Quantitative Finance

The times change: multivariate subordination. Empirical facts

Nicolas Huth [a] & Fréédééric Abergel [b]

[a] Natixis Investment Bank, Equity Derivatives and Arbitrage, Paris, France

[b] BNP Paribas Chair of Quantitative Finance, Ecole Centrale Paris, MAS Laboratory, Chââtenay Malabry, France

Available online: 19 Oct 2010

PLEASE SCROLL DOWN FOR ARTICLE

The times change: multivariate subordination. Empirical facts

NICOLAS HUTH† and FRÉDÉRIC ABERGEL*‡

†Natixis Investment Bank, Equity Derivatives and Arbitrage, Paris, France
‡BNP Paribas Chair of Quantitative Finance, Ecole Centrale Paris, MAS Laboratory,
Châtenay Malabry, France

The management of correlation risk is of the utmost importance in several areas of investment banking: multi-asset derivatives pricing and hedging, optimal asset allocation, risk management, statistical arbitrage and many others.§ However, the modeling of correlation as a time-dependent quantity—as opposed to that of the volatility, say—is still in its infancy. In this article, we present a mechanism for the presence of a stochastic covariance matrix in financial markets, and provide an explanation for the returns' heavy tailed joint distribution. This mechanism relies on sampling returns according to an appropriate event time. Using CAC 40 high-frequency data, we compare the predictions of our model with real-world statistics and find very good agreement.

Keywords: Correlation modelling; Econophysics; Equity options; Quantitative finance

1. Introduction

Since the celebrated Black–Scholes model started its career as a benchmark model in derivatives houses, many attempts have been made to check whether its assumptions are in agreement with real-world data. Among those assumptions, the most heavily rejected is that returns in calendar time are normally distributed, or, in other words, that volatility does not fluctuate randomly. Indeed, non-Gaussian tails (such as power-law or exponential tails) are well documented in the empirical literature (Bouchaud and Potters 2004).

However, one is prone to ask whether, as studied in the pioneering work of Clark (1973), returns may indeed be normally distributed *modulo* a stochastic change of time. In the one-dimensional case considered by Clark and several other authors thereafter (see, e.g., Ane and Geman 1996 and Carr *et al.* 2002), a stochastic clock is introduced, the so-called business time, speeding up the usual calendar time as market activity rushes. Therefore, the random intensity of the market activity flow, through

the arrival of trades, provides an explanation for the fluctuating aspect of volatility. In a more mathematical formulation: Returns are then driven by a subordinated Brownian motion, the subordinator being the business time, leading to heavy-tailed¶ distributions. Moreover, the degree of heaviness is easily seen to depend on the distribution of arrival times. It is noteworthy that, to the best of our knowledge, no such approach has previously been applied to the multi-dimensional case, and the aim of the work presented in this paper is precisely to do so. Inspired by a multidimensional version of the Central Limit Theorem, we introduce a suitable, rather natural multidimensional business time and study some of its elementary properties. We then determine the influence of this stochastic change of time on the joint distribution of returns, our main objective being to introduce a finer-grain, more accurate description of the origin of covariance between pairs of assets. Finally, we present an analytically tractable model that may be viewed as a simple mathematical tool possessing statistical properties similar to those of the multidimensional

---

trade time. Both the statistical estimates and the models that we present in this note are new, and we strongly believe that the use of multidimensional trade time for the modeling of correlated assets at the microstructure level is prone to many interesting new developments, some of which may already be found in the work of Huth (2010) and Muni Toke (2010).

The paper is organized as follows. In section 1 we provide a very simple theoretical background supporting Clark's model as well as the empirical and statistical studies presented by Christian Silva (2005) and extend it to the multivariate case. After a brief description, in section 2, of the data sets we use, we apply our approach to high-frequency data in section 3, unveiling a strikingly convincing agreement. Section 4 states our temporary conclusions as well as questions for future research, some of which are addressed by Huth (2010) and Muni Toke (2010).

2. Theoretical framework

2.1. Univariate case

Let us consider an asset whose price fluctuates randomly during trading hours. Between the $(i-1)$th and the ith trade, the price increment is simply $\delta P_i := P_i - P_{i-1}$. Then, N trades away from the opening price P_0, the total variation is given by the sum of these elementary increments

$$P_N - P_0 = \sum_{i=1}^{N} \delta P_i.$$

The elementary price increment can itself be decomposed into two parts, its sign and magnitude: $\delta P_i = \mathrm{sgn}(\delta P_i)|\delta P_i|$. In the simplest description of the price evolution in trade time, the sign of the increment is given by that of the corresponding market order, whether it is a buy or sell order, and its magnitude is then a function of the volume of the trade and of the available liquidity in the order book just before the trade. It is well known, and well documented in the literature (see, e.g., Bouchaud *et al.* 2009) that such a simplistic explanation may not always be accurate, but the mathematical decomposition of the price increment as above is obviously always legitimate. We are then left with the following expression for the relative price increment, i.e. the asset return:

$$R_N := \frac{P_N - P_0}{P_0} = \frac{1}{P_0} \sum_{i=1}^{N} \mathrm{sgn}(\delta P_i)|\delta P_i|.$$

The asset return is clearly the sum of random variables and we would like to apply a version of the Central Limit Theorem (CLT). The basic CLT is stated for independent and identically distributed random variables, a property which clearly fails to hold if one considers the returns of a

financial asset: it has been well documented, and is easily verified experimentally, that absolute values of returns are autocorrelated (Bouchaud *et al.* 2003). However, the CLT can be extended to the more general case of dependent variables $X_1, \ldots, X_n$ (Whitt 2002). The main condition for it to hold is the existence of the asymptotic variance

$$\lim_{n \to +\infty} \mathbb{V}\mathrm{ar}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i\right) = \mathbb{V}\mathrm{ar}(X_1) + 2 \sum_{k=1}^{+\infty} \mathbb{C}\mathrm{ov}(X_1, X_{1+k}),$$

assuming the same distribution for all X_i.

