The Fourier volatility estimation method and some applications

Maria Elvira Mancino

University of Firenze

Dimai & Disei Seminar
July 7th, 2021

Motivation

- Computation of **volatility/covariance** of financial asset returns plays a central role for many issues in finance: risk management, hedging strategies, forecasting...
- Black&Scholes model - **constant volatility** - does not account for (e.g.) volatility smile and covariance between asset returns and volatility (leverage effect) $\Rightarrow$ **stochastic volatility** models proposed to model asset price evolution and to price options (adding **risk factors** represented by Brownian motions [Heston, 1993, Hull and White, 1987, Stein and Stein, 1991], **jumps** [Bates, 1996], introducing **memory** [Hobson and Rogers, 1998], [Comte and Renault, 1998] (fBM with $H > 1/2$) to recent rough volatility [Alos et al., 2007, Fukasawa, 2011] (fBM with $H < 1/2$).
- Availability of **high frequency data** have the potential to improve the capability of computing volatility/covariances in an efficient way to many extend: forecasting, risk factor models, asset allocation....

Volatility: The problem

Volatility is **not observable**

Estimation

- **parametric**: the expected volatility is modeled through a functional form of variables observed in the market
- **non-parametric**: the computation of the historical volatility without assuming a functional form of the volatility

Outline

- Definition of Fourier estimator of spot and integrated volatility/covariance
- Asymptotic Normality
- Finite Sample properties of Fourier estimator with high frequency data

Some applications:

- Quarticity estimation
- Volatility of Volatility estimation
- Leverage estimation
- Forecasting Volatility
- Contingent claim pricing-hedging (i.e., stochastic derivation of volatility along the time evolution)
- Non-parametric calibration of the geometry of the Heath-Jarrow-Morton interest rates dynamics ($\Rightarrow$ measure of hypoellipticity of the infinitesimal generator)
- Dynamic Principal Component Analysis
- Optimal Portfolio Allocation
- Indicators of market instability (2020)

Outline

- Definition of Fourier estimator of spot and integrated volatility/covariance
- Asymptotic Normality
- Finite Sample properties of Fourier estimator with high frequency data

Some applications:

- Quarticity estimation
- Volatility of Volatility estimation
- Leverage estimation
- Forecasting Volatility
- Contingent claim pricing-hedging (i.e., stochastic derivation of volatility along the time evolution)
- Non-parametric calibration of the geometry of the Heath-Jarrow-Morton interest rates dynamics ($\Rightarrow$ measure of hypoellipticity of the infinitesimal generator)
- Dynamic Principal Component Analysis
- Optimal Portfolio Allocation
- Indicators of market instability (2020)

Continuous time model: Non-parametric/Model free

$$(\mathbf{B}) \qquad dp^j(t) = \sum_{i=1}^{d} \sigma_i^j(t)\, dW^i(t) + b^j(t)\, dt, \qquad j = 1, \ldots, n,$$

$W = (W^1, \ldots, W^d)$ are independent Brownian motions and σ_*^* and b^* are adapted random processes satisfying

$$E[\int_0^{2\pi} (b^j(t))^2 dt] < \infty, \quad E[\int_0^{2\pi} (\sigma_i^j(t))^4 dt] < \infty \qquad i = 1, \ldots, d,\ j = 1, \ldots, m$$

Objective: estimation of the time dependent *volatility matrix*:

$$\Sigma^{jk}(t) = \sum_{i=1}^{d} \sigma_i^j(t)\sigma_i^k(t) \qquad j, k = 1, \ldots, n$$

Main Issues

$p^*(t)$ asset log-price Brownian semimartingale $\Rightarrow$ integrated volatility/covariance

$$\int_0^t \Sigma^{ik}(s)ds = P - \lim_{n\to\infty} \sum_{0\leq j\leq n-1} \left(p^i(t_{j+1}) - p^i(t_j)\right)\left(p^k(t_{j+1}) - p^k(t_j)\right).$$

Nevertheless, when sampling high frequency returns, three difficulties arise:

1) the distortion from efficient prices due to the market microstructure noise such as price discreteness, infrequent trading,...[Roll, 1984].
2) instantaneous volatility computation involves a sort of numerical derivative, which gives rise to numerical instabilities [Foster and Nelson, 1996, Comte and Renault, 1998]

In the multivariate case also:

3) the **non-synchronicity** of the arrival times of trades across markets leads to a bias towards zero in correlations among stocks as the sampling frequency increases [Epps, 1979]

Fourier method

Mean Covariance Theorem [Malliavin and M. 2002, 2009]

Theorem

Consider a process p satisfying the assumption **(B)**. Then we have:

$$\frac{1}{2\pi}\,\mathcal{F}(\Sigma^{ij}) = \mathcal{F}(dp^i) *_B \mathcal{F}(dp^j). \tag{1}$$

The convergence of the convolution product (1) is attained in probability where, for $k \in \mathbf{Z}$

$$\mathcal{F}(dp^i)(k) := \frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}kt}\, dp^i(t)$$

$$(\Phi *_B \Psi)(k) := \lim_{N\to\infty} \frac{1}{2N+1}\sum_{s=-N}^{N} \Phi(s)\Psi(k-s)$$

$$\mathcal{F}(\Sigma^{ij})(k) := \frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}kt}\,\Sigma^{ij}(t)\,dt$$

Fourier instantaneous covariance computation

By the theorem we gather all the Fourier coefficients of the volatility matrix by means of the Fourier transform of the log-returns. Then reconstruct the **co-volatility functions** $\Sigma^{ij}(t)$ from its Fourier coefficients by the Fourier-Fejer summation:
let for $i, j = 1, 2$ and for any $|k| \leq N$,

$$c_N^{ij}(k) := \frac{1}{2N+1} \sum_{|s| \leq N} \mathcal{F}(dp^i)(s)\mathcal{F}(dp^i)(k-s),$$

then

$$\Sigma^{ij}(t) = \lim_{M \to \infty} \sum_{|k| \leq M} \left(1 - \frac{|k|}{M+1}\right) c_N^{ij}(k) e^{\mathrm{i}kt}.$$

Asymptotic Normality for (univariate) Fourier estimator

Let $\{0 \leq t_{0,n} \leq t_{1,n} \leq \cdots \leq t_{k_n,n} = 2\pi\}$ be the (possibly unequally-spaced) trading dates (we omit the second index), the returns are $\delta_i(p) := p(t_{i+1}) - p(t_i)$.
Given the discrete Fourier transform

$$c_k(dp_n) := \frac{1}{2\pi} \sum_{i=0}^{k_n-1} e^{-\mathrm{i}kt_i} \delta_i(p),$$

consider the following *convolution* formula

$$c_k(\sigma^2_{n,N}) := \frac{2\pi}{2N+1} \sum_{|h|\leq N} c_h(dp_n) c_{k-h}(dp_n)$$

For any $t \in (0, 2\pi)$, then $\sigma^2(t)$ from its Fourier coefficients by the Fourier-Fejer inversion formula

$$\widehat{\sigma}^2_{n,N,M}(t) = \sum_{|k|<M} \left(1 - \frac{|k|}{M}\right) \left[\frac{2\pi}{2N+1} \sum_{|h|\leq N} c_h(dp_n) c_{k-h}(dp_n)\right] e^{itk}$$

The definition of the estimator $\widehat{\sigma}^2_{n,N,M}(t)$ depends on three parameters.

Asymptotic Normality for (univariate) Fourier estimator

Let $\{0 \leq t_{0,n} \leq t_{1,n} \leq \cdots \leq t_{k_n,n} = 2\pi\}$ be the (possibly unequally-spaced) trading dates (we omit the second index), the returns are $\delta_i(p) := p(t_{i+1}) - p(t_i)$.
Given the discrete Fourier transform

$$c_k(dp_n) := \frac{1}{2\pi} \sum_{i=0}^{k_n-1} e^{-\mathrm{i}kt_i} \delta_i(p),$$

consider the following *convolution* formula

$$c_k(\sigma^2_{n,N}) := \frac{2\pi}{2N+1} \sum_{|h|\leq N} c_h(dp_n) c_{k-h}(dp_n)$$

For any $t \in (0, 2\pi)$, then $\sigma^2(t)$ from its Fourier coefficients by the Fourier-Fejer inversion formula

$$\widehat{\sigma}^2_{n,N,M}(t) = \sum_{|k|<M} \left(1 - \frac{|k|}{M}\right) \left[\frac{2\pi}{2N+1} \sum_{|h|\leq N} c_h(dp_n) c_{k-h}(dp_n)\right] e^{\mathrm{i}tk}$$

The definition of the estimator $\widehat{\sigma}^2_{n,N,M}(t)$ depends on three parameters.

