NBER WORKING PAPER SERIES

INFERENCE ON RISK PREMIA IN CONTINUOUS-TIME ASSET PRICING MODELS

Yacine Aït-Sahalia
Jean Jacod
Dacheng Xiu

Working Paper 28140
http://www.nber.org/papers/w28140

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2020

Inference on Risk Premia in Continuous-Time Asset Pricing Models
Yacine Aït-Sahalia, Jean Jacod, and Dacheng Xiu
NBER Working Paper No. 28140
November 2020
JEL No. C51,C52,C58,G12

ABSTRACT

We develop and implement asymptotic theory to conduct inference on continuous-time asset pricing models using individual equity returns sampled at high frequencies over an increasing time horizon. We study the identification and estimation of risk premia for the continuous and jump components of risks. Our results generalize the Fama-MacBeth two-pass regression approach from the classical discrete-time factor setting to a continuous-time factor model with general dynamics for the factors, idiosyncratic components and factor loadings, while accounting for the fact that the inputs of the second-pass regression are themselves estimated in the first pass.

Yacine Aït-Sahalia
Department of Economics
Bendheim Center for Finance
Princeton University
Princeton, NJ 08540
and NBER
yacine@princeton.edu

Jean Jacod
Institut de Mathématiques de Jussieu
CNRS UMR 7586
Université P. et M. Curie (Paris-6)
jean.jacod@gmail.com

Dacheng Xiu
Booth School of Business
University of Chicago
5807 South Woodlaswn Avenue
Chicago, IL 60637
dachxiu@chicagobooth.edu

1 Introduction

Factor models have been extensively employed to represent the cross-section of equity returns since the beginning of the empirical asset pricing literature. The two-pass regression approach of Fama and MacBeth (1973) is the standard inference method for such models. The first pass estimates individual factor loadings by regressing the time series of their returns onto the factors. In the second pass, the cross-section of average returns is regressed on the previously estimated loadings in order to estimate each factor's risk premium. While many refinements have been implemented over the years, the basic structure of the inference procedure for factor models remains largely unchanged.

Nonetheless, a simple discrete-time factor model is inadequate in two important ways. First, economic factors have complex dynamics, such as stochastic volatility and jumps, and moreover individual equity returns respond to these factors with time-varying risk exposures (or "betas"). Second, it is natural to expect that risk exposures to these dissimilar risk components are rewarded differently: for instance, investors can be expected to demand different premia for bearing the tail risks of systemic factors, see, e.g., momentum crashes in Daniel and Moskowitz (2016). Standard discrete-time factor models do not capture this finer structure of factor dynamics, risk exposures, and consequently of risk premia.[1]

Continuous-time models with high frequency observations are well understood by now to be useful in addressing the first issue, namely estimating factor loadings in richer models.[2] However, the second issue, estimating risk premia, requires an expansive time

[1]The literature has long been aware of the fact that individual equity returns feature time-varying risk exposures and rewards, dating back to as early as Rosenberg (1974), which prompted extensions of the baseline unconditional factor model to a conditional version. Gagliardini, Ossola, and Scaillet (2016) uses one characteristic and two common time-series variables to model these risk exposures and premia. Kelly, Pruitt, and Su (2019) investigate a list of 36 characteristics and provide evidence that these asset characteristics proxy for time-varying exposures to unobservable risk factors. Gu, Kelly, and Xiu (2019) model these risk exposures as nonlinear neural network functions of almost 100 characteristics. All these papers tackle the curse of dimensionality by imposing additional parametric assumptions. Raponi, Robotti, and Zaffaroni (2019) estimate risk premia in a linear factor model on a sequence of moving windows. Their asymptotic analysis allows for a small (and fixed) time window, but requires an increasing cross-section.

[2]The first pass regression is a continuous-time regression model, which can be estimated using a realized beta estimator, as the ratio of realized covariance to realized variance (see Barndorff-Nielsen and Shephard (2004) and Andersen, Bollerslev, Diebold, and Wu (2005)). These papers do not allow for jumps, and the implicit regression model has constant betas over the time interval considered. Todorov and Bollerslev (2010) also investigate a univariate model, but allow the continuous and jump betas to differ. Li, Todorov, and Tauchen (2017) study a regression model but focus only on the jump components. Aït-Sahalia and Xiu (2017) estimate a latent factor model for a large cross-section of equity returns (see also, Pelger (2019)), and use it to construct a large covariance matrix. In contrast to all these models

span, and consequently different tools. Although the two forms of asymptotics have been employed jointly in other contexts, such as estimating diffusion models (see Bandi and Phillips (2003)), they have not been combined to analyze two-pass regressions. Yet, such a combination is essential if we are to estimate risk premia in a model where factors have dynamics that possibly include stochastic volatility and jumps and where factor sensitivities are themselves stochastic.

So we develop in this paper a two-pass inference procedure for continuous-time factor models in a general setting, relying on both an increasing sampling frequency and an increasing time span. This development requires new assumptions and asymptotic results that have not yet been employed in the literature. The results we provide generalize the Fama-MacBeth two-pass regression approach to a continuous-time factor model with general dynamics for the factors, idiosyncratic components and factor loadings, including a proper accounting of the fact that the inputs of the second-pass regression are themselves estimated in the first pass.

The importance of the latter point was originally made in the classical discrete-time setting by Shanken (1992), who provided the first rigorous analysis of the asymptotic behavior of the two-pass regression for unconditional discrete-time factor models, when taking into account the first pass estimation error in the betas of test assets. In a continuous-time setting, Bollerslev, Li, and Todorov (2016) compute risk premia with respect to both the continuous and the jump components of market risk as a single factor, using cross-sectional regressions of high frequency beta estimates, but does so as if the betas were perfectly observed. By contrast, we provide inference for the risk premia in the second-pass regressions, while also allowing for multiple factors and stochastic betas in the first stage, and treat the betas in the second pass as components that were estimated in the first pass. Ang and Kristensen (2012) and Chang, Choi, Kim, and Park (2016) develop tests of alphas using nonparametric and parametric time-series regressions, respectively. Their tests cannot be used to distinguish different components of risk premia, however, and are only applicable to tradable factors. In addition, their model specification does not allow for jumps, and requires restrictive assumptions such as the existence of a specific time change that homogenizes the diffusion processes. Our approach is therefore the first

that assume a constant beta, Mykland and Zhang (2006) show how to perform ANOVA for a univariate regression model with a time-varying coefficient, and Aït-Sahalia, Kalnina, and Xiu (2020) estimate a multivariate regression model with time-varying continuous and jump betas, see also Li, Todorov, and Tauchen (2016). Reiß, Todorov, and Tauchen (2015) propose a nonparametric test for the null hypothesis of constant beta in a bivariate setting. All these papers rely on high frequency asymptotics only, in which an increasing number of observations is sampled within a fixed sample period.

complete counterpart to the Fama-MacBeth two-pass approach, in a general continuous-time model, combining high frequency and long span asymptotics, and one that is robust to the Shanken (1992) critique.

When we implement our approach on a large universe of intraday individual stock returns, we find that a statistically and economically significant part of the market equity premia are earned because of exposures to the market's jump risk component, that various jump risks in Fama-French and momentum factors supersede their continuous counterparts as the primary pricing factors, and that augmenting Fama-French factor models with jump risks increases the explanatory portion of the cross-section of expected returns in the second pass.

The paper is organized as follows. Section 2 formulates the general continuous-time factor model we will seek to estimate. Section 3 constructs the estimators and derives their properties. Section 4 examines their small sample properties. We estimate empirically the model on a large cross-section of U.S. equities in Section 5. Section 6 concludes. Proofs are in the Appendix.

2 A Continuous-Time Factor Model

We have an economy with a risk free asset and M risky assets, which are each driven by K common factors and an idiosyncratic component. All factors and prices are stochastic processes defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. We let $\mathcal{M}_{p,q}$ denote the set of all $p \times q$ matrices and $\mathcal{M}_p^+$ the class of all symmetric nonnegative elements of $\mathcal{M}_{p,p}$. For any matrix $A \in \mathcal{M}_{p,q}$ we denote as $A^\intercal$ its transpose and $\|A\|$ its operator norm. When $A \in \mathcal{M}_p^+$, we denote as $\zeta(A)$ its smallest eigenvalue, so A is invertible if and only $\zeta(A) > 0$, in which case $\|A^{-1}\| = 1/\zeta(A)$. The $p \times p$ identity matrix is denoted by I_p. We write $x_n \asymp y_n$ for two sequences of positive numbers if both sequences x_n/y_n and y_n/x_n are bounded.

2.1 Factors

The factors are driven by a K–dimensional Brownian motion W^F and a Poisson random measure $\underline{p}^F$ on $\mathbb{R}_+ \times E$ (for some Polish space E, which can be for example $E = \mathbb{R}^K$) with intensity (or, compensating) measure $\underline{q}(dt, dz) = dt \otimes \nu(dz)$ with ν a σ-finite measure on E. The model is set up as usual in terms of log-prices and log-factors. The dynamics of the vector of log-factors $F = (F^1, \dots, F^K)$ are as follows:[3]

$$F_t = F_0 + \int_0^t \mu_s^F \, ds + F_t^C + F_t^J, \tag{1}$$

$$F_t^C = \int_0^t \sigma_s^F \, dW_s^F, \quad F_t^J = \int_0^t \int_E \delta^F(s,z) \, \underline{p}^F(ds, dz). \tag{2}$$

The M log-prices are driven by the factors above, plus (possibly) some idiosyncratic part, according to the model described in the next section. However, it is useful to specify right away that we are interested in the risk premia, which are a kind of assessment of the global exposure of all prices to the various "risks" associated with the factors, which we now discuss. For any factor F^k the risks are of two types: the "continuous risk" associated with the continuous part $F_t^{C,k}$, and the "jump risk" associated with the jumps of F^k, with itself a priori two components: the times at which those jumps occur, and their sizes. So one might say that we have a jump risk (repeatedly occurring at some random times) for each possible jump size x.

With this interpretation, and unless the set of possible jump sizes is finite (a very special model, most likely to be empirically inadequate) the number of distinct risks is infinite, and evaluating the risk premia is of course an impossible task for two main reasons: first one has a large but finite number M of prices from which one typically cannot deduce an infinite number of risk premia; second, the data over a finite time span $[0, T]$ give no information about the risk premia corresponding to jump sizes of F^k that did not occur within $[0, T]$, although such jump sizes could occur after time T, so that the predictive value of our inference is absent in this case.

One way to overcome this problem is to introduce a finite partition $B_0^k, \dots, B_{L_k}^k$ of $\mathbb{R}$, with each B_j^k a non-empty interval and with either $B_0^k = \mathbb{R}$ (so $L_k = 0$) or $B_0^k = [-a_k, a_k]$ for a nonnegative real a_k (so $L_k \geq 1$). We define the partial jump processes

$$F_t^{J,k,l} = \sum_{s \leq t} \Delta F_s^k \, 1_{B_l^k}(\Delta F_s^k), \qquad \text{hence} \quad F_t^{J,k} = \sum_{l=0}^{L_k} F_t^{J,k,l} \tag{3}$$

with the usual notation $\Delta F_t^k = F_t^k - F_{t-}^k$ for the jump size of F^k at time t. Then, with F^k, we associate $L_k + 1$ distinct factors of risk which are $F^{J,k,l}$ for $l = 1, \dots, L_k$ and $\widetilde{F}^k$

[3] Here, F_0 is $\mathcal{F}_0$–measurable, $\mu^F = \mu_s^F(\omega)$ is optional $\mathbb{R}^K$–valued, $\sigma^F = \sigma_s^F(\omega)$ is optional $\mathcal{M}_{K,K}$–valued, $\delta^F = \delta^F(\omega, s, z)$ is predictable $\mathbb{R}^K$–valued. Those coefficients satisfy Assumption 1 below, which in particular ensures that the various integrals above are meaningful.

given by:

$$\widetilde{F}^k = F^C + F^{J,k,0}. \tag{4}$$

This approach allows for some flexibility. Considering a given factor F^k, on one end of the spectrum, if $L_k = 0$, one considers the whole of F^k as a "single" risk factor. On the other end, if $L_k \geq 1$ and $a_k = 0$ the risk factors are the continuous one $F^{C,k}$ and the purely discontinuous ones $F^{J,k,l}$. In the middle, if $L_k \geq 1$ and $a_k > 0$, the risk factor $\widetilde{F}^k$ contains the continuous part of F and its "small" jumps, to which the prices respond in the same way (i.e., the same beta), whereas the response of prices to the risk factors $F^{J,k,l}$ for $1 \leq l \leq L_k$ are (possibly) different. Put another way, we can consider that, with $H = \sum_{k=1}^K L_k$, we really have $K + H$ risk factors, namely the $\widetilde{F}^k$'s and the $F^{J,k,l}$'s, to each of which individual assets are exposed. It is therefore convenient to stack all pure jump factors in a single H-dimensional vector $\overline{F}_t$. Specifically, we stack $\overline{F}_t$ and the associated B^k_l according to the following rule (with an empty sum taken to be 0 below):

$$\begin{gathered}\overline{F}^h = F^{J,k(h),l(h)}, \qquad \overline{B}^h = B^{k(h)}_{l(h)} \\ \text{where} \quad k(h) = j, \quad l(h) = h - \sum_{i=1}^{j-1} L_i \quad \text{when} \quad \sum_{i=1}^{j-1} L_i < h \leq \sum_{i=1}^{j} L_i.\end{gathered} \tag{5}$$

We should emphasize that what is observable is the K-dimensional log-factor process F, whereas the risk factors, which are the components of $\widetilde{F}$ and $\overline{F}$ as given above, are typically not separately observable and our econometric procedure will account for this.

2.2 Asset Prices

For the log-prices, we assume a factor model: each log-price is the sum of a linear response[4] to each of the risk factors specified above, plus an idiosyncratic part. The dynamics of the log-prices vector $P = (P^1, \ldots, P^M)$ is then, in matrix notation,

$$P_t = P_0 + \int_0^t \beta^C_s \, d\widetilde{F}_s + \int_0^t \beta^J_s \, d\overline{F}_s + P^I_t, \tag{6}$$

where β^C_t and β^J_t are predictable processes, with values in $\mathcal{M}_{M,K}$ and $\mathcal{M}_{M,H}$, respectively (quite often in the literature these processes are assumed to be simply constants), and

[4] The linear relationship between the price and the factors in a continuous-time factor model is only local or instantaneous — unlike discrete-time models, P_t can be a highly nonlinear function of $\widetilde{F}_t$, $\overline{F}_t$ and P^I_t.

the idiosyncratic part P^I is

$$P_t^I = \int_0^t \mu_s^I \, ds + \int_0^t \sigma_s^I \, dW_s^I + \int_0^t \int_E \delta^I(s,z) \, \underline{p}^I(ds,dz), \tag{7}$$

with W^I an M–dimensional Brownian motion and $\underline{p}^I$ a Poisson random measure with the same intensity $\underline{q}$ as is (1) (the latter is not a restriction). To ensure that this is really an idiosyncratic part we assume that $(W^I, \underline{p}^I)$ and $(W^F, \underline{p}^F)$ are independent.[5] In the sequel, it is convenient to stack the two matrices β_t^C and β_t^J into a single $M \times (K+H)$ matrix by setting

$$\beta_t = \big(\beta_t^C, \beta_t^J\big). \tag{8}$$

Moreover, the risk free asset (whose log-price P_t^0 is not included into the vector P_t above) satisfies, with r_t an optional process:

$$P_t^0 = \int_0^t r_s \, ds. \tag{9}$$

A critical matter which is largely irrelevant in a high frequency setting with a finite time span, but becomes essential over a long time horizon, is the issue of survivorship: existing individual stocks typically have a finite lifetime for reasons such as bankruptcy, mergers, acquisitions, etc., while new stocks appear after the beginning of the sample as well. Stocks are not necessarily active over the whole interval of observation $[0, T_n]$, unlike the common factors which are present at all times. We model this as follows. We do not require that the initial observation P_0 in (6) be $\mathcal{F}_0$-measurable. Stock m enters the market at some (possibly random) finite time ζ_m (with the convention $\zeta_m = 0$ if it is present at time 0) and disappears at another random time $\theta_m > \zeta_m$, which may be before or after T_n. The time interval in which the stock is "active" is $\mathcal{L}_m = (\zeta_m, \theta_m)$. In our empirical application below, for over three quarters of the stocks the interval $\mathcal{L}_m$ is strictly included in $[0, T_n]$.

As a result, (6) for a particular component m describes the dynamics of the log-price P^m over the time interval $[\zeta_m, \theta_m]$ only. Equivalently, writing this equation component-wise, we have

$$\begin{aligned} P_t^m &= P_{\zeta_m}^m + \sum_{k=1}^K \int_0^t \beta_s^{C,m,k} \, d\widetilde{F}_s^k + \sum_{h=1}^H \int_0^t \beta_s^{J,m,h} \, d\overline{F}_s^h + P_t^{I,m}, \\ P_t^{I,m} &= \int_0^t \mu_s^{I,m} \, ds + \sum_{m'=1}^M \int_0^t \sigma_s^{I,m,m'} \, dW_s^{I,m'} + \int_0^t \int_E \delta^{I,m}(s,z) \, \underline{p}^I(ds,dz), \end{aligned} \tag{10}$$

[5] As above, μ^I is optional $\mathbb{R}$-valued, σ^I is optional $\mathcal{M}_{M,M}$–valued and δ^I is predictable $\mathbb{R}^M$–valued.

with $\beta_s^{C,m,k} = \beta_s^{J,m,h} = \mu_s^{I,m} = \sigma_s^{I,m,m'} = \delta^{I,m}(s,z) = 0$ for all $s \notin \mathcal{L}_m$, and where ζ_m, θ_m are stopping times and $P^m_{\zeta_m}$ is $\mathcal{F}_{\zeta_m}$-measurable. Thus P^m_t equals $P^m_{\zeta_m}$ for $t \leq \zeta_m$ and $P^m_{\theta_m}$ for $t \geq \theta_m$, but those are spurious prices added for mathematical convenience and not used in the estimation procedure below.

2.3 Risk Premia

Risk premia are associated with the structure of the drifts of the log-price processes P^m, or rather their excess drift from which we have subtracted r_t. Note that μ^I in (10) is not the true drift of P, which is the process μ_t such that $P_t - \int_0^t \mu_s\, ds$ is a martingale. The true drift process μ is connected with μ^I and the various coefficients in (1) and (10) as follows:

$$\begin{aligned} \mu_t^m = \mu_t^{I,m} &+ \sum_{k=1}^{K} \int_E \beta_t^{C,m,k} \delta^{F,k}(t,z) 1_{B_0^k}(\delta^{F,k}(t,z))\, \nu(dz) \\ &+ \sum_{k=1}^{K} \sum_{l=1}^{L_k} \beta_t^{J,m,l} \delta^{F,k}(t,z) 1_{B_l^k}(\delta^{F,k}(t,z))\, \nu(dz). \end{aligned} \tag{11}$$

What is commonly called risk premia is in connection with a no-arbitrage property, when the prices have no idiosyncratic part (except for a possible drift, see Ross (1976)), and in the case of $\zeta_m = 0$ and $\theta_m = \infty$ for all m for simplicity. Consider a portfolio with actualized value $Y_t = \sum_{m=1}^M \int_0^t \phi_s^m (dP_s^m - dP_s^0)$ at time t, with ϕ_t a predictable M-dimensional process. Upon using the compensated risk factors $\widetilde{F}^*$ and $\overline{F}^*$ (meaning, for example, that $\widetilde{F}^*$ is a martingale and $\widetilde{F} - \widetilde{F}^*$ is continuous with locally finite variation), we then have

$$Y_t = \sum_{m=1}^{M} \Big(\int_0^t \phi_s^m (\mu_s^m - r_s)\, ds + \sum_{k=1}^{K} \int_0^t \phi_s^m \beta_s^{C,m,k}\, d\widetilde{F}_s^{*k} + \sum_{h=1}^{H} \int_0^t \phi_s^m \beta_s^{J,m,h}\, d\overline{F}_s^{*h} \Big).$$

Then, an arbitrage is possible if we can find a process ϕ such that all martingale terms above vanish, but the drift part does not. A contrario, the no-arbitrage property implies the following (for Lebesgue-almost all t, hence for all t if we use a proper version for μ_t and r_t), in matrix notation (recall (8)) and with $\overline{r}_t$ the M-dimensional vector with all components equal to r_t:

$$\phi_t^\intercal \beta_t = 0 \quad \Longrightarrow \quad \phi_t^\intercal (\mu_t - \overline{r}_t) = 0,$$

which in turn implies that $\mu_t - \overline{r}_t$ is equal to $\beta_t \lambda_t$ for some $K+H$-dimensional vector λ_t, which is the vector of risk premia.

