Identifying financial instability using high frequency data

S. Sanfelici

Department of Economics and Business
University of Parma, Italy
simona.sanfelici@unipr.it

Intensive Research Programme on Quantitative Finance
Centre de Recerca Matematica, Campus de Bellaterra, Barcelona
July 8th, 2021

References

[1] Allaj, E. and Sanfelici, S., (2021) *An Early Warning System for Identifying financial instability*. Submitted.

[2] Mancino, M.E. and Sanfelici, S., (2020) *Identifying financial instability conditions using high frequency data*. J. Economic Interaction and Coordination, 15(1), 221-242.

[3] Malliavin P. and Mancino, M.E., (2002) *Instantaneous liquidity rate, its econometric measurement by volatility feedbacks*. Comptes Rendus de l'Academie des Sciences, Paris.

Introduction

Financial crises prediction is an essential topic in finance. We propose an Early Warning Indicator (EWI) for predicting possible financial crises or, more generally, market instability conditions. Our system is based on the so called *price-volatility feedback rate* [3], which is supposed to describe the ease of the market in absorbing small price perturbations.

The indicator combines non-linearly *volatility*, *leverage* and *covariance between leverage and price* and is model-free.
The rate of variation through time of an initial perturbation of a given high frequency financial time series enables us to understand if such a shock will be rapidly absorbed or, on the contrary, it will be amplified by the market.

Introduction

- Recent theoretical and empirical studies on volatility modeling have pointed out the existence of volatility feedback effects.

Introduction

- Recent theoretical and empirical studies on volatility modeling have pointed out the existence of volatility feedback effects.
- Volatility feedback and leverage effects are related to the same phenomena: the *leverage effect* explains why a negative return causes an increase in the volatility and was first discussed in (Black, 1976) and (Christie, 1982); conversely, the notion of *volatility feedback effect* is based on the argument that volatility is priced and an increase in the volatility raises the required return on the asset, which can only be produced by an immediate decline in the asset price as observed by (French et al., 1987).

Introduction

- Recent theoretical and empirical studies on volatility modeling have pointed out the existence of volatility feedback effects.
- Volatility feedback and leverage effects are related to the same phenomena: the *leverage effect* explains why a negative return causes an increase in the volatility and was first discussed in (Black, 1976) and (Christie, 1982); conversely, the notion of *volatility feedback effect* is based on the argument that volatility is priced and an increase in the volatility raises the required return on the asset, which can only be produced by an immediate decline in the asset price as observed by (French et al., 1987).
- The volatility feedback rate is a second order quantity which is supposed to describe the facility for the market to absorb small perturbations and can be used to explain the irregular behavior and instability of financial markets.

Outline of the talk

- We present our EWI of financial instability based on the computation of the decay rate for the propagation of a given market shock.

Outline of the talk

- We present our EWI of financial instability based on the computation of the decay rate for the propagation of a given market shock.
- We resume some consistency results and other properties of the indicator under the CEV model, investigated in [2].

Outline of the talk

- We present our EWI of financial instability based on the computation of the decay rate for the propagation of a given market shock.
- We resume some consistency results and other properties of the indicator under the CEV model, investigated in [2].
- A logit regression based Early Warning System is employed to predict future financial crises. Our study conducted in [1] on the S&P 500 index futures reveals that, while the RV may sometimes fail in predicting crises, the EWI employing the price-volatility feedback rate is always an important predictor of financial instability.

The model

- Let p_t be the logarithmic asset price process. For simplicity, assume that the drift coefficient is zero (i.e., we work under the risk neutral measure). Suppose p_t satisfies the model

$$dp_t = \sigma(p_t)\ dW_t - \frac{1}{2}\sigma^2(p_t)dt.$$

The model

- Let p_t be the logarithmic asset price process. For simplicity, assume that the drift coefficient is zero (i.e., we work under the risk neutral measure). Suppose p_t satisfies the model

$$dp_t = \sigma(p_t)\ dW_t - \frac{1}{2}\sigma^2(p_t)dt.$$

- Consider an infinitesimal perturbation $p_t + \varepsilon\zeta_t$ of the asset price. The pathwise sensitivity ζ_t is solution to the linearized stochastic differential equation

$$d\zeta_t = \zeta_t\,(\sigma'(p_t)\ dW_t - \sigma'(p_t)\sigma(p_t)dt\,).$$

The model

- Let p_t be the logarithmic asset price process. For simplicity, assume that the drift coefficient is zero (i.e., we work under the risk neutral measure). Suppose p_t satisfies the model

$$dp_t = \sigma(p_t)\ dW_t - \frac{1}{2}\sigma^2(p_t)dt.$$

- Consider an infinitesimal perturbation $p_t + \varepsilon\zeta_t$ of the asset price. The pathwise sensitivity ζ_t is solution to the linearized stochastic differential equation

$$d\zeta_t = \zeta_t\,(\sigma'(p_t)\ dW_t - \sigma'(p_t)\sigma(p_t)dt\,).$$

- We associate to ζ_t the *rescaled variation* defined as

$$z_t := \frac{\zeta_t}{\sigma_t}, \quad \text{where } \sigma_t = \sigma(p_t).$$

The model

- By applying Itô formula, in [3] it is proved that the rescaled variation is a differentiable function with respect to t and, for any $s < t$, it can be expressed as

$$z_t = z_s \; \exp\Big(\int_s^t \; \lambda_\tau \; d\tau\Big),$$

where

$$\lambda_t := -\frac{1}{2}(\sigma_t \sigma_t' + \sigma_t \sigma_t'').$$

The model

- By applying Itô formula, in [3] it is proved that the rescaled variation is a differentiable function with respect to t and, for any $s < t$, it can be expressed as

$$z_t = z_s \ \exp\Big(\int_s^t \ \lambda_\tau \ d\tau\Big),$$

where

$$\lambda_t := -\frac{1}{2}(\sigma_t \sigma_t' + \sigma_t \sigma_t'').$$

- The function λ_t is called the price-volatility feedback rate.