If the sum above is finite, i.e. if the autocorrelation function of δP_i decays fast enough,[†] then the CLT yields, as $N \to +\infty$,

$$\frac{R_N}{\sqrt{N}} \xrightarrow{d} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 := \lim_{N \to +\infty} \mathbb{V}\mathrm{ar}((1/\sqrt{N}P_0) \sum_{i=1}^{N} \delta P_i)$ and we have assumed[‡] $\mathbb{E}(\delta P_1) \approx 0$. Hence, there holds, for $N \sim \infty$,

$$R_N \sim \mathcal{N}(0, N\sigma^2),$$

i.e. returns are asymptotically normally distributed when they are sampled in trade time. Recast in the context of stochastic processes, the returns can therefore be viewed as a Brownian motion in a stochastic clock (such processes are called subordinated Brownian motions), the clock being the number of trades.

Now, the number of trades over a time period is obviously random. Therefore, the Δt-return $R_{\Delta t} := (P_{\Delta t} - P_0)/P_0$ in calendar time exhibits a random variance $N_{\Delta t}\sigma^2$ where $N_{\Delta t}$ is the number of trades occurring during a time period of length Δt. The distribution of calendar time returns $R_{\Delta t}$ can be recovered from trade time returns $R_{N_{\Delta t}}$ through the application of Bayes' formula[§]

$$P_{R_{\Delta t}}(r) = \int_0^{+\infty} P_{\mathcal{N}(0, n\sigma^2)}(r) P_{N_{\Delta t}}(n) \, \mathrm{d}n,$$

where $P_X(x)$ is the probability density function of X. As mentioned in the introduction, this mechanism has been extensively studied in the finance literature (Christian Silva 2005), but only in the univariate framework.

Slightly rephrasing the equations above, one has that $R_{\Delta t} = \sigma \sqrt{N_{\Delta t}}.X$ in distribution, where $X \sim \mathcal{N}(0, 1)$ independently from the value of $N_{\Delta t}$ and σ is a scaling constant. In order to derive the final expression for $P_{R_{\Delta t}}$, $P_{N_{\Delta t}}$ must be specified. Nevertheless, some model-free properties can be established, such as the computation of moments. Defining $\lambda_k(X) := \mathbb{E}(X^k)/\mathbb{E}(X^2)^{k/2}$ the kth dimensionless moment of X, there holds

$$\begin{aligned}
\mathbb{E}(R_{\Delta t}^k) &= \mathbb{E}(N_{\Delta t}^{k/2} \mathbb{E}(X^k \mid N_{\Delta t})) \\
&= \sigma^k \mathbb{E}(N_{\Delta t}^{k/2}) \mathbb{E}(\mathcal{N}(0, 1)^{k/2}) \\
&= \sigma^k \mathbb{E}(N_{\Delta t}^{k/2}) \left(\frac{k}{2}\right)!! \, \mathbb{1}_{\{\exists m \in \mathbb{N}^*: k=2m\}},
\end{aligned}$$

[†]In the case of price differences δP_i, the autocorrelation function decays avery fast and can be considered as statistically insignificant after some lag k close to one, even for small time scales (Abergel *et al.* 2009). Therefore, the above sum may reasonably be assured to converge.
[‡]It is a reasonable assumption since we are dealing with high-frequency data.
[§]We assume that the distribution of the trading activity $N_{\Delta t}$ can be approximated by a continuous distribution.

leading to

$$\lambda_k(R_{\Delta t}) = \frac{\mathbb{E}(N_{\Delta t}^{k/2})}{\mathbb{E}(N_{\Delta t})^{k/2}} \left(\frac{k}{2}\right)!! \, \mathbb{1}_{\{\exists m \in \mathbb{N}^*: k=2m\}},$$

where $n!! := 1 \times 3 \times \cdots \times (n-1)$ for n even (see appendix A). For $k = 4$, we obtain

$$\lambda_4(R_{\Delta t}) = 3 \frac{\mathbb{E}(N_{\Delta t}^2)}{\mathbb{E}(N_{\Delta t})^2} \geq 3 = \lambda_4(\mathcal{N}(0,1)),$$

thanks to Jensen's inequality.† As the fourth moment provides information on the heaviness of the tails, we see that such a model predicts distribution tails fatter than those of the normal distribution for $R_{\Delta t}$. Therefore, stochastic volatility, *via* the random fluctuations of the number of trades, provides an explanation for the fat tails of returns sampled in calendar time.‡

Moreover, the shape of the tails of $P_{R_{\Delta t}}$ is given by the asymptotic behavior of $P_{N_{\Delta t}}$. For instance, it can be shown that if $N_{\Delta t}$ obeys an inverse gamma (respectively gamma) distribution, which has heavy (respectively semi-heavy) tails, then $P_{R_{\Delta t}}$ is a Student (respectively Variance Gamma) law (Carr *et al.* 1998, Bouchaud and Potters 2004), therefore exhibiting heavy (respectively semi-heavy) tails. In fact, the tail behavior of $P_{N_{\Delta t}}$ propagates over to $P_{R_{\Delta t}}$ (see appendix B).

2.2. Multivariate case

We now turn to the more interesting case of $d \geq 2$ assets. How can we extend the previous framework to take several assets into account? Let us assume that an event time N is defined, and that we sample returns $R_N = (R_N^1, \ldots, R_N^d)$ according to this time. Using a multivariate CLT, we follow the same line of reasoning as in the previous section and obtain for $N \sim \infty$

$$R_N \sim \mathcal{N}(0, N\Sigma),$$

where§ $\Sigma = \lim_{N \to +\infty} \mathbb{V}\mathrm{ar}((1/\sqrt{N})\mathrm{diag}(P_0)^{-1} \sum_{i=1}^{N} \delta P_i)$ is now the covariance matrix.