Now, we come back to our situation, where each P^m may have an idiosyncratic part, and also non necessarily trivial birth and death times ζ_m and θ_m. In this case, if for all m we have

$$t \in \mathcal{L}_m \quad \Longrightarrow \quad \mu_t^m - r_t = \sum_{k=1}^{K} \beta_t^{C,m,k} \lambda_t^{C,k} + \sum_{h=1}^{H} \beta_t^{J,m,h} \lambda_t^{J,h} + \lambda_t^{I,m}, \tag{12}$$

then the (real-valued) process $\lambda_t^{C,k}$ is called the risk premium process relative to the risk factor $\widetilde{F}^k$, and $\lambda_t^{J,h}$ is analogously relative to the risk factor $\overline{F}^h$, and $\lambda_t^{I,m}$ is relative to the idiosyncratic risk (with the convention $\lambda_t^{I,m} = 0$ when $t \notin \mathcal{L}_m$).[6]

It will be convenient again to use matrix notation and stack the risk premia $\lambda^{C,k}$ and $\lambda^{J,h}$ as a single $(K+H)$-dimensional process λ with components:

$$\lambda_t^l = \begin{cases} \lambda_t^{C,l} & \text{if } 1 \leq l \leq K \\ \lambda_t^{J,l-k} & \text{if } K < l \leq K+H, \end{cases} \tag{13}$$

whereas λ_t^I is the vector process with components $\lambda_t^{I,m}$, and (12) becomes

$$\mu_t - \overline{r}_t = \beta_t \lambda_t + \lambda_t^I, \qquad \text{where} \quad \overline{r}_t^m = r_t \, 1_{\mathcal{L}_m}(t). \tag{14}$$

Our main aim in this paper is to estimate the risk premia, on the basis of observations on the factors and prices at times $i\Delta_n$ for $i = 0, 1, \ldots, n$, plus the introduction times ζ_m, and drop-out times θ_m when they are smaller than the horizon $T_n = n\Delta_n$.

2.4 Identification

The discretization problem being addressed in the high-frequency setting $\Delta_n \to 0$, we suppose for a while that the processes $\widetilde{F}$ and $\overline{F}$ and all P^m and r are fully observed on the time interval $[0, T_n]$, and determine first what we can identify about the risk premia. This question presents several aspects.

First, only long term time-averages can be estimated. Let us consider the case $M = K = 1$ with the simplistic model (with W a Brownian motion and β a known positive

[6] Of course, $\lambda_t^{I,m}$ is relevant only when the idiosyncratic part of P^m does not reduces to a pure drift.

constant):

$$F_t = W_t, \qquad P_t = P_0 + \int_0^t \mu_s\, ds + \beta F_t, \qquad r_t = 0,$$

so $\lambda_t^I = 0$ because P has no idiosyncratic part. Then (14) gives $\lambda_t = \mu_t/\beta$. We fully observe P_t over $[0, T_n]$, but there is no consistent estimator for μ_t. On the other hand, $\frac{1}{T_n} P_{T_n}$ is an estimator for the average $\frac{1}{T_n}\int_0^{T_n} \beta\lambda_t\, dt$, with an $\mathcal{N}(0, 1/T_n)$ estimation error. So in this case only the time-average of λ_t over an interval $[0, T_n]$ with $T_n \to \infty$ can be consistently estimated. This is typical for the situation at hand: the risk premia being effectively linear functions of the drifts, only time-averages of them can (at best) be estimated. Therefore, with the notation

$$\Lambda_T^C = \frac{1}{T}\int_0^T \lambda_t^C\, dt, \qquad \Lambda_T^J = \frac{1}{T}\int_0^T \lambda_t^J\, dt, \qquad \Lambda_T = \frac{1}{T}\int_0^T \lambda_t\, dt, \tag{15}$$

we focus our attention to the behavior of those averages for $T = T_n \to \infty$, whereas pointwise estimation of λ_t or its components for any fixed t is out of the question.

Next, let us return to the definition of the risk premia at time t, as given by (14). A natural question is then whether the knowledge of $\mu_t, \overline{r}_t, \beta_t$ yields the risk premia λ_t and λ_t^I, and the answer is no: indeed, if we choose arbitrarily the vector λ_t, this equation trivially gives the vector λ_t^I (no surprise here, since (14) is a system of M linear equations with $M + K + H$ unknowns). However, at least when the eigenvalues of the idiosyncratic diffusion matrix σ_t^I are bounded (independently of M) and M is large, diversification arguments suggest that λ_t^I vanishes, or at least that its average Λ_T^I is negligible as $T \to \infty$. In the first case, (14) is the no-arbitrage condition discussed at the beginning of the section, while in the second case it still is a form of asymptotic no-arbitrage condition, which will be more precisely stated as Assumption 5 in the next section.

Finally, going back again to (14), and since λ_t^I will be negligible, in order to evaluate Λ_T we need to estimate β_t as a preliminary step, the first pass regression. Recalling (10) and (8), it turns out that β_t^C can be estimated at any time t by using the truncated realized covariation process between each P^m and F^k.

For β_t^J, it is another matter. Indeed, considering for simplicity that $H = 1$, we have $\Delta P_t^m = \beta_t^{J,m} \Delta \overline{F}_t$, which (since the jumps of the observed processes are also fairly accurately estimated at high frequency) gives us a good estimate of $\beta_t^{J,m}$ at any jump time t of $\overline{F}$. On the other hand, we have simply no way of estimating $\beta_t^{J,m}$ for the other values of t at which points no jumps occur. Assuming that $t \mapsto \beta_t^{J,m}$ is smooth does not help because the jump times are isolated points. And, unfortunately, to compute Λ_T^J we need to know (or, have good estimators for) λ_t^J for *all* $t \in [0,T]$.

The only way to resolve this problem seems to impose an identification restriction on the jump beta's: namely, that each $\beta_t^{J,m,h}$ does not depend on t on the set $\mathcal{L}_m$ (recall that by convention this quantity is 0 when $t \notin \mathcal{L}_m$). So we have

$$\beta_t^{J,m,h} = \overline{\beta}^{J,m,h}\, 1_{\mathcal{L}_m}(t), \tag{16}$$

where $\overline{\beta}^{J,m,h}$ is a constant.[7]

2.5 Assumptions

We now state and discuss the various assumptions that will be imposed. Toward this aim, we first introduce the following property for a (possibly multidimensional) optional process Y_t and a random interval $\mathcal{L} = (R, R'] \cap (0,\infty)$, with $R < R'$ two stopping times: there is a constant C such that for all $s \geq 0$ and all finite stopping times S with $R < S \leq R'$ we have

$$\big\| \mathbb{E}(Y_{(S+s)\wedge R'} - Y_S \mid \mathcal{F}_S)\big\| \leq Cs, \qquad \mathbb{E}\left(\|Y_{(S+s)\wedge R'} - Y_S\|^2 \mid \mathcal{F}_S\right) \leq Cs. \tag{17}$$

This holds for example for $\mathcal{L} = (0,\infty)$ (so $R = 0$ and $R' = \infty$) and Y an Itô semimartingale with bounded spot characteristics and bounded jumps. Note that, apart from the constant C itself, (17) does not depend on the chosen norm, and it holds for a multidimensional process as soon as it holds for each of its components.

We first make assumptions on the coefficients of the equations defining the factors, the idiosyncratic components, and the risk free asset:

Assumption 1. *(i) The processes μ_t^F and σ_t^F are optional and bounded, and $c_t^F = \sigma_t^F(\sigma_t^F)^\intercal$ is invertible, with a bounded inverse (equivalently, $\zeta(c_t^F)$ is bounded away from 0).*

(ii) The function δ^F on $\Omega\times\mathbb{R}_+\times E$ is predictable and there are a Borel bounded function Υ on E and a number $\alpha \in [0,1)$ such that $\|\delta^F(\omega,t,z)\| \leq \Upsilon(z)$ and $\int_E \Upsilon(z)^\alpha\, \nu(dz) < \infty$.

Assumption 2. *(i) For some $\varepsilon > 0$ and each m the times ζ_m and $(\theta_m - \varepsilon) \vee \zeta_m$ are stopping times.*

(ii) The processes μ_t^I and σ_t^I are optional and bounded and vanishing for $t \notin \mathcal{L}_m$, the function δ^I is predictable vanishing for $t \notin \mathcal{L}_m$ and $\|\delta^I(\omega,t,z)\| \leq \Upsilon(z)$, with Υ as in Assumption 1.

(iii) The process r_t is optional, bounded, and satisfies (17) *on $\mathbb{R}_+$.*

[7] It could be, more generally, measurable with respect to $\mathcal{F}_{\zeta_m}$, but at this stage assuming that it is a constant is not a substantial additional restriction.

These assumptions are standard, except for the uniform boundedness of the coefficients, instead of the more usual local boundedness. The boundedness is required here since the usual localization procedure does not apply when the time horizon goes to infinity, in the absence of an ergodicity condition. Note also that, instead of the usual ν-integrability of the function $\Upsilon \wedge 1$ we impose that Υ^α is ν-integrable and bounded: this implies that the jump of F are bounded and, since $\alpha < 1$, they are (locally) summable. (i) of Assumption 2 means that the drop-out time θ_m can be exactly predicted ε ahead of time, which does not seem too strong a restriction in practice.

Next, we need hypotheses on the factor loadings, and also on the splitting of each $F^{J,k}$ according to (3), (4) and (5). For any two reals $\chi, \chi' > 0$ and with $\partial \overline{B}^h$ denoting the boundary of $\overline{B}^h$ we set

$$\overline{B}^h(\chi, \chi') = \big\{ x \in \overline{B}^h : \ \ \chi \leq |x| \leq \chi', \ \ d(x, \partial \overline{B}^h) \leq \chi' \big\}. \tag{18}$$

For each $m = 1, \ldots, M$ and $t > 0$ and $\rho' > 4\rho > 0$, we put

$$\begin{gathered} A(\rho,\rho')_t^m = (\zeta_m, \theta_m) \bigcap \Big(\bigcup_{1 \leq h \leq H} \{ s \in (0,t] : \ \Delta \overline{F}_s^h \in \overline{B}^h(2\rho, \rho'/2) \} \Big) \\ N(\rho,\rho')_t^m = \#(A(\rho,\rho')_t^m), \\ R(m,\rho,\rho',t) \ \text{ is the } N(\rho,\rho')_t^m \times H\text{-matrix with entries} \\ R(m,\rho,\rho',t)^{s,h} = \Delta \overline{F}_s^h \ \text{ for } s \in A(\rho,\rho')_t^m, \ 1 \leq h \leq H. \end{gathered} \tag{19}$$

With this notation, in addition to Assumptions 1 and 2 we assume the following:

Assumption 3. *In the case $H \geq 1$ we have $\mathbb{P}\big(\zeta(R(m,\rho,\rho',t)^\intercal R(m,\rho,\rho',t)) \geq \varepsilon\big) \to 1$ as $t \to \infty$, for each m and some $\varepsilon, \rho, \rho' > 0$ (implying $\mathbb{P}(N(\rho,\rho')_t^m \geq H) \to 1$ as $t \to \infty$).*

The implication of this assumption is basically that the linear space spanned by all jump vectors $\Delta \overline{F}_s$ for $s \in [0,t] \cap \mathcal{L}_m$ is the whole of $\mathbb{R}^H$ with a probability going to 1 as $t \to \infty$: this is clearly necessary if we want the $\beta^{J,m,h}$'s to be uniquely determined by the jumps of P^m and $\overline{F}$. Note that this assumption becomes weaker if ρ decreases and/or ρ' increases.

The assumptions on the factor loadings are:

Assumption 4. *(i) The process β_t^C is optional and bounded, and for each m,k the component $\beta_t^{C,m,k}$ vanishes outside $\mathcal{L}_m$ and satisfies* (17) *on the interval $\mathcal{L}_m$.*

(ii) The process β_t^J is given by (16)*, with each $\overline{\beta}^{J,m,h}$ being a constant.*

(iii) With the notation (8)*, the $\mathcal{M}_{K+H}^+$-valued process $b_t = \beta_t^\intercal \beta_t$ is invertible, with a bounded inverse (this implies, in particular, that $M \geq K + H$).*

Finally, the no-arbitrage condition is:

Assumption 5. *We have* (14), *and* $\sqrt{T}\,\Lambda_T^{I,m} \xrightarrow{\mathbb{P}} 0$ *(convergence in probability) as* $T \to \infty$*, for all* $m = 1, \dots, M$*, where* $\Lambda_T^I = \frac{1}{T}\int_0^T \lambda_t^I\, dt$*.*

One may think that Assumption 5 is very mild, or even always satisfied, since any stock will at the end have a finite lifetime θ_m, whereas $\lambda_t^{I,m} = 0$ for $t > \theta_m$, so $\int_0^\infty \lambda_t^{I,m}\, dt < \infty$. The assumption requires that the integral $\int_0^{T_n} \lambda_t^I\, dt$ is negligible in front of $\sqrt{T_n}$, a property that is in principle testable. Of course, if idiosyncratic risk is unpriced ($\lambda_t^I = 0$, as often assumed in financial models in the setting of Ross (1976)) the assumption will be satisfied, but this strong restriction is not required under Assumption 5.

3 Estimators

As already mentioned, throughout we assume $\Delta_n \to 0$ and $T_n \to \infty$, and suppose that the observations are not contaminated by noise, so in empirical applications Δ_n should probably not be smaller than 1 or perhaps 5 minutes. The ith time interval and the ith return of any process X at stage n are denoted as

$$\mathcal{I}_i^n = ((i-1)\Delta_n, i\Delta_n], \qquad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}.$$

We now construct in stages estimators for the various components of Λ_{T_n}, see (15). Note that Λ_{T_n} is a "moving" target, which converges to a limit Λ_∞ under appropriate conditions such as a form of ergodicity; but even if this is the case, since the observations available say nothing about what happens after time T_n, the only way to estimate Λ_∞ when it exists is really to estimate Λ_{T_n} and then make additional assumptions about how fast this converges to Λ_∞. We avoid this extra step here, which permits substantially weaker assumptions.

We start with some heuristic considerations explaining the construction of the estimators. This is effectively the Fama-MacBeth two-pass regression procedure adapted to this (substantially more general) setting. First, as already seen, we have the decomposition

$$P_t = P_0 + \int_0^t \mu_s\, ds + P_t^{\mathrm{Mart}}, \quad \text{with } P^{\mathrm{Mart}} \text{ a martingale.} \tag{20}$$

Besides $c_t^F = \sigma_t^F(\sigma_t^F)^\intercal$ and $b_t = \beta_t^\intercal \beta_t$ we define the processes

$$\gamma_t = \beta_t^C c_t^F, \qquad \eta_t = b_t^{-1}\beta_t^\intercal. \tag{21}$$

Note that the matrix b_t is invertible by Assumption 4. We have $c_t^F \in \mathcal{M}_K^+$ and $\gamma_t \in \mathcal{M}_{M,K}$ and $b_t, b_t^{-1} \in \mathcal{M}_{K+H}^+$ and $\eta_t \in \mathcal{M}_{K+H,M}$. (14) and (20) allow one to write the following key formula:

$$\begin{aligned} &\Lambda_T = U_T - U'_T - U''_T - \tfrac{1}{T}\,\overline{U}_T, \quad \text{where} \quad U_T = \tfrac{1}{T}\int_0^T \eta_s\, dP_s, \\ &U'_T = \tfrac{1}{T}\int_0^T \eta_s \overline{r}_s\, ds, \quad U''_T = \tfrac{1}{T}\int_0^T \eta_s \lambda_s^I\, ds, \quad \overline{U}_t = \int_0^t \eta_s\, dP_s^{\mathrm{Mart}}. \end{aligned} \tag{22}$$

In the above decomposition, U_{T_n} and U'_{T_n} will be estimated and U''_{T_n} will turn out to be negligible. The variable $\overline{U}_{T_n}$ cannot be estimated, because we do not observe the process P_t^{Mart}. However, one can use a central limit theorem for the local martingale $\overline{U}_t$ as $t \to \infty$, for which its quadratic variation process V plays a central role. The components of this quadratic variation are

$$V_t^{j,j'} = \sum_{m,m'=1}^{M} \int_0^t \eta_s^{j,m} \eta_s^{j',m'}\, d[P^m, P^{m'}]_s. \tag{23}$$

Hence we need to estimate U_{T_n}, U'_{T_n} and V_{T_n}, and the procedure requires a number of steps. Before starting, we emphasize once more that the estimators may depend on F, P, r, but not on $\widetilde{F}$ or $\overline{F}$.

To construct the estimators, we will need three sequences of tuning parameters: two sequences u_n and v_n of positive reals, and a sequence q_n of positive integers, subject to

$$u_n \asymp \Delta_n^{\varpi}, \qquad q_n \asymp \Delta_n^{-\varpi'}, \qquad v_n \asymp \log(1/(T_n\Delta_n)), \tag{24}$$

and for some $\varpi, \varpi' > 0$ to be specified later. q_n dictates the length of local windows, within which we estimates β; u_n, as usual in the literature, helps separate jumps from continuous components of returns; and finally, v_n is introduced to censor exploding matrix inverses for better finite sample performance.

The procedure is split into a number of steps, and we start by an estimation of the spot quantities occurring in (21), and will indeed estimate those at the times $iq_n\Delta_n$ for $i = 0, \ldots, n/q_n - 2$ only, denoting for example $\widehat{\eta}_{n,i}$ the estimator for $\eta_{iq_n\Delta_n}$.