The model

- By applying Itô formula, in [3] it is proved that the rescaled variation is a differentiable function with respect to t and, for any $s < t$, it can be expressed as

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

where

$$\lambda_t := -\frac{1}{2}(\sigma_t \sigma_t' + \sigma_t \sigma_t'').$$

- The function λ_t is called the price-volatility feedback rate.
- The price-volatility feedback rate can be understood as the appreciation rate of the rescaled variation.

The model

- By applying Itô formula, in [3] it is proved that the rescaled variation is a differentiable function with respect to t and, for any $s < t$, it can be expressed as

$$z_t = z_s \ \exp(\int_s^t \ \lambda_\tau \ d\tau),$$

where

$$\lambda_t := -\frac{1}{2}(\sigma_t\sigma_t' + \sigma_t\sigma_t'').$$

- The function λ_t is called the price-volatility feedback rate.
- The price-volatility feedback rate can be understood as the appreciation rate of the rescaled variation.
- *Remark.* Note that in the Black-Scholes framework it holds $\lambda = 0$.

The model

- Large positive values of λ indicate market instability, while negative values corresponds to stable market directions:
  - a negative λ would witness a period of stability, because $z_t \to 0$ as $t \to +\infty$;
  - a positive λ would signal instability.

The model

- Large positive values of λ indicate market instability, while negative values corresponds to stable market directions:
  - a negative λ would witness a period of stability, because $z_t \to 0$ as $t \to +\infty$;
  - a positive λ would signal instability.
- Large positive values of the feedback rate usually anticipate a significant decrease in the price level.

The model

- Large positive values of λ indicate market instability, while negative values corresponds to stable market directions:
  - a negative λ would witness a period of stability, because $z_t \to 0$ as $t \to +\infty$;
  - a positive λ would signal instability.
- Large positive values of the feedback rate usually anticipate a significant decrease in the price level.
- Values of λ around zero imply that the price level is remains stable.

The model

- 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.

The model

- 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.
- 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.

The model

- λ is a second order quantity whose estimation from observed price paths is challenging.

The model

- λ is a second order quantity whose estimation from observed price paths is challenging.
- It would be important to estimate the volatility feedback rate without assuming the knowledge of an explicit expression of the volatility function.

The model

- λ is a second order quantity whose estimation from observed price paths is challenging.
- It would be important to estimate the volatility feedback rate without assuming the knowledge of an explicit expression of the volatility function.
- This is made possible by using iterated cross-volatilities that can be estimated using the Fourier estimation method.

- Malliavin and Mancino (2002), *Fourier Series method for measurement of Multivariate Volatilities*, Finance and Stochastics.
- Malliavin and Mancino (2009), *A Fourier transform method for nonparametric estimation of volatility*, The Annals of Statistics.

Advertising

A gentle introduction to the Fourier Malliavin methodology

Mancino, Recchioni, and Sanfelici, (2017) *Fourier-Malliavin Volatility Estimation: Theory and Practice*, Springer Briefs in Quantitative Finance Series.

Computation of λ

The important result of [3] is that

$$\lambda_\tau = -\frac{1}{2}(\sigma'(p_\tau)\sigma(p_\tau) + \sigma''(p_\tau)\sigma(p_\tau)),$$

where $p_t = \log S_t$ can be expressed by means of terms which are iterated cross volatilities

Theorem

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

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

Then the feedback rate function λ_t has the following expression

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

Computation of λ

Assume that the variance and the leverage processes are given by

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

where W_t^A and W_t^B are Brownian motions, possibly correlated with W_t.

Computation of λ

Assume that the variance and the leverage processes are given by

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

where W_t^A and W_t^B are Brownian motions, possibly correlated with W_t.

Define, for $|k| \leq N$,

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

where

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

Computation of λ

Assume that the variance and the leverage processes are given by

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

where W_t^A and W_t^B are Brownian motions, possibly correlated with W_t.

Define, for $|k| \leq N$,

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

where

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

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

Computation of λ

Similarly

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

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

Computation of λ

Similarly

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

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

The statistical properties (consistency) of the estimators of B, C and λ are studied in [1].

Computation of λ

Similarly

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

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

The statistical properties (consistency) of the estimators of B, C and λ are studied in [1].

Very important: the choice of the cutting frequencies.

Computation of λ

Similarly

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

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

The statistical properties (consistency) of the estimators of B, C and λ are studied in [1].
Very important: the choice of the cutting frequencies. $N/n \to 1/2$, $M^2/n \to 0$, $L^2 M^2/N \to 0$ for $n, N, M \to \infty$.

Computation of λ

The estimation of these higher order moments is admittedly challenging and requires the use of **high frequency data**. The use of the Fourier estimator of co-variation allows us to deal with **microstructure noise effects** that contaminate high frequency data. In fact, the Fourier estimator needs no correction in order to be statistically efficient and robust to various type of market frictions.

