(\tau)\right|^2 \le K\Delta_n k_n\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\mathbb{E}\left|\zeta_i'^n\right|^2 \le Kk_n^2\Delta_n\to 0.$$

This finishes the proof of (6.2) for $j=4$.

Step 3. In this step, we show (6.3). The finite-dimensional convergence is obtained by a straightforward extension of the proof of Lemma 4.5 in Jacod and Rosenbaum (2013). It remains to show that the process $V_{n,5}(\tau)$ is stochastic equicontinuous.

We set $\alpha_{lm,i}'^n=\mathbb{E}[\alpha_{lm,i}^n|\mathcal{F}_{i\Delta_n}]$ and $\alpha_{lm,i}''^n=\alpha_{lm,i}^n-\alpha_{lm,i}'^n$. We can decompose

$$V_{n,5}(\tau)=V_{n,5}'(\tau)+V_{n,5}''(\tau),$$

where

$$\begin{aligned}
V_{n,5}'(\tau) &\equiv \Delta_n^{-1/2}k_n^{-1}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\left(\sum_{l,m=1}^{d}\partial_{lm}g(c_{i\Delta_n})\sum_{u=1}^{k_n}\alpha_{lm,i+u-1}'^n\right)w_\tau(i\Delta_n),\\
V_{n,5}''(\tau) &\equiv \Delta_n^{-1/2}k_n^{-1}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\left(\sum_{l,m=1}^{d}\partial_{lm}g(c_{i\Delta_n})\sum_{u=1}^{k_n}\alpha_{lm,i+u-1}''^n\right)w_\tau(i\Delta_n).
\end{aligned}$$

Note that

$$\sup_{\tau\in\mathcal{T}}\left\|V_{n,5}'(\tau)\right\|\le K\Delta_n^{-1/2}k_n^{-1}\sum_{i=0}^{\lfloor T/\Delta_n\rfloor-k_n}\left(\sum_{l,m=1}^{d}\sum_{u=1}^{k_n}\left\|\alpha_{lm,i+u-1}'^n\right\|\right)\xrightarrow{\mathbb{P}}0,$$

where the inequality follows from (6.5) and the convergence is derived by using (4.12) and Lemma 4.2 in Jacod and Rosenbaum (2013).

It remains to show that $V''_{n,5}(\tau)$ is stochastically equicontinuous. Note that $(\alpha''^{n}_{lm,i}, \mathcal{F}_{(i+1)\Delta_n})$ forms an array of martingale differences. Hence, $\sum_{u=1}^{k_n} \alpha''^{n}_{lm,i+u}$ and $\sum_{u=1}^{k_n} \alpha''^{n}_{lm,j+u}$ are uncorrelated whenever $|i-j| \geq k_n$. Moreover, $\mathbb{E}\,\|\alpha''^{n}_{i}\|^2 \leq \mathbb{E}\,\|\alpha^{n}_{i}\|^2 \leq K\Delta_n^2$, where the second inequality is by (4.10) in Jacod and Rosenbaum (2013). Therefore,

$$\mathbb{E}\left\|\sum_{u=1}^{k_n} \alpha''^{n}_{lm,i+u}\right\|^2 \leq K k_n \Delta_n^2. \tag{6.18}$$

For any $\tau, \eta \in \mathcal{T}$,

$$\begin{aligned}
& \mathbb{E}\left\|V''_{n,5}(\tau) - V''_{n,5}(\eta)\right\|^2 \\
= \;& \Delta_n^{-1} k_n^{-2} \mathbb{E}\left[\left(\sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n}) \sum_{u=1}^{k_n} \alpha''^{n}_{lm,i+u-1}\right)(w_\tau(i\Delta_n) - w_\eta(i\Delta_n))\right)^2\right] \\
\leq \;& K\Delta_n^{-1} k_n^{-1} \mathbb{E}\left[\sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n}) \sum_{u=1}^{k_n} \alpha''^{n}_{lm,i+u-1}\right)^2 (w_\tau(i\Delta_n) - w_\eta(i\Delta_n))^2\right] \\
\leq \;& K\left|\tau - \eta\right|^2 \Delta_n^{-1} k_n^{-1} \mathbb{E}\left[\sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(\sum_{l,m=1}^{d} \partial_{lm} g(c_{i\Delta_n}) \sum_{u=1}^{k_n} \alpha''^{n}_{lm,i+u-1}\right)^2\right] \\
\leq \;& K\left|\tau - \eta\right|^2,
\end{aligned}$$

where the first inequality follows from the Cauchy–Schwarz inequality, the second inequality follows from Assumption WF, and the third inequality follows from (6.18) and the boundedness of $\|\partial g(c_t)\|$. It readily follows that $V''_{n,5}(\tau)$ is stochastically equicontinuous. This finishes the proof of (6.3). The proof of the theorem is complete. *Q.E.D.*