3.1 Estimation of Jump Loadings

Since β^J only occurs when $H \geq 1$, we assume this in this part. For any given m we first make a global estimation of the constants $(\overline{\beta}^{J,m,h})_{1\leq h\leq H}$, on the basis of the observation within $[0, T_n] \cap \mathcal{L}_m$. We choose two numbers $\rho, \rho' > 0$ (typically ρ is very small and ρ' is very large), such that Assumption 3 holds with those ρ, ρ' and some $\varepsilon > 0$. Set

$$I_n = \{1, \dots, n\}, \qquad I_n^m = \{i \in I_n : \ \zeta_m \le (i-1)\Delta_n \le \theta_m - 2\Delta_n\}, \tag{25}$$

and, recalling the notation (18), define the $\mathcal{M}_{n,H}$–valued $\widehat{R}_{n,m}$, the $\mathcal{M}_H^+$–valued $\widehat{R}'_{n,m}$, and the $\mathcal{M}_{H,n}$–valued $\widehat{R}''_{n,m}$, as follows:

$$\begin{gathered}\widehat{R}_{n,m}^{i,h} = \Delta_i^n F^{k(h)}\, 1_{\overline{B}^h(\rho,\rho')}(\Delta_i^n F^{k(h)})\, 1_{I_n^m}(i) \\ \widehat{R}'_{n,m} = (\widehat{R}_{n,m})^\intercal\, \widehat{R}_{n,m}, \qquad \widehat{R}''_{n,m} = \begin{cases} (\widehat{R}'_{n,m})^{-1}\, (\widehat{R}_{n,m})^\intercal & \text{if } \ \zeta(\widehat{R}'_{n,m}) \ge v_n \\ 0 & \text{otherwise.}\end{cases}\end{gathered} \tag{26}$$

The estimator $\widehat{\beta}_{n,i}^J$ for $\beta_{iq_n\Delta_n}^J$ of equation (16) is then defined, component-wise for stock m and risk factor h, as

$$\widehat{\beta}_{n,i}^{J,m,h} = \begin{cases} \sum\limits_{j \in I_n^m} \widehat{R}''^{h,j}_{n,m}\, \Delta_j^n P^m & \text{if } \ iq_n\Delta_n \in \mathcal{L}_m \\ 0 & \text{otherwise.}\end{cases}$$

3.2 Estimation of Continuous Loadings

The processes c_t^F and γ_t are spot volatilities or covolatilities, easily estimated by considering q_n successive returns after time t. Upon truncating for the jumps by using u_n, we estimate the components of $c_{iq_n\Delta_n}^F$ by

$$\widehat{c}_{n,i}^{F,k,k'} = \frac{1}{q_n\Delta_n} \sum_{j=1}^{q_n} \Delta_{iq_n+j}^n F^k\, \Delta_{iq_n+j}^n F^{k'}\, 1_{\{|\Delta_{iq_n+j}^n F^k| \le u_n,\, |\Delta_{iq_n+j}^n F^{k'}| \le u_n\}}. \tag{27}$$

For $\gamma_{iq_n\Delta_n}$, being aware of the birth and death times of the stocks, we construct the estimator as, component-wise,

$$\widehat{\gamma}_{n,i}^{m,k} = \begin{cases} \frac{1}{q_n\Delta_n} \sum\limits_{j=1}^{q_n} \Delta_{iq_n+j}^n P^m\, \Delta_{iq_n+j}^n F^k \\ \qquad \times 1_{\{|\Delta_{iq_n+j}^n F^k| \le u_n,\, |\Delta_{iq_n+j}^n P^m| \le u_n\}} & \text{if } \zeta_m < iq_n\Delta_n \le \theta_m - q_n\Delta_n \\ 0 & \text{otherwise.}\end{cases} \tag{28}$$

Since $\beta_t^C = (c_t^F)^{-1}\gamma_t$ one can in principle deduce from the above an estimator for β_t^C. However, although c_t^F is invertible with a bounded inverse, the estimator $\widehat{c}_{n,i}^F$, although belonging to $\mathcal{M}_K^+$ by construction, might be not invertible or might have an unbounded inverse. This is why we introduce some additional truncation procedure, based on the sequence v_n and, recalling that $\zeta(A)$ is the smallest eigenvalue of a nonnegative semi-definite matrix A, we estimate $\beta^C_{iq_n\Delta_n}$ by

$$\widehat{\beta}^C_{n,i} = \begin{cases} \widehat{\gamma}_{n,i}\,(\widehat{c}^F_{n,i})^{-1} & \text{if } \ \zeta(\widehat{c}^F_{n,i}) > 1/v_n \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

Estimation of η. In view of what precedes, plus (8), the estimator for the $\mathcal{M}_{M,K+H}$-valued $\beta_{iq_n\Delta_n}$ and the $\mathcal{M}^+_{K+H}$-valued $b_{iq_n\Delta_n}$ are naturally given by

$$\widehat{\beta}_{n,i} = \big(\widehat{\beta}^C_{n,i}, \widehat{\beta}^J_{n,i}\big), \qquad \widehat{b}_{n,i} = (\widehat{\beta}_{n,i})^\intercal \widehat{\beta}_{n,i}, \tag{30}$$

and the one for $\eta_{iq_n\Delta_n}$ by

$$\widehat{b^{-1}_{n,i}} = \begin{cases} (\widehat{b}_{n,i})^{-1} & \text{if } \ \zeta(\widehat{b}_{n,i}) > 1/v_n \\ 0 & \text{otherwise,} \end{cases} \qquad \widehat{\eta}_{n,i} = \widehat{b^{-1}_{n,i}}\,(\widehat{\beta}_{n,i})^\intercal. \tag{31}$$

3.3 Estimation of Risk Premia

We are now ready to construct the estimators of Λ_{T_n} and V_{T_n}. Recalling that, since ζ_m and θ_m are observed, the process $\overline{r}_t$ is observed at each time $t = i\Delta_n$, they are, respectively, and still in matrix notation:

$$\widehat{\Lambda}_n = \widehat{U}_n - \widehat{U}'_n, \quad \text{where} \quad \begin{cases} \widehat{U}_n = \frac{1}{T_n} \sum\limits_{i=0}^{[n/q_n]-2} \widehat{\eta}_{n,i}\,(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}) \\ \widehat{U}'_n = \frac{q_n\Delta_n}{T_n} \sum\limits_{i=0}^{[n/q_n]-2} \widehat{\eta}_{n,i}\,\overline{r}_{iq_n\Delta_n}, \end{cases} \tag{32}$$

$$\widehat{V}_n = \sum_{i=0}^{[n/q_n]-2} \widehat{\eta}_{n,i}\,(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n})\,(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n})^\intercal\,(\widehat{\eta}_{n,i})^\intercal. \tag{33}$$

In light of equation (21), the estimator (32) is reminiscent of the classical Fama-MacBeth procedure, in that we first regress excess returns $P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n} - q_n\Delta_n\overline{r}_{iq_n\Delta_n}$ onto the betas estimated over the previous time window $((i-1)q_n\Delta_n, iq_n\Delta_n]$, and then aggregate the estimated "local risk premia" to construct the final estimator.

3.4 Asymptotics for Risk Premia

We can now state the main result, which characterizes the asymptotic distribution of the risk premia estimators. We need to first specify conditions on the relative asymptotic behavior of the two sequences $\Delta_n \to 0$ and $T_n \to \infty$ (or equivalently Δ_n and n), and on the values of ϖ, ϖ' in (31). Depending on whether $H = 0$ (corresponding to the case where the factors are continuous, or may be discontinuous but the beta's are the same for the continuous and discontinuous parts), or $H \geq 1$ (in which case we need more stringent conditions), we require the following:

$$\begin{cases} T_n \Delta_n^{\tau} \to 0 \\ T_n \Delta_n^{\tau'} \to \infty \end{cases} \quad \text{for some } \tau, \tau' \text{ with } \begin{cases} 0 < \frac{\tau}{2} < \tau' < \tau < 1 & \text{if } H = 0 \\ 0 < \tau' < \tau < \frac{2}{11} & \text{if } H \geq 1. \end{cases} \tag{34}$$

The main theorem has two parts. In part (a), the assumptions of Section 2.5 only ensure that the estimators $\widehat{\Lambda}_n$ converge at rate $\sqrt{T_n}$, as should be expected since we need to estimate various drift components. An additional assumption, given in part (b) below, allows us to get a proper, and importantly feasible, central limit theorem for the risk premia:

Theorem 1. *Suppose that Assumptions 1, 2, 3, 4 and 5, plus* (34)*, hold true and choose the tuning parameters ϖ, ϖ' in* (24) *as follows:*

$$\begin{cases} \frac{\tau}{2} \bigvee (1-\tau') < \varpi' < 1-\frac{\tau}{2}, & 0 < \frac{1}{2} - \varpi < \frac{1-\alpha}{64-2\alpha} & \textit{if } H = 0 \\ (5\tau) \bigvee (1-\tau') < \varpi' < 1-\frac{\tau}{2}, & 0 < \frac{1}{2} - \varpi < \frac{1-\alpha}{64-2\alpha} \bigvee (\varpi' - 5\tau) \bigvee \frac{1-\varpi'}{6} & \textit{if } H \geq 1 \end{cases}. \tag{35}$$

a) The sequence $\sqrt{T_n}(\widehat{\Lambda}_n - \Lambda_{T_n})$ is bounded in probability.

b) If we additionally suppose that as $t \to \infty$, the variables $(1/t)V_t$ converge in probability to a limit V_∞, there is a sequence Ω_n of subsets of Ω such that $\mathbb{P}(\Omega_n) \to 1$, $\widehat{V}_n$ is invertible on Ω_n, and the variables

$$Z_n = \begin{cases} T_n \widehat{V}_n^{-1/2} \, (\widehat{\Lambda}_n - \Lambda_{T_n}) & \textit{on } \Omega_n. \\ 0 & \textit{on } \Omega \backslash \Omega_n \end{cases} \tag{36}$$

converge in law to $\mathcal{N}(0, I_{K+H})$. Further, the sequences $T_n^{1/2} \, \widehat{V}_n^{-1/2} \, 1_{\Omega_n}$ and $T_n^{-1/2} \, \widehat{V}_n^{1/2} \, 1_{\Omega_n}$ are bounded in probability.

Part (a) above tells us that the rate of convergence is at least $\sqrt{T_n}$, and (b) that $\sqrt{T_n}$ is indeed the genuine rate. Importantly, the result in part (b) is feasible, in the sense that it directly allows for constructing asymptotic variance estimators and confidence regions, for example. Note that throughout the target Λ_{T_n} is random and moving with n. And $\widehat{V}_n$ is not necessarily invertible everywhere so we need the dummy value 0 on $\Omega \backslash \Omega_n$ to define Z_n properly. However, when $\widehat{V}_n$ is invertible, the matrix $\widehat{V}_n^{-1/2}$ is well defined because $\widehat{V}_n$ belongs to $\mathcal{M}_K^+$ by construction, and this happens on the set Ω_n whose probability goes to 1, so the dummy value has no impact on the asymptotic theory. The limit V_∞ in part (b) is of course $\mathcal{M}^+_{M+K}$–valued, and the other assumptions imply that it is necessarily invertible. When it is assumed we can as well assume that the variables Λ_{T_n} converge in probability to some limiting variable Λ_∞ (this, however, does not tell us at which rate this convergence takes place, so in the theorem we cannot replace Λ_{T_n} by Λ_∞). It also implies, because the limit of $(1/T_n)V_{T_n}$ is non-degenerate, that there are at least $K+H$ stocks that have an infinite lifetime. This is due to our formulation of the problem, where we have M stocks and do not care about whether they will be alive after the time horizon T_n.[8]

4 Monte Carlo Simulations

This section investigates the finite sample performance of the estimators.

4.1 Data Generating Process

We simulate the cross-section of stock returns using a $K(=3)$-dimensional vector of log-factors F. To do so, it is convenient to directly simulate the martingale components of these factors and the idiosyncratic components, i.e., $\widetilde{F}^*$ and $\overline{F}^*$, and then construct the "genuine" drift terms of individual stocks from simulated risk premia using (14). Specifically, the volatility of the martingale component of factors F^C in (1) is driven by three CIR processes, for $i = 1, 2$, and 3:

$$d(\sigma_t^{F^i})^2 = \kappa^\sigma(\theta^\sigma - (\sigma_t^{F^i})^2)dt + \gamma^\sigma \sigma_t^{F^i} dW_t^{\sigma,i},$$

where $W^{\sigma,i}$s are independent Brownian motions. The correlation between $W^{\sigma,i}$ and W^{F^i} in (1), ρ_i, is set to be negative, which contributes to the leverage effect.[9]

With respect to the jumps, we simulate a compound Poisson process for each F^i,

[8] If we introduced an ergodic requirement for the model (which is what the additional assumption on the convergence of $(1/t)V_t$ in part (b) partially does), we could consider a model with potentially infinitely many stocks, of which M_t are alive at time t, and with the condition that $K+H \le M_t \le M$ for all t and some finite bound M.

[9] Note that our assumptions need bounded volatility processes and their inverse, so the actual volatilities we use are truncated versions of the simulated $\sigma_t^{F^i}$s above (we recycle the same notation for convenience), i.e., $\sigma_t^{F^i}$s will be truncated at C^{-1} or C, for some predetermined C, if they go beyond the range of $[C^{-1}, C]$. In practice, we use a large value for C. The same procedure is adopted for all bounded processes throughout the simulations.

Parameters	Description	Value or Distribution
T_n	sampling span	5; 10
Δ_n	sampling interval	$1/(252 \times 78)$; $1/(252 \times 26)$
M	number of stocks	1,000
K	number of factors	3
H	number of jump factors	3×2
q_n	size of local window	936; 1,872; 6,552
v_n	eigenvalue bound	$100 \times \log(1/(T_n\Delta_n))$
$\sigma^{I,m}$	idiosyncratic variance	$\mathrm{U}\,(0.05, 0.15)$
$\kappa^{\sigma,k}$	CIR parameter for $(\sigma^{F^k})^2$	3
$\theta^{\sigma,k}$	CIR parameter for $(\sigma^{F^k})^2$	0.06
$\gamma^{\sigma,k}$	CIR parameter for $(\sigma^{F^k})^2$	0.3
ρ	correlation of factor and volatility	-0.5
$\kappa^{\lambda,l}$	OU parameter for λ^l	1
g	distribution parameter for $\theta^{\lambda,l}$	0.05
$\gamma^{\lambda,l}$	OU parameter for λ^l	0.1
$\kappa^{\beta,k}$	OU parameter for $\beta^{C,m,k}$	2
$\theta^{\beta,k}$	OU parameter for $\beta^{C,m,k}$	1 for $k = 1$, and 0 otherwise
$\gamma^{\beta,k}$	OU parameter for $\beta^{C,m,k}$	0.5
ν^F	jump intensity of F	63
a	distribution parameter for δ^{F^i}	0.02
$B_{k,l}$	partitions of large jumps	$[-a, -a/2]$, $[a/2, a]$
B_0	the partition of small jumps	$[-a/2, a/2]$
$\bar{\beta}^{J,m,h}$	jump betas	$\mathrm{U}(1-b, 1+b)$ for $h = 1, 2$
		$\mathrm{U}(-b, b)$ for $h = 3, 4, 5, 6$
b	distribution parameter for $\bar{\beta}^{J,m,h}$	0.5
ν^m	jump intensity of P^I	$\mathrm{U}(2\nu_F/3, 4\nu_F/3)$
$a^{I,m}$	distribution parameter for $\delta^{I,m}$	$\mathrm{U}(2a/3, 4a/3)$

Table 1: Parameter Values in Monte Carlo Simulations

Note: This table reports the parameter values used in the data generating process of the Monte Carlo simulations.

$i = 1, 2, 3$, with a constant intensity ν^F and a uniform distribution for their jump sizes:

$$\delta^{F^i} \sim \mathrm{U}(-a, a).$$

Next, we decompose each jump process into three components according to their realized jump sizes. Jumps with respective sizes in $[-a, -a/2]$ and $[a/2, a]$ are deemed separate risk factors with their respective betas, whereas the remaining jumps with sizes realized in $(-a/2, a/2)$ will share the same beta with their corresponding continuous component. As a result, we have $L_i = 2$ for each $i = 1, 2, 3$, $K = 3$, and $H = 6$. All these risk factors are compensated to form martingales.

Next, we simulate risk premia, (13), for continuous and jump risks from Ornstein-

Uhlenbeck (OU) processes:

$$d\lambda_t^l = \kappa^{\lambda,l}(\theta^{\lambda,l} - \lambda_t^l)dt + \gamma^\lambda dW_t^{\lambda,l}, \quad l = 1, 2, \ldots, K + H, \tag{37}$$

where W^λ's are independent Brownian motions. For continuous risks, we draw their corresponding $\theta^{\lambda,l}$s from $\mathrm{U}(-g, g)$, $\mathrm{U}(-g, 0)$ for positive jump risks, and $\mathrm{U}(0, g)$ for negative jumps. There are no risk premia for the idiosyncratic components: $\lambda_t^{I,m} = 0$.

As for the betas, we simulate $\beta_t^{C,m,k}$ driven by OU processes, for $k = 1, 2, 3$, and $m = 1, 2, \ldots, M$:

$$d\beta_t^{C,m,k} = \kappa^\beta(\theta^\beta - \beta_t^{C,m,k})dt + \gamma^\beta dW_t^{\beta,m,k}, \tag{38}$$

where $W^{\beta,m,k}$s are independent Brownian motions. With respect to $\beta_t^{J,m,h}$ given by (16), we draw the constants $\overline{\beta}^{J,m,h}$ once from a uniform distribution $\mathrm{U}(c-b, c+b)$ and fix them afterwards. For the first two factors, which are market risk related, we fix $c = 1$; otherwise $c = 0$. To finalize the drift part of individual stocks, we set $\overline{r}_t$ to zero, and simulate μ_t using (14).

Finally, the martingale component of the idiosyncratic part in (4), P^I, features constant idiosyncratic volatilities $\sigma^{I,m}$s and compensated Poisson jump processes with intensities $\nu^{I,m}$s. These jumps are independent, and their sizes follow independent uniform distributions:

$$\delta^{I,m} \sim \mathrm{U}(-a^{I,m}, a^{I,m}).$$

Across $m = 1, 2, \ldots, M$, $\nu^{I,m}$ and $a^{I,m}$ are constants drawn randomly from an exponential distribution.

Throughout we choose a sample of $M = 1,000$ stocks spanning $T = 5$ or 10 years with sampling frequencies $\Delta_n = 5$ and 15 minutes. To simulate the births and deaths of stocks, we randomly allocate stock m to one of four groups: $\mathcal{L}_m = [0, T]$, $[\zeta_m, T]$, $[0, \theta_m]$, and $[\zeta_m, \theta_m]$, where the first group accounts for 70% of all stocks and the other three groups contain 10% each. θ_m and ζ_m are randomly drawn from $[0, T]$. All other parameter values are summarized in Table 1.

To construct the estimators, we need three tuning parameters: u_n, v_n and q_n, subject to (24) and (35). We set $u_n = a\Delta_n^b(\widehat{IV}_t)^{1/2}$, where we choose $a = 3$ and $b = 0.47$: see, e.g., Aït-Sahalia and Jacod (2014). $\widehat{IV}_t$ is a pilot estimate of the daily annualized variation. For q_n, to improve the finite sample performance, we choose the divisor of n that is closest to $1/(\Delta_n T_n^{0.5})$. Finally, we choose v_n to be an arbitrary large number relative to $\log(1/(T_n\Delta_n))$, say, $100 \times \log(1/(T_n\Delta_n))$ such that $1/v_n$ gives a reasonable lower bound for the minimum eigenvalue of the spot covariance estimates.

4.2 Simulation Results

We report the biases and root-mean-squared errors (RMSEs) in Table 2 for a variety of settings, including cases with a lower sampling frequency ($\Delta_n = 15$ minutes) and a shorter time span ($T_n = 5$ years). A few patterns emerge. First, estimating risk premia is a challenging problem. Even with a time span of 10 years, the estimates remain noisy relative to the magnitude of the signal. Second, the reported numbers in Columns "Bias", "Stdev", and "RMSE" do not deteriorate much as the sampling frequency drops from 5 minutes (top panel) to 15 minutes (middle panel), whereas they become worse as the time span drops from 10 years (top panel) to 5 years (bottom panel). This suggests that it is the time span that plays a critical role in determining the precision and accuracy of the risk premia estimates, rather than the sampling frequency (at least once sampling is sufficiently frequent to establish reasonably accurate first pass estimates of the loadings). Third, the choices of q_ns roughly match the size of a biweekly, monthly, and quarterly local window, respectively, among which a monthly window is most commonly employed in the empirical asset pricing literature. The results exhibit the usual bias-variance trade-off: larger q_ns lead to a bigger bias and a smaller variance.