For a given time interval Δt, the respective numbers of trades for each asset $N_{\Delta t}^1, \ldots, N_{\Delta t}^d$ are obviously different, and we need to define a global event time $N_{\Delta t} = f(N_{\Delta t}^1, \ldots, N_{\Delta t}^d)$. We suggest using $N_{\Delta t} = \sum_{i=1}^{d} N_{\Delta t}^i$,

which amounts to increment time as soon as a trade occurs on any one of the d assets. This choice seems to us the simplest and most intuitive generalization of the univariate case, since it amounts to considering a single asset with dimension d: we aggregate the time series of the returns of each asset in chronological order, and then count trades as in the univariate case.¶

In many areas of the finance industry—derivatives pricing, asset management, risk management, statistical arbitrage, etc.—the (calendar time) correlation matrix arises as a fundamental input in the determination of profit and loss. It is then clearly important to determine whether the sampling in trade time has an effect on the correlation matrix itself. Using our model, we can easily recover the calendar time covariance matrix by Bayes' calculus. Since $R_{\Delta t} = \Sigma^{1/2} \sqrt{N_{\Delta t}} X$ in distribution, with $X \sim \mathcal{N}_d(0, I_d)$ independently from the value of $N_{\Delta t}$, we obtain

$$\begin{aligned}
\mathbb{E}(R_{\Delta t} R_{\Delta t}') &= \mathbb{E}(N_{\Delta t} \Sigma^{1/2} XX' (\Sigma^{1/2})') \\
&= \Sigma^{1/2} \mathbb{E}(N_{\Delta t} \mathbb{E}(XX' \mid N_{\Delta t}))(\Sigma^{1/2})' \\
&= \mathbb{E}(N_{\Delta t}) \Sigma,
\end{aligned}$$

which implies that the calendar time correlation $\rho_{\Delta t}^{ij}$ between asset i and j is given by

$$\begin{aligned}
\rho_{\Delta t}^{ij} &:= \frac{\mathbb{C}\mathrm{ov}(R_{\Delta t}^i, R_{\Delta t}^j)}{\sqrt{\mathbb{V}\mathrm{ar}(R_{\Delta t}^i)\mathbb{V}\mathrm{ar}(R_{\Delta t}^i)}} \\
&= \frac{\mathbb{E}(N_{\Delta t})\Sigma_{ij}}{\sqrt{\mathbb{E}(N_{\Delta t})\Sigma_{ii}\mathbb{E}(N_{\Delta t})\Sigma_{jj}}} \\
&= \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}} \\
&= \frac{\Sigma_{ij} N_{\Delta t}}{\sqrt{\Sigma_{ii} N_{\Delta t} \Sigma_{jj} N_{\Delta t}}} \\
&= \rho_{N_{\Delta t}}^{ij} = \rho^{ij}.
\end{aligned}$$

Therefore, the calendar time correlation is equal to the trade time correlation, and does not depend on the time interval length Δt. We will check this property in the empirical part of our study below.

---

†The equality is reached iff $N_{\Delta t}$ is almost surely constant.

‡As pointed out by Gillemot *et al.* (2006), there might be other factors explaining the volatility of financial markets on a microscopic time scale, such as the shape of the order book. However, the number of trades tends to be a reliable proxy for variance and covariance according to Christian Silva (2005) and our study.

§We note

$$\mathrm{diag}(P_0) := \begin{pmatrix} P_0^1 & 0 & \ldots & 0 \\ 0 & P_0^2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & P_0^d \end{pmatrix}.$$

¶Such a representation makes sense when the orders of magnitude of the liquidity of each of the d assets are similar: consider the extreme case of one heavily traded asset and another one with only one trade per day. In this case, the normality of the couple as $N_{\Delta t}$ may fail to appear before an unrealistically large number of trades.
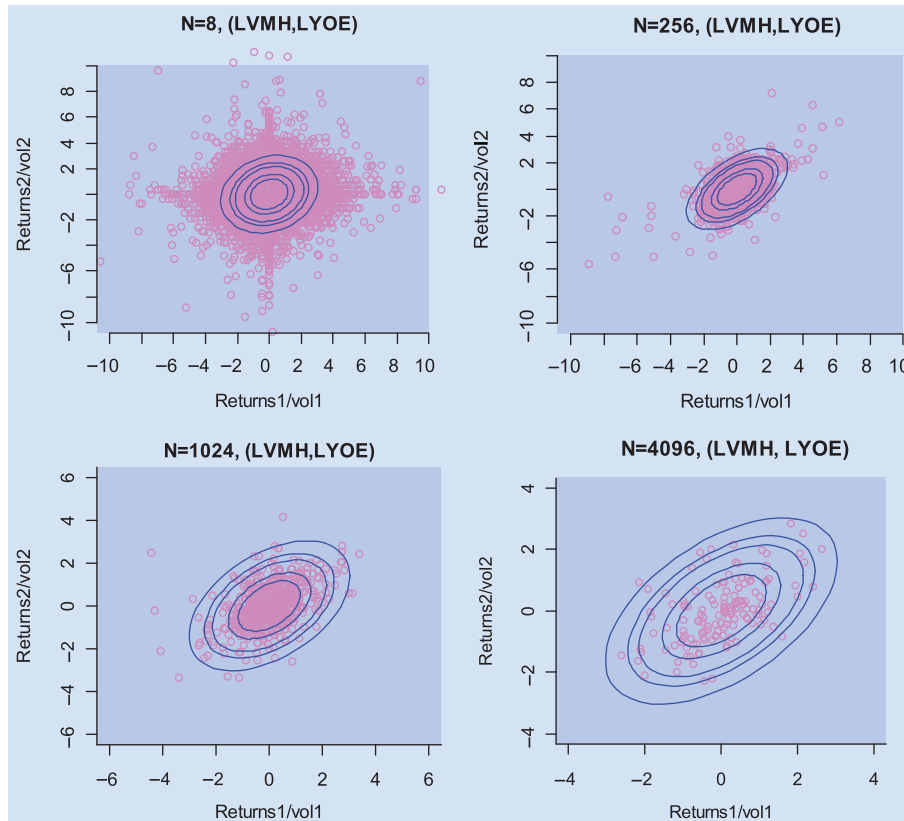
Figure 1. Scatter plot of normalized returns of (LVMH, LYOE) with Gaussian confidence ellipses (levels $\alpha = 50\%$, 70%, 90%, 95% and 99%) for $N = 8$, 256, 1024 and 4096.

3. Data description and methodology

In this section, we test† our theory against high-frequency multivariate data. The main statements are as follows.

- Do returns become jointly normal when sampled in trade time as N grows?
- Is the empirical covariance matrix of returns scaling linearly with N?
- Is the calendar time correlation equal to the trade time correlation and independent of Δt?
- What does $P_{N_{\Delta t}}$ look like?

We use CAC 40 tick-by-tick data starting from 04/02/2008 to 20/03/2008, that is 34 trading days. We study four pairs of stocks, namely‡ (BNPP, SOGN), (RENA, VLOF), (LVMH, LYOE) and (FTE, ISPA). Empirical results are found to be quite stable over these stocks.