Proof of Theorem 3. Theorem 2, when applied to the constant weight function, implies that $\bar{x}_n \xrightarrow{\mathbb{P}} \bar{x}$ and $\bar{y}_n \xrightarrow{\mathbb{P}} \bar{y}$. The assertion then follows from Assumption ID. *Q.E.D.*

Proof of Theorem 4. Recall that $\bar{x}_n \equiv \mathbb{F}^n x$, $\bar{y}_n \equiv \mathbb{F}^n y$, $\bar{x} \equiv \mathbb{F}x$ and $\bar{y} \equiv \mathbb{F}y$. We set

$$\xi_{x,n} = \Delta_n^{-1/2}(\bar{x}_n - \bar{x}), \quad \xi_{y,n} = \Delta_n^{-1/2}(\bar{y}_n - \bar{y}), \quad \xi_{\Psi,n} = \Delta_n^{-1/2}(\Psi_n - \Psi). \tag{6.19}$$

We observe

$$\begin{aligned}
&\Delta_n^{-1/2}\bar{x}_n^{\intercal}\Psi_n\left(\bar{y}_n-\bar{x}_n\theta_0^*\right)\\
&\quad=\Delta_n^{-1/2}(\bar{x}+\Delta_n^{1/2}\xi_{x,n})^{\intercal}\left(\Psi+\Delta_n^{1/2}\xi_{\Psi,n}\right)\left(\bar{y}-\bar{x}\theta_0^*+\Delta_n^{1/2}\left(\xi_{y,n}-\xi_{x,n}\theta_0^*\right)\right)\\
&\quad=\Delta_n^{-1/2}\bar{x}^{\intercal}\Psi\left(\bar{y}-\bar{x}\theta_0^*\right)+\xi_{x,n}^{\intercal}\Psi(\bar{y}-\bar{x}\theta_0^*)+\bar{x}^{\intercal}\xi_{\Psi,n}\left(\bar{y}-\bar{x}\theta_0^*\right)+\bar{x}^{\intercal}\Psi\left(\xi_{y,n}-\xi_{x,n}\theta_0^*\right)+o_p(1)\\
&\quad=\xi_{x,n}^{\intercal}\Psi(\bar{y}-\bar{x}\theta_0^*)+\bar{x}^{\intercal}\xi_{\Psi,n}\left(\bar{y}-\bar{x}\theta_0^*\right)+\bar{x}^{\intercal}\Psi\left(\xi_{y,n}-\xi_{x,n}\theta_0^*\right)+o_p(1),
\end{aligned} \tag{6.20}$$

where the first equality is by definition, the second equality follows from the fact that $\xi_{x,n}$, $\xi_{y,n}$ and $\xi_{\Psi,n}$ are $O_p(1)$, and the third equality is obtained by using $\bar{x}^{\intercal}\Psi\left(\bar{y}-\bar{x}\theta_0^*\right)=0$.

Note that $\Delta_n^{-1/2}(\theta_n-\theta_0^*)=(\bar{x}_n^{\intercal}\Psi_n\bar{x}_n)^{-1}\Delta_n^{-1/2}\bar{x}_n^{\intercal}\Psi_n\left(\bar{y}_n-\bar{x}_n\theta_0^*\right)$. We derive the assertion of the theorem by combining (6.20) with Assumption SC. *Q.E.D.*

Proof of Proposition 1. (a) Recall (6.19). Let $g=(y^{\intercal},vec(x)^{\intercal})^{\intercal}$. Since g is three times continuously differentiable, by Theorem 2, Assumption SC is verified with $(\xi_{y,n},vec(\xi_{x,n}))\xrightarrow{\mathcal{L}\text{-}s}\mathcal{MN}\left(0,\mathbb{F}\mathbb{V}\left(g,g\right)\right)$ and $\xi_{\Psi,n}=\xi_{\Psi}=0$. Note that $\xi_x\theta_0^*=(\theta_0^{*\intercal}\otimes I_q)vec\left(\xi_x\right)$. The assertion of part (a) is then implied by Theorem 4.