Finally, Figure 1 reports histograms of the standardized risk premia estimates, $\widehat{\Lambda}_n$, for continuous and positive/negative jump components. These histograms match the standard normal density, suggesting that the asymptotic theory in Theorem 1 approximates the finite sample performance of the estimators well.

5 Risk Premia of U.S. Equities

We next conduct a large-scale stock-level analysis of risk premia of various factors that are known to be possible determinants of the cross-section of stock returns. By combining high frequency data and long span estimation of risk premia, our results provide a new perspective on the pricing of Brownian and jump risks.

We collect data the entire cross-section of individual stocks at the 15-minute frequency from the constituents of the S&P 500 (LargeCap), 400 (MidCap), and 600 (SmallCap) indices for a long sample spanning from January 1, 1996 to May 31, 2020. The number of tickers is about 1,500 per day, which, after filtering by share codes and exchange codes, is reduced to a minimum of 1,300 stocks per day. Our sample covers several business cycles,

$T_n = 10$ yrs $\Delta_n = 5$ mins

	Truth	$q_n = 936$ (12 days)			$q_n = 1{,}638$ (21 days)			$q_n = 4{,}914$ (63 days)		
		Bias	Stdev	RMSE	Bias	Stdev	RMSE	Bias	Stdev	RMSE
$\widehat{\Lambda}^{C,1}$	2.16	-0.08	7.74	7.73	-0.10	7.54	7.54	-0.28	6.76	6.76
$\widehat{\Lambda}^{C,2}$	5.75	-0.75	7.17	7.21	-0.93	6.94	7.00	-1.88	5.74	6.04
$\widehat{\Lambda}^{C,3}$	-5.63	0.43	7.17	7.18	0.67	6.95	6.98	1.78	5.67	5.93
$\widehat{\Lambda}^{J,1}$	3.18	0.11	2.31	2.31	0.13	2.32	2.33	0.21	2.39	2.40
$\widehat{\Lambda}^{J,2}$	-4.55	0.08	2.39	2.39	0.17	2.40	2.40	0.56	2.46	2.52
$\widehat{\Lambda}^{J,3}$	0.48	-0.01	2.33	2.33	0.00	2.32	2.32	0.08	2.29	2.29
$\widehat{\Lambda}^{J,4}$	-3.06	-0.26	2.36	2.37	-0.24	2.36	2.37	-0.19	2.35	2.36
$\widehat{\Lambda}^{J,5}$	1.43	0.04	2.43	2.43	0.04	2.44	2.43	0.02	2.43	2.42
$\widehat{\Lambda}^{J,6}$	-3.91	-0.24	2.51	2.52	-0.23	2.50	2.51	-0.18	2.48	2.49

$T_n = 10$ yrs $\Delta_n = 15$ mins

	Truth	$q_n = 312$ (12 days)			$q_n = 546$ (21 days)			$q_n = 4{,}914$ (63 days)		
		Bias	Stdev	RMSE	Bias	Stdev	RMSE	Bias	Stdev	RMSE
$\widehat{\Lambda}^{C,1}$	2.20	0.00	7.78	7.77	0.06	7.81	7.80	-0.22	7.03	7.03
$\widehat{\Lambda}^{C,2}$	5.84	-0.59	7.49	7.51	-0.68	7.46	7.49	-1.71	6.29	6.52
$\widehat{\Lambda}^{C,3}$	-5.76	0.59	7.42	7.44	0.63	7.27	7.30	1.50	6.05	6.23
$\widehat{\Lambda}^{J,1}$	3.03	0.28	2.27	2.28	0.25	2.25	2.26	0.35	2.32	2.35
$\widehat{\Lambda}^{J,2}$	-4.59	-0.01	2.22	2.22	0.02	2.22	2.22	0.43	2.27	2.31
$\widehat{\Lambda}^{J,3}$	0.49	0.14	2.52	2.53	0.16	2.48	2.48	0.28	2.45	2.46
$\widehat{\Lambda}^{J,4}$	-2.90	-0.21	2.39	2.39	-0.19	2.36	2.37	-0.08	2.33	2.33
$\widehat{\Lambda}^{J,5}$	1.23	0.17	2.49	2.50	0.17	2.46	2.47	0.15	2.43	2.43
$\widehat{\Lambda}^{J,6}$	-3.84	0.02	2.47	2.47	0.04	2.46	2.46	0.06	2.39	2.39

$T_n = 5$ yrs $\Delta_n = 5$ mins

	Truth	$q_n = 936$ (12 days)			$q_n = 1{,}638$ (21 days)			$q_n = 4{,}914$ (63 days)		
		Bias	Stdev	RMSE	Bias	Stdev	RMSE	Bias	Stdev	RMSE
$\widehat{\Lambda}^{C,1}$	1.51	0.49	10.54	10.54	0.45	10.36	10.36	0.18	9.07	9.07
$\widehat{\Lambda}^{C,2}$	6.57	-0.74	10.36	10.38	-1.07	10.00	10.05	-2.53	8.18	8.56
$\widehat{\Lambda}^{C,3}$	-6.52	0.78	10.72	10.74	1.05	10.31	10.36	2.39	8.61	8.93
$\widehat{\Lambda}^{J,1}$	2.44	0.27	3.10	3.11	0.30	3.11	3.13	0.40	3.21	3.24
$\widehat{\Lambda}^{J,2}$	-6.19	0.05	3.22	3.21	0.20	3.23	3.23	0.89	3.31	3.42
$\widehat{\Lambda}^{J,3}$	0.00	0.10	3.45	3.45	0.13	3.45	3.45	0.29	3.43	3.44
$\widehat{\Lambda}^{J,4}$	-2.92	-0.14	3.41	3.41	-0.13	3.41	3.41	0.03	3.37	3.36
$\widehat{\Lambda}^{J,5}$	0.95	-0.14	3.44	3.44	-0.14	3.42	3.42	-0.16	3.38	3.39
$\widehat{\Lambda}^{J,6}$	-3.62	-0.21	3.55	3.55	-0.17	3.54	3.54	-0.02	3.47	3.47

Table 2: Simulation Results

Note: This table reports simulation results for various scenarios of T_n and Δ_n for different choices of q_n. Column "Truth" provides the average of $\widehat{\Lambda}_T$ over 1,000 Monte Carlo samples. Columns "Bias", "Stdev", and "RMSE" provide the bias, the standard deviation, and the root-mean-squared-error of each of the 9 components of $\widehat{\Lambda}_T - \Lambda_T$. All numbers in this table are multiplied by 100.

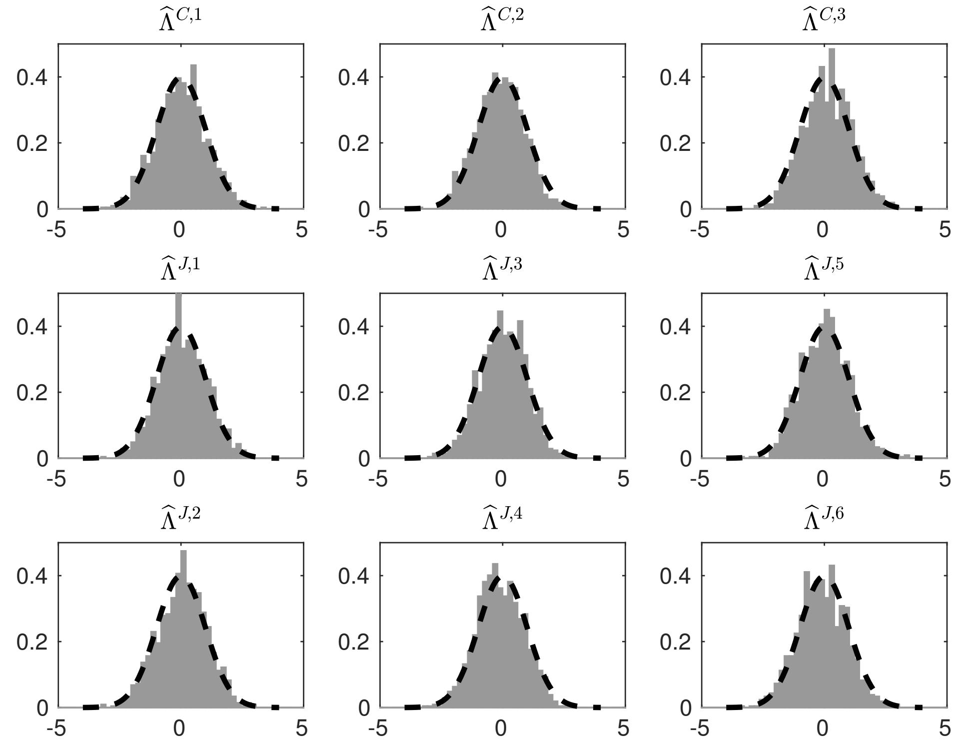

Figure 1: Standardized Estimates of $\widehat{\Lambda}_n$

Note: In this figure, we report the standardized estimates for the risk premia estimator $\widehat{\Lambda}_n$. The first row plots the estimates for the continuous components, the second row risk premia estimates for negative jump factors, and the bottom for positive jump factors. The first, second, and third columns correspond to F^1, F^2, and F^3, respectively. The span is 10 years long and the sampling frequency is every 5 minutes. The local window size q_n is fixed at 1,638 (21 days).

including the dot-com bubble, the financial crisis, and the recent and ongoing covid-19 pandemic.

To preprocess the data downloaded from NYSE TAQ, we follow standard procedure. We remove trades and quotes with condition codes Z, B, U, T, L, G, W, K, J, and the corresponding odd lot trades with an additional letter I, as well as those with non-empty suffix codes for preferred shares. We identify the opening and closing trades with their unique identifiers, and remove all trades beyond this window. We only keep trades with correction indicator 00 or 01. We then construct the national best bid and offer (NBBO) data using quotes from all exchanges every second, with which we match the trades and eliminator, among them, that are outside the range of NBBOs. Finally, we remove all trades from TAQ whose prices exceed the daily minimum and maximum prices from CRSP.

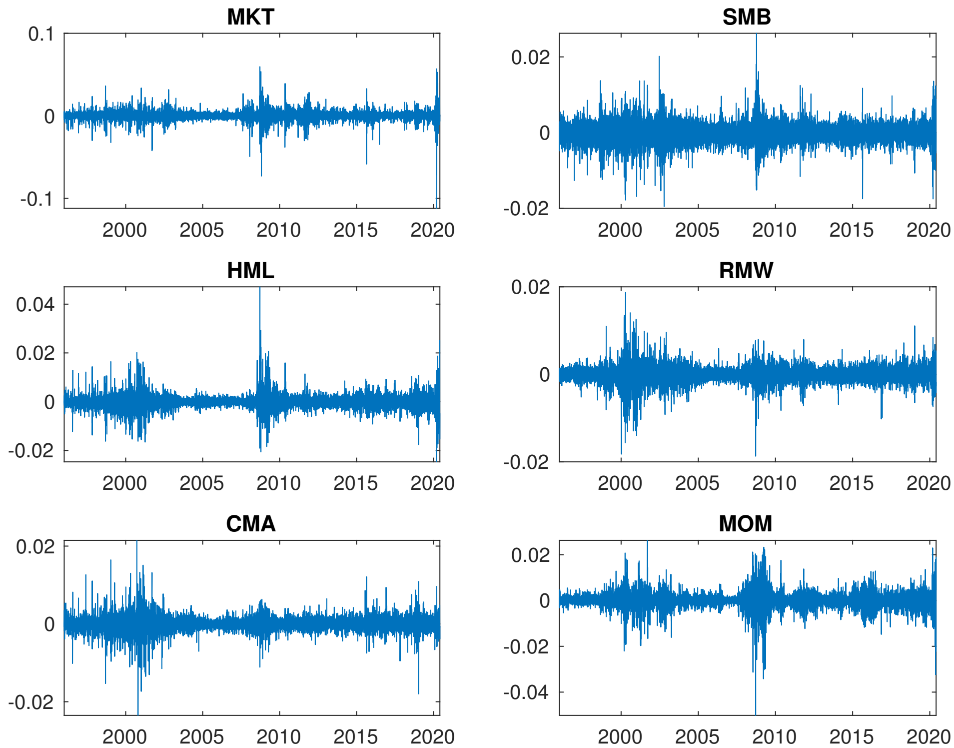

Figure 2: 15-minute Factor Returns

Note: In this figure, we plot the time series of 15-minute returns for the Fama-French five factors (MKT, SMB, HML, RMW, CMA) plus momentum (MOM) from January 1, 1996 to May 31, 2020.

We employ 15-minute snapshots of the Fama-French five factors (Fama and French (2015)) plus momentum (MOM), which were constructed at high frequency by Aït-Sahalia, Kalnina, and Xiu (2020). The five factors are the market return (MKT), and mimicking portfolios for size (SMB, small minus big), value (HML, high minus low), profitability (RMW, robust minus weak) and investment (CMA, conservative minus aggressive). The mimicking portfolio for MOM is UMD (up minus down). The high-frequency factor returns are plotted in Figure 2.[10] Large returns, or possibly jumps, are clearly present for all factors, particularly in earlier sample periods. Figure 3 plots time series of factor jumps while Figure 4 compares their magnitudes.

We estimate augmented versions of the CAPM, Fama-French three-factor (FF3),

[10] "MKT" denotes market return in excess of the one-month T-bill rate, which serves as the proxy for risk free rate in our model.

		CAPM		FF3		FF4		FF5		FF6	
MKT	Cont. + Small Jumps	4.73	(1.55)	2.24	(0.96)	2.48	(1.22)	3.70	(1.83)	3.41	(1.81)
	Large Neg. Jumps	9.02	(2.96)	9.22	(3.78)	8.78	(4.00)	8.18	(3.75)	8.66	(4.13)
	Large Pos. Jumps	-2.36	(−1.05)	0.70	(0.32)	5.11	(2.64)	2.35	(1.07)	4.11	(2.02)
SMB	Continuous			3.12	(2.16)	3.53	(2.62)	2.31	(1.65)	3.03	(2.31)
	Jumps			-0.79	(−0.86)	-1.83	(−2.00)	-0.76	(−0.82)	-1.42	(−1.55)
HML	Continuous			0.43	(0.28)	0.99	(0.76)	0.08	(0.06)	0.91	(0.73)
	Jumps			-5.63	(−3.44)	-5.55	(−3.47)	-7.02	(−4.53)	-5.54	(−3.64)
RMW	Continuous							0.52	(0.48)	0.62	(0.60)
	Jumps							1.74	(1.58)	-0.33	(−0.29)
CMA	Continuous							-0.22	(−0.24)	0.31	(0.37)
	Jumps							-3.13	(−2.43)	-3.59	(−2.87)
MOM	Continuous					-2.37	(−0.98)			-2.48	(−1.06)
	Jumps					14.57	(6.75)			14.27	(6.05)
R^2 (%)		18.21		21.57		22.89		23.21		24.14	

Table 3: Estimation Results using High Frequency/Long Horizon Data

Note: This table reports the estimated risk premia for a variety of models. "CAPM" includes only MKT and its jumps as factors, "FF3" adds to that HML, SMB and their jumps, "FF4" further adds MOM to "FF3", "FF5" adds RMW and CMA to "FF3", and finally "FF6" includes all six portfolios and their jumps. For MKT, we separately estimate the premia of its large positive and large negative jump components, whereas for other factors we estimate their continuous and jump risk premia. Large jumps are those whose sizes exceed 1%. The risk premia estimates are multiplied by 100. All t-statistics are reported in parentheses. The reported R^2 is the time series average of the cross-sectional R^2s. The individual stock returns are sampled every 15 minutes from January 1, 1996 to May 31, 2020.

	CAPM		FF3		FF4		FF5		FF6	
MKT	7.00	(1.99)	7.45	(2.23)	7.44	(2.23)	7.59	(2.28)	7.44	(2.23)
SMB			0.30	(0.28)	0.44	(0.42)	0.32	(0.31)	0.38	(0.36)
HML			-0.76	(−0.57)	-0.82	(−0.62)	-0.84	(−0.63)	-1.00	(−0.76)
RMW							-0.84	(−0.63)	-0.81	(−0.61)
CMA							-0.20	(−0.15)	-0.17	(−0.13)
MOM					0.73	(0.35)			0.67	(0.32)
R^2 (%)	14.33		16.09		16.75		16.94		17.40	

Table 4: Estimation Results using Low Frequency Data and the Fama-McBeth Two-Pass Procedure

Note: This table reports the estimated risk premia for a variety of models. "CAPM" includes only the MKT factor, "FF3" adds HML and SMB, "FF4" further adds MOM to "FF3", "FF5" adds RMW and CMA to "FF3", and finally "FF6" includes all six portfolios. The risk premia are estimated by the standard Fama-MacBeth procedure in which betas are estimated on a monthly rolling window. The estimated risk premia are multiplied by 100. All t-statistics are reported in parentheses. The reported R^2 is the time series average of the cross-sectional R^2s. The individual stock returns are sampled daily from January 1, 1996 to May 31, 2020.

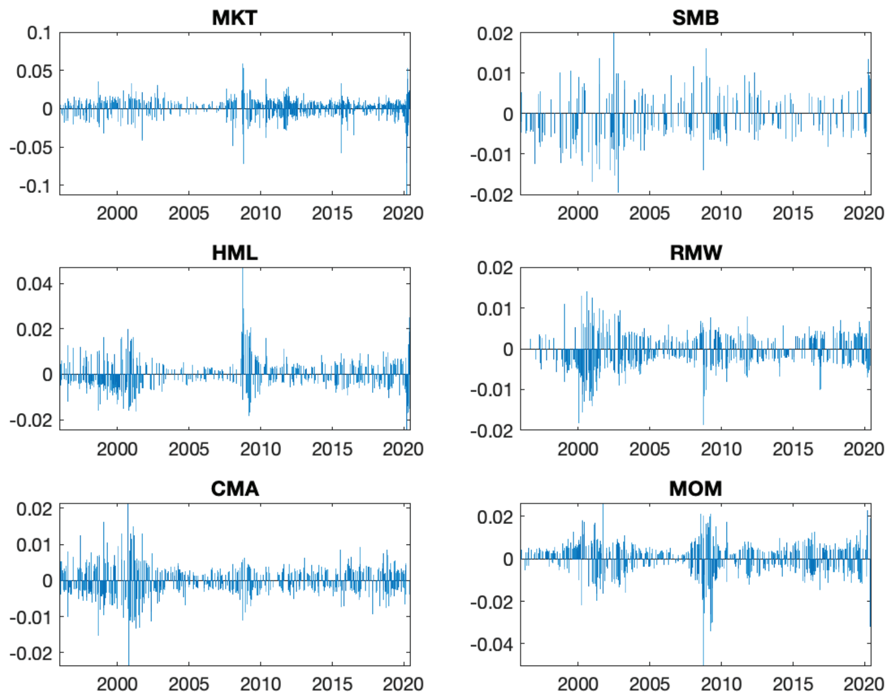

Figure 3: 15-minute Factor Jumps

Note: In this figure, we plot the time series of 15-minute return jumps for Fama-French five factors (MKT, SMB, HML, RMW, CMA) plus momentum (MOM) from January 1, 1996 to May 31, 2020. We set the threshold for jump truncation as $u_n = 3\Delta_n^{0.47}(\widehat{IV}_t)^{1/2}$, where $\widehat{IV}_t$ is a pilot estimate of daily volatility based on 15-minute returns.

four-factor (FF4, i.e., FF3 +MOM), five-factor (FF5), and six-factor models (FF6, i.e., FF5+MOM), in which we employ three factors for MKT (large positive jumps, large negative jumps, continuous component plus small jumps) and two factors for all other factors (continuous components and jumps) for parsimony and ease of interpretation; the theory allows for versatile choices of risk factors. For comparison with the discrete-time low frequency approach, we follow Lewellen and Nagel (2006) and implement Fama-MacBeth regressions on a monthly rolling window for each model with daily data, which yields estimates of the total risk premia (the validity of this approach requires much stronger assumptions on the data generating process, such as constancy of factor loadings, absence of jumps, etc., than what we have assumed).