We can write the estimated Fourier coefficients as

$$c_k(\sigma^2_{n,N}) = \frac{1}{2\pi}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} D_N(t_j - t_i)e^{-\mathrm{i}kt_j}\delta_i(p)\delta_j(p),$$

where D_N is the rescaled Dirichlet kernel

$$D_N(x) = \frac{1}{2N+1}\sum_{|h|\leq N} e^{\mathrm{i}hx} = \frac{1}{2N+1}\frac{\sin(2N+1)\frac{x}{2}}{\sin\frac{x}{2}}. \tag{2}$$

Thus, the Fourier estimator of spot volatility can be expressed as follows

$$\widehat{\sigma}^2_{n,N,M}(t) = \frac{1}{2\pi}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} F_M(t - t_j)D_N(t_j - t_i)\delta_i(p)\delta_j(p), \tag{3}$$

where F_M is the Fejer kernel

$$F_M(x) = \sum_{|k|\leq M}\left(1 - \frac{|k|}{M+1}\right)e^{\mathrm{i}kx}$$

The estimator (3) contains two terms: the quadratic part

$$\frac{1}{2\pi}\sum_{j=0}^{n-1} F_M(t-t_j)(\delta_j(p))^2 \tag{4}$$

and the cross terms

$$\frac{1}{2\pi}\sum_{i=0}^{n-1}\sum_{\substack{j=0\\ j\neq i}}^{n-1} F_M(t-t_j)D_N(t_j-t_i)\delta_i(p)\delta_j(p). \tag{5}$$

The quadratic term (4) behaves like the Kernel-based spot volatility estimators seen in [Fan and Wang, 2008].
The second addend (5) is crucial in terms of robustness of the estimator in the presence of microstructure noise, through the choice of the frequency N
[Barndorff-Nielsen, Hansen, Lunde and Shephard, 2008].

CLT

Assume **(B)** and σ is a.s. Hölder continuous with parameter $\nu \in (0, \frac{1}{2})$ (e.g., driven by a second Brownian semimartingale), and the following conditions hold: $\lim_{n,M\to\infty} \frac{M^\gamma}{n} = a > 0$, for some $\gamma > 1$, and $\lim_{n,N\to\infty} \frac{N}{n} = c > 0$. Then, for any fixed $t \in (0, 2\pi)$, as $n, N, M \to \infty$,

$$\sqrt{\frac{n}{M}} \left(\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t) \right) \to \mathcal{N} \left(0, \frac{4}{3}(1 + 2\eta(c))\, \sigma^4(t) \right),$$

where the convergence is stable in law. The constant $\eta(c)$ is equal to $\frac{1}{2\tilde{c}^2} r(\tilde{c})(1 - r(\tilde{c}))$, where $\tilde{c} = 2c$ and $r(x) = x - [x]$, with $[x]$ the integer part of x.

- Rate of convergence: The convergence rate is of order $n^{\frac{\gamma-1}{2\gamma}}$. It appears in the proof that $1 < \gamma < 2\nu + 1$: for $\gamma \approx 2$ the rate of convergence becomes $\frac{1}{4}$, which is the optimal rate of convergence for a non-parametric spot volatility estimator.
- Optimal variance: $\eta(c)$ is nonnegative for any positive c and equal to zero when $c = \frac{1}{2}k$, $k = 1, 2, \ldots$. The case $\eta(c) = 0$ provides the optimal asymptotic variance $\frac{4}{3}\sigma^4(t)$, [Cuchiero and Teichmann, 2015]. The optimal asymptotic variance is obtained for $c = \frac{1}{2}k$, $k = 1, 2, \ldots$ and the choice $k = 1$ (i.e., $c = 1/2$) corresponds to the natural choice of the Nyquist frequency for the Fourier estimator.

CLT

Assume **(B)** and σ is a.s. Hölder continuous with parameter $\nu \in (0, \frac{1}{2})$ (e.g., driven by a second Brownian semimartingale), and the following conditions hold: $\lim_{n,M\to\infty} \frac{M^\gamma}{n} = a > 0$, for some $\gamma > 1$, and $\lim_{n,N\to\infty} \frac{N}{n} = c > 0$. Then, for any fixed $t \in (0, 2\pi)$, as $n, N, M \to \infty$,

$$\sqrt{\frac{n}{M}} \left(\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t) \right) \to \mathcal{N} \left(0, \frac{4}{3}(1 + 2\eta(c))\, \sigma^4(t) \right),$$

where the convergence is stable in law. The constant $\eta(c)$ is equal to $\frac{1}{2\tilde{c}^2} r(\tilde{c})(1 - r(\tilde{c}))$, where $\tilde{c} = 2c$ and $r(x) = x - [x]$, with $[x]$ the integer part of x.

- Rate of convergence: The convergence rate is of order $n^{\frac{\gamma-1}{2\gamma}}$. It appears in the proof that $1 < \gamma < 2\nu + 1$: for $\gamma \approx 2$ the rate of convergence becomes $\frac{1}{4}$, which is the optimal rate of convergence for a non-parametric spot volatility estimator.
- Optimal variance: $\eta(c)$ is nonnegative for any positive c and equal to zero when $c = \frac{1}{2}k$, $k = 1, 2, \ldots$. The case $\eta(c) = 0$ provides the optimal asymptotic variance $\frac{4}{3}\sigma^4(t)$, [Cuchiero and Teichmann, 2015]. The optimal asymptotic variance is obtained for $c = \frac{1}{2}k$, $k = 1, 2, \ldots$ and the choice $k = 1$ (i.e., $c = 1/2$) corresponds to the natural choice of the Nyquist frequency for the Fourier estimator.

With this choice of c (in other words of N/n) the Fourier estimator has the same rate of convergence and asymptotic variance of the Fejer kernel-based realized spot volatility.

With an appropriate choice of N/n, the effect of adding the cross terms in (3), which is essential in order to get an estimator robust to microstructure noise, is also not detrimental in view of the asymptotic efficiency.

Model with microstructure

Market microstructure effects (discreteness of prices, bid/ask spreads, etc.) cause the discrepancy between asset pricing theory based on semi-martingales and the data at very fine intervals
The logarithm of the observed price process is given by

$$\widetilde{p}(t) = p(t) + \eta(t), \tag{6}$$

$p(t) =$ the efficient log-price process
$\eta(t) =$ the microstructure noise
$\tilde{p}$ is the observed price at a discrete unevenly spaced grid

- The random shocks $\eta(t_{j,n})$, for $\{0 \leq t_{0,n} \leq t_{1,n} \leq \cdots \leq t_{k_n,n} \leq 2\pi\}$ are i.i.d. with mean zero and bounded fourth moment.
- The true return process $\delta_{j,n}(p) := p(t_{j,n}) - p(t_{j-1,n})$ is independent of $\eta(t_{j,n})$ for any j, n.

Model with microstructure

Market microstructure effects (discreteness of prices, bid/ask spreads, etc.) cause the discrepancy between asset pricing theory based on semi-martingales and the data at very fine intervals
The logarithm of the observed price process is given by

$$\widetilde{p}(t) = p(t) + \eta(t), \tag{6}$$

$p(t) =$ the efficient log-price process
$\eta(t) =$ the microstructure noise
$\tilde{p}$ is the observed price at a discrete unevenly spaced grid

• The random shocks $\eta(t_{j,n})$, for $\{0 \leq t_{0,n} \leq t_{1,n} \leq \cdots \leq t_{k_n,n} \leq 2\pi\}$ are i.i.d. with mean zero and bounded fourth moment.
• The true return process $\delta_{j,n}(p) := p(t_{j,n}) - p(t_{j-1,n})$ is independent of $\eta(t_{j,n})$ for any j, n.

Bias

$$V := \int_0^{2\pi} \sigma^2(t)dt \text{ integrated volatility}$$

$$\hat{V}_n := \sum_{j=1}^{n} (\delta_j(\tilde{p}))^2, \text{ realized volatility estimator}$$

where n is the number of observations in the trading interval $[0, 2\pi]$.
Under the hypothesis $2\pi/n$ is the time between adjacent logarithmic prices:

Realized Volatility Bias

$$E[\hat{V}_n - V] = 2nE[\eta^2]$$

Fourier Estimator Bias

For any fixed integers n, N the following identity holds

$$E[\hat{\sigma}^2_{n,N} - V] = 2n\ E[\eta^2] \left(1 - \frac{1}{2N+1} \frac{\sin[(2N+1)\frac{\pi}{n}]}{\sin(\frac{\pi}{n})}\right) \quad (7)$$

Under the condition $N^2/n \to 0$

$$\lim_{n,N\to\infty} 2n\ E[\eta^2]\left(1 - \frac{1}{2N+1}\frac{\sin[(2N+1)\frac{\pi}{n}]}{\sin(\frac{\pi}{n})}\right) = 0.$$

Conclusion

Fourier estimator is asymptotically unbiased under the condition N^2/n goes to 0. Moreover the result (7) shows that for fixed n, i.e. for a finite sample, a suitable choice of N allows for lower bias with respect to the realized volatility estimator.

MSE

Realized Volatility MSE

$$E[(\hat{V}_n - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + \Lambda_n,$$

where Q is the so-called *integrated quarticity* $\int_0^{2\pi} \sigma^4(s)ds$, and

$$\Lambda_n := n^2\alpha + n\beta + \gamma,$$

with

$$\alpha = (E[\varepsilon^2])^2, \quad \beta = E[\varepsilon^4] + 2E[\varepsilon^2\varepsilon_{-1}^2] - 3(E[\varepsilon^2])^2, \tag{8}$$

$$\gamma = 4E[\varepsilon^2]V - 2E[\varepsilon^2\varepsilon_{-1}^2] + 2(E[\varepsilon^2])^2. \tag{9}$$

with the notation ε for $\eta_j - \eta_{j-1}$ for a generic j and ε_{-1} for $\eta_{j-1} - \eta_{j-2}$ for the same j.