4. Empirical results

We now address the four questions of the previous section. We present our results for the pair (LVMH,

LYOE) since the findings are similar for the others. To obtain a rough idea of the daily liquidity on these two stocks, the average daily number of trades for LVHM (respectively LYOE) is 7317 (respectively 10,236), so that the average daily number of trades for the couple (LVMH, LYOE) is 17,553.

4.1. Joint normality of returns

Figure 1 shows a scatter plot of the normalized trade time returns $(R_N^1/\sqrt{N\Sigma_{11}}, R_N^2/\sqrt{N\Sigma_{22}})$ for several N, together with bivariate Gaussian confidence ellipses (see appendix C) for the confidence levels $\alpha = 50\%$, 70%, 90%, 95% and 99%.

First, we note that, as N grows, the correlation between the two returns appears. This effect is known as the Epps effect in the financial literature (Epps 1979, Toth and Kertesz 2006, 2009), and we will also document it in section 3.3. The bivariate normality, not a likely property at all for $N = 8$, reaches better and better agreement with the data as N grows.

In order to confirm normality, figure 2(left) plots the empirical probability of being inside the α confidence ellipse as a function of N. Figure 2(right) shows a plot of

†All the statistical computations were performed using the free statistical software R, available at http://cran.r-project.org.
‡BNPP, BNP Paribas; SOGN, Société Générale; RENA, Renault; VLOF, Valeo; LVMH, Louis Vuitton Moët Hennessy; LYOE, Suez; FTE, France Telecom; ISPA, Arcelor-Mittal.
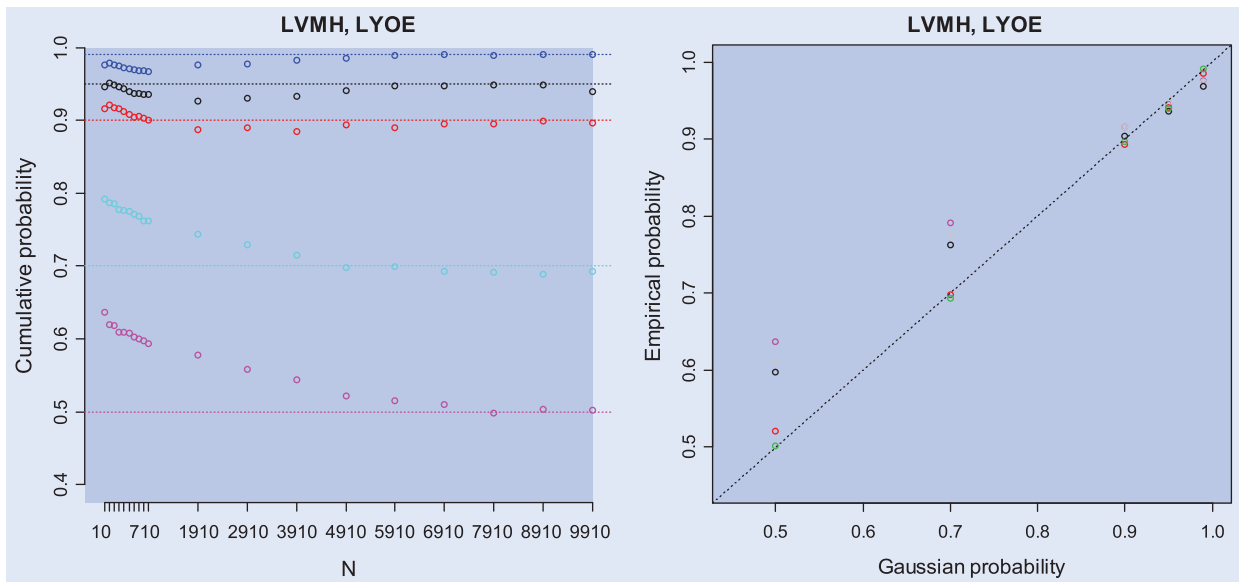
Figure 2. Left: Plot of the empirical probability of lying inside the Gaussian confidence ellipse of level $\alpha = 50\%$, 70%, 90%, 95% and 99% versus N. Right: Empirical probability versus theoretical probability of lying inside Gaussian confidence ellipses for $N = 10$, 310, 810, 5910 and 9910 (pink, grey, black, red and green).
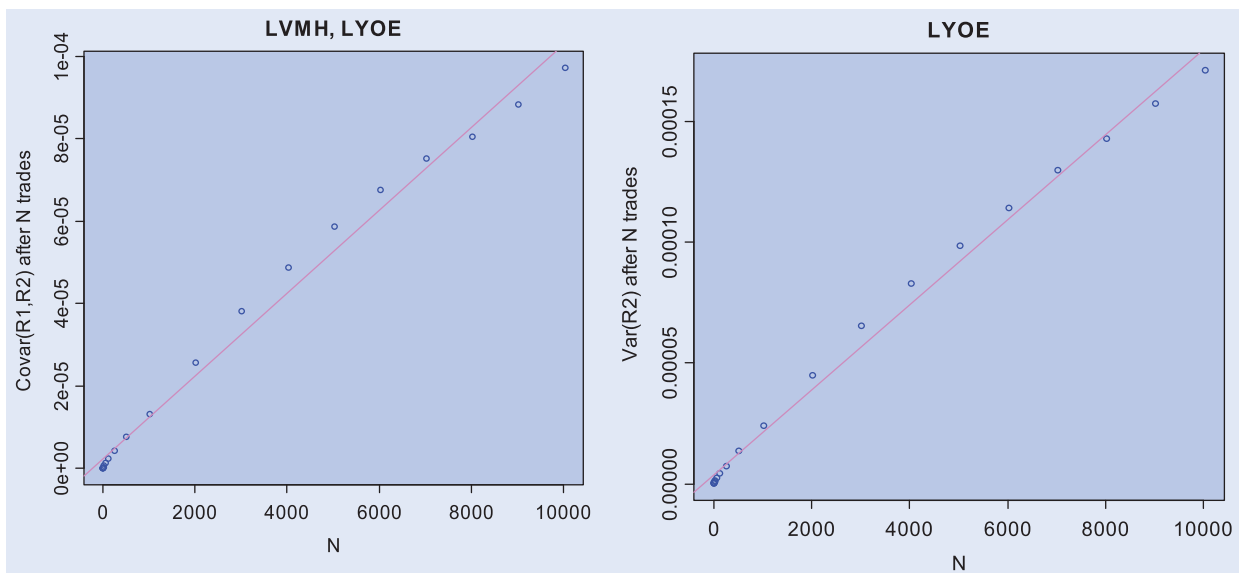