Table 3 provides the estimates based on high frequency data, and Table 4 for daily data. We summarize the main findings as follows. First, the negative jumps of MKT are significantly priced with positive premia (12.5% per year), whereas its continuous compo-

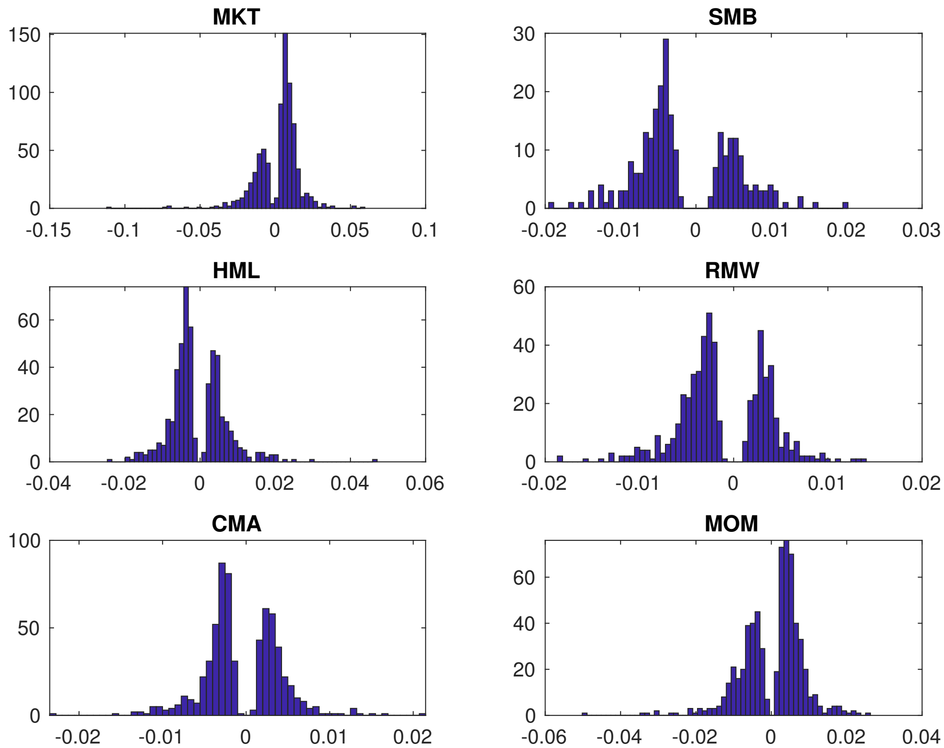

Figure 4: Histograms of 15-minute Factor Jumps

Note: In this figure, we plot the histograms of 15-minute return jumps for Fama-French five factors (MKT, SMB, HML, RMW, CMA) plus momentum (MOM) from January 1, 1996 to May 31, 2020. We set the threshold for jump truncation as $u_n = 3\Delta_n^{0.47}(\widehat{IV}_t)^{1/2}$, where $\widehat{IV}_t$ is a pilot estimate of daily volatility based on 15-minute returns.

nent and smaller jumps are statistically and economically insignificant. Our results seem to suggest that negative market jump beta is a stronger proxy for market risk, regardless of the choice of the benchmark model. Second, the continuous component of SMB, the jump components of HML, CMA, and MOM are also significant with economically meaningful magnitudes, and robustly so across different models. Neither the continuous part nor the jumps of RMW appear significant. In contrast, the results based on low frequency Fama-MacBeth regressions are insignificant throughout. This is not surprising given the potential model misspecification and generally low signal-to-noise ratio in individual stock returns. Third, the time series average of cross-sectional R^2s increases from 18.21% to 24.14%, as the model expands. By contrast, the low frequency R^2s are substantially lower – the largest model's R^2 is smaller than 18%.

		CAPM		FF3		FF4		FF5		FF6	
MKT	Cont. + Small Jumps	6.87	(1.99)	3.00	(1.02)	4.34	(1.68)	4.09	(1.58)	4.01	(1.66)
	Large Neg. Jumps	4.22	(1.24)	6.30	(2.21)	6.94	(2.65)	7.07	(2.59)	7.92	(3.04)
	Large Pos. Jumps	-2.13	(−0.69)	5.04	(1.76)	6.39	(2.36)	4.16	(1.53)	5.47	(2.11)
SMB	Continuous			3.08	(1.9)	3.35	(2.21)	3.14	(1.99)	3.78	(2.56)
	Jumps			-5.17	(−4.04)	-5.31	(−4.16)	-5.38	(−4.27)	-5.52	(−4.4)
HML	Continuous			0.16	(0.09)	1.58	(0.98)	0.82	(0.45)	2.19	(1.32)
	Jumps			-7.12	(−4.61)	-4.91	(−3.55)	-7.51	(−4.76)	-5.17	(−3.73)
RMW	Continuous							1.17	(1.02)	1.22	(1.10)
	Jumps							2.19	(2.45)	-0.40	(−0.44)
CMA	Continuous							-0.90	(−0.85)	-0.06	(−0.06)
	Jumps							-4.10	(−3.38)	-4.05	(−3.31)
MOM	Continuous					-2.88	(−0.98)			-3.16	(−1.08)
	Jumps					16.96	(6.56)			18.29	(6.57)
R^2 (%)		19.88		22.95		24.31		24.75		25.76	

Table 5: Robustness Check: Alternative Sampling Period

Note: The content of this table is comparable to Table 3, except that the sample period in this case starts from January 1, 2004.

		CAPM		FF3		FF4		FF5		FF6	
MKT	Cont. + Small Jumps	4.73	(1.55)	5.02	(1.87)	5.62	(2.24)	7.19	(2.41)	7.32	(2.48)
	Large Neg. Jumps	9.02	(2.96)	10.02	(4.27)	8.36	(4.02)	9.68	(4.06)	9.83	(4.4)
	Large Pos. Jumps	-2.36	(−1.05)	-2.45	(−1.16)	0.95	(0.5)	-2.88	(−1.24)	-1.73	(−0.78)
SMB	Cont. + Small Jumps			3.46	(2.29)	2.90	(1.98)	2.14	(1.45)	2.25	(1.56)
	Large Neg. Jumps			-0.09	(−0.18)	0.12	(0.27)	0.57	(0.87)	0.30	(0.46)
	Large Pos. Jumps			-1.19	(−2.18)	-1.02	(−1.94)	-0.66	(−1.12)	-0.47	(−0.81)
HML	Cont. + Small Jumps			-1.49	(−0.92)	-1.04	(−0.72)	-2.0	(−1.19)	-1.34	(−0.84)
	Large Neg. Jumps			-1.67	(−1.52)	-2.48	(−2.41)	-2.78	(−2.62)	-2.55	(−2.46)
	Large Pos. Jumps			-1.62	(−1.67)	-1.41	(−1.5)	-1.10	(−1.11)	-0.78	(−0.83)
RMW	Cont. + Small Jumps							0.66	(0.51)	0.48	(0.39)
	Large Neg. Jumps							-1.06	(−1.51)	-1.36	(−1.98)
	Large Pos. Jumps							-0.21	(−0.71)	-0.16	(−0.54)
CMA	Cont. + Small Jumps							-0.57	(−0.52)	-0.27	(−0.27)
	Large Neg. Jumps							-1.12	(−1.57)	-1.28	(−1.88)
	Large Pos. Jumps							0.70	(1.10)	0.20	(0.33)
MOM	Cont. + Small Jumps					-1.41	(−0.56)			-1.55	(−0.58)
	Large Neg. Jumps					3.36	(2.85)			4.33	(3.15)
	Large Pos. Jumps					3.62	(3.05)			3.59	(2.29)
R^2 (%)		18.21		21.35		22.61		22.60		23.49	

Table 6: Robustness Check: Alternative Specification of Risk Factors

Note: The content of this table is comparable to Table 3, except that all factors' large positive and negative jumps are treated as separate risk factors.

		CAPM		FF3		FF4		FF5		FF6	
MKT	Cont. + Small Jumps	4.39	(1.44)	1.89	(0.82)	1.78	(0.88)	3.33	(1.68)	2.97	(1.60)
	Large Neg. Jumps	7.44	(2.40)	8.55	(3.4)	8.04	(3.52)	7.43	(3.28)	8.12	(3.71)
	Large Pos. Jumps	-0.44	(-0.18)	1.85	(0.78)	6.96	(3.27)	3.75	(1.57)	5.70	(2.59)
SMB	Continuous			3.13	(2.16)	3.76	(2.79)	2.50	(1.79)	3.25	(2.50)
	Jumps			-0.84	(-0.91)	-2.31	(-2.46)	-1.19	(-1.25)	-1.90	(-2.00)
HML	Continuous			0.48	(0.32)	1.11	(0.85)	0.02	(0.01)	0.87	(0.70)
	Jumps			-5.76	(-3.47)	-5.59	(-3.42)	-7.07	(-4.52)	-5.44	(-3.55)
RMW	Continuous							0.54	(0.50)	0.63	(0.61)
	Jumps							2.02	(1.80)	-0.10	(-0.09)
CMA	Continuous							-0.24	(-0.26)	0.31	(0.37)
	Jumps							-3.07	(-2.38)	-3.57	(-2.87)
MOM	Continuous					-2.44	(-1.01)			-2.49	(-1.06)
	Jumps					15.19	(6.91)			14.84	(6.22)
Cross-Sectional R^2 (%)		18.2		21.59		22.93		23.24		24.17	

Table 7: Robustness Check: Alternative Threshold for "Large" Jumps

Note: The content of this table is comparable to Table 3, except that the threshold for "large" jumps in the MKT factor is set at 0.5%.

We finally conduct a battery of robustness checks. In light of the potential concern on the presence of microstructure noise, we provide results in Table 5 with a smaller sample excluding the period of 1996 - 2003, during which microstructure noise may be prevalent in small- and mid-cap stocks even at 15-minute frequencies. In addition, we consider an alternative specification of jump factors in Table 6, and a different threshold for "large" jumps in Table 7. These results are consistent with the primary findings summarized above.

6 Conclusions

High frequency econometrics has made much progress over the years characterizing asset returns dynamics, with a particular emphasis on functions of second order "moments" in the form of quadratic (co)variations. The empirical asset pricing literature, however, retains an important focus on first order "moments", corresponding to the drift components in semimartingales. This paper provides new econometric techniques for inference on this drift component, allowing standard asset pricing models to be estimated using a generalization to continuous-time factor models of the Fama-MacBeth two-pass regression procedure that combines high frequency and long horizon methods. These techniques can be employed in future work to test the specification of the canonical asset pricing factor model, estimate individual alphas, conduct event studies, all of which are standard issues in empirical asset pricing that have not yet been fully investigated in a continuous-time setting.

References

Aït-Sahalia, Y., and J. Jacod (2014): *High Frequency Financial Econometrics*. Princeton University Press.

Aït-Sahalia, Y., I. Kalnina, and D. Xiu (2020): "High Frequency Factor Models and Regressions," *Journal of Econometrics*, 216, 86–105.

Aït-Sahalia, Y., and D. Xiu (2017): "Using Principal Component Analysis to Estimate a High Dimensional Factor Model with High-Frequency Data," *Journal of Econometrics*, 201, 388–399.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and J. Wu (2005): "A Framework for Exploring the Macroeconomic Determinants of Systematic Risk," *American Economic Review*, 95, 398–404.

Ang, A., and D. Kristensen (2012): "Testing Conditional Factor Models," *Journal of Financial Economics*, 106, 132–156.

Bandi, F. M., and P. C. B. Phillips (2003): "Fully Nonparametric Estimation of Scalar Diffusion Models," *Econometrica*, 71, 241–283.

Barndorff-Nielsen, O. E., and N. Shephard (2004): "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72(3), 885–925.

Bollerslev, T., S. Z. Li, and V. Todorov (2016): "Roughing up beta: Continuous versus discontinuous betas and the cross section of expected stock returns," *Journal of Financial Economics*, 120, 464–490.

Chang, Y., Y. Choi, H. Kim, and J. Y. Park (2016): "Evaluating Factor Pricing Models using High-Frequency Panels," *Quantitative Economics*, 7, 889–933.

Crimaldi, I., and L. Pratelli (2005): "Convergence results for multivariate martingales," *Stochastic Processes and their Applications*, 115, 571–577.

Daniel, K., and T. J. Moskowitz (2016): "Momentum Crashes," *Journal of Financial Economics*, 122, 221–247.

Fama, E. F., and K. R. French (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116(1), 1–22.

Fama, E. F., and J. D. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

GAGLIARDINI, P., E. OSSOLA, AND O. SCAILLET (2016): "Time-varying risk premium in large cross-sectional equity datasets," *Econometrica*, 84, 985–1046.

GU, S., B. KELLY, AND D. XIU (2019): "Autoencoder Asset Pricing Models," Discussion paper, University of Chicago.

JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Springer-Verlag.

KELLY, B., S. PRUITT, AND Y. SU (2019): "Characteristics are Covariances: A Unified Model of Risk and Return," *Journal of Financial Economics, forthcoming.*

LEWELLEN, J., AND S. NAGEL (2006): "The conditional CAPM does not explain asset-pricing anomalies," *Journal of Financial Economics*, 82(2), 289–314.

LI, J., V. TODOROV, AND G. TAUCHEN (2016): "Inference Theory on Volatility Functional Dependencies," *Journal of Econometrics*, 193, 17–34.

——— (2017): "Jump Regressions," *Econometrica*, 85, 173–195.

MYKLAND, P. A., AND L. ZHANG (2006): "ANOVA for Diffusions and Itô Processes," *Annals of Statistics*, 34, 1931–1963.

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

RAPONI, V., C. ROBOTTI, AND P. ZAFFARONI (2019): "Testing Beta-Pricing Models Using Large Cross-Sections," *Review of Financial Studies, forthcoming.*

REISS, M., V. TODOROV, AND G. E. TAUCHEN (2015): "Nonparametric Test for a Constant Beta between Itô Semi-martingales based on High-Frequency Data," *Stochastic Processes and their Applications*, 125, issue 8, pages, 2955–2988.

ROSENBERG, B. (1974): "Extra-Market Components of Covariance in Security Returns," *Journal of Financial and Quantitative Analysis*, 9, 263–274.

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

SHANKEN, J. (1992): "On the Estimation of Beta Pricing Models," *The Review of Financial Studies*, 5, 1–33.

TODOROV, V., AND T. BOLLERSLEV (2010): "Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks," *Journal of Econometrics*, 157, 220–235.

Appendix: Proofs

Throughout the Appendix, we suppose that all assumptions 1–5 hold, and C denotes a constant which may change from line to line and may depend on the bounds in our assumptions, but not on n or the component indices k, h, m or the index i when we use $\mathcal{I}_i^n$ or $\Delta_i^n X$ for example; when it depends on an extra parameter p we call it C_p. The integers k, h, m used below are always supposed to belong to $\{1,, \dots, K\}$ or $\{1, \dots, H\}$ or $\{1, \dots, M\}$, respectively. We use the simplifying notation $\mathbb{E}_i^n$ for the conditional expectation with respect to $\mathcal{F}_{i\Delta_n}$, and recall that $I_n = \{1, \dots, n\}$.

Since we look at asymptotic properties, without loss of generality and in view of assumption 2, we always assume

$$\Delta_n < 1, \qquad \theta_m - 2\Delta_n \text{ is a stopping time for all } m. \tag{A.1}$$

A.1 Preliminaries

Let us first state a few – more or less well known – facts about matrices, starting with the following trivial estimate (on matrices with the proper dimensions):

$$\|(A+B)(A'+B') - AA'\| \leq \|A\|\,\|B'\| + \|B\|\,\|A'\| + \|B\|\,\|B'\|. \tag{A.2}$$

Next, we have:

Lemma 1. *Let $A, A' \in M_d^+$ and $B = A' - A$, suppose that $\zeta(A) \geq 1/a$ for some $a \geq 1$, and let $v \geq 2a$. The matrix $\overline{A}$ equal to A'^{-1} if $\zeta(A') > 1/v$ and to 0 otherwise satisfies:*

$$\|\overline{A} - A^{-1}\| \leq 3av\,\|B\|, \qquad \|\overline{A} - A^{-1} - A^{-1}BA^{-1}\| \leq 7a^2v\,\|B\|^2. \tag{A.3}$$

Proof. We write $C = \overline{A} - A^{-1}$ and $D = C - A^{-1}BA^{-1}$, and single out three cases.

Suppose first that $\zeta(A') \leq 1/v$, so there is a unit vector x in R^q with $\|A'x\| \leq 1/2a$, whereas $\|Ax\| \geq 1/a$ by hypothesis, so $\|Bx\| \geq 1/2a$ and thus $\|B\| \geq 1/2a \geq 1/v$. Therefore $\|A^{-1}\| \leq a \leq av\|B\| \leq 2a^2v\|B\|^2$ and $\|A^{-1}BA^{-1}\| \leq a^2\|B\| \leq a^2v\|B\|^2$ and, since $\overline{A} = 0$, (A.3) holds.

Next, suppose that $\zeta(A') > 1/v$ and $\|A^{-1}B\| > 1/2$. We have $\|B\| > 1/2a \geq 1/v$ and $\|\overline{A}\| = \|A'^{-1}\| < v \leq 2av\|B\| \leq 4a^2v\|B\|^2$ and $\|A^{-1}\| \leq a \leq av\|B\| \leq 2a^2v\|B\|^2$ and $\|A^{-1}BA^{-1}\| \leq a^2\|B\| \leq a^2v\|B\|^2$, which clearly imply (A.3).

Finally, suppose that $\zeta(A') > 1/v$ and $\|A^{-1}B\| \leq 1/2$. The matrix $G = \sum_{j\geq 0}(A^{-1}B)^jA^{-1}$ is then well defined, with $\|G\| \leq a$, and with I the $q \times q$ identity matrix and since $\overline{A} = A'^{-1}$

now, we have

$$C=\big((I+A^{-1}B)^{-1}-I\big)\,A^{-1}=A^{-1}BG,\qquad D=C-A^{-1}BA^{-1}=(A^{-1}B)^2G.$$

Therefore $\|C\|\leq a^2\|B\|\leq av\|B\|/2v$ and $\|D\|\leq a^3\|B\|^2\leq a^2v\|B\|^2/2$, and again (A.3) holds. $\square$

Next, we state some properties connected with (17):

Lemma 2. *If Y_t is a bounded adapted $\mathcal{M}_{p,q}$–valued process satisfying (17) on some interval $\mathcal{L}$ and f is a C^2 function on $\mathcal{M}_{p,q}$, then the process $f(Y_t)$ also satisfies (17) on $\mathcal{L}$. In particular:*

(a) If $p=q$ and Y_t^{-1} exists and is also bounded on $\mathcal{L}$, it satisfies (17) on $\mathcal{L}$ as well.