Computation of λ

The estimation of these higher order moments is admittedly challenging and requires the use of **high frequency data**. The use of the Fourier estimator of co-variation allows us to deal with **microstructure noise effects** that contaminate high frequency data. In fact, the Fourier estimator needs no correction in order to be statistically efficient and robust to various type of market frictions.

Remark: When we combine estimates of A_t, B_t and C_t non-linearly to obtain λ_t, the resulting estimator is usually biased. Furthermore, the potential oscillations in A_t, B_t and C_t are amplified and the resulting approximation of λ_t is very unstable. Therefore, in order to build a daily indicator of market instability, we consider the integrated value $\int_I \lambda_t dt$, where I is the trading period, and **we combine daily integrated quantities to get a more stable indicator** of market instability.

Constant Elasticity of Variance model

$$dS(t) = \sigma S^{\delta}(t) dW_t, \quad S(0) = s_0$$

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

Constant Elasticity of Variance model

$$dS(t) = \sigma S^{\delta}(t)dW_t, \qquad S(0) = s_0$$

- we obtain an analytical formula for the indicator;
- we use this explicit formula to perform a simulated analysis showing the effectiveness in estimating λ and the relation occurring between the feedback rate sign and the price variations.
- The first variation process is solution to

$$\begin{aligned} &d\zeta(t) = \zeta(t)\ \sigma\delta S^{\delta-1}(t)dW_t \\ &\zeta(0) = 1. \end{aligned}$$

Constant Elasticity of Variance model

$$dS(t) = \sigma S^{\delta}(t) dW_t, \qquad S(0) = s_0$$

- we obtain an analytical formula for the indicator;
- we use this explicit formula to perform a simulated analysis showing the effectiveness in estimating λ and the relation occurring between the feedback rate sign and the price variations.
- The first variation process is solution to

$$\begin{aligned} d\zeta(t) &= \zeta(t)\ \sigma\delta S^{\delta-1}(t) dW_t \\ \zeta(0) &= 1. \end{aligned}$$

- The logarithmic-price $p_t := \ln S_t$ satisfies

$$dp_t = \sigma e^{(\delta-1)p_t} dW_t - \frac{1}{2}\sigma^2 e^{2(\delta-1)p_t} dt.$$

Constant Elasticity of Variance model

- It is possible to compute the feedback rate explicitly

$$\lambda_t = -\frac{1}{2}\sigma^2\delta(\delta-1)S_t^{2(\delta-1)} =$$

$$= -\frac{1}{2}\sigma^2\delta(\delta-1)e^{2(\delta-1)p_t} = -\frac{1}{2}\delta(\delta-1)A_t.$$

Constant Elasticity of Variance model

- It is possible to compute the feedback rate explicitly

$$\lambda_t = -\frac{1}{2}\sigma^2\delta(\delta-1)S_t^{2(\delta-1)} =$$

$$= -\frac{1}{2}\sigma^2\delta(\delta-1)e^{2(\delta-1)p_t} = -\frac{1}{2}\delta(\delta-1)A_t.$$

Remark:

if $0 < \delta < 1$ then $\lambda_t > 0$;

if $\delta = 1$ then $\lambda_t = 0$ (Black-Scholes);

if $\delta > 1$ then $\lambda_t < 0$.

Constant Elasticity of Variance model

Forward Euler method over $[0,1]$ with $n = 21,600$ equispaced nodes. Effect of a 5% perturbation of the initial price $S_0 = 100$.

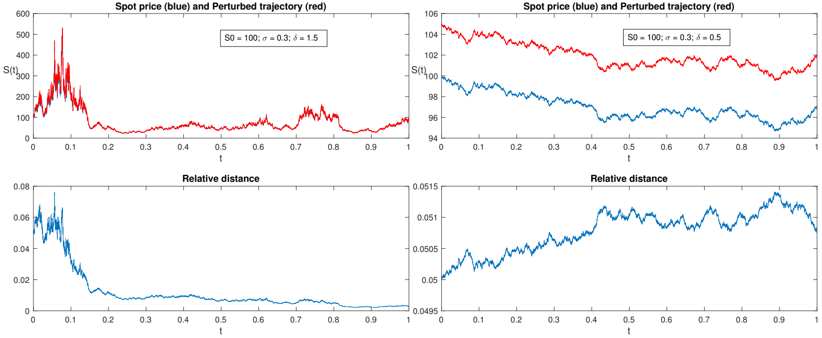


Figure: Left panels: spot price trajectory (blue) and perturbed trajectory (red) in the upper panel and relative distance between the two trajectories in the lower panel. Parameter values: $S_0 = 100$, $\sigma = 0.3$, $\delta = 1.5$. Right panels: same format with $\delta = 0.5$.

Constant Elasticity of Variance model

Case $\delta > 1$: The feedback rate λ_t is negative, the leverage B_t positive, the volatility A_t gets larger as the price increases, and viceversa (*inverse leverage effect*).