Fourier Estimator MSE

Fourier Estimator MSE

For any fixed n, N the following relation holds

$$E[(\hat{\sigma}^2_{n,N} - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}, \tag{10}$$

where

$$\hat{\alpha} = \alpha\ (1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})),$$

$$\hat{\beta} = \beta\ (1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})),$$

$$\hat{\gamma} = \gamma + 4Q\frac{2\pi}{2N+1} + 4(E[\eta^2]^2 + E[\eta^4])(2D_N(\frac{2\pi}{n}) - D_N^2(\frac{2\pi}{n})), \tag{11}$$

with α, β, γ as in (8)

$$MSE_{RV} = 2\frac{2\pi}{n}(Q + o(1))$$

$$MSE_{RVm} = MSE_{RV} + n^2\alpha + n\beta + \gamma \qquad \to \infty.$$

$$MSE_F = MSE_{RV} + c(n, N),$$

where $c(n, N) \to 0$ as $N, n \to \infty$,[1]

$$MSE_{Fm} = MSE_F + n^2\hat{\alpha}(n, N) + n\hat{\beta}(n, N) + \widetilde{\gamma}(n, N)$$

If $N^2/n \to 0$ then

$$\lim_{n,N\to\infty} n^2\hat{\alpha}(n, N) + n\hat{\beta}(n, N) = 0$$

and

$$\lim_{n,N\to\infty} \widetilde{\gamma}(n, N) = 8E[\eta^2]V + 2E[\eta^4] + 6E[\eta^2]^2.$$

It follows that the MSE of the Fourier estimator does not diverge and by choosing N conveniently we obtain that MSE_F and MSE_{Fm} differ about the positive constant term.

[1] We prove that $c(n, N)$ is less or equal than $4Q\ 2\pi/(2N + 1)$

$$MSE_{RV} = 2\frac{2\pi}{n}(Q + o(1))$$

$$MSE_{RVm} = MSE_{RV} + n^2\alpha + n\beta + \gamma \qquad \to \infty.$$

$$MSE_F = MSE_{RV} + c(n, N),$$

where $c(n, N) \to 0$ as $N, n \to \infty$,[1]

$$MSE_{Fm} = MSE_F + n^2\hat{\alpha}(n, N) + n\hat{\beta}(n, N) + \tilde{\gamma}(n, N)$$

If $N^2/n \to 0$ then

$$\lim_{n,N\to\infty} n^2\hat{\alpha}(n, N) + n\hat{\beta}(n, N) = 0$$

and

$$\lim_{n,N\to\infty} \tilde{\gamma}(n, N) = 8E[\eta^2]V + 2E[\eta^4] + 6E[\eta^2]^2.$$

It follows that the MSE of the Fourier estimator does not diverge and by choosing N conveniently we obtain that MSE_F and MSE_{Fm} differ about the positive constant term.

[1] We prove that $c(n, N)$ is less or equal than $4Q\ 2\pi/(2N + 1)$.

Optimized Fourier estimator

A single optimization over the number of frequencies N renders the Fourier estimator very efficient even in presence of microstructure noise

Formulae (7) and (10) provide a way to optimize the finite sample performance of the Fourier estimator as a function of the number of frequencies by the minimization of the bias and the MSE, for a given number of intra-daily observations.

For simulation and empirical applications, see [Mancino and Sanfelici, 2008].

Application: market instability indicators

Identifying financial instability conditions using high frequency data.
Maria Elvira Mancino - Simona Sanfelici *Journal of Economic Interaction and Coordination*, 2020

Let x_t be the logarithm of the price

$$dx_t = a(x_t)dW_t - \frac{1}{2}a^2(x_t)dt. \tag{12}$$

The variation process is the solution of the linearized SDE

$$\frac{d\zeta_t}{\zeta_t} = a'(x_t)dW_t - a'(x_t)a(x_t)dt.$$

The rescaled variation is defined

$$z_t := \frac{\zeta_t}{a(x_t)}. \tag{13}$$

The following result shows that the perturbations of asset prices evolve through time according to an ODE, see [Malliavin and Mancino, 2002] for the proof.

The rescaled variation (13) is differentiable with respect to t and it holds that

$$z_t = z_s \exp\left(\int_s^t \lambda_\tau d\tau\right) \qquad s \leq t,$$

where

$$\lambda_\tau = -\frac{1}{2}(a'(x_\tau)a(x_\tau) + a''(x_\tau)a(x_\tau)) \tag{14}$$

is called the price-volatility feedback rate.

The feedback rate defined by (14) is a decay rate:

- the negative sign of λ entails a damping effect $\Rightarrow$ negative values correspond to stable market directions
- the positive sign of λ implies that any perturbation is amplified $\Rightarrow$ large positive values indicate market instability and usually anticipate a significant decrease in the price level
- values of λ around zero imply that the price level remains stable.

When λ is large then perturbations of price are more likely to propagate, so that an increase in volatility in the presence of a large λ value may trigger a volatility feedback effect and cause large price movements, as theory suggests (see also the empirical analysis in [Inkaya and Ocur, 2014]).
Thus, the volatility feedback rate can reveal conditions that may facilitate the propagation of perturbations in the market.
This may help to discriminate between stable market conditions and conditions when the price process has a potential risk of being affected by an increase in the volatility.

Estimation

In order to effectively compute the feedback rate, we need to express the rescaled variation using quantities which can be empirically estimated using the price process observations: we use *iterated cross volatilities* that can be estimated using the Fourier estimation method.

Denoting by $\langle\,,\rangle$ the quadratic (co-)variation operation, define:

$$\langle dx_t\,,\,dx_t\rangle := A_t\,dt\;,\;\langle dA_t\,,\,dx_t\rangle := B_t\,dt\;,\;\langle dB_t\,,\,dx_t\rangle := C_t\,dt\,.$$

Then:

$$\lambda_t = \frac{3}{8}\frac{B_t^2}{A_t^3} - \frac{1}{4}\frac{B_t}{A_t} - \frac{1}{4}\frac{C_t}{A_t^2}\,. \tag{15}$$

(Proof [Malliavin and Mancino, 2002])

$$dA_t = \alpha_t dt + \gamma_t dW_t^A, \qquad dB_t = \mu_t dt + \beta_t dW_t^B,$$

where W^A and W^B are Brownian motions, possible correlated with W. Define, for $|k| \leq N$,

$$c_k(A_{n,N}) := \frac{2\pi}{2N+1} \sum_{|s| \leq N} c_s(dx_n) c_{k-s}(dx_n), \tag{16}$$

where

$$c_k(dx_n) = \frac{1}{2\pi} \sum_{j=0}^{k_n - 1} e^{-\mathrm{i}kt_{j,n}} \delta_j(x),$$

$$\widehat{A}_{n,N,N_A}(t) := \sum_{|k| \leq N_A} \left(1 - \frac{|k|}{N_A + 1}\right) c_k(A_{n,N}) e^{\mathrm{i}kt},$$

$$c_k(B_{n,N,M}) := \frac{2\pi}{2M+1} \sum_{|j| \leq M} c_j(dx_n) c_{k-j}(dA_{n,N}),$$

$$c_k(C_{n,N,M,L}) := \frac{2\pi}{2L+1} \sum_{|j| \leq L} c_j(dx_{n,N}) c_{k-j}(dB_{n,N,M}).$$[2]

[2] The statistical properties of the estimators of B, C and λ are studied in [Mancino and Toscano, 2021], [Mancino et al., 2021].

$$dA_t = \alpha_t dt + \gamma_t dW_t^A, \qquad dB_t = \mu_t dt + \beta_t dW_t^B,$$

where W^A and W^B are Brownian motions, possible correlated with W. Define, for $|k| \leq N$,

$$c_k(A_{n,N}) := \frac{2\pi}{2N+1} \sum_{|s| \leq N} c_s(dx_n) c_{k-s}(dx_n), \tag{16}$$

where

$$c_k(dx_n) = \frac{1}{2\pi} \sum_{j=0}^{k_n - 1} e^{-\mathrm{i}kt_{j,n}} \delta_j(x),$$

$$\widehat{A}_{n,N,N_A}(t) := \sum_{|k| \leq N_A} \left(1 - \frac{|k|}{N_A + 1}\right) c_k(A_{n,N}) e^{\mathrm{i}kt},$$

$$c_k(B_{n,N,M}) := \frac{2\pi}{2M+1} \sum_{|j| \leq M} c_j(dx_n) c_{k-j}(dA_{n,N}),$$

$$c_k(C_{n,N,M,L}) := \frac{2\pi}{2L+1} \sum_{|j| \leq L} c_j(dx_{n,N}) c_{k-j}(dB_{n,N,M}).$$

[2] The statistical properties of the estimators of B, C and λ are studied in [Mancino and Toscano, 2021], [Mancino et al., 2021].