(b) If Y'_t is $\mathcal{M}_{p,q}$–valued and satisfies the same conditions as Y_t, then $Y_t+Y'_t$ satisfies (17) on $\mathcal{L}$.

(c) If Y'_t is $\mathcal{M}_{q,r}$–valued and satisfies the same conditions as Y_t, then $Y_tY'_t$ satisfies (17) on $\mathcal{L}$.

Proof. (a) follows from the first claim because for any $\alpha>1$ there is a C^2 function f on $\mathcal{M}_{p,p}$ such that $f(x)=x^{-1}$ on the set of all $x\in\mathcal{M}_{p,p}$ such that $\|x\|,\ \|x^{-1}\|\leq\alpha$. Since as soon as two processes Y_t and Y'_t satisfy (17) the same holds for the pair (Y_t,Y'_t), (b) and (c) also follow from the first claim, with the function $f(x,y)=x+y$ on $(\mathcal{M}_{p,q})^2$ for (b), and $f(x,y)=xy$ on $\mathcal{M}_{p,q}\times\mathcal{M}_{q,r}$ for (c).

For the first claim, we can suppose without restriction that f is one-dimensional and also bounded, as well as its partial derivatives of order 1 and 2, because Y_t takes its values in a compact subset of $\mathcal{M}_{p,q}$. We have

$$|f(Y_{t+s}-f(Y_t)|\leq C\|Y_{t+s}-Y_t\|,\ |f(Y_{t+s}-f(Y_t)-\nabla f(Y_t)(Y_{t+s}-Y_t)|\leq C\|Y_{t+s}-Y_t\|^2.$$

Then the claim is obvious, upon using the fact that $\nabla f(Y_t)$ is $\mathcal{F}_t$–measurable. $\square$

In the next lemma, we consider a possibly multi-dimensional Itô semimartingale X of the form

$$X_t=X_0+\int_0^t a_s\,ds+\int_0^t b_s\,dW_s+\int_0^t\int_E\delta(s,z)\,\underline{p}(ds,dz),$$

where W is a Brownian motion and $\underline{p}$ a Poisson measure with compensator $\underline{q}$, and where a_t,b_t are bounded and $\|\delta(t,z)\|\leq\Upsilon(z)$, with Υ as in Assumption 1. With $\psi\in\big(0,\frac{1}{2}\big)$ and $\varphi>0$ we set

$$X(\psi,\varphi)^n_t=X_t-\sum_{s\leq t}\Delta X_s\,1_{\{\|\Delta X_t\|>w_n\}},\ \text{where } w_n=\begin{cases}0 & \text{if } \alpha=0\\ \Delta_n^{\psi+\varphi/2\alpha} & \text{otherwise,}\end{cases}\tag{A.4}$$

Lemma 3. *In the previous setting there is a constant* Γ *(depending on the bounds on* a_t, b_t *and on* Υ *and* ψ, η*) such that, if* $A_n = \{t > 0 : \|\Delta X_t\| > w_n\}$*, we have*

$$\mathbb{P}(\Omega_i^n) \le \Gamma \Delta_n^{2-2\alpha\psi-\varphi}, \ \textit{where } \Omega_i^n = \big\{\#(A_n \cap \mathcal{I}_i^n) \ge 2\big\} \bigcup \big\{\|\Delta_i^n X(\psi,\varphi)^n\| > \Delta_n^{\psi}\big\}.$$

Proof. One has the decomposition $X(\psi,\varphi) = X_0 + X^C + X'^n$, where

$$X_t^C = \int_0^t a_s\, ds + \int_0^t b_s\, dW_s, \qquad X_t'^n = \sum_{s \le t} \Delta X_s\, 1_{\{\|\Delta X_s\| \le w_n\}}.$$

The boundedness of a_t, b_t yields $\mathbb{E}(\|\Delta_i^n X^C\|^p) \le C_p \Delta_n^{p/2}$ for any $p > 0$, so by Markov inequality and upon taking $p > \frac{4}{1-2\psi}$ we have

$$\mathbb{P}\big(\|\Delta_i^n X^C\| > \frac{1}{2}\, \Delta_n^{\psi}\big) \le \frac{2^p C_p}{\Delta_n^{p\psi}}\, \Delta_n^{p/2} \le C \Delta_n^2.$$

On the other hand, $X'^n \equiv 0$ when $\alpha = 0$, and otherwise by (2.1.41) of Jacod and Protter (2012) we have for any $p \ge 1$:

$$\mathbb{E}(\|\Delta_i^n X'^n\|^p) \le C_p\big(\Delta_n \int_E \Upsilon(z)^p\, 1_{\{\Upsilon(z) \le w_n\}}\, \nu(dz) + \Delta_n^p\big) \le C_p\big(\Delta_n^{1+(p-\alpha)(\psi+\varphi/2\alpha)} + \Delta_n^p\big),$$

and by Markov's inequality again we deduce, upon taking p large enough:

$$\mathbb{P}\big(\|\Delta_i^n X'^n\| > \frac{1}{2}\, \Delta_n^{\psi}\big) \le C_p\big(\Delta_n^{1+(p-\alpha)(\psi+\varphi/2\alpha)-p\psi} + \Delta_n^{p(1-\psi)}\big) \le C \Delta_n^2.$$

Therefore $\Omega_i'^n = \{\|\Delta_i^n X(\psi,\varphi)^n\| > \Delta_n^{\psi}\}$ satisfies $\mathbb{P}(\Omega_i'^n) \le C \Delta_n^2$.

Finally, any jump of X with size bigger than w_n occurs at a jump time of the process $\underline{p}\big((0,t]] \times \{z : \Upsilon(z) > w_n\}\big)$, which is a Poisson process with parameter $\chi_n = \nu(\{z : \Upsilon(z) > w_n\})$. By Markov's inequality once more, we have $\chi_n \le C w_n^{-\alpha}$ because Υ^α is ν-integrable, so

$$\mathbb{P}\big(\{\#(A_n \cap \mathcal{I}_i^n) \ge 2\}\big) \le (\chi_n \Delta_n)^2 \le C \Delta_n^{2-2\alpha\psi-\varphi},$$

and the claim follows. □

A.2 The case where $H \ge 1$

In this subsection we prove a number of results that are specifically needed in the case $H \ge 1$. We begin with the following one, with ψ, φ as in (A.4):

Lemma 4. *As soon as* $T_n \Delta_n^{1-2\psi-\varphi} \to 0$ *we have*

$$\mathbb{P}\big(\sup_{i \in I_n} \|\widehat{\beta}_{n,i}^J - \beta_{iq_n\Delta_n}^J\| \le v_n^2 T_n^{3/2} \Delta_n^{\psi}\big) \to 1.$$

Proof. 1) Since $\widehat{\beta}^{J,m,h}_{n,i} = 0$ by construction and $\beta^{J,m,h}_{iq_n\Delta_n} = 0$ by hypothesis when $iq_n\Delta_n \notin \mathcal{L}_m$, whereas $v_n \to \infty$, it is enough to prove that for all m, h there is a constant C such that

$$\mathbb{P}(\Omega^{m,h}_n) \to 1, \text{ where } \Omega^{m,h}_n = \big\{|Y^{m,h}_n - \overline{\beta}^{J,m,h}| \leq C v_n T_n^{3/2} \Delta_n^{\psi}\big\}, \quad Y^{m,h}_n = \sum_{j \in I^m_n} \widehat{R}''^{h,j}_{n,m}\, \Delta P^m_j.$$

Therefore, below we fix m, h and basically drop them from our notation. We apply Lemma 3 with the $K+1$-dimensional process $X = (F, P^m)$ and the associated sets Ω^n_i, to get that the set $\Omega_{n,1} = \cap_{i=1}^n (\Omega^n_i)^c$ satisfies

$$\mathbb{P}((\Omega_{n,1})^c) \leq \sum_{i=1}^n \mathbb{P}(\Omega^n_i) \leq C T_n \Delta_n^{1-2\psi-\varphi} \to 0.$$

Next, for $\chi, \chi' > 0$ let $D_n(\chi, \chi')$ be the (random) set of all times in $(\zeta_m, \theta_m] \cap (0, T_n]$ at which at least one of the components $\overline{F}^h$ has a jump with size in $\overline{B}^h(\chi, \chi')$, and write $N_n(\chi, \chi') = \#(D_n(\chi, \chi'))$. The set $D_n(\chi, \chi')$ is included into the set of jump times of the Poisson process $Y_t = \underline{p}([0,t] \times \{z : \Upsilon(z) \geq \chi\})$, whose parameter (depending on χ) is finite, hence for some constant C_χ we have $\mathbb{P}(Y_{T_n} > C_\chi T_n) \to 0$ as $n \to \infty$ and also $\mathbb{P}(Y_{S+2\Delta_n} > Y_S) \leq C_\chi \Delta_n$ for any finite stopping time S. This with $\chi = \rho/2$, and since $\zeta_m \wedge T_n$ and $(\theta_m - 2\Delta_n) \wedge T_n$ are stopping times (recall (A.1)), yields

$$j = 2,3,4 \Rightarrow \mathbb{P}(\Omega_{n,j}) \to 1, \text{ where } \begin{cases} \Omega_{n,2} = \{N_n(\rho/2, 2\rho') \leq C_{\rho/2} T_n\} \\ \Omega_{n,3} = \{D_n(\rho/2, \rho') \cap (\zeta_m, \zeta_m + \Delta_n] = \emptyset\} \\ \Omega_{n,4} = \{D_n(\rho/2, \rho') \cap (\theta_m - 2\Delta_n, \theta_m] = \emptyset\}. \end{cases}$$

Moreover, Assumption 3 with ε, ρ, ρ' yields

$$\mathbb{P}(\Omega_{n,5}) \to 1, \qquad \text{where} \quad \Omega_{n,5} = \{\zeta(R(m,\rho,\rho',T_n)^\intercal\, R(m,\rho,\rho',T_n)) \geq \varepsilon\},$$

hence the set $\Omega'_n = \cap_{j=1}^5 \Omega_{n,j}$ satisfies $\mathbb{P}(\Omega'_n) \to 1$, and we are left to proving that $\Omega_n \subset \Omega^{m,h}_n$.

2) For proving $\Omega_n \subset \Omega^{m,h}_n$ we argue ω-wise, with n fixed and ω fixed inside Ω_n. Without loss of generality we may also assume $\Delta_n^\psi < \rho/2$, hence a fortiori $w_n < \rho/2$ with w_n as in (A.4). Coming back to (26), we let A be the set of all $j \in I_n$ for which the vector $(\widehat{R}^{j,h}_{n,m})_{1 \leq h \leq H}$ is not vanishing, and $A' = I_n \backslash A$. Since $\omega \in (\Omega^n_j)^c$ for each $j \in I_n$ and $w_n < \rho/2$, we see that $A \subset I^n_m$ and

$$\begin{array}{ll} j \in A & \Rightarrow\ \mathcal{I}^n_j \cap D_n(\rho/2, 2\rho') = \{t_i\},\ \|\Delta^n_j \overline{F} - \Delta \overline{F}_{t_i}\| + \big|\Delta^n_j P^m - \sum\limits_{h'=1}^H \overline{\beta}^{J,m,h'} \Delta \overline{F}^{h'}_{t_j}\big| \leq \Delta_n^\psi \\ j \in A' & \Rightarrow\ \mathcal{I}^n_j \cap D_n(2\rho, \rho'/2) = \emptyset, \end{array} \tag{A.5}$$

and below we define the $n \times H$ matrix R by $R^{j,h'} = \Delta \overline{F}^{h'}_{t_j}$ if $j \in A$ and $R^{j,h} = 0$ if $j \in A'$.

With the notation (19), we have $D_n(2\rho,\rho'/2) = A(\rho,\rho')^m_{T_n}$, hence $N_n(2\rho,\rho'/2) = N(\rho,\rho')^m_{T_n}$, which is not bigger that $N_n(\rho/2,2\rho')$, and the second part of (A.5) implies that each column of the matrix $R(m,\rho,\rho',T_n)$ is also a column of the matrix R. Thus $\zeta(R^\intercal R) \geq \zeta(R(m,\rho,\rho',T_n)^\intercal R(m,\rho,\rho',T_n)) \geq \varepsilon$ (recall $\omega\in\Omega_{n,5}$), implying $\|(R^\intercal R)^{-1}\| \leq 1/\varepsilon$.

Set $R' = R^\intercal R$ and $R'' = (R')^{-1} R^\intercal$. Observing that R has at most $N_n(\rho/2,2\rho') \leq C_{\rho/2}T_n$ non vanishing columns and that each entry $R^{i,h}$ is smaller than $2\rho'$, we easily check that $\|R\| \leq C\sqrt{T_n}$ and $\|R^\intercal\| \leq C\sqrt{T_n}$, so $\|R'\| \leq CT_n$. The same argument using the first part of (A.5) yields $\|\widehat{R}_{n,m} - R\| \leq C\sqrt{T_n}\,\Delta_n^\psi$ and $\|\widehat{R}^\intercal_{n,m} - R^\intercal\| \leq C\sqrt{T_n}\,\Delta_n^\psi$, hence (A.2) gives us $\|\widehat{R}'_{n,m} - R'\| \leq \Gamma T_n\Delta_n^\psi$ for some constant Γ. Since $\zeta(R') \geq \varepsilon$ we then deduce from (A.3) with $a = 1/\varepsilon$ and $v = v_n$ as given by (24) and once more (A.2) that $\|\widehat{R}''_{n,m} - R''\| \leq Cv_nT_n^{3/2}\Delta_n^\psi$.

Since $R^{j,h'} = \widehat{R}^{j,h'}_{n,m} = 0$ for $j\in A'$, the jth column of R'' and $\widehat{R}''_{n,m}$ are vanishing for $j\in A'$ and thus

$$Y_n^{m,h} = \sum_{j\in A} \widehat{R}''^{h,j}_{n,m}\,\Delta_j^n P^m, \qquad \overline{\beta}^{J,m,h} = \sum_{j\in A} R''^{h,j}\sum_{h'=1}^H \overline{\beta}^{J,m,h'}\Delta\overline{F}^{h'}_{t_j},$$

where we have used $\Delta\overline{F}^{h'}_{t_j} = R^{j,h'}$ for $j\in A$ and the definition of R'' for the second equality above. Then, an application of (A.5) yields

$$|Y_n^{m,h} - \overline{\beta}^{J,m,h}| \leq \|\widehat{R}''_{n,m}\|\,\Delta_n^\psi + \Big|\sum_{j\in A}\big((\widehat{R}''_{n,m})^{h,j} - R''^{h,j}\big)\sum_{h'=1}^H \overline{\beta}^{J,m,h'}\Delta\overline{F}^{h'}_{t_j}\Big|.$$

Observe that $\|R''\| \leq C\sqrt{T_n}$, because $\|R'^{-1}\| \leq 1/\varepsilon$ and $\|R^\intercal\| \leq C\sqrt{T_n}$. Then $\|\widehat{R}''_{n,m} - R''\| \leq Cv_nT_n^{3/2}\Delta_n^\psi$ implies $\|\widehat{R}''_{n,m}\| \leq C\sqrt{T_n}$ as well, and we deduce $|Y_n^{m,h} - \overline{\beta}^{J,m,h}| \leq Cv_nT_n^{3/2}\Delta_n^\psi$. In other words, we have proved that $\omega\in\Omega_n^{m,h}$, and the proof is complete. □

A crucial remark is in order here: whereas $\widehat{\beta}^C_{n,i}$ is by construction measurable with respect to $\mathcal{F}_{(i+1)q_n\Delta_n}$, this is not the case of $\widehat{\beta}^J_{n,i}$. This induces some difficulties because the analysis of the limiting behavior of $\widehat{U}_n$, for example, strongly uses the fact that it is a kind of "discrete" stochastic integral with respect to the Itô semimartingale P, and having a non adapted integrand makes the analysis almost impossible to do. This is why we replace $\widehat{\eta}_{n,i}$ by $\mathcal{F}^n_{(i+1)q_n}$–measurable variables $\widehat{\eta}^*_{n,i}$, according to the following procedure:

$$\begin{gathered} \beta^*_{n,i} = (\widehat{\beta}^C_{n,i}, \beta^J_{iq_n\Delta_n}), \qquad b^*_{n,i} = (\beta^*_{n,i})^\intercal\,\beta^*_{n,i} \\ (b^{-1})^*_{n,i} = \begin{cases} (b^*_{n,i})^{-1} & \text{if } \zeta(b^*_{n,i}) > 1/v_n \\ 0 & \text{otherwise,} \end{cases} \qquad \eta^*_{n,i} = (b^{-1})^*_{n,i}\,(\beta^*_{n,i})^\intercal. \end{gathered} \tag{A.6}$$

With Y_t being any of the processes β_t, b_t, b_t^{-1}, η_t, we have the genuine estimators $\widehat{Y}_{n,i}$ as given by (30) and (31), and the "fake" ones $Y^*_{n,i}$ as given by (A.6). We set

$$\mathcal{E}(Y)^*_{n,i} = \widehat{Y}_{n,i} - Y^*_{n,i}, \qquad \mathcal{E}(Y)_{n,i} = Y^*_{n,i} - Y_{iq_n\Delta_n}, \tag{A.7}$$

and will now study the "error" $\mathcal{E}(\eta)^*_{n,i}$. To this effect, since we assume $H \geq 1$ here we have the second parts of (34) and (35) and we define $\psi \in (0, 1/2)$ and choose some $\varphi > 0$ as follows:

$$\psi = \frac{1}{2} - \frac{\varpi'}{10}, \qquad 0 < \varphi < 1 - 2\psi - \tau.$$

As easily checked, these imply for any $p > 0$:

$$T_n \Delta_n^{1-2\psi-\varphi} \to 0, \qquad v_n^p T_n^3 \Delta_n^{2\psi} \to 0, \qquad v_n^p T_n^4 \Big(\frac{u_n^8 \Delta_n^{2\psi-5}}{q_n} + u_n^{12} \Delta_n^{2\psi-6} \Big) \to 0. \tag{A.8}$$

Lemma 5. *In restriction to the set* $\Omega_n = \{\sup_{i \in I_n} \|\widehat{\beta}^J_{n,i} - \beta^J_{iq_n\Delta_n}\| \leq v_n^2 T_n^{3/2} \Delta_n^{\psi}\}$, *we have*

$$\|\mathcal{E}(\eta)^*_{n,i}\| \leq C z_n, \quad \textit{with } z_n = v_n^6 T_n^{3/2} u_n^4 \Delta_n^{\psi-2}.$$

Proof. For Y_t being any of β_t, b_t, b_t^{-1}, η_t, we consider the following property, for some sequence z_n:

$$\mathcal{P}^*_{z_n}: \quad \|\mathcal{E}(Y)^*_{n,i}\| \leq C z_n, \qquad \forall n \geq 1, \;\; i \in I_n, \;\; \omega \in \Omega_n.$$

Below, we argue ω-wise and fix $\omega \in \Omega_n$.

By the very definition of $\beta^*_{n,i}$ the process β_t satisfies $\mathcal{P}^*_{z_{n,1}}$ with $z_{n,1} = v_n^2 T_n^{3/2} \Delta_n^{\psi}$.