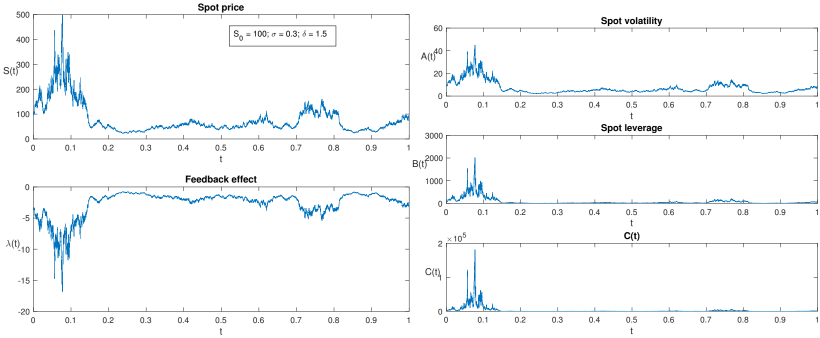


Figure: Analytically computed trajectories of S_t, A_t, B_t, C_t and λ_t in the case of a stable market ($\delta = 1.5$).

Constant Elasticity of Variance model

Case $0 < \delta < 1$**:** The feedback rate λ_t is positive, the leverage B_t negative, the volatility A_t gets larger as the price decreases, and viceversa (*leverage effect*).

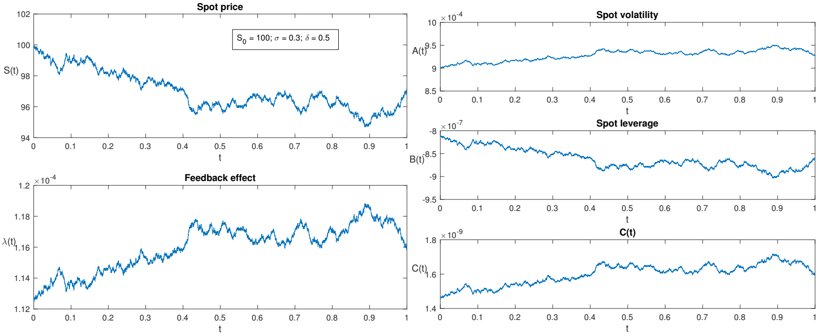
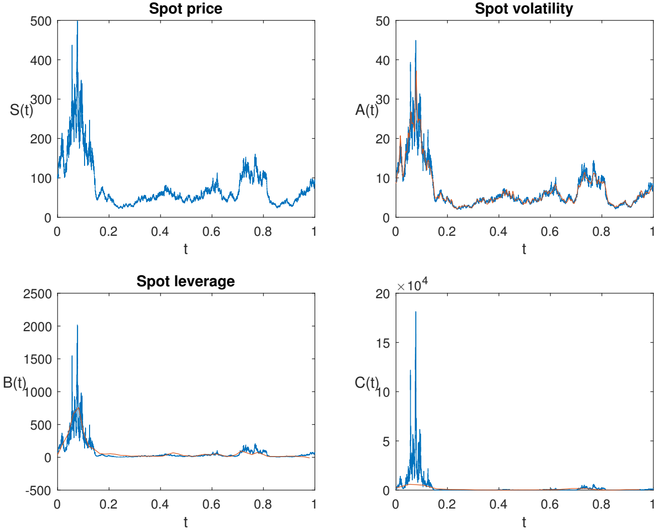


Figure: Analytically computed trajectories of S_t, A_t, B_t, C_t and λ_t in the case of a unstable market ($\delta = 0.5$).

Constant Elasticity of Variance model

Analytical S_t, A_t, B_t, C_t in blue versus Estimated values in red



Constant Elasticity of Variance model

- The choice of the cut-off frequencies is crucial for the quality of the estimation.

Constant Elasticity of Variance model

- The choice of the cut-off frequencies is crucial for the quality of the estimation.
- At the first level we set $N = n/2$, $N_A = n^{0.5}$; at the second level we set $M = N_A$, $M_B = (4N_A)^{0.5}$; at the third level we set $L = M_B$, $L_C = M_B^{0.5}$.

Constant Elasticity of Variance model

- The choice of the cut-off frequencies is crucial for the quality of the estimation.
- At the first level we set $N = n/2$, $N_A = n^{0.5}$; at the second level we set $M = N_A$, $M_B = (4N_A)^{0.5}$; at the third level we 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.

Constant Elasticity of Variance model

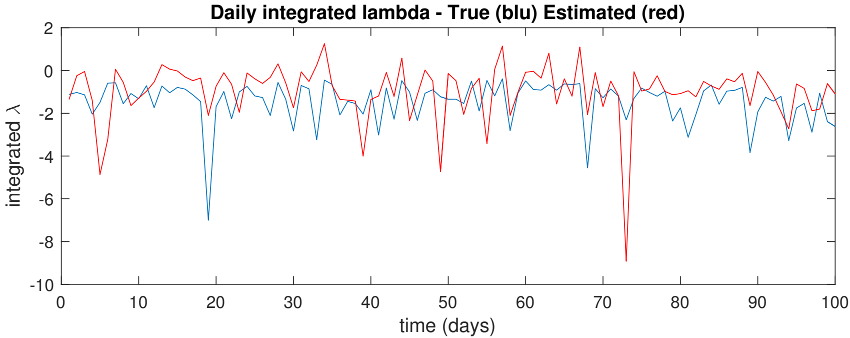


Figure: Real integrated feedback $\int_0^T \lambda_\tau d\tau$ (blue) and its Fourier estimate (red) over 100 days with $\delta = 1.5$.

The approximation can be considered good, as it catches the correct negative sign with a hit rate of 88%.

Testing the feedback rate as EWI

- The volatility feedback rate can be used as an EWI that can help to predict whether large price variations of a given asset or index are likely to occur within a specific time horizon.