(b) By Theorem 3, $\theta_n\xrightarrow{\mathbb{P}}\theta_0^*$. Since x is continuous, by Theorem 1, we have $\widehat{\mathbb{F}}^n x\xrightarrow{\mathbb{P}}\bar{x}$. Furthermore, with $g=(y^{\intercal},vec(x)^{\intercal})^{\intercal}$, $\mathbb{V}\left(g,g\right)$ is also continuous. Hence, we can apply Theorem 1 to the function $\mathbb{V}\left(g,g\right)$ and derive that $\widehat{\mathbb{F}}^n\mathbb{V}\left(g,g\right)\xrightarrow{\mathbb{P}}\mathbb{F}\mathbb{V}\left(g,g\right)$. From here we derive $A_n\xrightarrow{\mathbb{P}}A^*$. The assertion of part (b) readily follows. *Q.E.D.*

Proof of Proposition 2. It is obvious that $\Omega_{0,T}\subseteq\{\mathbb{F}_\tau\left(y^\star-x^\star\theta_0^*\right)=0\text{ for all }\tau\in\mathcal{T}\}$. It remains to show the inclusion in the other direction. Fix a path $\omega\in\Omega$ on which $\mathbb{F}_\tau\left(y^\star-x^\star\theta_0^*\right)=0$ for all $\tau\in\mathcal{T}$. On this path, we consider a càdlàg function $f\left(s\right)=y^\star(c_s)-x^\star(c_s)\theta_0^*$ and observe that $\int_0^T f\left(s\right)w_\tau(s)ds=0$ for all $\tau\in\mathcal{T}$. Since the weight functions form a complete family, we have $f(s)=0$ for Lebesgue almost every $s\in[0,T]$. Since $z_m^\star\left(\cdot\right)$ is strictly positive, this further implies that $\tilde{y}_m(c_s)-\tilde{x}_m(c_s)^{\intercal}\theta_{m,0}^*=0$ for almost every $s\in[0,T]$ and all $1\leq m\leq\bar{m}$. Hence, $\omega\in\Omega_{0,T}$, which completes the proof. *Q.E.D.*

Proof of Proposition 3. Fix an arbitrary càdlàg function $f:[0,T]\mapsto\mathbb{R}$. The proof requires showing the claim that, if $\int_0^T f\left(s\right)w_\tau(s)ds=0$ for all $\tau\in\mathcal{T}$, then $f\left(s\right)=0$ for almost every $s\in[0,T]$. Below, we suppose the condition in this claim holds. Let S be a random variable uniformly distributed on $[0,T]$ and $U=f(S)$. It remains to show that $\mathbb{E}[U|S]=0$ almost surely.

Suppose the claim were not true, that is, $\mathbb{P}(\mathbb{E}\left[U|S\right]=0)<1$. Note that, since f is càdlàg, U is bounded. Then, by Theorem 1 of Bierens and Ploberger (1997), the set $\mathcal{T}_0\equiv\{\tau\in\mathbb{R}:$

$\mathbb{E}\left[Uw(\tau S)\right]=0\}$ would have zero Lebesgue measure. Since the Lebesgue measure of $\mathcal{T}$ is positive, we then would have $0\neq\mathbb{E}\left[Uw(\tau S)\right]=\int_0^T f\left(s\right)w\left(\tau s\right)ds$ for some $\tau\in\mathcal{T}$, a contradiction. The proof is now complete. *Q.E.D.*

Proof of Theorem 5. Part (a). By Theorem 1, $\mathbb{F}_\tau^n y^\star=\mathbb{F}_\tau y^\star+o_{pu}(1)$, $\mathbb{F}_\tau^n x^\star=\mathbb{F}_\tau x^\star+o_{pu}(1)$ and $\theta_n=\theta_0^*+o_{pu}\left(1\right)$. The assertion in part (a) readily follows from the definition of $\zeta_n\left(\tau\right)$.