$$dA_t = \alpha_t dt + \gamma_t dW_t^A, \qquad dB_t = \mu_t dt + \beta_t dW_t^B,$$

where W^A and W^B are Brownian motions, possible correlated with W. Define, for $|k| \leq N$,

$$c_k(A_{n,N}) := \frac{2\pi}{2N+1} \sum_{|s| \leq N} c_s(dx_n) c_{k-s}(dx_n), \tag{16}$$

where

$$c_k(dx_n) = \frac{1}{2\pi} \sum_{j=0}^{k_n - 1} e^{-\mathrm{i}kt_{j,n}} \delta_j(x),$$

$$\widehat{A}_{n,N,N_A}(t) := \sum_{|k| \leq N_A} \left(1 - \frac{|k|}{N_A + 1}\right) c_k(A_{n,N}) e^{\mathrm{i}kt},$$

$$c_k(B_{n,N,M}) := \frac{2\pi}{2M+1} \sum_{|j| \leq M} c_j(dx_n) c_{k-j}(dA_{n,N}),$$

$$c_k(C_{n,N,M,L}) := \frac{2\pi}{2L+1} \sum_{|j| \leq L} c_j(dx_{n,N}) c_{k-j}(dB_{n,N,M}).^2$$

[2] The statistical properties of the estimators of B, C and λ are studied in [Mancino and Toscano, 2021], [Mancino et al., 2021].

Application I: Market volatility and financial crisis

- Empirical and theoretical studies have investigated the relationship between market volatility and financial crisis. However, although there is a general consensus that unusual levels of financial market volatility imply an increased likelihood of a subsequent financial crisis, the use of high market volatility as an indicator of financial instability is not completely reliable, in [Danielson et al., 2016].
- As long as policymakers intervene just when risks materialize, this may lead to a too little too late response. If we were able to go deeper in assessing the regularity of the latent volatility dynamics by enhancing the discontinuities characterizing its evolution, we could detect relevant differences between volatility paths that, under standard scrutiny methods, would seem similar.

Characteristics of the indicator

- The indicator λ_t is designed to amplify relevant fluctuations in financial data, and thus increase our ability to predict large price variations even at short-time horizons.
- Instead of relying on a first-order measure (such as the volatility), the indicator considers second order quantities
- The indicator quantifies the rate of propagation of a perturbation of the price process, thus it can be interpreted as a market breakdown index. When the coefficient is positive, the perturbation propagates over a trajectory, when the coefficient is negative, the perturbation is attenuated over a trajectory.
- The rate of variation through time of an initial perturbation of the price process enables us to understand if such a shock will be rapidly absorbed or, on the contrary, it will be amplified by the market.
- The indicator combines non-linearly volatility, leverage and covariance between leverage and price and is model-free. It provides an early warning indicator of instability for a given high frequency financial time series.

Simulation study: CEV model

$$dx_t = \sigma e^{(\delta-1)x_t} dW_t - \frac{1}{2}\sigma^2 e^{2(\delta-1)x_t} dt. \tag{17}$$

1) we obtain an analytical formula for the indicator;
2) we use this explicit formula to perform a simulated analysis showing the effectiveness in estimating the feedback effect and the relation occurring between the feedback rate sign and the price variations.

$A(x_t) = \sigma^2 e^{2(\delta-1)x_t}$: if $\delta < 1$, then $A(x_t)$ decreases if x_t increases, and viceversa.
From (15), the process λ_t for the CEV model:

$$\lambda_t = -\frac{1}{2}\sigma^2\delta(\delta-1)e^{2(\delta-1)x_t} = -\frac{1}{2}\delta(\delta-1)A_t. \tag{18}$$

The relation between λ and the spot volatility highlights the importance of using the feedback effect to discriminate days with high volatility regimes:
If $\delta > 1$ large volatility values do not necessarily denote market instability, since the feedback effect is negative and the price tends to increase. If $\delta > 1$, volatility is larger as the price increases, and viceversa.

Simulation study: CEV model

$$dx_t = \sigma e^{(\delta-1)x_t} dW_t - \frac{1}{2}\sigma^2 e^{2(\delta-1)x_t} dt. \tag{17}$$

1) we obtain an analytical formula for the indicator;
2) we use this explicit formula to perform a simulated analysis showing the effectiveness in estimating the feedback effect and the relation occurring between the feedback rate sign and the price variations.

$A(x_t) = \sigma^2 e^{2(\delta-1)x_t}$: if $\delta < 1$, then $A(x_t)$ decreases if x_t increases, and viceversa.
From (15), the process λ_t for the CEV model:

$$\lambda_t = -\frac{1}{2}\sigma^2\delta(\delta-1)e^{2(\delta-1)x_t} = -\frac{1}{2}\delta(\delta-1)A_t. \tag{18}$$

The relation between λ and the spot volatility highlights the importance of using the feedback effect to discriminate days with high volatility regimes:
If $\delta > 1$ large volatility values do not necessarily denote market instability, since the feedback effect is negative and the price tends to increase. If $\delta > 1$, volatility is larger as the price increases, and viceversa.

CEV discretized by the forward Euler method over $[0, 1]$ with $n = 21600$ equispaced nodes

Fig. 1: Effect of a 5% perturbation of the initial price $S_0 = 100$, $\sigma = 0.3$

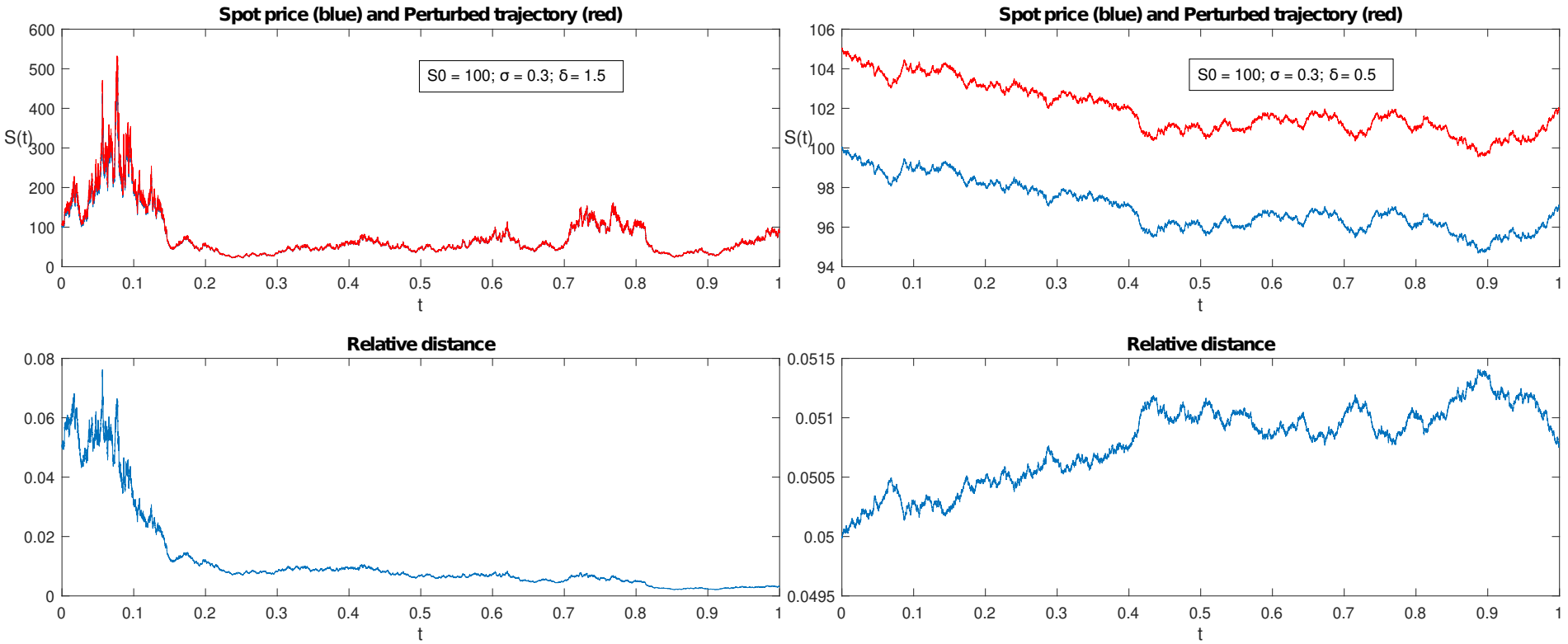
Left panels $\delta = 1.5$ Right panels, $\delta = 0.5$

Upper panels: spot price trajectory (blue) and perturbed trajectory (red)

Lower panels: relative distance between the two trajectories plotted as function of time

Left panels: the distance reduces as time passes by, the initial perturbation is attenuated

Right panels: the two trajectories remain far apart, their relative distance tends to increase over time



The feedback rate λ_t *is negative*, the leverage B_t positive, the volatility A_t is larger as the price increases, and viceversa (inverse leverage effect).