Testing the feedback rate as EWI

- The volatility feedback rate can be used as an EWI that can help to predict whether large price variations of a given asset or index are likely to occur within a specific time horizon.
- To test the efficacy of this indicator, we use a *logit regression model* where different predictors or EWIs based on the realized variance (RV) and the integrated feedback rate are considered.

Testing the feedback rate as EWI

- The volatility feedback rate can be used as an EWI that can help to predict whether large price variations of a given asset or index are likely to occur within a specific time horizon.
- To test the efficacy of this indicator, we use a *logit regression model* where different predictors or EWIs based on the realized variance (RV) and the integrated feedback rate are considered.
- The dependent variable is an indicator of financial crises. It is defined in terms of losses on daily returns exceeding a given threshold measured by the *Value-at-risk* (VaR) or, alternatively, in terms of average returns conditional on the event that the losses on daily returns exceeding the VaR is less than the *Expected shortfall* (ES).

Testing the feedback rate as EWI

- The volatility feedback rate can be used as an EWI that can help to predict whether large price variations of a given asset or index are likely to occur within a specific time horizon.
- To test the efficacy of this indicator, we use a *logit regression model* where different predictors or EWIs based on the realized variance (RV) and the integrated feedback rate are considered.
- The dependent variable is an indicator of financial crises. It is defined in terms of losses on daily returns exceeding a given threshold measured by the *Value-at-risk* (VaR) or, alternatively, in terms of average returns conditional on the event that the losses on daily returns exceeding the VaR is less than the *Expected shortfall* (ES).
- Our dataset includes tick-by-tick prices of the S&P 500 index futures covering the period from 3 January 2000 to 31 December 2002 and a total of 752 trading days.

Crisis indicators

Realized Variance

$$\sum_{j=0}^{n-1}(p_{j+1}-p_j)^2, \text{ using intraday data.}$$

Crisis indicators

Realized Variance

$$\sum_{j=0}^{n-1}(p_{j+1}-p_j)^2, \text{ using intraday data.}$$

Let $R_t = P_{t+1} - P_t$ be the daily return over the time interval $[t, t+1]$, $t = 1, 2, ..., T-1$.

Value at Risk

The VaR of the loss $L_t = -R_t$ at the confidence level $1-\alpha$ is defined by the following quantile

$$VaR_{1-\alpha}(L_t) = -\inf\{r \in \mathbb{R} : \mathbb{P}(R_t \leq r) > \alpha\}.$$

Crisis indicators

Realized Variance

$$\sum_{j=0}^{n-1}(p_{j+1}-p_j)^2, \text{ using intraday data.}$$

Let $R_t = P_{t+1} - P_t$ be the daily return over the time interval $[t, t+1]$, $t = 1, 2, ..., T-1$.

Value at Risk

The VaR of the loss $L_t = -R_t$ at the confidence level $1-\alpha$ is defined by the following quantile

$$VaR_{1-\alpha}(L_t) = -\inf\{r \in \mathbb{R} : \mathbb{P}(R_t \leq r) > \alpha\}.$$

VaR measures the maximum loss in the value of an asset over the time interval $[t, t+1]$ for a fixed confidence level, giving thus an indication of the risk of loss for a given asset.

Value at Risk

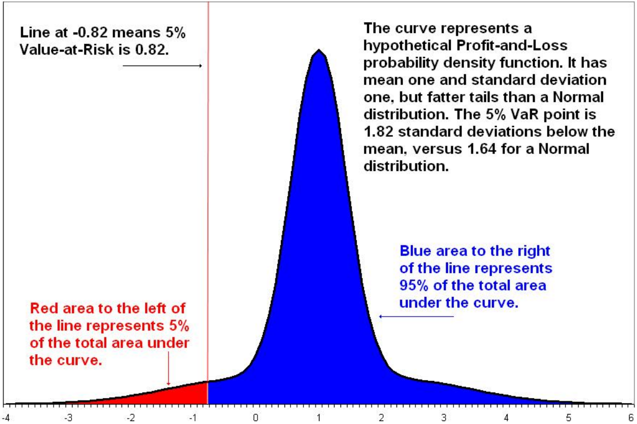


Figure: Probability density function.

Expected shortfall

Another classical risk measure used to assess the risk of an investment is the ES

Expected shortfall

The ES of L_t at confidence level $1-\alpha$ over the time interval t to $t+1$ can be defined as

$$ES_{1-\alpha}(L_t) = \mathbb{E}[L_t | L_t > VaR_{1-\alpha}(L_t)].$$

Expected shortfall

Another classical risk measure used to assess the risk of an investment is the ES

Expected shortfall

The ES of L_t at confidence level $1-\alpha$ over the time interval t to $t+1$ can be defined as

$$ES_{1-\alpha}(L_t) = \mathbb{E}[L_t | L_t > VaR_{1-\alpha}(L_t)].$$

ES measures with a certain confidence level and over a given interval t to $t+1$ the expected loss in the value of an asset when the loss is greater or equal than the VaR risk measure.