Part (b). Recall that $h^\star=(y^{\star\intercal},vec\left(x^\star\right)^\intercal)^\intercal$ and $h=(y^\intercal,vec\left(x\right)^\intercal)^\intercal$. By a slight extension of Theorem 2, we see that the sequence of processes

$$\xi_n\left(\tau\right)\equiv\left(\Delta_n^{-1/2}\left(\mathbb{F}_\tau^n h^\star-\mathbb{F}_\tau h^\star\right)^\intercal,\Delta_n^{-1/2}\left(\mathbb{F}^n h-\mathbb{F}h\right)^\intercal\right)^\intercal \tag{6.21}$$

converges stably in law to a process $\xi\left(\tau\right)$ which, conditional on $\mathcal{F}$, is centered Gaussian with covariance function

$$Q(\tau,\eta)\equiv\left(\begin{array}{cc}\mathbb{F}_{\tau,\eta}\mathbb{V}\left(h^\star,h^\star\right) & \mathbb{F}_\tau\mathbb{V}\left(h^\star,h\right)\\ \mathbb{F}_\eta\mathbb{V}\left(h,h^\star\right) & \mathbb{F}\mathbb{V}\left(h,h\right)\end{array}\right).$$

Note that $\xi_n(\tau)$ consists of the following components:

$$\left\{\begin{array}{l}\xi_{y,n}^\star(\tau)=\Delta_n^{-1/2}(\mathbb{F}_\tau^n y^\star-\mathbb{F}_\tau y^\star),\quad \xi_{x,n}^\star(\tau)=\Delta_n^{-1/2}\left(\mathbb{F}_\tau^n x^\star-\mathbb{F}_\tau x^\star\right),\\ \xi_{y,n}=\Delta_n^{-1/2}(\mathbb{F}^n y-\mathbb{F}y),\quad \xi_{x,n}=\Delta_n^{-1/2}\left(\mathbb{F}^n x-\mathbb{F}x\right).\end{array}\right.$$

In restriction to $\Omega_{0,T}$,

$$\begin{aligned}\zeta_n\left(\tau\right) &= \Delta_n^{-1/2}(\mathbb{F}_\tau^n y^\star-\left(\mathbb{F}_\tau^n x^\star\right)\theta_n)\\ &= \Delta_n^{-1/2}(\mathbb{F}_\tau^n y^\star-\left(\mathbb{F}_\tau^n x^\star\right)\theta_0)-\Delta_n^{-1/2}\left(\mathbb{F}_\tau^n x^\star\right)\left(\theta_n-\theta_0\right)\\ &= \xi_{y,n}^\star(\tau)-\xi_{x,n}^\star\left(\tau\right)\theta_0-\left(\mathbb{F}_\tau x^\star\right)\Xi\left(\xi_{y,n}-\xi_{x,n}\theta_0\right)+o_{pu}\left(1\right)\\ &= \xi_{y,n}^\star(\tau)-\left(\theta_0^\intercal\otimes I_{\bar{m}}\right)vec\left(\xi_{x,n}^\star\left(\tau\right)\right)-\left(\mathbb{F}_\tau x^\star\right)\Xi\left(\xi_{y,n}-\left(\theta_0^\intercal\otimes I_q\right)vec\left(\xi_{x,n}\right)\right)+o_{pu}\left(1\right).\end{aligned}$$

Recall that $\kappa\left(\tau;\theta\right)=\left(I_{\bar{m}},-\left(\theta^\intercal\otimes I_{\bar{m}}\right),-\left(\mathbb{F}_\tau x^\star\right)\Xi,\left(\mathbb{F}_\tau x^\star\right)\Xi\left(\theta^\intercal\otimes I_q\right)\right)$. Hence, $\zeta_n\left(\tau\right)=\kappa(\tau;\theta_0)\xi_n(\tau)+o_{pu}(1)$. Therefore, $\zeta_n(\cdot)$ converges stably in law to a mixture centered Gaussian process with $\mathcal{F}$-conditional covariance function $\kappa\left(\tau;\theta_0\right)Q\left(\tau,\eta\right)\kappa\left(\eta;\theta_0\right)^\intercal$ for $\tau,\eta\in\mathcal{T}$.

Part (c). By Theorem 1, we have $Q_n(\tau,\eta)\xrightarrow{\mathbb{P}}Q(\tau,\eta)$ uniformly in $\tau,\eta\in\mathcal{T}$. Applying Theorem 1 again, we derive $\widehat{\mathbb{F}}_\tau^n x^\star=\mathbb{F}_\tau x^\star+o_{pu}(1)$ and $\Xi_n\xrightarrow{\mathbb{P}}\Xi$. Since $\theta_n\xrightarrow{\mathbb{P}}\theta_0^*$, we have $\kappa_n(\tau;\theta_n)\xrightarrow{\mathbb{P}}\kappa\left(\tau;\theta_0^*\right)$ uniformly in $\tau\in\mathcal{T}$. From here, we readily derive the assertions of part (c). *Q.E.D.*

References

Aït-Sahalia, Y. and J. Jacod (2012). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic Literature 50*, 1007–1050.

