International Journal of Forecasting 39 (2023) 1777–1803

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Early Warning Systems for identifying financial instability

Erindi Allaj [*], Simona Sanfelici [1]

Department of Economics and Management, University of Parma, Italy

ARTICLE INFO

Keywords:
Early Warning System
Market instability
Non-parametric estimation
Price-volatility feedback rate
Realized variance

ABSTRACT

Financial crises prediction is an essential topic in finance. Designing an efficient Early Warning System (EWS) can help prevent catastrophic losses resulting from financial crises. We propose different EWSs for predicting potential market instability conditions, where market instability refers to large asset price declines. The EWSs are based on the logit regression and employ Early Warning Indicators (EWIs) based on the realized variance (RV) and/or price-volatility feedback rate. The latter EWI is supposed to describe the ease of the market in absorbing small price perturbations. Our study reveals that, while RV is important in predicting future price losses in a given time series, the EWI employing the price-volatility feedback rate can improve prediction further.

1. Introduction

Detecting financial crises is extremely important for both policy-makers and practitioners. This has been shown to be particularly crucial during the financial crisis of 2007–2008 suggesting that a forward-looking framework to financial risk management is more indispensable than ever. This fact has, for example, been explicitly pointed out by Breden (2008) and Jorion (2009). Rochette (2009) also argues that the organization with a strong risk culture is the one that is forward-looking. This allows the organisation to learn from its mistakes and most importantly to respond sufficiently quickly to emerging threats and opportunities.

In this context, as noted by Kunze et al. (2020) and Gonzalez et al. (2021), Early Warning Indicators (EWIs) can be extremely helpful for asset managers to ameliorate forward-looking Early Warning Systems (EWSs). More specifically, using housing market data from UK and time-series econometric techniques, Kunze et al. (2020) find that the sentiment indicator can improve forecast accuracy of property prices in the UK. On the other side, Gonzalez et al. (2021) through the combination of a machine learning approach and the concept of transfer entropy find that sentiment indicators are valid predictors of real estate assets in the US.

The question of preventing possible future financial crises was already important before the financial crisis of 2007–2008. Kaminsky et al. (1998) use a signalling approach to construct EWSs capable of predicting a possible financial crisis. These systems are based on observing the behaviour of several economic and political indicators in the periods preceding a determinate currency crisis. Values of these indicators exceeding a given threshold are an indication of a currency crisis within the next 24 months. The signalling approach was also adopted later, for example, by Borio and Drehmann (2009) and Drehmann and Juselius (2014).

Berg and Pattillo (1999) show that probit models provide better forecasts than the Kaminsky et al. (1998) approach. The currency crisis is defined as in Kaminsky et al. (1998) and the probit models are run using a categorical variable assuming the value of 1 if a crisis occurs within the next T_m periods and 0 otherwise. The economic regressors of Kaminsky et al. (1998) and others are employed in running the respective regressions. The probit model has been used, among others, by Kauppi and Saikkonen (2008) and Antunes et al. (2018).

Bussiere and Fratzscher (2006), Kumar et al. (2003), Schnatz (1998) and Candelon et al. (2014) use logit regression models to build EWSs capable of predicting currency

* Corresponding author.
E-mail addresses: erindi.allaj@unipr.it (E. Allaj), simona.sanfelici@unipr.it (S. Sanfelici).

1 Member of the INdAM-GNCS Research Group, Italy

https://doi.org/10.1016/j.ijforecast.2022.08.004

crises. Unlike the other authors, in order to deal with the so called "post-crisis bias", Bussiere and Fratzscher (2006) suggest the use of the multinomial logit regression EWS as a possible substitute of the binary logit EWS. They conclude that their proposed EWS increases the ability to forecast financial crises. Manasse et al. (2003) and Dawood et al. (2017) also propose logit regression EWSs for predicting sovereign debt crises. In particular, Dawood et al. (2017) utilize both a binary and a multinomial logit regression while constructing their EWS. Caggiano et al. (2014, 2016), focusing on low income countries, propose a multinomial logit regression EWS for predicting banking crises. Their model is able to address, what they call, the "crisis duration bias". Also, Barrell et al. (2010) predict banking crises by estimating binary logit regression models for OECDs countries.