Figure 3. Left: $\mathbb{C}ov(R_N^1, R_N^2)$ versus N. Right: $\mathbb{V}ar(R_N^2)$ (LYOE) versus N; LVMH is not shown but displays the same linear pattern.

the empirical probability versus the theoretical probability for several values of N. Both graphs confirm the normality of the returns as N grows. After 5910 trades, roughly a third of the average daily number of trades for this pair, the normal fit is quite satisfactory.†

$\mathbb{V}ar(R_N^2)$ as a function of N. The linear fit is quite convincing. When switching to calendar time, the linearity still holds, but with respect to calendar time. As we shall see in section 3.4, $\mathbb{E}(N_{\Delta t}) \propto \Delta t$, leading to $\mathbb{V}ar(R_{\Delta t}) \propto \Delta t$. One can see from figure 4 that the linearity is indeed experimentally satisfied.

4.2. Linear scaling of the covariance matrix

We now check the linear growth of $\mathbb{V}ar(R_N)$ as N grows. Figure 3 shows graphs of $\mathbb{C}ov(R_N^1, R_N^2)$, $\mathbb{V}ar(R_N^1)$ and

4.3. Trade time and calendar time correlations

We now test the equality between trade time and calendar time correlations, and the independence of

---

†Even if the graphical fit seems to suggest that the normal hypothesis is reasonable for large N, we computed a bivariate Shapiro–Wilk normality test using R. The average *p*-value for N varying from 910 to 9910 is 25%, indicating that the normality hypothesis cannot be rejected at any usual confidence level.
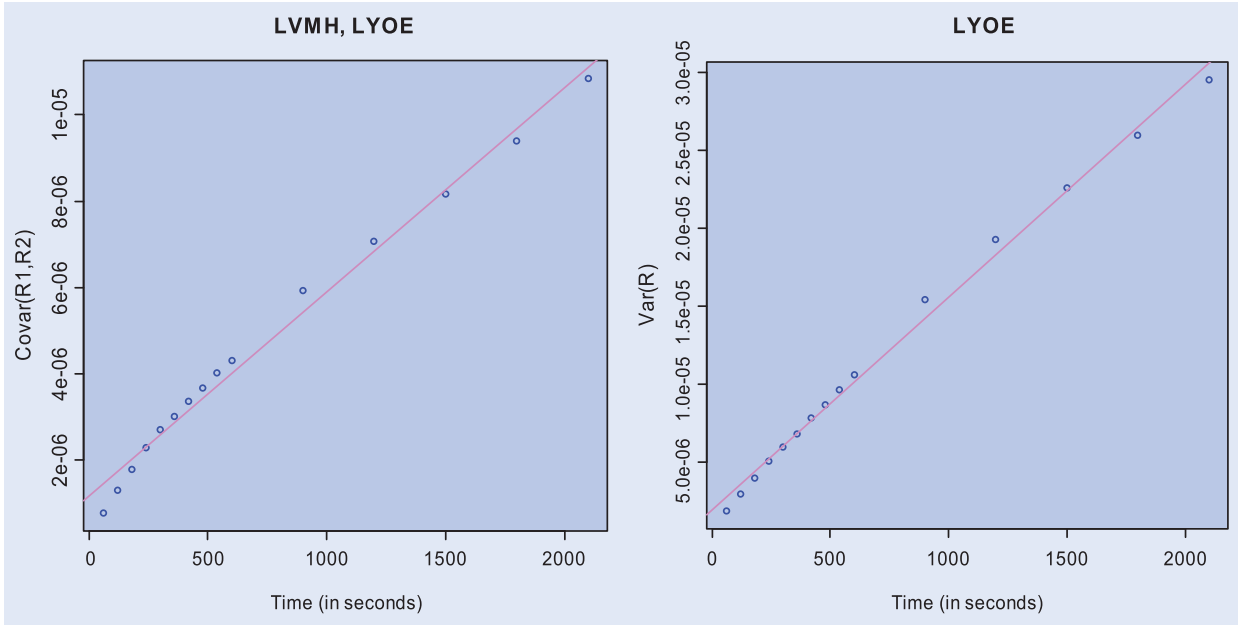
Figure 4. Left: $\mathbb{C}\text{ov}(R^1_{\Delta t}, R^2_{\Delta t})$ versus Δt. Right: $\mathbb{V}\text{ar}(R^2_{\Delta t})$ (LYOE) versus Δt; LVMH is not shown but displays the same linear pattern.
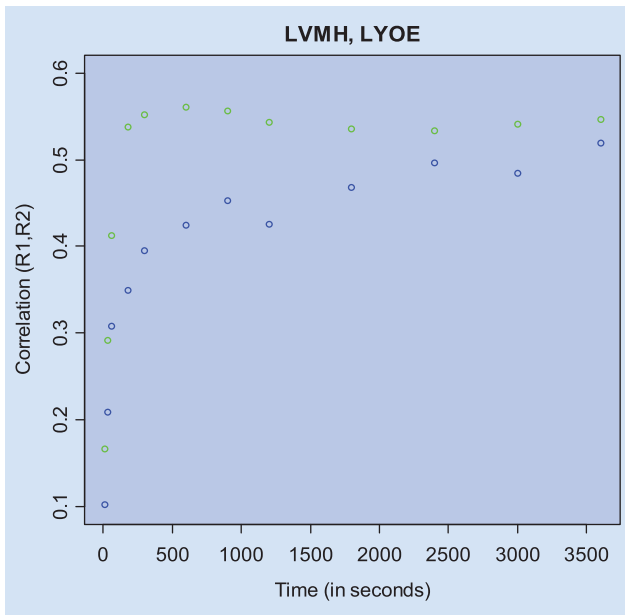

Figure 5. Calendar time (blue) and trade time (green) correlation versus Δt. The number of trades chosen to compute the trade time correlation is such that $\mathbb{E}(N) = \Delta t$ for each Δt.

the latter with respect to Δt. Figure 5 plots the trade time and calendar time correlation. The equality between trade time and calendar time correlations holds when both reach their stationary regime, i.e. when the Epps effect—or other effects at the microstructure level—vanishes. It is noteworthy that the stationary regime appears to be reached much earlier in trade time, a fact that strengthens our preference for multidimensional trade time over the usual calendar time.