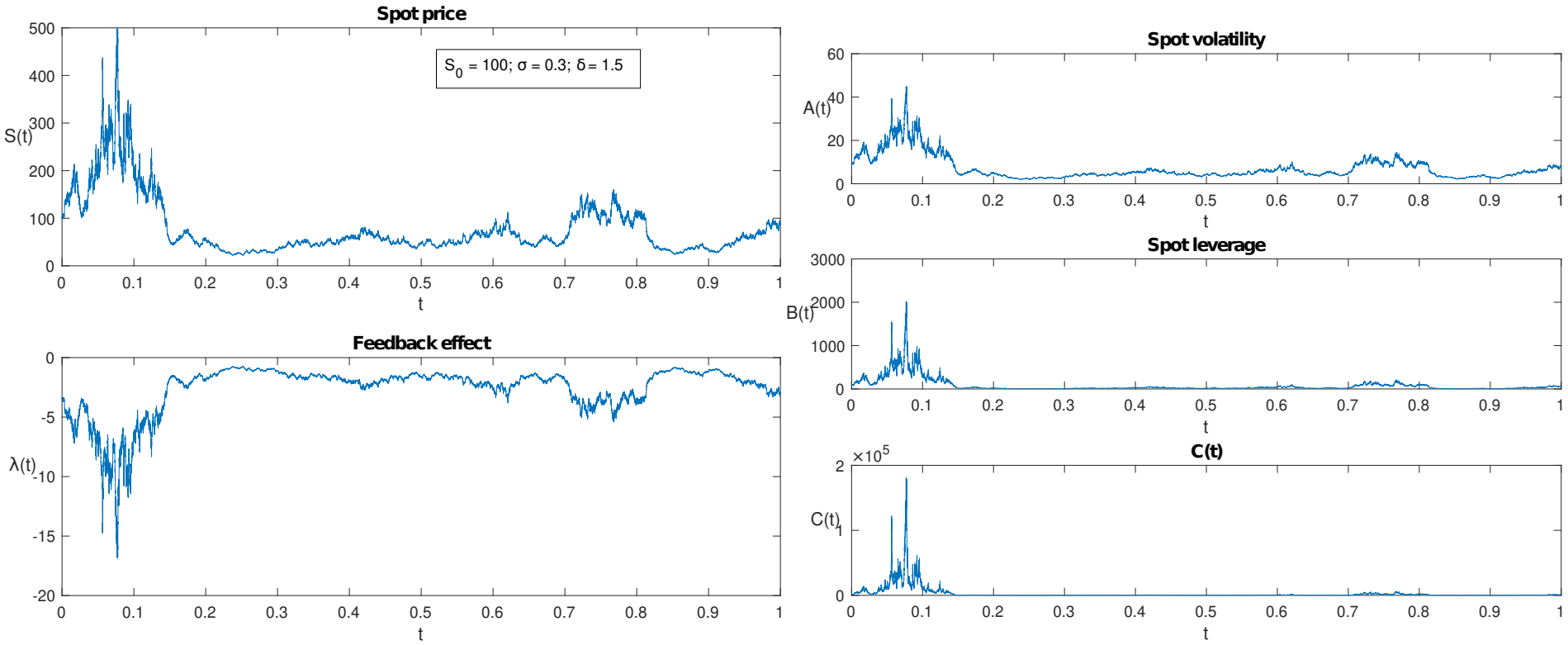


Figura: Analytically computed trajectories of S_t, A_t, B_t, C_t and λ_t in the case of a stable market ($\delta = 1.5$).

Opposite situation: $\delta = 0.5$ and the *feedback rate is positive*

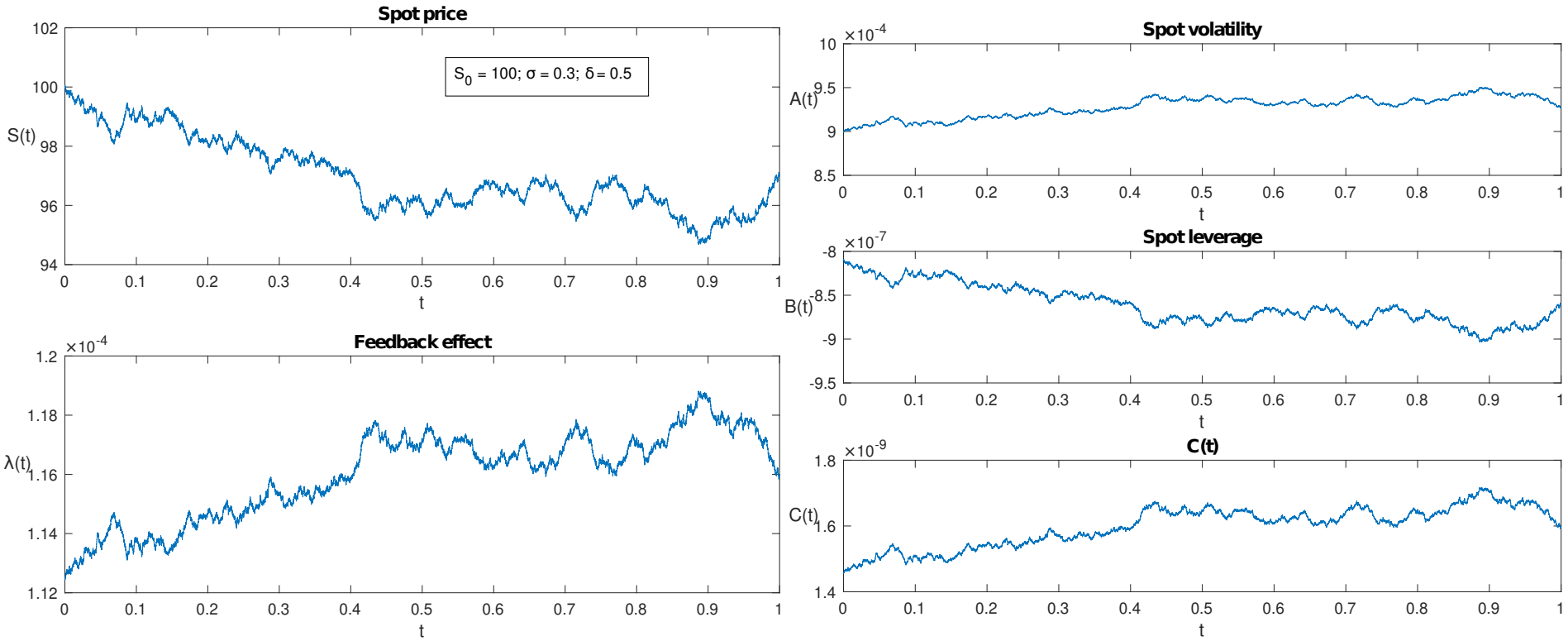


Figura: Analytically computed trajectories of S_t, A_t, B_t, C_t and λ_t in the case of a unstable market ($\delta = 0.5$).

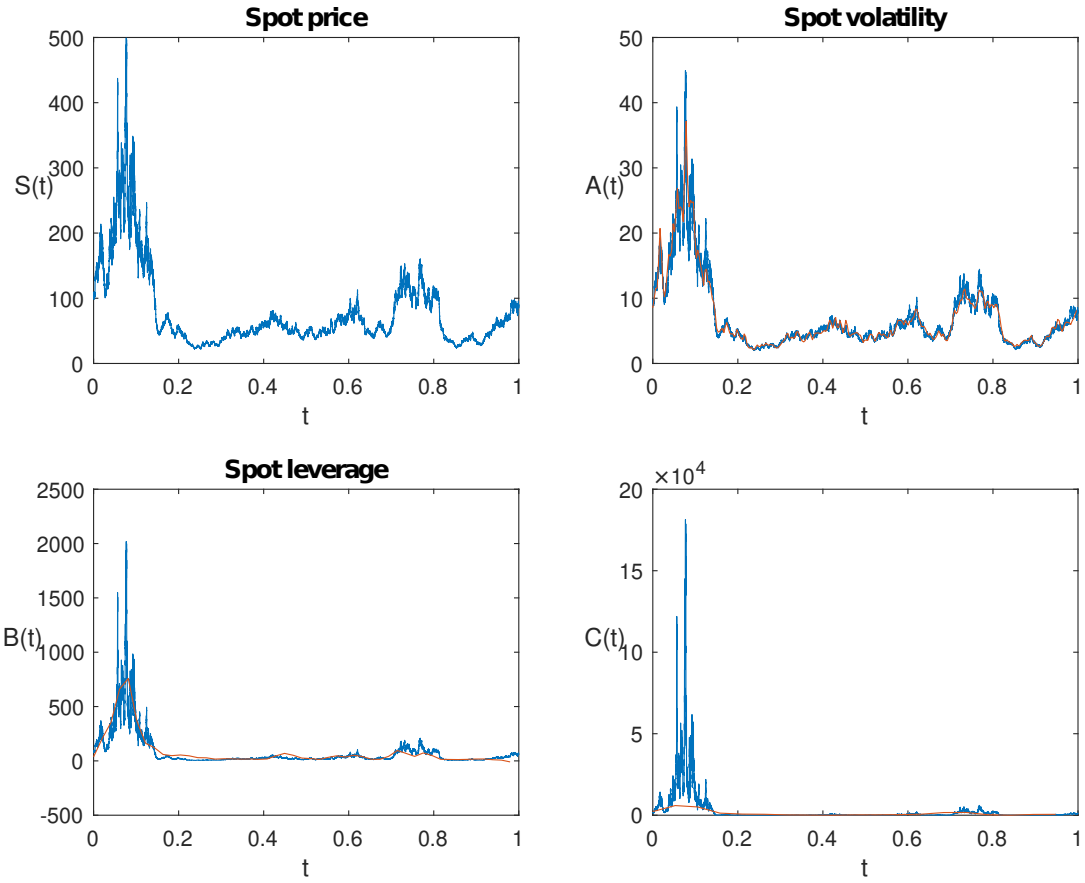


Figura: $\delta = 1.5$: analytically computed spot functions S_t, A_t, B_t, C_t in blue. In red are A_t, B_t, C_t. Cut-off frequencies: at the first level (A_t) set $N = n/2$, $N_A = n^{0.5}$; at the second level (B_t) set $M = N_A$, $M_B = (4N_A)^{0.5}$; at the third level (C_t) set $L = M_B$, $L_C = M_B^{0.5}$.