Ghosh and Ghosh (2003) and Manasse et al. (2003) propose the use of binary recursive trees, while Fuertes and Kalotychou (2007) the use of K-clustering approach to detect financial crises. Billio et al. (2016) analyse the forecasting ability of the entropy measures to predict financial crises by applying systemic risk measures such as Marginal Expected Shortfall (MES), ΔCoVaR and network connectedness. They show that the ability of the entropy measures is remarkable.

A common denominator of the majority of the approaches outlined is the use of macrodata and financial variables. Good overviews on the EWSs can be found for example in Abiad (2003), Babecký et al. (2014) and Douady and Kornprobst (2018).

Many empirical and theoretical studies have also investigated the relationship between market volatility and financial crisis. For instance, Douady and Kornprobst (2018) propose financial crisis indicators based on the distribution of the eigenvalues and the trace of the empirical covariance matrix, while (Li et al., 2015) provide evidence that the VIX index can be useful in predicting financial crises. Assessment of the degree of market uncertainty and potential financial loss or gain may be measured using variability in prices. In statistics and econometrics, the most common measure to determine variability is the standard deviation, i.e., the variability of returns from the mean. In particular, realized variance (RV), also called realized volatility, is a non-parametric measure of the return variation within a defined period of time.

The recent literature on volatility modeling has pointed out the existence of feedback effects of asset prices on volatilities and viceversa in financial markets. In a recent paper, Mancino and Sanfelici (2020) propose an Early Warning Indicator (EWI) based on the estimation of a second order quantity (named *price-volatility feedback rate*), which is supposed to describe the ease of the market in absorbing small price perturbations. The feedback rate relies on the computation of the decay rate for the propagation of a given market shock. The rate of variation through time of an initial perturbation enables us to understand if such a shock will be rapidly absorbed or, on the opposite, it will be amplified by the market. When the feedback rate is positive, the perturbation propagates over a trajectory, whereas, when the feedback rate is negative, the perturbation is attenuated over a trajectory. More precisely, the authors show 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. This financial instability index combines in a non-linear way volatility, leverage and covariance between leverage and price and can be non-parametrically estimated from high frequency financial data using the Fourier method developed in Malliavin and Mancino (2002a) and Malliavin and Mancino (2009).

Starting from their study, the volatility feedback rate can be used as an EWI that can help 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 binary logit regression model where different predictors or EWIs based on RV and the integrated feedback rate (over a trading day) are considered. According to Beutel et al. (2018), the binary logit approach provides better out-of-sample performance than other machine learning methods. These predictors are given by the logarithm of the RV, average of the logarithms of the RVs over a certain period of time and the average of another variable taking the value of the sample mean of the daily integrated feedback rate derived over the whole temporal period if the daily feedback rate exceeds the sample mean, 0 otherwise. The dependent variable, also known as the forward-looking 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). This is how the term "financial crises" should be intended throughout the present paper. The forward-looking variable is alternatively defined by comparing the average returns, conditional on having losses exceeding the corresponding VaR measure, with the Expected shortfall (ES).

Our definition of financial crisis includes the classical crisis indicators used in the literature, where a crisis usually happens when the variable used to determine it is more than a multiple of the standard deviation above its average. However, here, we are offering more general definitions and, to the best of our knowledge, this is the first study to use these definitions. This is also true for the EWIs.