Crisis indicators

- $CI_t = \begin{cases} 1 & \text{if} \quad R_t < -VaR_{1-\alpha} \\ 0 & \text{else} \end{cases}$
  "A crisis at day $t+1$ occurs when the returns over the interval $[t, t+1]$ is less than the VaR"

Crisis indicators

- $CI_t = \begin{cases} 1 & \text{if} \quad R_t < -VaR_{1-\alpha} \\ 0 & \text{else} \end{cases}$
  "A crisis at day $t+1$ occurs when the returns over the interval $[t, t+1]$ is less than the VaR"

- $CEI_M = \begin{cases} 1 & \text{if} \quad A_M < -ES_{1-\alpha} \\ 0 & \text{else,} \end{cases}$

  where
  $$A_M = \frac{\sum_{j=M-T_k+1}^{M} R_j \mathbf{1}_{\{-R_j > VaR^0_{1-\alpha}\}}}{\sum_{j=M-T_k+1}^{M} \mathbf{1}_{\{-R_j > VaR^0_{1-\alpha}\}}}$$
  "A crisis at day $M+1$ occurs whenever in the last T_k days the average return conditional on the event that the (negative) returns exceed $VaR_{1-\alpha}$ is less than $ES_{1-\alpha}$"

Forward-looking variable

- $$Y_s = \begin{cases} 1 & \text{if} \quad \exists \quad k = 1, 2, ..., T_m \quad \text{s.t.} \quad CI_{s+k-1} = 1 \\ 0 & \text{else} \end{cases},$$
  where $s = T_w, T_w + 1, ..., T - T_m$.
  "A crisis at day s occurs whenever the indicator CI is equal to one in one of the next T_m trading days"

Forward-looking variable

- $Y_s = \begin{cases} 1 & \text{if} \quad \exists \quad k = 1, 2, ..., T_m \quad \text{s.t.} \quad CI_{s+k-1} = 1 \\ 0 & \text{else} \end{cases}$,
  where $s = T_w, T_w + 1, ..., T - T_m$.
  "A crisis at day s occurs whenever the indicator CI is equal to one in one of the next T_m trading days"

- $Y_s = \begin{cases} 1 & \text{if} \quad CEI_{s+T_m-1} = 1 \\ 0 & \text{else} \end{cases}$
  where again $s = T_w, T_w + 1, ..., T - T_m$ *"A crisis at day s whenever the indicator CEI takes the value of one at day $s + T_m - 1$"*

Logit model

We suppose the probability of having financial crises can be modelled through a logit regression model

$$\mathbb{P}(Y_s = 1) = \frac{\exp\left(\beta_0 + \mathbf{x}_s \boldsymbol{\beta}\right)}{1 + \exp\left(\beta_0 + \mathbf{x}_s \boldsymbol{\beta}\right)}$$

where $\mathbf{x}_s$ is a vector of possible regressors and $\mathbf{x}_s \boldsymbol{\beta} = x_{1,s}\beta_1 + x_{2,s}\beta_2, \dots, x_{k,s}\beta_k$.

First and second regressions

- $\mathbf{x}_s = \ln RV_s$ and Y_s given by the two forward-looking variables.
- RV estimated with sparse sampling using 1, 5, 10 or 15-minutes returns.
- $T_w = 1$ and $T_m = 22$ or $T_m = 33$.

Logit model

Define a new variable

$$Y_{\lambda,i} = \begin{cases} \overline{\Lambda}_{n,N,M,L} & \text{if} \quad \Lambda^{i}_{n,N,M,L} > \overline{\Lambda}_{n,N,M,L} \\ 0 & \text{else} \end{cases}$$

where $\Lambda^{i}_{n,N,M,L}$ gives the daily feedback rate at day i and $\overline{\Lambda}_{n,N,M,L}$ the sample mean of the daily $\Lambda_{n,N,M,L}$ computed over the whole time period.

Logit model

Define a new variable

$$Y_{\lambda,i} = \begin{cases} \overline{\Lambda}_{n,N,M,L} & \text{if} \quad \Lambda^{i}_{n,N,M,L} > \overline{\Lambda}_{n,N,M,L} \\ 0 & \text{else} \end{cases}$$

where $\Lambda^{i}_{n,N,M,L}$ gives the daily feedback rate at day i and $\overline{\Lambda}_{n,N,M,L}$ the sample mean of the daily $\Lambda_{n,N,M,L}$ computed over the whole time period.

Then, define the regressors $\lambda_{av,s} = \frac{\sum_{i=s-T_w+1}^{s} Y_{\lambda,i}}{T_w}$ and $RV_{av,s} = \frac{\sum_{i=s-T_w+1}^{s} \ln RV_i}{T_w}$, where $s = T_w, T_w + 1, \ldots, T - T_m$.

Third and fourth regressions

- $\mathbf{x}_s = RV_{av,s}$ or $\mathbf{x}_s = \lambda_{av,s}$ and Y_s given by the two forward-looking variables.
- RV estimated with sparse sampling using 1, 5, 10 or 15-minutes returns. $T_w = 22, 33$ and $T_m = 22, 33$.

Empirical application

S&P 500 index futures from January 3, 2000 to December 31, 2002, a period of 752 trading days, having 2267,121 tick-by-tick observations and characterized by 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 application