At each approximation step the resolution is lower, nevertheless the reconstruction of the trajectories is very good.

Comment: Although the approximation of the Fourier estimates of A_t, B_t and C_t is very good, when combining them non-linearly to obtain λ_t, the resulting estimator is usually biased (this fact is well recognized in the non-parametric estimation literature). We consider the daily integrated value $\int_0^T \lambda_\tau d\tau$ as an indicator of market instability. The approximation can be considered good, as it catches the correct negative sign with a hit rate of 88%.

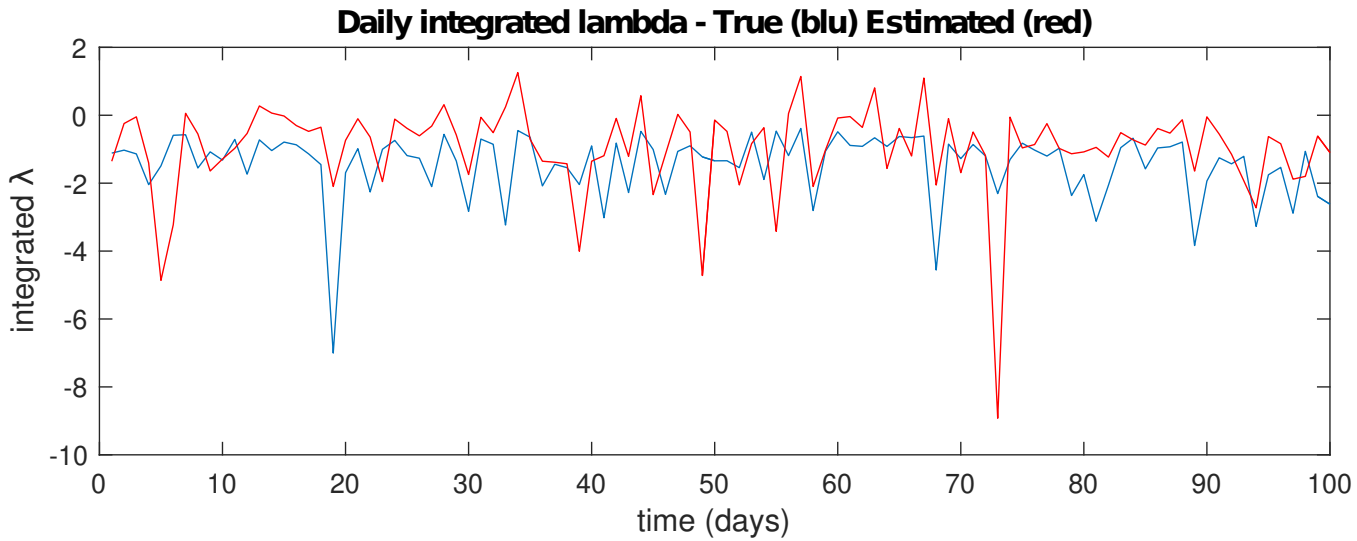


Figura: Real integrated feedback $\int_0^T \lambda_\tau d\tau$ (blue) and its Fourier estimate (red) over 100 days with $\delta = 1.5$.

Comment: Although the approximation of the Fourier estimates of A_t, B_t and C_t is very good, when combining them non-linearly to obtain λ_t, the resulting estimator is usually biased (this fact is well recognized in the non-parametric estimation literature). We consider the daily integrated value $\int_0^T \lambda_\tau d\tau$ as an indicator of market instability. The approximation can be considered good, as it catches the correct negative sign with a hit rate of 88%.

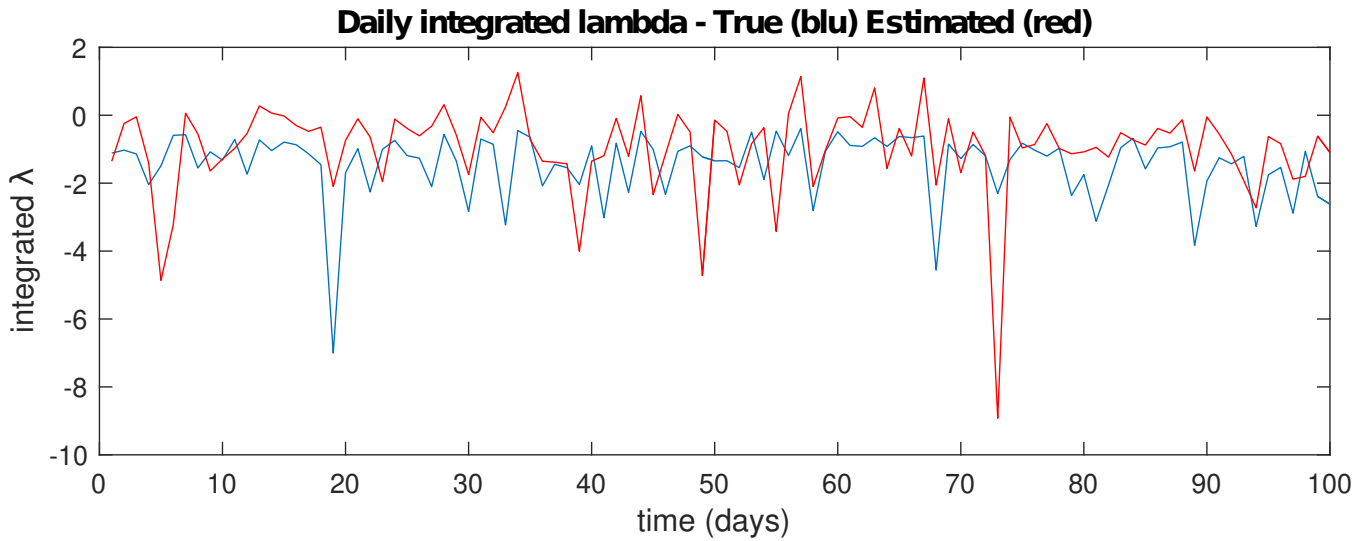


Figura: Real integrated feedback $\int_0^T \lambda_\tau d\tau$ (blue) and its Fourier estimate (red) over 100 days with $\delta = 1.5$.

Empirical analysis

The potentiality of the indicator is tested on a sample of the S&P500 Futures including the following financial market crashes:

- 2000-03-10: NASDAQ Crash (dot-com Bubble)
- 2001-02-19: Turkish Crisis
- 2001-09-11: Twin Tower Attacks
- 2001-12-27: Argentine Default

Those crises are very different in nature to each other and had quite different impact on the S&P500 Futures series.

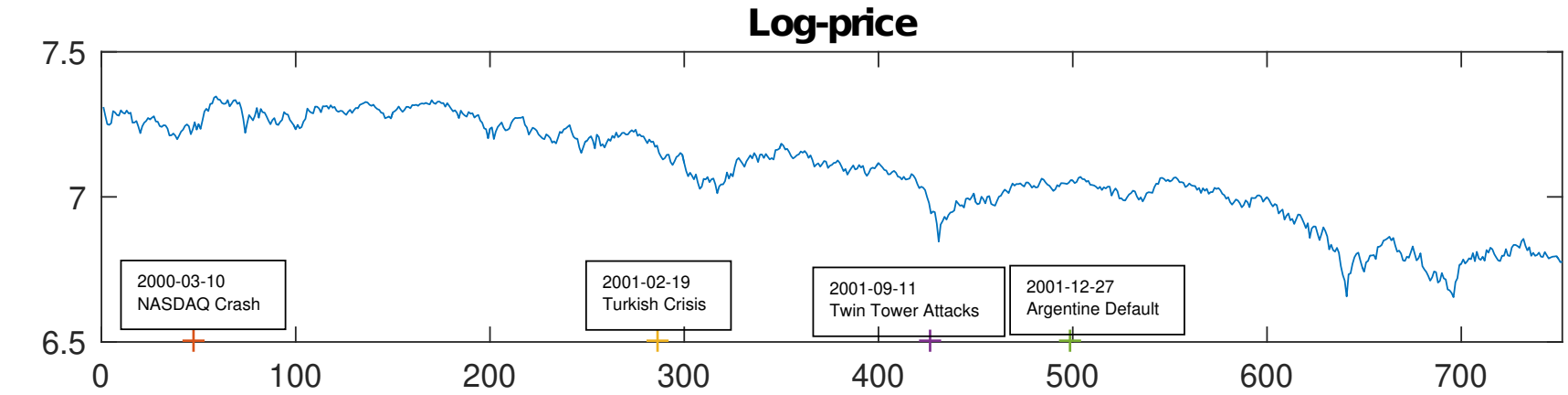
We provide empirical evidence that large values of the feedback rate reveal conditions in the market where perturbations in the price level may evolve in large price declines or changes in general.