Next, $|\widehat{\gamma}^{m,k}_{n,i}| \leq u_n^2/\Delta_n$ by construction, so (29) yields $\|\widehat{\beta}^C_{n,i}\| \leq C v_n u_n^2/\Delta_n$, hence $\|\beta^*_{n,i}\| \leq C v_n u_n^2/\Delta_n$ as well because $u_n^2/\Delta_n \geq 1$ and β^J_t is bounded. Using (A.2), we deduce that b_t satisfies $(\mathcal{P}^*_{z_{n,2}})$ with $z_{n,2} = v_n u_n^2 z_{n,1}/\Delta_n + z_{n,1}^2$, smaller than $C v_n u_n^2 z_{n,1}/\Delta_n$ by (A.8).

If $\zeta(b^*_{n,i}) > 1/v_n$, Lemma 1 with $A = b^*_{n,i}$ and $A' = \widehat{b}_{n,i}$ and the property $(\mathcal{P}^*_{z_{n,2}})$ for b_t yield $\|\mathcal{E}(b^{-1})_{n,i}\| \leq z_{n,3} = 3 v_n^2 z_{n,2}$. If $\zeta(\widehat{b}_{n,i}) > 1/v_n$ we get the same inequality by exchanging $b^*_{n,i}$ and $\widehat{b}_{n,i}$, whereas if both $\zeta(b^*_{n,i}) \leq 1/v_n$ and $\zeta(\widehat{b}_{n,i}) \leq 1/v_n$ we obviously have $\mathcal{E}(b^{-1})_{n,i} = 0$. Therefore b_t^{-1} satisfies $(\mathcal{P}^*_{z_{n,3}})$, and we have $z_{n,3} \leq C v_n^3 u_n^2 z_{n,1}/\Delta_n$.

Finally, using again (A.2) and since by construction $\|(b^{-1})^*_{n,i}\| \leq v_n$, we deduce that η_t satisfies $(\mathcal{P}^*_{z'_n})$ with $z'_n = v_n z_{n,1} + v_n u_n^2 z_{n,3}/\Delta_n + z_{n,1} z_{n,3}$, smaller than $C v_n^4 u_n^4 z_{n,1}/\Delta_n^2$, hence the claim. ∎

With the variables $\eta^*_{n,i}$ we associate the processes

$$
\begin{aligned}
U^*_n &= \tfrac{1}{T_n} \sum_{i=0}^{[n/q_n]-2} \eta^*_{n,i} \left(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}\right) \\
U'^*_n &= \tfrac{q_n\Delta_n}{T_n} \sum_{i=0}^{[n/q_n]-2} \eta^*_{n,i}\, \overline{r}_{iq_n\Delta_n} \\
V^*_n &= \sum_{i=0}^{[n/q_n]-2} \eta^*_{n,i} \left(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}\right)\left(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}\right)^\intercal (\eta^*_{n,i})^\intercal.
\end{aligned}
\tag{A.9}
$$

Lemma 6. *We have*

$$
\sqrt{T_n}\, \|\widehat{U}_n - U^*_n\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad \sqrt{T_n}\, \|\widehat{U}'_n - U'^*_n\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad \frac{1}{\sqrt{T_n}}\, \|\widehat{V}_n - V^*_n\| \stackrel{\mathbb{P}}{\longrightarrow} 0.
$$

Proof. First, we have

$$
a_n := \sqrt{T_n}\,(\widehat{U}_n - U^*_n) = \frac{1}{\sqrt{T_n}} \sum_{i=0}^{[n/q_n]-2} \zeta^n_i, \quad \zeta^n_i = \mathcal{E}(\eta)^*_{n,i}\, \rho^n_i, \quad \rho^n_i = P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}.
$$

Then in restriction to the set Ω_n the previous lemma yields

$$
\|a_n\| \leq C\, \frac{z_n}{\sqrt{T_n}} \sum_{i=0}^{[n/q_n]-2} \|\rho^n_i\|.
$$

Since P is an Itô semimartingale with bounded spot characteristics and bounded jumps, we have $\mathbb{E}(\|\rho^n_i\|^2) \leq C q_n\Delta_n$. Therefore

$$
\mathbb{E}\big(\|a_n\|\, 1_{\Omega_n}\big) \leq C\, \frac{z_n}{\sqrt{T_n}}\, \frac{n}{q_n}\, \sqrt{q_n\Delta_n} \leq C\, \frac{z_n\sqrt{T_n}}{\sqrt{q_n\Delta_n}}.
$$

Secondly, we have

$$
a'_n := \frac{1}{\sqrt{T_n}}\,(\widehat{U}_n - U'^*_n) = \frac{1}{\sqrt{T_n}} \sum_{i=0}^{[n/q_n]-2} \zeta'^n_i, \qquad \zeta'^n_i = \mathcal{E}(\eta)^*_{n,i}\, \rho'^n_i, \quad \rho'^n_i = q_n\Delta_n \overline{r}_{iq_n\Delta_n}.
$$

Since $\overline{r}_t$ is bounded, using the previous lemma yields

$$
\mathbb{E}\big(\|a'_n\|\, 1_{\Omega_n}\big) \leq C\, \frac{z_n}{\sqrt{T_n}}\, \frac{n}{q_n}\, q_n\Delta_n \leq C\, z_n\, \sqrt{T_n}.
$$

Thirdly, we have

$$
a''_n := \frac{1}{\sqrt{T_n}}\,(\widehat{V}_n - V^*_n) = \frac{1}{\sqrt{T_n}} \sum_{i=0}^{[n/q_n]-2} \big(\zeta^n_i(\zeta^n_i)^\intercal + \eta^*_{n,i}\rho^n_i(\zeta^n_i)^\intercal + \zeta^n_i(\rho^n_i)^\intercal(\eta^*_{n,i})^\intercal\big).
$$

Recall $\|\beta^*_{n,i}\| \leq C v_n u_n^2/\Delta_n$ from the previous proof, so $\|\eta^*_{n,i}\| \leq C v_n^2 u_n^2/\Delta_n$ and we thus have in restriction to the set Ω_n:

$$\|a''_n\| \leq C\, \frac{z_n^2 + z_n v_n^2 u_n^2 \Delta_n^{-1}}{\sqrt{T_n}} \sum_{i=0}^{[n/q_n]-2} \|\rho_i^n\|^2,$$

implying

$$\mathbb{E}\big(\|a''_n\|\, 1_{\Omega_n}\big) \leq C\, \frac{z_n^2 + z_n v_n^2 u_n^2 \Delta_n^{-1}}{\sqrt{T_n}}\, \frac{n}{q_n}\, q_n \Delta_n \leq C\big(z_n^2 + z_n v_n^2 u_n^2 \Delta_n^{-1}\big) \sqrt{T_n}.$$

Now, we apply (A.8). First, together with Lemma 4, it implies $\mathbb{P}(\Omega_n) \to 1$. Second, it also implies

$$z_n^2 T_n\Big(\frac{1}{q_n \Delta_n} + 1 + z_n^2 + \frac{v_n^4 u_n^4}{\Delta_n^2}\Big) \to 0.$$

Therefore the three sequences a_n, a'_n, a''_n go to 0 in probability. □

A.3 Properties of U_n^*, $U_n'^*$ and V_n^*

From now on we no longer require $H \geq 1$. Then we have the first parts of (34) and (35) (which are in any case weaker than their second parts), and we easily see that, for any $p \geq 0$,

$$\text{with } \rho_n = q_n \Delta_n + \frac{1}{q_n}: \qquad v_n^p\Big(\rho_n^2 T_n + \frac{1}{q_n \Delta_n T_n}\Big) \to 0. \tag{A.10}$$

By virtue of Lemma 2, the processes η_t and $\eta_t \overline{r}_t$ satisfy (17), a property which will be essential in the sequel, on any interval $\mathcal{L}$ on which both β_t and $\overline{r}_t$ satisfy (17) themselves. With $\varepsilon > 0$ such that each $(\theta_m - \varepsilon)^+$ is a stopping time, we introduce the (predictable) sets, with $(a, a'] = \emptyset$ when $a \geq a'$:

$$D_m = (0, \zeta_m] \cup (\zeta_m + \varepsilon, \theta_m - \varepsilon] \cup (\theta_m, \infty), \qquad D = \cap_{1 \leq m \leq M} D_m.$$

Since D is, for each ω, the union of at most $3M$ disjoint intervals of $\mathbb{R}$, separated by intervals of size less than $M\varepsilon$, there is an increasing sequence $(\tau_j)_{0 \leq 1 \leq J}$ of stopping times with $\tau_0 = 0$ and $\tau_J = \infty$ and $\tau_j < \tau_{j+1}$ if $\tau_j < \infty$ for some (non random) integer $J \leq 3M$, such that:

$$\begin{aligned} & D = \textstyle\bigcup_{j=1}^J \mathcal{L}^{(j)}, \quad \text{where } \mathcal{L}^{(j)} = (\tau_{j-1}, \tau_j] \cap \mathbb{R}_+ \\ & \mathbb{R}_+ \backslash D \text{ is the union of at most } 3M \text{ intervals of length not bigger than } \varepsilon \\ & \text{Any } \mathcal{L}^{(j)} \text{ is contained, for all } m, \text{ in either } (0, \zeta_m] \text{ or } (\zeta_m + \varepsilon, \theta_m - \varepsilon] \text{ or } (\theta_m, \infty). \end{aligned} \tag{A.11}$$

Then, by the last property above, plus (P-1), Lemma 2 and the definition (14) for $\overline{r}_t$, we have that the processes

$$c_t^F,\ \gamma_t,\ \beta_t,\ \overline{r}_t,\ \eta_t,\ \eta_t \overline{r}_t \text{ satisfy (17) on each } \mathcal{L}'^{(j)} = (\tau_{j-1}, \tau_j + 2 q_n \Delta_n], \tag{A.12}$$

and the estimation errors are easily tracked only when the integer i belongs to the random set B_n given by

$$B_n = \bigcup_{j=1}^{J} B_n^j, \qquad B_n^j = \big\{i:\, 1 \le i \le [n/q_n]-2,\ iq_n\Delta_n \in \mathcal{L}^{(j)}\big\}. \tag{A.13}$$

We use the notation $\mathcal{E}(Y)_{n,i}$ of (A.6) for Y_t being one of the processes β_t, b_t, $(b_t)^{-1}$, η_t, and also when Y_t is c_t^F or γ_t or β_t^C, so by convention we write $(c^F)^*_{n,i} = \widehat{c}^F_{n,i}$ and $\gamma^*_{n,i} = \widehat{\gamma}_{n,i}$ and $(\beta^C)^*_{n,i} = \widehat{\beta}^C_{n,i}$. Then $\mathcal{E}(\beta)^{m,l}_{n,i} = \mathcal{E}(\beta^C)^{m,l}_{n,i}$ if $1 \le l \le K$, and $\mathcal{E}(\beta)^{m,l}_{n,i} = 0$ if $K < l \le K+H$.

Lemma 7. *We have*

$$i \in B_n \ \Rightarrow \qquad \|\,\mathbb{E}^n_{iq_n}(\mathcal{E}(\eta)_{n,i})\| \le C\,\rho_n v_n^4, \qquad \mathbb{E}^n_{iq_n}\big(\|\mathcal{E}(\eta)_{n,i}\|^2\big) \le C_p\,\rho_n v_n^8, \tag{A.14}$$

$$\mathbb{E}^n_{iq_n}(\|\eta^*_{n,i}\|^6) \le C v_n^{12}. \tag{A.15}$$

Proof. Since $\rho_n \to 0$, we can and will assume below $\rho_n \le 1$.

1) In a first step, we consider the $K+M$–dimensional Itô semimartingale $Z=(F,P)$ whose spot volatility is denoted by c_t, and its continuous part by $X_t = Z_t - \sum_{s\le t}\Delta Z_s$. We set

$$\widehat{c}^{l,l'}_{n,i} = \tfrac{1}{q_n\Delta_n}\sum_{j=1}^{q_n} \Delta^n_{iq_n+j}Z^l\,\Delta^n_{iq_n+j}Z^{l'}\,1_{\{|\Delta^n_{iq_n+j}Z^l|\le u_n,\ |\Delta^n_{iq_n+j}Z^{l'}|\le u_n\}}$$
$$\widehat{c}'^{l,l'}_{n,i} = \tfrac{1}{q_n\Delta_n}\sum_{j=1}^{q_n} \Delta^n_{iq_n+j}X^l\,\Delta^n_{iq_n+j}X^{l'}.$$

Those are estimators for $c^{l,l'}_{iq_n\Delta_n}$, although the second one is not feasible because X is not observed, and only serves us as a technical tool.

By Lemma B.5 of Aït-Sahalia and Jacod (2014) and $\alpha<1$ and since $1-2\varpi < \frac{1-\alpha}{32-\alpha}$ by (35), for $p \in [1,16]$ and any l,l', we have

$$\mathbb{E}^n_{iq_n}\big(|\widehat{c}^{l,l'}_{i,n} - \widehat{c}'^{l,l'}_{i,n}|^p\big) \le C_p\Delta_n^{(2p-\alpha)\varpi-p+1} \le C_p\sqrt{\Delta_n}. \tag{A.16}$$

2) Here, we estimate the difference $\widehat{c}'^{l,l'}_{n,i} - c^{l,l'}_{iq_n\Delta_n}$. Let $i \ge 0$ and l,l' be fixed, and for simplicity write $t = iq_n\Delta_n$ and often omit the index n. We have

$$\widehat{c}'^{l,l'}_{n,i} - c^{l,l'}_{iq_n\Delta_n} = \frac{1}{q_n\Delta_n}\sum_{j=1}^{q_n}\zeta_j, \qquad \zeta_j = \Delta^n_{iq_n+j}X^l\Delta^n_{iq_n+j}X^{l'} - \Delta_n c^{l,l'}_t.$$

Itô's formula yields $\zeta_j = \overline{\zeta}_j + \widetilde{\zeta}_j$, where, with $\mathcal{I}_j = (t+(j-1)\Delta_n, t+j\Delta_n]$,

$$\overline{\zeta}_j = \int_{\mathcal{I}_j}(X^l_s - X^l_{t+(j-1)\Delta_n})\,dX^{l'}_s + \int_{\mathcal{I}_j}(X^{l'}_t - X^{l'}_{t+(j-1)\Delta_n})\,dX^l_s, \qquad \widetilde{\zeta}_j = \int_{\mathcal{I}_j}(c^{l,l'}_s - c^{l,l'}_t)\,ds,$$

hence

$$\widehat{c}_{n,i}^{l,l'} - c_{iq_n\Delta_n}^{l,l'} = \overline{A} + \widetilde{A}, \quad \overline{A} = \frac{1}{q_n\Delta_n}\sum_{j=1}^{q_n}\overline{\zeta}_j, \quad \widetilde{A} = \frac{1}{q_n\Delta_n}\sum_{j=1}^{q_n}\overline{\zeta}_j = \frac{1}{q_n\Delta_n}\int_t^{t+q_n\Delta_n}(c_s^{l,l'} - c_t^{l,l'})\,ds.$$

On the one hand, standard computations using Burkholder-Davis-Gundy inequality and the boundedness of c_t and of the drift of X yield for any $p > 0$:

$$\big|\,\mathbb{E}^n_{iq_n+j-1}(\overline{\zeta}_j)\big| \leq C\Delta_n^2, \qquad \mathbb{E}^n_{iq_n+j-1}\big(|\overline{\zeta}_j|^p\big) \leq C_p\Delta_n^p,$$

and by another classical martingale argument, we deduce

$$\big|\,\mathbb{E}^n_{iq_n}(\overline{A})\big| \leq C\Delta_n, \qquad \mathbb{E}^n_{iq_n}\big(|\overline{A}|^p\big) \leq C_p(\Delta_n^{p/2} + q_n^{-p/2}) \leq C_p/q_n.$$

On the other hand, observe that if $1 \leq l' \leq K$ the variable $c_t^{l,l'}$ is $c_t^{F,l,l'}$ if $1 \leq l \leq K$ and $\gamma_t^{l-K,l'}$ if $K < l \leq K+M$. Therefore, if $l' \leq K$, the process $c_t^{l,l'}$ satisfies (17) on each interval $\mathcal{L}'^{(j)}$ by (A.12). This, the boundedness of c_t and the $\mathcal{F}^n_{iq_n}$–measurability of the set $\{i \in B_n\}$ yield for any $p \geq 2$:

$$\text{on the set } \{i \in B_n\}, \text{ if } l' \leq K: \qquad \big|\,\mathbb{E}^n_{iq_n}(\widetilde{A})\big| \leq Cq_n\Delta_n, \qquad \mathbb{E}^n_{iq_n}\big(|\widetilde{A}|^p\big) \leq C_pq_n\Delta_n.$$

Putting together the estimates for $\overline{A}$ and $\widetilde{A}$, plus (A.16), we deduce as soon as $l' \leq K$ and $p \in [2, 16]$:

$$\text{on the set } \{i \in B_n\}: \qquad \big|\,\mathbb{E}^n_{iq_n}(\widehat{c}_{n,i}^{l,l'} - c_{iq_n\Delta_n}^{l,l'})\big| \leq C\rho_n, \quad \mathbb{E}^n_{iq_n}\big(|\widehat{c}_{n,i}^{l,l'} - c_{iq_n\Delta_n}^{l,l'}|^p\big) \leq C_p\rho_n. \quad \text{(A.17)}$$

3) At this stage, the proof of (A.14) follows the same route as the proof of Lemma 5. For a process Y such as those described after (A.13) and any sequences $z_n, \overline{z}_n \geq 1$ and $p \geq 2$, we consider the property

$$(\mathcal{P}_{z_n,\overline{z}_n,p}): \quad \begin{cases} \text{on the set } \{i \in B_n\} \text{ and for all } p' \in [2,p] \text{ we have} \\ \|\,\mathbb{E}^n_{iq_n}(\mathcal{E}(Y)_{n,i})\| \leq C\,\rho_n z_n, \qquad \mathbb{E}^n_{iq_n}\big(\|\mathcal{E}(Y)_{n,i}\|^{p'}\big) \leq C_{p'}\,\rho_n\overline{z}_n^{p'}. \end{cases}$$

Upon using (A.2) and (A.3), we easily see the following: Let Y and Y' be two adapted bounded matrix–valued processes with the proper dimensions, satisfying $(\mathcal{P}_{z_n,\overline{z}_n,p})$ and $(\mathcal{P}_{z'_n,\overline{z}'_n,p'})$ respectively, and $Y^*_{n,i}$ and $Y'^*_{n,i}$ are estimators of $Y_{iq_n\Delta_n}$ and $Y'_{iq_n\Delta_n}$; we use below $(YY'^*)_{i,n} = Y^*_{n,i}Y'^*_{n,i}$ for estimating the product $(YY)'_{iq_n\Delta_n}$; in the second case below Y is $\mathcal{M}_d^+$–valued with a bounded inverse, and we use $(Y^{-1})^*_{i,n} = (Y^*_{n,i})^{-1}\,1_{\{\zeta(Y^*_{n,i}\geq 1/v_n\}}$ for estimating the inverse $Y^{-1}_{iq_n\Delta_n}$; then we have the following: If $p\wedge p' \geq 4$, the process YY' satisfies $(\mathcal{P}_{z''_n,\overline{z}''_n,p''})$ with $z''_n = \max(z_n, z'_n, \overline{z}_n\overline{z}'_n)$ and $\overline{z}''_n = \overline{z}_n\overline{z}'_n$ and $p'' = \frac{p\wedge p'}{2}$. And the process Y^{-1} satisfies $(\mathcal{P}_{z''_n,\overline{z}''_n,p''})$ with $z''_n = \max(z_n, v_n\overline{z}_n^2)$ and $\overline{z}''_n = v_n\overline{z}_n$ and $p'' = p$.