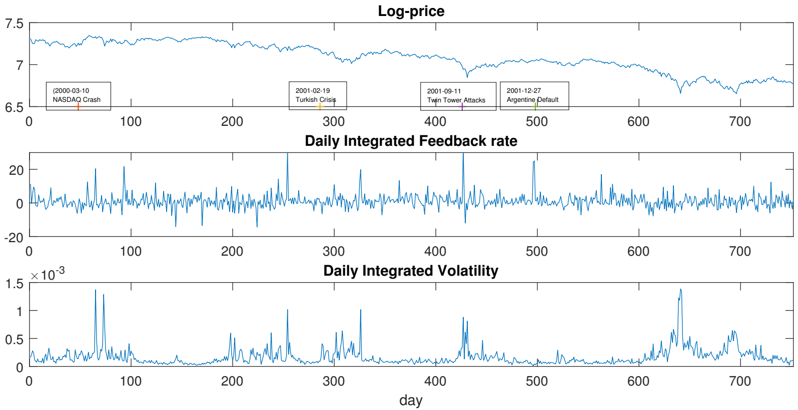
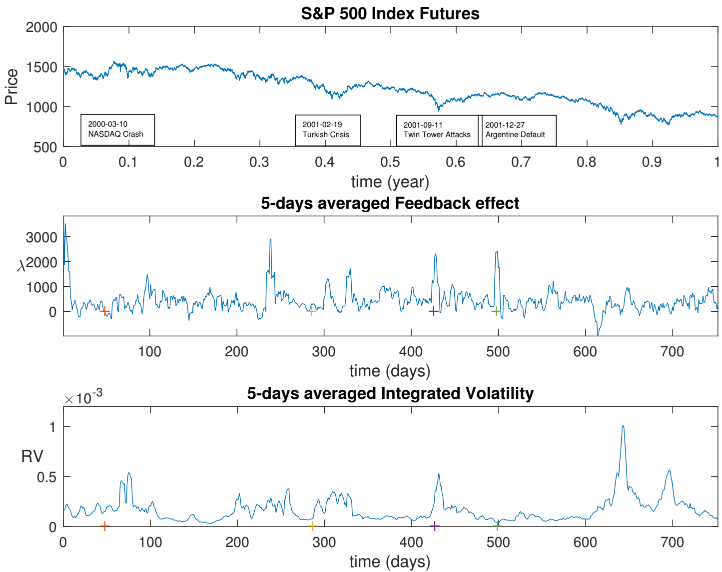


Figure: S&P 500 index futures from January 3rd, 2000 to December 31st, 2002. Upper panel: log-price trajectory; middle panel: daily integrated feedback rate; lower panel: daily integrated volatility.

Qualitative analysis



Statistical analysis:first and second regressions results

- Both the intercept and slope of the logit regressions are highly significant in all the regressions.

Statistical analysis:first and second regressions results

- Both the intercept and slope of the logit regressions are highly significant in all the regressions.
- Take, for example, the regression using $VaR_{0.90}$ and $\ln RV^1$:

$$\ln \frac{p}{1-p} = 18.0851 + 1.9070 \ln RV^1.$$

Then

$$\frac{p}{1-p} = exp(18.0851 + 1.9070 \ln RV^1).$$

The exponential of the slope coefficient is approximately equal to 6.7329 which says that for one unit change in $\ln RV^1$, we expect about 573% increase in the odds of having a financial crisis in the next $T_m = 22$ trading days.

Statistical analysis:first and second regressions results

- Both the intercept and slope of the logit regressions are highly significant in all the regressions.
- Take, for example, the regression using $VaR_{0.90}$ and $\ln RV^1$:

$$\ln \frac{p}{1-p} = 18.0851 + 1.9070 \ln RV^1.$$

  Then

$$\frac{p}{1-p} = exp(18.0851 + 1.9070 \ln RV^1).$$

  The exponential of the slope coefficient is approximately equal to 6.7329 which says that for one unit change in $\ln RV^1$, we expect about 573% increase in the odds of having a financial crisis in the next $T_m = 22$ trading days.
- Note also that the odds of having a financial crisis in the next $T_m = 22$ trading days are also very high when $\ln RV^1$ is 0. This can be easily seen by taking the exponential of the intercept estimate.

Statistical analysis: first and second regressions results

- The regressions using a VaR computed with 0.90 level of confidence have the lowest p-values of the t-statistic tests and chi-square test. In addition, they have the highest values of the slope, intercept and R-squared. We can therefore assert that less extreme crises are much more likely than more extreme ones.

Statistical analysis: first and second regressions results

- The regressions using a VaR computed with 0.90 level of confidence have the lowest p-values of the t-statistic tests and chi-square test. In addition, they have the highest values of the slope, intercept and R-squared. We can therefore assert that less extreme crises are much more likely than more extreme ones.
- The regression using $\ln RV^5$ as independent variable and $VaR_{0.90}$ has the lowest p-values of the t-statistic tests and chi-square test, and the highest R-squared among all the other regressions.

Statistical analysis: first and second regressions results

- The regressions using a VaR computed with 0.90 level of confidence have the lowest p-values of the t-statistic tests and chi-square test. In addition, they have the highest values of the slope, intercept and R-squared. We can therefore assert that less extreme crises are much more likely than more extreme ones.
- The regression using $\ln RV^5$ as independent variable and $VaR_{0.90}$ has the lowest p-values of the t-statistic tests and chi-square test, and the highest R-squared among all the other regressions.
- When the $\ln RV$ is computed with 1 and 5-minutes returns and $T_m = 22$ the outputs of the logit regressions are quite better than those of the regressions using lower sampling frequencies and a longer time horizon.

Statistical analysis: first and second regressions results

- The second type of regression performs better than the first type.

Statistical analysis: first and second regressions results

- The second type of regression performs better than the first type.
- Assuming Student's t losses rather than normal losses, we get better results, for almost every output regression result, when we take $\alpha = 0.10$ and $\ln RV^1$.