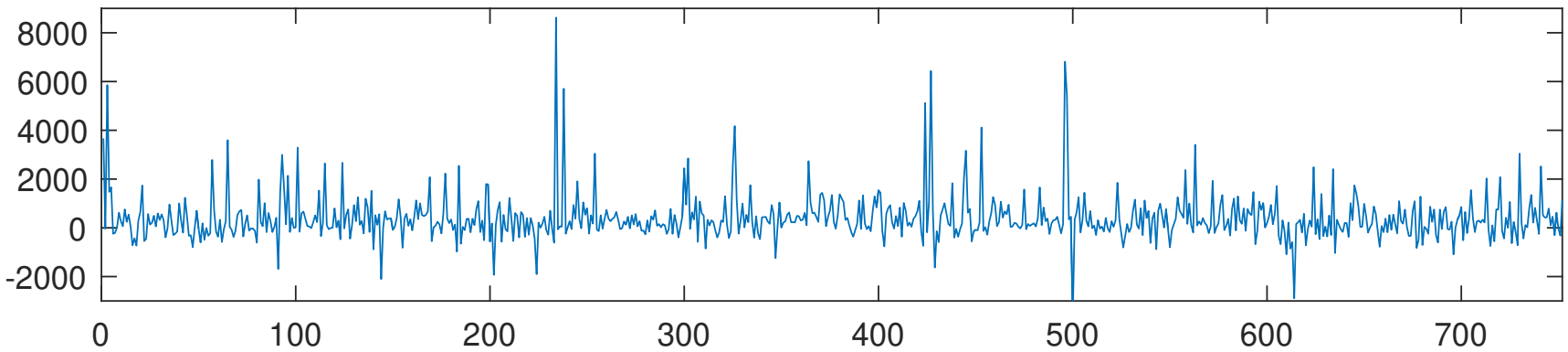
Empirical analysis

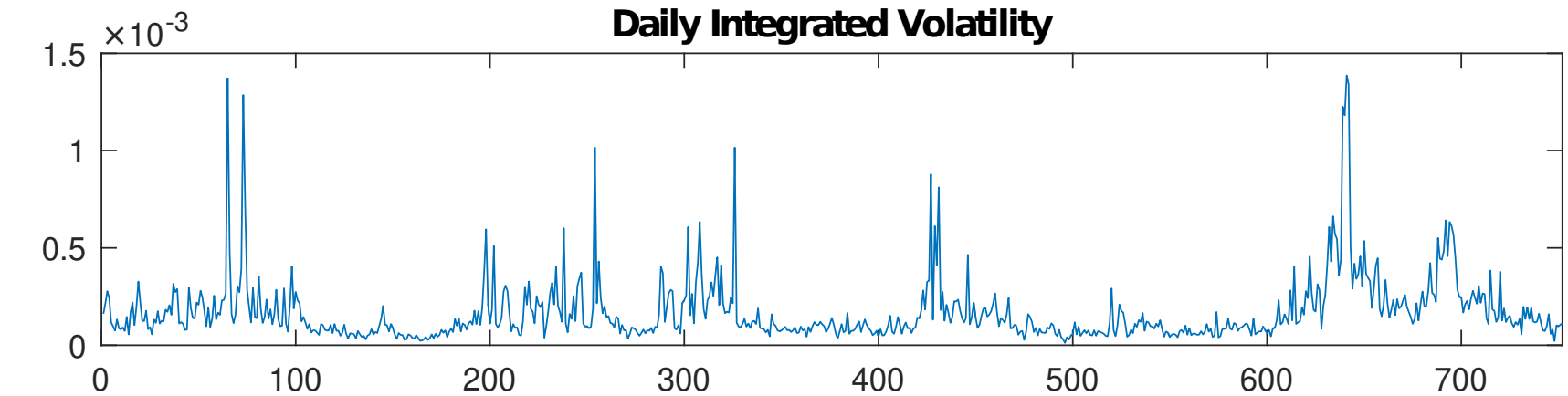
The potentiality of the indicator is tested on a sample of the S&P500 Futures including the following financial market crashes:

- 2000-03-10: NASDAQ Crash (dot-com Bubble)
- 2001-02-19: Turkish Crisis
- 2001-09-11: Twin Tower Attacks
- 2001-12-27: Argentine Default

Those crises are very different in nature to each other and had quite different impact on the S&P500 Futures series.
We provide empirical evidence that large values of the feedback rate reveal conditions in the market where perturbations in the price level may evolve in large price declines or changes in general.


Log-price
7.5
7
6.5
2000-03-10
NASDAQ Crash
2001-02-19
Turkish Crisis
2001-09-11
Twin Tower Attacks
2001-12-27
Argentine Default
0
100
200
300
400
500
600
700


Daily Integrated Feedback rate
8000
6000
4000
2000
0
-2000
0
100
200
300
400
500
600
700


Daily Integrated Volatility
×10-3
1.5
1
0.5
0
0
100
200
300
400
500
600
700

λ_t, A_t, B_t, C_t from day 253 to day 255 (i.e., January 2-4, 2001), during the turbulent period leading to the Turkish crisis, when the largest peak of λ occurs.

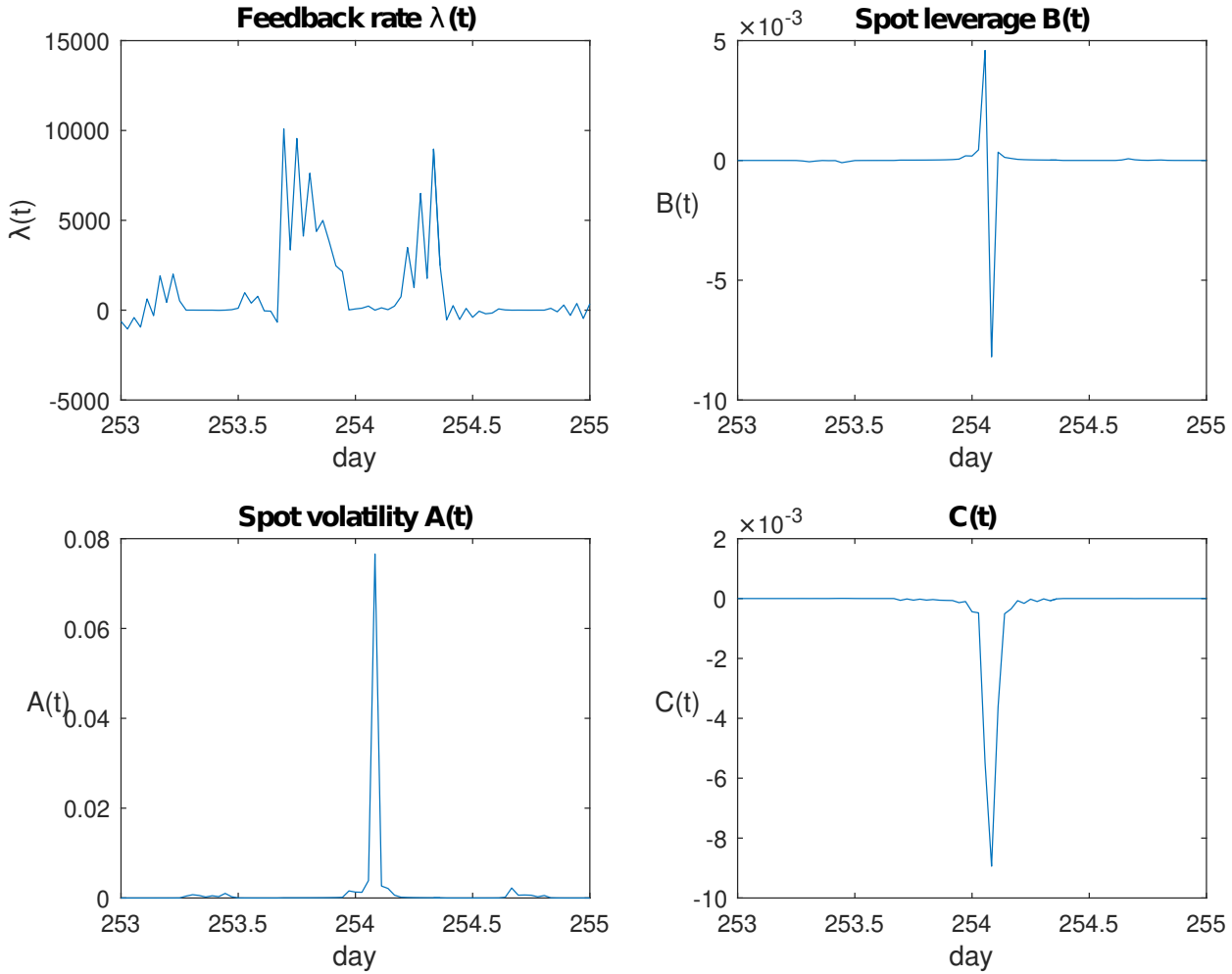


Figura: S&P 500 index futures from January 2nd, 2001 to January 3rd, 2001 (days 253-255 in our sample). λ_t, B_t, A_t and C_t trajectories.