We have $(c^{F,k,k'})^*_{n,i} = \widehat{c}^{k,k'}_{n,i}$ and $(\gamma^{m,k})^*_{n,i} = \widehat{c}^{K+m,k}_{n,i}$ and (A.17) tells us that the two processes c^F_t and γ_t satisfy $(\mathcal{P}_{z_n,\overline{z}_n,p})$ with $p = 16$ and $z_n = \overline{z}_n = 1$. Then, what precedes shows us, successively, the following properties (for the second one, the property is obvious for β^C_t, and it extends to β by virtue of the fact stated after (A.13)): $(c^F_t)^{-1}$ satisfies $(\mathcal{P}_{v_n,v_n,16})$; β_t satisfies $(\mathcal{P}_{v_n,v_n,8})$; b_t satisfies $(\mathcal{P}_{v_n^2,v_n^2,4})$; $(b_t)^{-1}$ satisfies $(\mathcal{P}_{v_n^3,v_n^3,4})$; η_t satisfies $(\mathcal{P}_{v_n^4,v_n^4,2})$. The last property above is exactly (A.14).

4) It remains to show (A.15). Since c^F_t and γ_t are bounded, (A.17) with $p = 6$ yields $\mathbb{E}_{iq_n}(\|\widehat{c}^F_{n,i}\|^6 + \|\widehat{\gamma}_{n,i}\|^6) \leq C$. Thus (29) implies that $\mathbb{E}^n_{iq_n}(\|\widehat{\beta}^C_{n,i}\|^6) \leq Cv_n^6$, and the same holds for $\beta^*_{n,i}$ because β^J_t is bounded, and by (A.6) we readily deduce the claim. □

Lemma 8. *We have* $\sqrt{T_n}\,\|U^*_n - U_{T_n}\| \stackrel{\mathbb{P}}{\longrightarrow} 0$.

Proof. Recalling (A.13), we set $B'_n = \{0, 1, \ldots, [n/q_n] - 2\}\backslash B_n$. We have $\sqrt{T_n}(U^*_n - U_{T_n}) = \sum_{l=1}^5 G^l_n$, where

$$\begin{gathered}
\zeta^n_i = \mathcal{E}(\eta)_{n,i}(P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}), \quad G^1_n = \tfrac{1}{\sqrt{T_n}} \sum_{i\in B_n} \zeta^n_i, \quad G^2_n = \tfrac{1}{\sqrt{T_n}} \sum_{i\in B'_n} \zeta^n_i \\
\xi^n_i = \int_{(i+1)q_n\Delta_n}^{(i+2)q_n\Delta_n} \big(\eta_{iq_n\Delta_n} - \eta_s\big)\, dP_s, \quad G^3_n = \tfrac{1}{\sqrt{T_n}} \sum_{i\in B_n} \xi^n_i, \quad G^4_n = \tfrac{1}{\sqrt{T_n}} \sum_{i\in B'_n} \xi^n_i \\
A_n = \big(([n/q_n] - 2)q_n\Delta_n, T_n\big], \quad G^5_n = -\tfrac{1}{\sqrt{T_n}} \int_{A_n} \eta_s\, dP_s.
\end{gathered}$$

The set B_n is random, but $\{i \in B_n\} \in \mathcal{F}^n_{iq_n}$, hence $\mathcal{E}(\eta)_{n,i}\, 1_{\{i\in B_n\}}$ is $\mathcal{F}^n_{(i+1)q_n}$–measurable. Therefore, taking advantage of (P-1) and (A.14), we have

$$\|\mathbb{E}^n_{iq_n}(\zeta^n_i)\| \leq C\,\rho_n v_n^4 q_n\Delta_n, \quad \mathbb{E}^n_{iq_n}(\|\zeta^n_i\|^2) \leq C\,\rho_n v_n^8 q_n\Delta_n \ \text{ on the set } \{i \in B_n\}. \tag{A.18}$$

Using the $\mathcal{F}_{(i+2)q_n\Delta_n}$–measurability of $\zeta^n_i\, 1_{B_n}(i)$ and decomposing G^1_n into the sum for all i even and the sum for all i odd, plus the fact that $\#(B_n) \leq n/q_n$, we deduce by a classical martingale argument and (24), plus (A.10), that

$$\mathbb{E}(\|G^1_n\|) \leq \frac{C}{\sqrt{T_n}}\Big(\frac{n}{q_n}\,\rho_n v_n^4 q_n\Delta_n + \sqrt{\frac{n}{q_n}}\,\sqrt{\rho_n v_n^8 q_n\Delta_n}\Big) \leq C v_n^4\,\big(\rho_n\sqrt{T_n} + \sqrt{\rho_n}\big) \to 0.$$

Next, (A.15) and $\{i \in B'_n\} \in \mathcal{F}^n_{iq_n}$, plus the property $\mathbb{E}^n_{(i+1)q_n}\big(\|P_{(i+2)q_n\Delta_n} - P_{(i+1)q_n\Delta_n}\|^2\big) \leq Cq_n\Delta_n$, imply by successive conditioning and the Cauchy-Schwarz inequality that

$$\mathbb{E}(\|\zeta^n_i\|\, 1_{B'_n}(i)) \leq C\, v_n^2\sqrt{q_n\Delta_n}\ \mathbb{P}(i \in B'_n). \tag{A.19}$$

By (A.11) we have $\#(B'_n) \leq 3M\varepsilon/q_n\Delta_n$, hence we have by (24) and (A.10) again:

$$\mathbb{E}\big(\|G^2_n\|\big) \leq C\, v_n^2\frac{\sqrt{q_n\Delta_n}}{\sqrt{T_n}}\,\mathbb{E}\Big(\sum_{i\geq 0} 1_{B'_n}(i)\Big) \leq C\,\frac{v_n^2}{\sqrt{q_n\Delta_n T_n}} \to 0.$$

For G_n^3, we recall (A.12) and observe that, on the $\mathcal{F}_{iq_n}^n$–measurable set $\{i\in B_n^j\}$, the interval $((i{+}1)q_n\Delta_n,(i{+}2)q_n\Delta_n]$ is contained in $\mathcal{L}'^{(j)}$. Upon writing P as in (20) and since the predictable quadratic variation $\overline{V}_t$ of any component of P^{Mart} is such that $at-\overline{V}_t$ is increasing, for some constant a, we easily get

$$\|\mathbb{E}_{iq_n}^n(\xi_i^n)\|+\mathbb{E}_{iq_n}^n(\|\xi_i^n\|^2)\leq C(q_n\Delta_n)^2$$

on each set $\{i\in B_n^j\}$, hence on the set $\{i\in B_n\}$ as well. Then, exactly as for G_n^1 above, by a martingale argument we deduce

$$\mathbb{E}(\|G_n^3\|)\leq C\left(q_n\Delta_n\sqrt{T_n}+\sqrt{q_n\Delta_n}\right)\to 0. \tag{A.20}$$

Next, using only the boundedness of η_t, for any $p\geq 2$ we have $\mathbb{E}_{iq_n}^n(\|\xi_i^n\|)\leq C\sqrt{q_n\Delta_n}$. Then, exactly as for G_n^2, we obtain

$$\mathbb{E}(\|G_n^4\|)\leq C\,\frac{1}{\sqrt{q_n\Delta_nT_n}}\to 0. \tag{A.21}$$

Finally, A_n being a non random interval with length smaller than $3q_n\Delta_n$, so

$$\mathbb{E}(\|G_n^5\|)\leq C\,\frac{\sqrt{q_n\Delta_n}}{\sqrt{T_n}}\to 0 \tag{A.22}$$

by the boundedness of η_t again. This completes the proof. □

Lemma 9. *We have* $\sqrt{T_n}\,\|U_n'^*-U_{T_n}'\|\xrightarrow{\mathbb{P}}0.$

Proof. We have $\sqrt{T_n}(\widehat{U}_n'-U_{T_n}')=\sum_{j=1}^5 G_n'^j$, where (with B_n' and A_n as in the previous proof):

$$\begin{gathered}
\zeta_i'^n=q_n\Delta_n\mathcal{E}(\eta)_{n,i}\,\overline{r}_{iq_n\Delta_n},\quad G_n'^1=\tfrac{1}{\sqrt{T_n}}\sum_{i\in B_n}\zeta_i'^n,\quad G_n'^2=\tfrac{1}{\sqrt{T_n}}\sum_{i\in B_n'}\zeta_i'^n\\
\xi_i'^n=\int_{(i+1)q_n\Delta_n}^{(i+2)q_n\Delta_n}\left(\eta_{iq_n\Delta_n}\overline{r}_{iq_n\Delta_n}-\eta_s\overline{r}_s\right)ds,\quad G_n'^3=\tfrac{1}{\sqrt{T_n}}\sum_{i\in B_n}\xi_i'^n,\\
G_n'^4=\tfrac{1}{\sqrt{T_n}}\sum_{i\in B_n'}\xi_i'^n,\quad G_n'^5=-\tfrac{1}{\sqrt{T_n}}\int_{A_n}\eta_s\overline{r}_s\,ds.
\end{gathered}$$

This is exactly the same as the decomposition $\sqrt{T_n}(\widehat{U}_n-U_{T_n})=\sum_{j=1}^5 G_n^j$ of the previous proof, except that we replace η_t by $\eta_t\overline{r}_t$ and $\mathcal{E}(\eta)_{n,i}$ by $\mathcal{E}(\eta)_{n,i}\overline{r}_{iq_n\Delta_n}$, and the M–dimensional semimartingale P by the one-dimensional "process" t which is again a (continuous) Itô semimartingale with bounded spot characteristics.

Since $\overline{r}_{iq_n\Delta_n}$ is $\mathcal{F}_{iq_n}^n$–measurable and bounded, the variables $\zeta_i'^n$ satisfy (A.18) and (A.19) (actually, much sharper bounds would be available in this case). Then, one can reproduce word for word the previous proof, to get the same bounds for the expectations $\mathbb{E}(\|G_n'^j\|)$, and the claim follows. □

Lemma 10. *We have* $\frac{1}{\sqrt{T_n}}\,\|V_n^*-V_{T_n}\|\xrightarrow{\mathbb{P}}0.$

Proof. It suffices to show that, for any fixed $l, l' \in \{1, \dots, K+H\}$ and $m, m' \in \{1, \dots, M\}$, we have

$$\overline{V}_n := \frac{1}{\sqrt{T_n}} \sum_{i=0}^{[n/q_n]-2} \eta_{n,i}^{*l,m} \eta_{n,i}^{*l',m'} P_i^{n,m} P_i^{n,m'} - \frac{1}{\sqrt{T_n}} \int_0^{T_n} \eta_t^{l,m} \eta_t^{l',m'} \, d[P^m, P^{m'}]_t \stackrel{\mathbb{P}}{\longrightarrow} 0$$
$$\text{where} \quad P_i^{n,m} = P^m_{(i+2)q_n\Delta_n} - P^m_{(i+1)q_n\Delta_n}.$$

We have the decomposition $\overline{V}_n = \sum_{l=1}^5 \overline{G}_n^l$, where (with B'_n and A_n as in the proof of Lemma 8):

$$\overline{\zeta}_i^n = \big(\mathcal{E}(\eta)_{n,i}^{l,m} \mathcal{E}(\eta)_{n,i}^{l',m'} + \mathcal{E}(\eta)_{n,i}^{l,m} \, \eta_{iq_n\Delta_n}^{l',m'} + \eta_{iq_n\Delta_n}^{l,m} \mathcal{E}(\eta)_{n,i}^{l',m'}\big) P_i^{n,m} P_i^{n,m'}$$
$$\overline{G}_n^1 = \frac{1}{\sqrt{T_n}} \sum_{i \in B_n} \overline{\zeta}_i^n, \quad \overline{G}_n^2 = \frac{1}{\sqrt{T_n}} \sum_{i \in B'_n} \overline{\zeta}_i^n$$
$$\overline{\xi}_i^n = \int_{(i+1)q_n\Delta_n}^{(i+2)q_n\Delta_n} \big(\eta_{iq_n\Delta_n}^{l,m} \eta_{iq_n\Delta_n}^{l',m'} - \eta_s^{l,m} \eta_s^{l',m'}\big) \, d[P^m, P^{m'}]_s, \quad \overline{G}_n^3 = \frac{1}{\sqrt{T_n}} \sum_{i \in B_n} \overline{\xi}_i^n$$
$$\overline{G}_n^4 = \frac{1}{\sqrt{T_n}} \sum_{i \in B'_n} \overline{\xi}_i^n, \quad \overline{G}_n^5 = -\frac{1}{T_n} \int_{A_n} \eta_s^{l,m} \eta_s^{k',m'} \, d[P^m, P^{m'}]_s.$$

Here again, this decomposition is the same as $\sqrt{T_n}(U_n^* - U_{T_n}) = \sum_{j=1}^5 G_n^j$ in the proof of Lemma 8, with the following changes:

- First, for $\overline{G}_n^1$ and $\overline{G}_n^2$: we replace ζ_i^n by $\overline{\zeta}_i^n$. Combining (A.14) and (A.15), we see that on the set $\{i \in B_n\}$:

$$\mathbb{E}_{iq_n}^n\big(\|\mathcal{E}(\eta)_{n,i}\|^4\big) \leq \sqrt{\mathbb{E}_{iq_n}^n\big(\|\mathcal{E}(\eta)_{n,i}\|^2\big) \, \mathbb{E}_{iq_n}^n\big(\|\mathcal{E}(\eta)_{n,i}\|^6\big)} \leq C v_n^{10} \sqrt{\rho_n}.$$

Since $\mathbb{E}_{(i+1)q_n}^n(|P_i^{n,m}|^p) \leq C q_n \Delta_n$ for $p = 2, 4$, by successive conditioning and again (A.14) and (A.15) we see that (A.18) and (A.19) should be replaced with $\mathbb{E}\big(|\overline{\zeta}_i^n| \, 1_{B'_n}(i)\big) \leq C v_n^4 q_n \Delta_n \, \mathbb{P}(i \in B'_n)$, and on the set $\{i \in B_n\}$:

$$\begin{cases} |\mathbb{E}_{iq_n}^n(\overline{\zeta}_i^n)| \leq C v_n^8 q_n \Delta_n \rho_n \\ \mathbb{E}_{iq_n}^n(|\overline{\zeta}_i^n|^2) \leq C v_n^{10} q_n \Delta_n \sqrt{\rho_n}. \end{cases}$$

Then, exactly the same proof as in Lemma 8 gives, under (24):

$$\mathbb{E}(|\overline{G}_n^1|) \leq C v_n^8 \big(\rho_n \sqrt{T_n} + \rho_n^{1/4}\big) \to 0, \qquad \mathbb{E}\big(|\overline{G}_n^2|\big) \leq C \, \frac{v_n^4}{\sqrt{q_n \Delta_n T_n}} \to 0.$$

- Second, for $\overline{G}_n^3$, $\overline{G}_n^4$ and $\overline{G}_n^5$: we replace η_t by $\eta_t^{k,m} \eta_t^{k',m'}$ and P by $Y = [P^m, P^{m'}]$, which is a process of (locally) finite variation. We have $\mathbb{E}\big(\int_t^{t+s} |dY_s|\big) \leq Cs$, so $\overline{G}_n^4$ and $\overline{G}_n^5$ enjoy the bound (A.21) and (A.22) (indeed, sharper estimates are available here). Since further $Y_t = M_t + \int_0^t a_s \, ds$ for some bounded process a_t and a martingale M_t such that $a' t - \langle M, M \rangle_t$ is increasing for some constant a', we also obtain that $\overline{G}_n^3$ satisfies (A.20), and the proof is complete. □

A.4 Proof of Theorem 1

Recalling (22), we readily deduce from Assumption 5 and the boundedness of η_t, plus Lemmas 6, 8 and 9, that

$$\sqrt{T_n}(\widehat{\Lambda}_n - \Lambda_{T_n}) - \frac{1}{\sqrt{T_n}}\,\overline{U}_{T_n} \xrightarrow{\mathbb{P}} 0, \tag{A.23}$$

as $n \to \infty$. Note also that our assumptions imply that the martingale $\overline{U}_t$ has bounded jumps and that its quadratic variation process V given component-wise by (23) satisfies, for some constant $C_0 > 1$:

$$\frac{t}{C_0} \le \zeta(V_t) \le \|V_t\|, \qquad \mathbb{E}(\|V_t\|) \le C_0 t. \tag{A.24}$$

This implies in particular that $\mathbb{E}\big(\frac{1}{T_n}\,\overline{U}^2_{T_n}\big) \le C_0$, yielding that the sequence $\frac{1}{\sqrt{T_n}}\,\overline{U}_{T_n}$ is bounded in probability. Then (a) of Theorem 1 follows from (A.23).

Let us now turn to (b). The additional assumption that the variables $(1/t)V_t$ converge in probability to a limit V_∞ and (A.24) imply that all hypotheses of Corollary 2.3 of Crimaldi and Pratelli (2005) are satisfied (with $a_t = (1/\sqrt{t})\,I_{K+H}$ and $A = \Omega$ in this corollary), so we have the following multivariate CLT:

$$(V_{T_n})^{-1/2}\,\overline{U}_{T_n} \text{ converges in law to } \mathcal{N}(0, I_{K+H}). \tag{A.25}$$

Therefore, it is enough to show the existence of subsets $\Omega_n \subset \Omega$ satisfying $\mathbb{P}(\Omega_n) \to 1$,

- (i) $(\widehat{V}_n)^{-1}$ exists on Ω_n
- (ii) $T_n(\widehat{V}_n)^{-1}\,1_{\Omega_n}$ and $\frac{1}{T_n}\,\widehat{V}_n\,1_{\Omega_n}$ are bounded in probability
- (iii) $G_n := (\widehat{V}_n)^{-1/2}(V_{T_n})^{1/2}$ converges in probability to I_{K+H}, in restriction to Ω_n.

Without loss of generality, we can and will suppose $T_n \ge 1$. By Lemmas 6 and 10, the sequence $\xi_n = \frac{1}{\sqrt{T_n}}\,\|\widehat{V}_n - V_{T_n}\|$ goes to 0 in probability, implying that the set $\Omega_n = \{\xi_n \le 1/2C_0\}$ satisfies $\mathbb{P}(\Omega_n) \to 1$. Observing that for any two matrices A, B in $\mathcal{M}^+_{K+H}$ we have $\zeta(A) \ge \zeta(B) - \|A - B\|$, and in view of (A.24) we see that, on the set Ω_n,

$$\zeta(\widehat{V}_n) \ge \zeta(V_{T_n}) - \frac{\sqrt{T_n}}{2C_0} \ge \frac{T_n}{2C_0}.$$

Thus on Ω_n the inverse $(\widehat{V}_n)^{-1}$ exists and $T_n\|(\widehat{V}_n)^{-1}\| \le 2C_0$, whereas $\frac{1}{T_n}\mathbb{E}\big(\|\widehat{V}_n\|\,1_{\Omega_n}\big) \le C_0 + 1$ by (A.24) again. We thus have (i) and (ii) above.

Finally, for any $A \in \mathcal{M}^+_{K+H}$ with $\zeta(A) \ge 1/2C_0$ we have $A^{-1/2} = f(A)$ and $A^{1/2} = g(A)$ for some continuous $\mathcal{M}^+_{K+H}$–valued functions f and g on $\mathcal{M}^+_{K+H}$, so on Ω_n we have $G_n = f(\widehat{V}_n/T_n)g(V_{T_n}/T_n)$. Since $\widehat{V}_n/T_n - V_{T_n}/T_n \xrightarrow{\mathbb{P}} 0$ and V_{T_n}/T_n is bounded in probability, we deduce (iii), and the proof of Theorem 1 is complete.