We ran our logit regressions by assuming different estimation intervals for the EWIs based on the RV and price-volatility feedback rate and different crises prediction intervals. In addition, the risk measures VaR and ES are derived under different levels of significance and distributions, namely, normal and t location-scale distribution. However, in most cases, we focus on the latter distribution since this fits our data better. Our dataset includes tick-by-tick prices of the S&P 500 index futures covering the period from 3 January 2000 to 31 December 2008 and a total of 2248 trading days. This index is often used by practitioners and researchers as a proxy for the financial and real economy. Examples might include Andersen et al. (2007), McCarthy and Orlov (2012) and Li et al. (2015). In particular, Andersen et al. (2007) find that S&P 500 index futures are related to fundamentals, while McCarthy and Orlov (2012) use the movements in the S&P 500 index futures to proxy for financial volatility and investigate the link between the latter and the oil market. Finally, Li et al. (2015) conclude that information contained in the S&P 500 index futures and options improve prediction of equity market crises. It is also worth pointing out that this study is similar to ours in the sense that it builds its EWIs employing the information in the S&P 500 indexes. As a result, predicting its future losses is very useful both from a practical and theoretical point of view. We would like however to emphasize the fact that our EWSs can of course be easily extended by introducing other EWIs such as macroeconomic indicators.

We find statistically significant results for the different combinations of the estimation, prediction intervals and levels of significance. Focusing on the t location-scale distribution, the EWIs using RV have a larger effect on the odds of having financial crises when the level of significance is equal to 0.90 or 0.95 suggesting that RV is more connected to less extreme financial crises. As for the EWI employing the price-volatility feedback rate, we observe that in almost all the cases the estimated regression coefficients of this EWI are greater when the level of confidence is equal to 0.99. This property of the feedback rate remains true both in the case when this EWI is used in the conjunction with the other EWI that uses the average of the logarithms of the RVs or when used alone. This means that when the EWI increases with one unit, the odds of having more extreme financial crises increases more than the odds of less extreme crises. Further, as the estimation and the prediction interval increase, the slope magnitude of the latter EWI increases in almost all the cases considered.

The EWIs considered are very capable in predicting notable financial crises like the Lehman Brothers Collapse. Additionally, the EWSs run with both the EWIs using the feedback rate and RV are preferred over the reduced EWSs ran with the single EWIs alone. Our findings remain intact even when a cross-validation exercise is run. We can therefore affirm the capacity of these EWIs in predicting future financial crises and in particular of the feedback rate to improve this prediction.

The paper is organized as follows. In Section 2 we introduce the proposed EWS by discussing the EWIs, forward-looking variables and the logit regression model. We report the output results of the various regressions in Section 3. Section 4 contains some further considerations on the EWSs proposed. Finally, Section 5 concludes.

2. Our model

In this section, we discuss the EWSs we use to predict financial crises. Specifically, we propose different logit EWSs that includes different EWIs. The proposed EWSs use the VaR and ES risk measure when constructing the forward-looking variables. The well-known RV (see, e.g., Andersen et al. (2003)) and the price-volatility feedback rate λ (Mancino & Sanfelici, 2020) are used to predict these variables. To begin with, we first recall the definition of the indicator of shock propagation through the market, called price-volatility feedback rate, an indicator which was proposed by Malliavin and Mancino (2002b) and Barucci et al. (2003). Afterwards, we introduce the RV estimator.

2.1. Definition of the market instability indicator $\Lambda_{n,N,M,L}$ and RV

Assume that the asset price S_t satisfies the following stochastic differential equation without drift[2]

$$\frac{dS_t}{S_t} = a_1(S_t)dW_t, \tag{1}$$

where $a_1(\cdot)$ is a smooth deterministic function of the asset price and W_t is a Brownian motion on a filtered probability space. Then the logarithm of the price, i.e. $P_t = \log S_t$, satisfies the following stochastic differential equation

$$dP_t = a_1(S_t)dW_t - \frac{1}{2}a_1^2(S_t)dt = a(P_t)dW_t - \frac{1}{2}a^2(P_t)dt, \tag{2}$$

where $a(P) := a_1(S) = a_1(\exp(P))$.

Consider the associated variation process, defined as the solution to the linearized stochastic differential equation[3]

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

This equation describes how perturbations of asset prices evolve through time. Then, it can be proved (see Malliavin and Mancino (2002b)) that the rescaled variation defined by

$$z_t := \frac{\zeta_t}{a(P_t)} \tag{3}$$

is differentiable with respect to t and it holds that

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

where

$$\lambda_\tau = -\frac{1}{2}(a'(P_\tau)a(P_\tau) + a''(P_\tau)a(P_\tau)), \tag{4}$$

is called the price-volatility feedback rate.