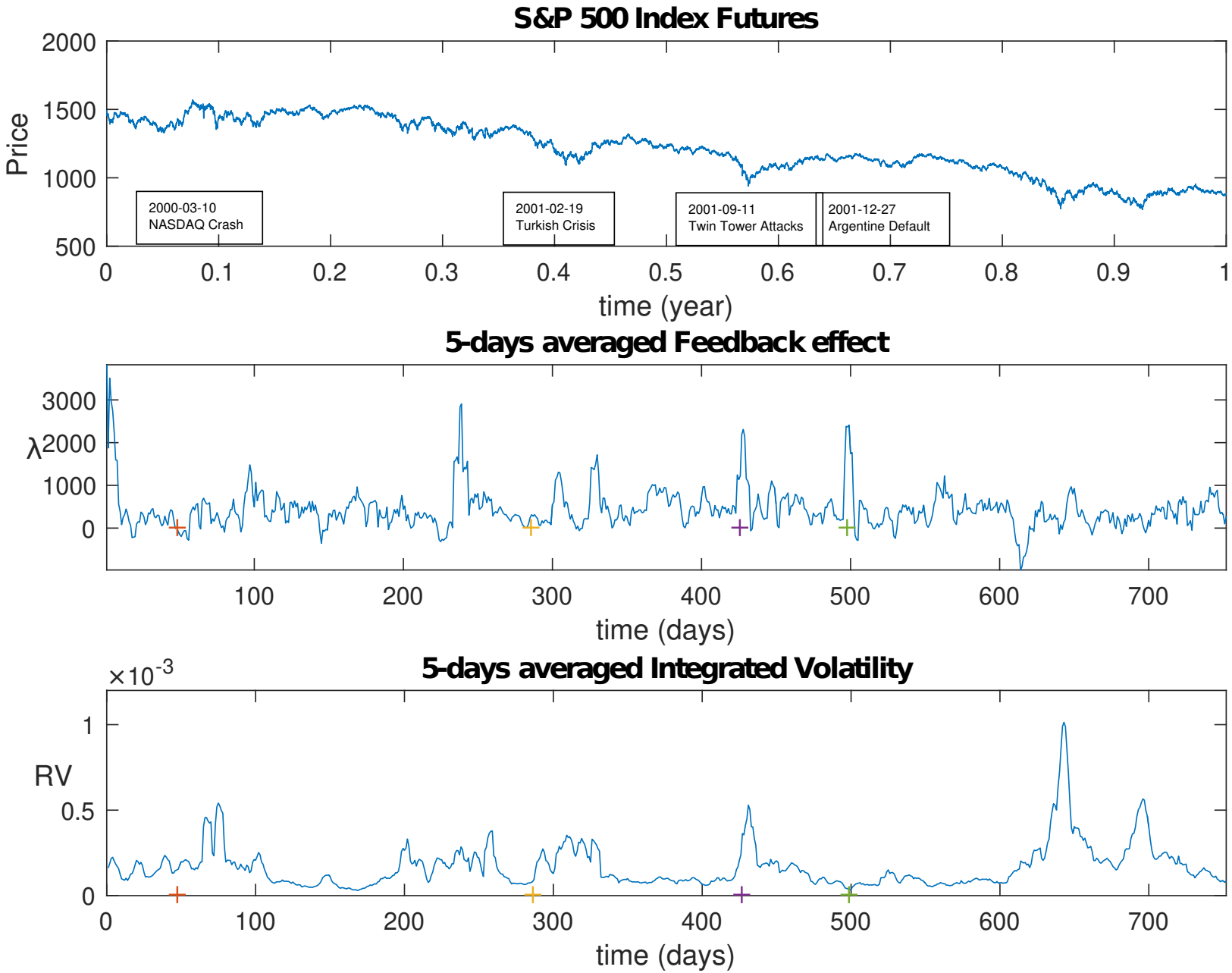


Figura: S&P 500 index futures from January 3rd, 2000 to December 31st, 2002. Upper

References

Allaj, E. and Sanfelici, S. (2021).

An Early Warning System for Identifying Financial Instability. *https* : *//papers.ssrn.com/sol3/papers.cfm?abstract_id* = 3738936

Alos, E., León, J.A. and Vives, J. (2007)

On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11 (4), 571–589. *Journal of Financial Econometrics*, 4, 1–30.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2008)

Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76, 1481–1536.

Bates D. (1996)

Jumps and stochastic volatility: exchange rate processes implicit in Deutschemark options, *Review of Financial Studies*, 9: 69-107.

Bollerslev, T. and Zhang, L. (2003)

Measuring and modeling systematic risk in factor pricing models using high-frequency data. *Journal of Empirical Finance*, **10**, 533–558.

Comte, F. and Renault, E. (1998)

Long memory in continuous time stochastic volatility models. *Math. Finance,* 8: 291-323.

Cuchiero, C. and Teichmann, J. (2015) Fourier transform methods for pathwise covariance estimation in the presence of jumps. *Stochastic Processes and Their Applications*, 125(1):116ñ-160.

Danielsson, J., Valenzuela, M. and Zer, I. (2016).

Learning from History: Volatility and Financial Crises. Finance and Economics Discussion Series 2016-093. Washington: Board of Governors of the Federal Reserve System. *https* : *//doi.org/*10.17016/*FEDS*.2016.093

Davydov, D., and Linetsky, V. (2001).

Pricing and Hedging Path-Dependent Options Under the CEV Process. *Management Science*, 47: 949–65.

Engle, R. and Colacito, R. (2006)

Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24(2), 238–253.

References

Epps, T. (1979)

Comovements in stock prices in the very short run. *Journal of the American Statistical Association,* **74**, 291–298.

Fan, J. and Wang, Y. (2008)

Spot volatility estimation for high frequency data. *Statistics and its Interface*, **1**, 279–288.

Foster, D.P. and Nelson, D.B. (1996)

Continuous record asymptotics for rolling sample variance estimators. *Econometrica*, 64: 139-174.

Fukasawa, M. (2011)

Asymptotic analysis for stochastic volatility: martingale expansion. *Finance Stoch.*, 15, 635ñ-654.

Ghysels, E. and Sinko, A. (2007).

Volatility forecasting and microstructure noise, Working Paper.

Heston S. (1993)

A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies*, 6: 327-343.

Hobson D., Rogers L. (1998)

Complete models with stochastic volatility, *Mathematical Finance*, 8: 27-48.

Hull J. and White A. (1987)

The pricing of options on assets with stcohastic volatilities, *Journal of Finance*, 42: 281-300.

Inkaya, A. and Yolcu Ocur, Y. (2014).

Analysis of volatility feedback and leverage effects on the ISE30 index using frequency data. *Journal of Computational and Applied Mathematics*, 259, 377–384.

Malliavin, P. and Mancino, M.E. (2002).

Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 4, 49–61.

References

Malliavin, P. and Mancino, M.E. (2009).

A Fourier transform method for nonparametric estimation of multivariate volatility. *The Annals of Statistics,* 37 (4), 1983–2010.

Malliavin, P. (1997).

Stochastic Analysis. A Series of Comprehensive Studies in Mathematics. Berlin/Heildelberg: Springer, vol. 313.

Malliavin, P. and Mancino, M.E. (2002)

Instantaneous liquidity rate, its econometric measurement by volatility feedbacks. *Comptes Rendus de l'Academie des Sciences, Paris,* Ser.I 334.

Mancino, M.E. and Sanfelici, S. (2008a)

Robustness of Fourier Estimator of Integrated Volatility in the Presence of Microstructure Noise. *Computational Statistics and Data Analysis*, 52(6), 2966– 2989.

Mancino, M.E. and Sanfelici, S. (2011).

Estimating covariance via Fourier method in the presence of asynchronous trading and microstructure noise. *J. of Fin. Econometrics*, 9(2), 367-408

Mancino, M.E., Oliva, I. and Toscano, G. (2021).

Asymptotic Efficiency for the Fourier Estimator of the Feedback Rate. *In preparation*.

Mancino, M.E. and Toscano, G. (2021).

Rate Efficient Asymptotic for the Fourier Estimator of the Leverage Process. *Statistics and Its Interface*.

Roll, R. (1984).

A simple measure of the bid-ask spread in an efficient market. *Journal of Finance*, 39, 1127–1139.

Stein E., Stein J. (1991)

Stock price distributions with stochastic volatility: an analytic approach, *Review of Financial Studies*, 4: 727-752.

Zhou, B. (1996)

High frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics,* **14**(1), 45–52.