4.4. Distribution of the event time

Finally, we investigate the best fit for $P_{N_{\Delta t}}$ for different Δt. Figure 6 displays the empirical survival function (average of number of trades greater than a given threshold) for $\Delta t = 180, 600, 1800$ and 6000 seconds on a log-linear scale. Different parametric fits are suggested: Poisson, gamma, inverse gamma and log-normal, with parameters chosen so as to match the first two empirical moments. The empirical distribution starts from inverse gamma (heavy tails) for small Δt and moves towards a gamma law (semi-heavy tails) for large Δt. An inverse gamma law for N leads to a Student's distribution for $P_{R_{\Delta t}}$, which is of power-law type, while a gamma law leads to a Variance Gamma model having exponential tails (Carr et al. 1998, Nicolato 2009).

In figure 7 we plot the two first moments of $N_{\Delta t}$: its mean, which impacts the variance of returns, and its second moment, which impacts the kurtosis of returns, against Δt. We find that $\mathbb{E}(N_{\Delta t}) \propto \Delta t$ and $\mathbb{E}(N^2_{\Delta t}) \propto \Delta t^\beta$ with $\beta \approx 1.85$ for the four pairs of stocks under consideration. The linearity of the average number of trades essentially states that trade time is comparable to calendar time. The sub-quadratic scaling of the second moment implies a slow decay for the kurtosis of the returns in calendar time, i.e. a slow convergence of the returns distribution in calendar time towards a Gaussian limit. Actually, we have

$$\frac{\mathbb{E}(R^4_{\Delta t})}{\mathbb{E}(R^2_{\Delta t})^2} \propto \frac{\mathbb{E}(N^2_{\Delta t})}{\mathbb{E}(N_{\Delta t})^2} \propto \Delta t^{\beta-2},$$

with $(\beta - 2) \approx -0.15$. This is in agreement with the existing literature (Bouchaud and Potters 2004), as well as the fact that an i.i.d. modeling of returns in calendar time is inappropriate, since it forecasts a scaling of the kurtosis in Δt^{-1}.
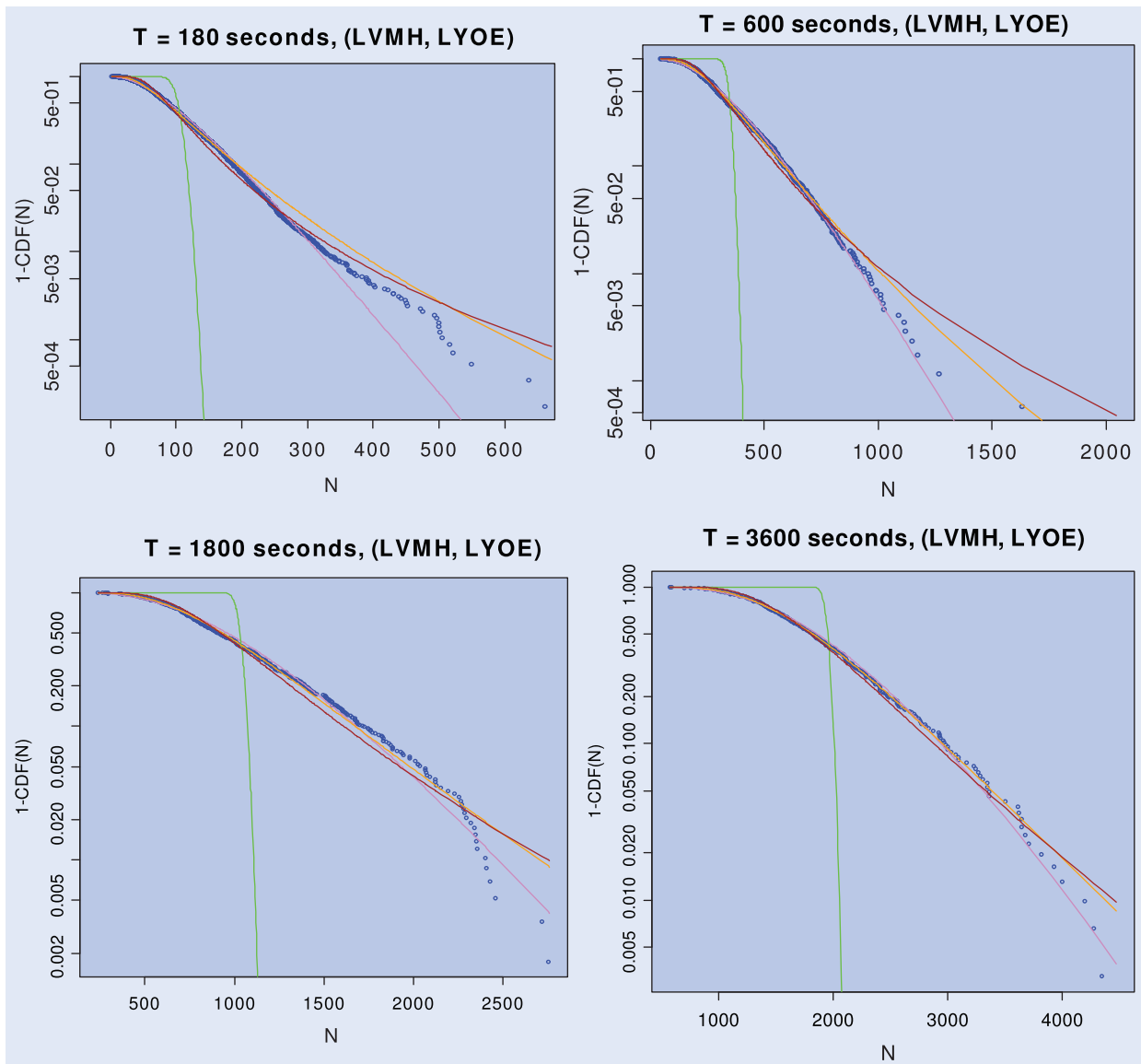
Figure 6. Log–linear plot of the empirical survival function of N_T (blue), Poisson (green), gamma (pink), inverse gamma (brown) and log-normal (orange) for $T = 180$, 600, 1800 and 3600 seconds.
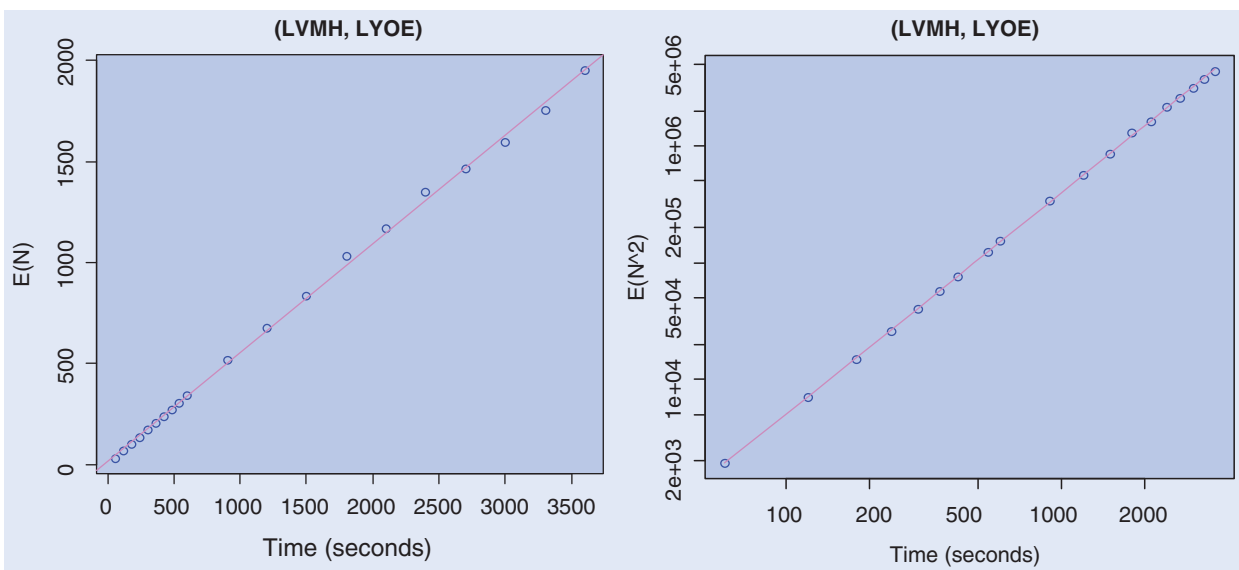