Basically, the price-volatility feedback rate can be understood as the appreciation rate of the rescaled variation. We note that the feedback rate defined by (4) is a decay rate: its negative sign entails a damping effect, while a positive sign implies that any perturbation is amplified. Therefore, large positive values of λ indicate market instability and usually anticipate a significant decrease in the price level, while negative values or around zero correspond to stable market directions. Thus, the volatility feedback rate can reveal conditions that may facilitate the propagation of perturbations in the market. This may help discriminate between stable market conditions and conditions where 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 (see also the empirical analysis in Inkaya and Okur (2014)).

Finally, in order to build a daily indicator of market instability, we consider the integrated value $\int_I \lambda_\tau d\tau$, where I is the daily trading period.

[2] The drift term can be added without any difficulty as done in Malliavin and Thalmaier (2006) and Inkaya and Okur (2014).

[3] The prime stands here for the first derivative with respect to the level P_t.

The rescaled variation contains terms which are latent, thus they should be empirically estimated from the data. The following theorem shows how to express the feedback rate by means of terms which are *iterated cross volatilities* that can be estimated using the Fourier estimation method developed by Malliavin and Mancino (2009). The proof of the theorem is given in Malliavin and Mancino (2002b).

Theorem 1. *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

$$\lambda_t = \frac{3}{8}\frac{B_t^2}{A_t^3} - \frac{1}{4}\frac{B_t}{A_t} - \frac{1}{4}\frac{C_t}{A_t^2}\,. \tag{5}$$

This theorem makes it possible to estimate the feedback rate in a model-free way. In fact, Mancino and Sanfelici (2020) show how to estimate non-parametrically the functions A_t, B_t, C_t by using the Fourier method. We highlight that these functions represent the instantaneous variance, leverage and covariation between leverage and asset price, respectively, and can be non-parametrically estimated under the more general assumption that the log-price, variance and leverage processes are continuous semimartingales

$$dP_t = \nu_t dt + \sigma_t dW_t, \quad dA_t = \alpha_t dt + \gamma_t dW_t^A,$$
$$dB_t = \mu_t dt + \beta_t dW_t^B,$$

where $A_t = \sigma_t^2$, W_t^A and W_t^B are Brownian motions possibly correlated with W_t, and the adapted processes $\sigma_t, \nu_t, \alpha_t, \gamma_t, \mu_t, \beta_t$ are absolutely bounded. It is worth mentioning that the choice of a continuous process for P_t is not restrictive. Indeed, working with intraday high frequency data, the question of jumps presence is debatable. For instance, Christensen et al. (2014) analyze the presence of jumps in tick data and observe that jumps in financial asset prices are often erroneously identified and are, in fact, rare events. For example, a short-lived burst of volatility is likely to be identified as a jump when working with data sampled at a frequency lower than 5 min but it is compatible with a continuous path when working with tick data. 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 types of market frictions (see Mancino et al. (2017)).

As already remarked in Mancino and Sanfelici (2020), when we combine non-linearly estimates of A_t, B_t and C_t to obtain λ_t as in Eq. (5), the resulting estimator is usually biased (this fact is well recognized in the non-parametric estimation literature). Furthermore, the potential oscillations in A_t, B_t, C_t are amplified and the resulting approximation of λ_t is very unstable. Therefore, we prefer to combine daily mean value quantities to get a more stable indicator of market instability. Being interested only in daily mean value quantities, for the reader's convenience, we only recall the Fourier estimators of $\frac{1}{I}\int_I A_t dt$, $\frac{1}{I}\int_I B_t dt$, $\frac{1}{I}\int_I C_t dt$ and refer the reader to Mancino and Sanfelici (2020) for further technical details.

Consider a discrete unevenly spaced sampling of the log-price process P_t. For notational ease, we assume that the log-price process P_t is observed on $I = [0, 2\pi]$. Fix a set of observation times $\mathcal{S}_n := \{0 = t_{0,n} \le t_{1,n} \le \cdots \le t_{k_n,n} = 2\pi\}$ for any $n \ge 1$ such that the mesh size of the partition goes to 0, that is $\rho(n) := \max_{0\le h\le k_n-1} |t_{h+1,n} - t_{h,n}| \to 0$ as $n \to \infty$. For any j, denote $\Delta P_{j,n} := P(t_{j+1,n}) - P(t_{j,n})$.