Aït-Sahalia, Y. and D. Xiu (2015). Principal component analysis of high frequency data. Technical report, University of Chicago.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), pp. 579–625.

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

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics 91*(1), 1 – 23.

Barndorff-Nielsen, O., S. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2005). A Central Limit Theorem for Realised Power and Bipower Variations of Continuous Semimartingales. In Y. Kabanov and R. Lipster (Eds.), *From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev*. Springer.

Barndorff-Nielsen, O. and N. Shephard (2002). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society Series B, 64*, 253–280.

Barndorff-Nielsen, O. and N. Shephard (2004a). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics 2*, 1–37.

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

Bierens, H. (1982). Consistent model specification tests. *Journal of Econometrics 20*, 105–134.

Bierens, H. J. and W. Ploberger (1997). Asymptotic theory of integrated conditional moment tests. *Econometrica 65*(5), pp. 1129–1151.

Bollerslev, T. (1990). Modelling the Coherence in Short Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *Review of Economics and Statistics 72*, 498–505.

Bollerslev, T., N. Sizova, and G. Tauchen (2012). Volatility in equilibrium: Asymmetries and dynamic dependencies. *Review of Finance 16*(1), 31–80.

Bollerslev, T. and H. Zhou (2002). Estimating Stochastic Volatility Diffusion using Conditional Moments of Integrated Volatility. *Journal of Econometrics 109*, 33–65.

Comte, F. and E. Renault (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance 8*, 291–323.

Corradi, V. and W. Distaso (2006). Semiparametric Comparision of Stochastic Volatility Models Using Realized Measures. *Review of Economic Studies 73*, 635–667.

Drechsler, I. and A. Yaron (2011). What's vol got to do with it. *Review of Financial Studies 24*(1), 1–45.

Engle, R. and B. Kelly (2012). Dynamic Equicorrelation. *Journal of Business and Economic Statistics 30*, 212–228.

Fan, J., A. Furger, and D. Xiu (2014). Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high frequency data. Technical report, University of Chicago.

Foster, D. and D. B. Nelson (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica 64*, 139–174.

Geman, D. and J. Horowitz (1980). Occupation Densities. *The Annals of Probability 8*, 1–67.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*, 1029–1054.

Ibragimov, I. and R. Has'minskii (1981). *Statistical Estimation: Asymptotic Theory.* Berlin: Springer.

Jacod, J. (2008). Asymptotic Properties of Power Variations and Associated Functionals of Semimartingales. *Stochastic Processes and their Applications 118*, 517–559.

Jacod, J. and P. Protter (2012). *Discretization of Processes.* Springer.

Jacod, J. and M. Rosenbaum (2013). Quarticity and Other Functionals of Volatility: Efficient Estimation. *Annals of Statistics 118*, 1462–1484.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (Second ed.). New York: Springer-Verlag.

Kalnina, I. and D. Xiu (2014). Nonparametric estimation of the leverage effect using information from derivatives markets. Technical report, University of Chicago.

King, M., E. Sentana, and S. Wadhwani (1994). Volatility and Links between National Stock Markets. *Econometrica 62*, 901–933.

Lehmann, E. L. and J. P. Romano (2005). *Testing Statistical Hypothesis.* Springer.

Li, J., V. Todorov, and G. Tauchen (2013). Volatility occupation times. *Annals of Statistics 41*, 1865–1891.

Li, J., V. Todorov, and G. Tauchen (2014). Adaptive estimation of continuous-time regression models using high-frequency data. Technical report, Duke University and Northwestern University.

Li, J. and D. Xiu (2015). Generalized methods of integrated moments for high-frequency data. Technical report, Duke University and University of Chicago.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari LXIV*, 19–47.

Mykland, P. and L. Zhang (2009). Inference for Continuous Semimartingales Observed at High Frequency. *Econometrica 77*, 1403–1445.

Oh, D. W. and A. J. Patton (2013). Time-varying systematic risk: Evidence from a dynamic copula model of cds spreads. Technical report, Duke University.

Singleton, K. (2006). *Empirical Dynamic Asset Pricing.* Princeton University Press.

Stock, J. H. and M. W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), pp. 147–162.

Todorov, V. (2009). Estimation of Coninuous-time Stochastic Volatility Models with Jumps using High-Frequency Data. *Journal of Econometrics 148*, 131–148.

Todorov, V. and G. Tauchen (2012). The realized Laplace transform of volatility. *Econometrica 80*, 1105–1127.