Figure 7. Left: $\mathbb{E}(N_{\Delta t})$ versus Δt. Right: log–log plot of $\mathbb{E}(N_{\Delta t}^2)$ versus Δt; the slope of the linear fit is ≈ 1.85.

5. Conclusion and further research

In this paper, we have presented empirical evidence that demonstrates the stochastic behavior of the covariance matrix in financial markets. A simple mechanism has been described that accounts for this stochastic behavior: as in the classical subordination approach of Clark (1973), the randomness of the covariance matrix stems from that of the arrival times of market orders. Going a step further and moving to continuous-time finance, one can think of returns being driven by a subordinated multivariate Brownian motion. The stochastic clock, i.e. the trade time, is distributed according to a gamma law for large enough time periods. The resulting distribution for stock returns in calendar time is known as the multivariate Variance Gamma model (Leoni and Schoutens 2008) (see also Nicolato 2009 for interesting extensions):

$$S_t^i = S_0^i \, e^{X_t^i} e^{BS_t^i} e^{\omega_t^i},$$
$$X_t^i = \theta^i G_t + \eta^i W^i(G_t),$$
$$G_t \sim \text{Gamma}\left(\frac{t}{\nu}, \frac{1}{\nu}\right),$$
$$BS_t^i = \int_0^t \sigma^i(s) \, dW_s^{i,\perp} - \frac{1}{2}\int_0^t \sigma^i(s)^2 \, ds,$$
$$d\langle W^i, W^j\rangle_t = \rho^{ij} \, dt.$$

The common stochastic clock G refers to the multivariate number of trades. This model has several advantages such as:

- an intuitive and parsimonious parameterization: θ^i provides control over the skewness of $\ln(S_t^i/S_0^i)$ while ν^i impacts its kurtosis;
- a straightforward marginal calibration: the Variance Gamma belongs to the class of Lévy processes, leading to an analytical formula for the characteristic function of $\ln(S_t^i/S_0^i)$. Therefore, vanillas can be priced quickly using FFT techniques; and
- correlation calibration: since the joint characteristic function of returns is available in closed form in this model, the Brownian correlation matrix can easily be calibrated onto the historical correlation matrix of the returns (Leoni and Schoutens 2008).

It also exhibits very serious shortcomings:

- a lack of spot/covariance correlation: for instance, the only correlation between spot and volatility in this model goes through the parameter θ. This might not be sufficient to capture the strong functional dependency between spot and volatility, especially during market crashes;
- there is no real covariance dynamics. The model cannot explain long-range covariance dependency or covariance clustering;
- the volatilities of the assets are perfectly correlated due to the common stochastic clock;

- the variance of G_t scales as t, contrary to empirical data scaling as $t^{1.85}$; and
- full independence cannot be reached in this model even if Brownian correlations are set to zero.

In our opinion, one of the main questions for future work (Huth 2010, Muni Toke 2010) is to understand, from the microstructure point of view and using order-book models, the spot/covariance correlation and covariance clustering. Other questions, more specific to equity derivatives management, include the following.

- Should equity derivatives traders manage their books in trade time or in calendar time? Which approach produces smoother P&L's and Greeks?
- Since correlation risk is not really hedgeable, due to the illiquidity of contracts providing pure correlation exposure, what is the best factor to project correlation onto? A common feature of correlation is that equity correlations rally towards 100% when the market drops. It is also well known that market drops are synonymous with very volatile periods. Therefore, the question is the following: Is it the index level or the index volatility that has the highest predictive power for correlation? The answer would yield important consequences in terms of the correlation hedge, as it would help us determine the respective parts of the correlation risk that can be hedged through delta and vega hedging.

References

Abergel, F., Chakraborti, A., Muni Toke, I. and Patriarca, M., Econophysics: empirical facts and agent-based models, Working Paper, 2009.

Ane, T. and Geman, H., Stochastic subordination. *Risk*, 1996, **9**, 145–149.

Bouchaud, J.-P., Farmer, J.D. and Lillo, F., How markets slowly digest changes in supply and demand. In *Handbook of Financial Markets: Dynamics and Evolution*, edited by K. Schenke-Hoppe and T. Hens, pp. 57–160, 2009 (North-Holland Publishers: Amsterdam).