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

$$c_k(A_{n,N}) := \frac{2\pi}{2N+1}\sum_{|s|\le N} c_s(dP_n)c_{k-s}(dP_n), \tag{6}$$

where for any integer k, $|k| \le 2N$, $c_k(dP_n)$ is the kth (discrete) Fourier coefficient of the return[4] process, namely

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

where the symbol i denotes the imaginary unit, $\mathrm{i} = \sqrt{-1}$. It is proved that (6) is a consistent estimator of the kth Fourier coefficient of the spot variance $c_k(A)$. Therefore, a consistent estimator of $\frac{1}{2\pi}\int_0^{2\pi} A_t dt$ is given by

$$\widehat{A}_{n,N} := c_0(A_{n,N}) = \frac{2\pi}{2N+1}\sum_{|s|\le N} c_s(dP_n)c_{-s}(dP_n),$$

and has been studied in Mancino and Sanfelici (2008).

The knowledge of the Fourier coefficients of the latent instantaneous volatility A_t allows us to handle this process as an observable variable and we can iterate the procedure in order to compute the covariation B_t between the stochastic variance process and the asset price process (i.e., the leverage) (Curato & Sanfelici, 2015 and Curato, 2019). The estimator of the kth Fourier coefficients of the leverage B_t is defined by

$$c_k(B_{n,N,M}) := \frac{2\pi}{2M+1}\sum_{|j|\le M} c_j(dP_n)c_{k-j}(dA_{n,N}), \tag{7}$$

where $c_j(dA_{n,N})$ is computed as

$$c_j(dA_{n,N}) := \mathrm{i}\, j\, c_j(A_{n,N}).$$

Note that the estimator (7) depends only on the Fourier coefficients of the asset return $c_j(dP_n)$ that are computed from the real data, and on the Fourier coefficients of the variance $c_j(A_{n,N})$, which have been estimated in the previous step. Therefore, a consistent estimator of $\frac{1}{2\pi}\int_0^{2\pi} B_t dt$ (i.e., the integrated leverage) is given by

$$\widehat{B}_{n,N,M} := c_0(B_{n,N,M}) = \frac{2\pi}{2M+1}\sum_{|j|\le M} c_j(dP_n)c_{-j}(dA_{n,N}).$$

Finally, as in the previous step, once its Fourier coefficients have been estimated, we can handle B_t as an

[4] In this paper return(s) refer to log-return(s) unless stated otherwise.

observable variable and exploit the multivariate version of the Fourier method to estimate the covariance C_t between the process B_t and the price return. More precisely, a consistent estimator of the kth Fourier coefficient of the function C_t can be obtained by

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

where $c_j(dB_{n,M,N}) := \mathrm{i}\, j\, c_j(B_{n,M,N})$. Therefore, a consistent estimator of $\frac{1}{2\pi}\int_0^{2\pi} C_t dt$ is given by

$$\widehat{C}_{n,N,M,L} := c_0(C_{n,N,M,L}) = \frac{2\pi}{2L+1} \sum_{|j|\le L} c_j(dP_{n,N}) c_{-j}(dB_{n,N,M}).$$

Finally, combining the previous mentioned estimators and considering Theorem 1, we build our estimator of the daily integrated feedback rate $\int_0^{2\pi} \lambda_t dt$

$$\Lambda_{n,N,M,L} = 2\pi \hat{\lambda}_t,$$

where

$$\hat{\lambda}_t = \frac{3}{8}\frac{\widehat{B}^2_{n,N,M}}{\widehat{A}^3_{n,N}} - \frac{1}{4}\frac{\widehat{B}_{n,N,M}}{\widehat{A}_{n,N}} - \frac{1}{4}\frac{\widehat{C}_{n,N,M,L}}{\widehat{A}^2_{n,N}}.$$

When we use high frequency data to estimate the feedback rate, possible microstructure noise contained in the data has to be filtered out. For the Fourier estimator, this can be achieved by a suitable choice of the cutting frequencies. In particular, the first cut-off frequency N is fixed according to the analysis developed in Mancino and Sanfelici (2008) by minimizing the estimated MSE, while we set $M = k_n^{0.5}$ and $L = M/2$.