Statistical analysis: first and second regressions results

- The second type of regression performs better than the first type.
- Assuming Student's t losses rather than normal losses, we get better results, for almost every output regression result, when we take $\alpha = 0.10$ and $\ln RV^1$.
- Significant regression coefficients are also found for $\alpha = 0.05$. Less significant coefficient results are found for $\alpha = 0.01$.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, **using** CI

- As α decreases, the intercept and coefficient of RV_{av} decrease. Furthermore, these become statistically insignificant at all the significance levels when VaR is computed with $\alpha = 0.01$ indicating that RV_{av} does not predict the forward-looking variable Y.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CI

- As α decreases, the intercept and coefficient of RV_{av} decrease. Furthermore, these become statistically insignificant at all the significance levels when VaR is computed with $\alpha = 0.01$ indicating that RV_{av} does not predict the forward-looking variable Y.
- On the opposite, the coefficient of λ_{av} is significant and increases its magnitude with the decrease of α. This suggests a stronger increase in the odds of having a more extreme financial crisis in the next T_m days as λ_{av} increases of one unit.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CI

- As α decreases, the intercept and coefficient of RV_{av} decrease. Furthermore, these become statistically insignificant at all the significance levels when VaR is computed with $\alpha = 0.01$ indicating that RV_{av} does not predict the forward-looking variable Y.
- On the opposite, the coefficient of λ_{av} is significant and increases its magnitude with the decrease of α. This suggests a stronger increase in the odds of having a more extreme financial crisis in the next T_m days as λ_{av} increases of one unit.
- The reduced regressions obtained with the EWI λ_{av} show a positive association between this EWI and the log-odds that becomes stronger as α decreases. Note also how the sign of the intercept is always negative for these regressions which indicates that a crisis event is less likely when $\lambda_{av} = 0$.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CI

- The values of the coefficient attached to λ_{av} are robust to changes in T_w and T_m. Indeed, they have almost the same values for $T_w = T_m = 22$ and $T_w = T_m = 33$. We can therefore conclude that λ preserves its predictive capabilities even for longer time horizons.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- The estimated coefficient of λ_{av} has more or less the same value among all regressions and is statistically significant in all the situations. This shows the reliability of λ_{av} to predicting Y.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- The estimated coefficient of λ_{av} has more or less the same value among all regressions and is statistically significant in all the situations. This shows the reliability of λ_{av} to predicting Y.
- The intercept in the reduced regression using only λ_{av} is negative implying a reduction in the odds of a having a financial crisis as defined by Y using CEI when λ_{av} is equal to zero.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- The estimated coefficient of λ_{av} has more or less the same value among all regressions and is statistically significant in all the situations. This shows the reliability of λ_{av} to predicting Y.
- The intercept in the reduced regression using only λ_{av} is negative implying a reduction in the odds of a having a financial crisis as defined by Y using CEI when λ_{av} is equal to zero.
- When $\alpha = 0.05$, the reduced model with λ_{av} is preferred over the full model. No significant coefficients for RV_{av} were found for the full and reduced models.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- The estimated coefficient of λ_{av} has more or less the same value among all regressions and is statistically significant in all the situations. This shows the reliability of λ_{av} to predicting Y.
- The intercept in the reduced regression using only λ_{av} is negative implying a reduction in the odds of a having a financial crisis as defined by Y using CEI when λ_{av} is equal to zero.
- When $\alpha = 0.05$, the reduced model with λ_{av} is preferred over the full model. No significant coefficients for RV_{av} were found for the full and reduced models.
- When $\alpha = 0.10$, the regressions show a negative and significant coefficient of RV_{av} indicating a negative relationship with the odds of having a financial crisis.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- When we assume Student's t distribution for the returns and $T_w = T_m = 22$, the full model is always preferred over the reduced model using RV_{av}, but this is not always true for the other reduced model. The regression coefficients of RV_{av} in the full and reduced model are not always significant at conventional significance levels. In addition, they are all negative.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- When we assume Student's t distribution for the returns and $T_w = T_m = 22$, the full model is always preferred over the reduced model using RV_{av}, but this is not always true for the other reduced model. The regression coefficients of RV_{av} in the full and reduced model are not always significant at conventional significance levels. In addition, they are all negative.
- When $\alpha = 0.05$, the full model performs better than the reduced model employing RV_{av} and the reduced model with λ_{av} better then the full model.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- When we assume Student's t distribution for the returns and $T_w = T_m = 22$, the full model is always preferred over the reduced model using RV_{av}, but this is not always true for the other reduced model. The regression coefficients of RV_{av} in the full and reduced model are not always significant at conventional significance levels. In addition, they are all negative.
- When $\alpha = 0.05$, the full model performs better than the reduced model employing RV_{av} and the reduced model with λ_{av} better then the full model.
- Finally, when $\alpha = 0.10$, significant results are found for both RV_{av} and λ_{av} and the full model as well.

Statistical analysis: $\ln \frac{p}{1-p} = \beta_0 + \beta_1 RV_{av} + \beta_2 \lambda_{av}$, using CEI

- When we assume Student's t distribution for the returns and $T_w = T_m = 22$, the full model is always preferred over the reduced model using RV_{av}, but this is not always true for the other reduced model. The regression coefficients of RV_{av} in the full and reduced model are not always significant at conventional significance levels. In addition, they are all negative.
- When $\alpha = 0.05$, the full model performs better than the reduced model employing RV_{av} and the reduced model with λ_{av} better then the full model.
- Finally, when $\alpha = 0.10$, significant results are found for both RV_{av} and λ_{av} and the full model as well.
- Similar results for $T_w = T_m = 33$.

THANK YOU

for your attention