Bouchaud, J.-P. and Potters, M., Fluctuations and response in financial markets: The subtle nature of 'Random' price changes. *Quant. Finance*, 2004, **4**, 176–190.

Bouchaud, J.-P., Potters, M., Gefen, Y. and Wyart, M., *Fluctuations and Response in Financial Markets: The Subtle Nature of 'Random' Price Changes*, 2003.

Carr, P., Geman, H., Madan, D. and Yor, M., The fine structure of asset returns: an empirical investigation. *J. Business*, 2002, **75**, 305–332.

Carr, P., Madan, D. and Chang, E.C., The variance gamma process and option pricing. *Eur. Finance Rev.*, 1998, **2**, 79–105.

Christian Silva, A., Applications of physics to finance and economics: returns, trading activity and income. PhD thesis, University of Maryland, 2005.

Clark, P.K., A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 1973, **41**, 135–155.

Epps, T.W., Comovements in stock prices in the very short-run. *J. Am. Statist. Assoc.*, 1979, **74**, 291–298.

Gillemot, L., Farmer, J.D. and Lillo, F., There's more to volatility than volume. *Quant. Finance*, 2006, **6**(5), 371–384.

Huth, N., High frequency correlation modelling, in *International Workshop on Econophysics of Order-Driven Markets*, Kolkata, 2010.

Leoni, P. and Schoutens, W., Multivariate smiling. *Wilmott Mag.*, 2008, **34**.

Muni Toke, I., Market making behaviour in an electronic order book and its impact on the bid–ask spread, in *International Workshop on Econophysics of Order-Driven Markets*, Kolkata, 2010.

Nicolato, E., Multivariate modelling via matrix subordination, in *Quantitative Methods in Finance Conference*, Sydney, 2009.

Toth, B. and Kertesz, J., Increasing market efficiency: evolution of cross-correlations of stock returns. *Physica A*, 2006, **360**, 505–515.

Toth, B. and Kertesz, J., The Epps effect revisited. *Quant. Finance*, 2009, **9**(7), 793–802.

Whitt, W., *Stochastic-Process Limits*, 2002 (Springer: Berlin).

Appendix A: Moments of the Gaussian distribution

Let $X \sim \mathcal{N}(0,1)$. We state that $\mathbb{E}(X^{2m}) = (2m)!!$ and that $\mathbb{E}(X^{2m-1}) = 0$ for $m \in \mathbb{N}^*$. We prove it by forward induction. It is clear that, for $m = 1$, $\mathbb{E}(X) = 0$ and $\mathbb{E}(X^2) = 1 = 2!!$. Then assume it is true for $m - 1$. We obtain

$$
\begin{aligned}
\mathbb{E}(X^{2m}) &= \int_{\mathbb{R}} x^{2m-1} (x e^{-x^2/2}) \frac{\mathrm{d}x}{\sqrt{2\pi}} \\
&= (2m-1) \int_{\mathbb{R}} x^{2m-2} e^{-x^2/2} \frac{\mathrm{d}x}{\sqrt{2\pi}} \\
&= (2m-1)\mathbb{E}(X^{2(m-1)}) = (2m-1)(2(m-1))!! = (2m)!!
\end{aligned}
$$

The same computations lead to

$$
\mathbb{E}(X^{2m-1}) = (2m-2)\mathbb{E}(X^{2(m-1)-1}) = 0,
$$

so that the initial statement is true for any $m \in \mathbb{N}^*$.

Appendix B: Probability distribution of returns in calendar time

Given the probability distribution of the number of trades over the time period Δt, the probability distribution of returns in calendar time can be computed as

$$
\begin{aligned}
P_{R_{\Delta t}}(r) &= \int_0^{+\infty} P_{\mathcal{N}(0, n.\sigma^2)}(r) P_{N_{\Delta t}}(n) \, \mathrm{d}n \\
&= \int_0^{+\infty} \frac{e^{-(r^2/2\sigma^2 n) + \ln(P_{N_{\Delta t}}(n))}}{\sqrt{2\pi\sigma^2 n}} \, \mathrm{d}n.
\end{aligned}
$$

An approximation of this integral can be provided for $|r| \sim +\infty$ by performing the integration around the maximizer $n^* = n^*(r)$ of the exponential function. n^* is defined through the following equation:

$$
\frac{r^2}{2(\sigma n^\star)^2} + \frac{P'_{N_{\Delta t}}(n^\star)}{P_{N_{\Delta t}}(n^\star)} = 0.
$$

Let us consider two cases for $P_{N_{\Delta t}}(n)$: exponential and power-law tails, i.e. for $n \sim +\infty$,

$$
P_{N_{\Delta t}}(n) \propto e^{-\lambda n}, \tag{B1}
$$

$$
P_{N_{\Delta t}}(n) \propto n^{-\mu}. \tag{B2}
$$

In the first case, $n^* = b|r|$, while in the second case, $n^* = cr^2$. This yields, for $|r| \sim +\infty$,

$$
P_{R_{\Delta t}}(r) \propto e^{-d|r|}, \quad \text{for case 1,}
$$

$$
P_{R_{\Delta t}}(r) \propto \frac{1}{|r|^{2\mu}}, \quad \text{for case 2.}
$$

therefore, in both cases, the tail behavior of the number of trades, i.e. of the stochastic volatility, is propagated over the tails of returns with adjusted shape coefficients, leading to non-Gaussian tails.

Appendix C: Gaussian confidence ellipses

Let $X = (X^1, \ldots, X^d)' \sim \mathcal{N}(\mu, \Sigma)$. Then $Y = \Sigma^{-1/2}(X - \mu) \sim \mathcal{N}(0, Id)$ and therefore $Y'Y \sim \chi^2(d)$. For a given $\alpha \in (0, 1)$, we are looking for a such that

$$
\mathbb{P}(Y'Y \leq a) = \alpha.
$$

Setting F as the cumulative density function of a $\chi^2(d)$-distributed variable, we obtain $a = F^{-1}(\alpha) = a(\alpha)$. We call the region $\{Y'Y \leq a(\alpha)\}$ the Gaussian confidence ellipse with level α, since

$$
Y'Y - a = (X - \mu)'\Sigma^{-1}(X - \mu) - a
$$

gives an ellipsoid in the Gaussian hyperplane $(X^1, \ldots, X^d)$.