We conclude this section by recalling the well-known RV estimator. Analysts make use of high frequency intra-day data to determine measures of daily volatility. RV is simply defined as the sum of squared (intra-day) returns

$$\sum_{j=0}^{k_n-1} \Delta P^2_{j,n}$$

and provides an estimator of the integrated variance over the (daily) trading period $[0, 2\pi]$. This estimator is not robust to market microstructure effects that appear at high frequency level and must be combined with sparse sampling or other devices (Bandi & Russell, 2011).

2.2. Crisis indicators

Let $R_t = P_{t+1} - P_t$ denote the daily return over the time interval t to $t+1$, where P_t gives the log-price of a specific asset at day t and $t = 1, 2, \ldots, T-1$. The VaR of the loss $L_t = -R_t$ at the confidence level $1-\alpha$ (Acerbi & Tasche, 2002; Artzner et al., 1999) and McNeil et al. (2015) is defined by the following quantile

$$\begin{aligned} VaR_{1-\alpha}(L_t) &= \inf\{r \in \mathbb{R} : F_{L_t}(r) \ge 1-\alpha\} \\ &= -\inf\{r \in \mathbb{R} : F_{R_t}(r) > \alpha\} \\ &= -\inf\{r \in \mathbb{R} : \mathbb{P}(R_t \le r) > \alpha\}, \end{aligned} \tag{8}$$

where $F_X(x)$ denotes the cumulative distribution function of X and the values of α are usually set equal to 1%, 5% and 10%.

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

VaR can be computed easily if one assumes that L_t follows a normal or a t location-scale distribution. Indeed, VaR takes one of the following two closed forms

$$VaR^N_{1-\alpha}(L_t) = \mu_{t+1} + Z_{1-\alpha}\sigma_{t+1}, \tag{9}$$

$$VaR^t_{1-\alpha}(L_t) = \mu_{t+1} + t_{1-\alpha,\nu_{t+1}}\sigma_{t+1}, \tag{10}$$

where $Z_{1-\alpha}$ and $t_{1-\alpha,\nu_{t+1}}$ denote the $1-\alpha$ quantiles of the standard normal distribution and the Student's t-distribution with ν_{t+1} degrees of freedom. In the case of the normal distribution, the parameters μ_{t+1} and σ_{t+1} give the mean and the standard deviation of L_t, while in the case of the t location-scale distribution give respectively its location and scale parameter. It follows that $\frac{L_t-\mu_{t+1}}{\sigma_{t+1}}$, where L_t has a t-location scale distribution with mean μ_{t+1}, degrees of freedom ν_{t+1} and variance equal to $\sigma_t^2 \frac{\nu_{t+1}}{\nu_{t+1}-2}$ is a Student's t-distribution with ν_{t+1} degrees of freedom. Note that μ_{t+1} and the variance are well-defined when $\nu_{t+1} > 2$.

Another classical risk measure used to assess the risk of an investment is the ES (Acerbi & Tasche, 2002). The ES of L_t at confidence level $1-\alpha$ over the time interval t to $t+1$ is defined (whenever this is well-defined) as

$$\begin{aligned} ES_{1-\alpha}(L_t) = -\frac{1}{\alpha}\{&E[-L_t \mathbf{1}_{\{L_t \ge -q_\alpha(-L_t)\}}] + q_\alpha(-L_t)\alpha \\ &- q_\alpha(-L_t)\mathbb{P}[L_t \ge -q_\alpha(-L_t)]\}, \end{aligned} \tag{11}$$

where $q_\alpha(X) = \inf\{x \in \mathbb{R} | \mathbb{P}(X \le x) \ge \alpha\}$ and $\mathbf{1}_{\{\cdot\}}$ gives the indicator function.

Assuming that the cumulative distribution function of L_t is continuous and strictly increasing, $ES_{1-\alpha}(L_t)$ can be written as

$$ES_{1-\alpha}(L_t) = \mathbb{E}[L_t | L_t - VaR_{1-\alpha}(L_t) > 0]. \tag{12}$$

Differently from the VaR risk measure, ES can be shown to satisfy the axioms of coherent risk measures (Artzner et al., 1999). From a financial point of view, 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 than the VaR risk measure.

In the particular case where L_t follows a normal or a t location-scale distribution, ES assumes, respectively, the following form

$$ES^N_{1-\alpha}(L_t) = \mu_{t+1} + \frac{\psi(Z_{1-\alpha})}{\alpha}\sigma_{t+1}, \tag{13}$$

$$\begin{aligned} ES^t_{1-\alpha}(L_t) = \mu_{t+1} + &\frac{f_{\nu_{t+1}}(t_{1-\alpha,\nu_{t+1}})}{\alpha} \\ &\times \left(\frac{\nu_{t+1} + t^2_{1-\alpha,\nu_{t+1}}}{\nu_{t+1}-1}\right)\sigma_{t+1}, \end{aligned} \tag{14}$$

where $\psi(\cdot)$ and $f(\cdot)$ give the density distribution of the normal and Student's t-distribution. The reader is referred to Bertsimas et al. (2004) and McNeil et al. (2015) for a derivation of these formulas.

Therefore, VaR and ES are good candidates for detecting losses or suffering of a specific asset. When VaR and

ES are applied to a stock market index like the S&P 500 index futures, they can also be used as indicators of a financial crisis, the latter can also affect the real economy. Therefore, we can define the following crisis indicators

$$CI_t = \begin{cases} 1 & \text{if} \quad R_t < -VaR_{1-\alpha} \\ 0 & \text{else}, \end{cases} \tag{15}$$

or

$$CEI_M = \begin{cases} 1 & \text{if} \quad A_M < -ES_{1-\alpha} \\ 0 & \text{else}, \end{cases} \tag{16}$$

where the thresholds $VaR_{1-\alpha}$ and $ES_{1-\alpha}$ are computed using Eqs. (9), (10), (13) and (14).

These are the definitions of asset crisis that will be used in our econometric analysis. Note that for ease of exposition, we have suppressed the dependence of $VaR_{1-\alpha}$ and $ES_{1-\alpha}$ on the losses L.

The quantity A_M is given by

$$A_M = \frac{\sum_{j=M-T_k+1}^{M} R_j \mathbf{1}_{\{-R_j > VaR_{1-\alpha}\}}}{\alpha T_k}, \tag{17}$$

where, for fixed T_k, $M = T_k, T_k + 1, \dots, T - 1$.

A_M can also be seen as the natural estimator of ES as defined by Eq. (12). This follows by writing ES as an unconditional expectation (see Acerbi and Szekey (2014)).

The first indicator in Eq. (15) says that a crisis at day $t + 1$ occurs when the returns over the interval t to $t + 1$ is less than the negative of the VaR risk measure. On the other hand, the second indicator, as defined by Eq. (16), defines a crisis at day $M + 1$ 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 negative of $ES_{1-\alpha}$.

However, the goal of an EWS is to predict future financial crises within a certain time horizon. In order to determine whether a financial crisis occurs within a given horizon of time, we transform our crisis indicator CI into a forward-looking variable (Berg & Pattillo, 1999 and Bussiere & Fratzscher, 2006). This new variable is given by

$$Y_s = \begin{cases} 1 & \text{if} \quad \exists \quad k = 1, 2, \dots, T_m \quad \text{s.t.} \quad CI_{s+k-1} = 1 \\ 0 & \text{else}, \end{cases} \tag{18}$$

where $s = T_w, T_w + 1, \dots, T - T_m$.

In the case of the second indicator, we assume henceforth throughout our paper that T_k is equal to T_m. The forward-looking variable can then be defined as follows

$$Y_s = \begin{cases} 1 & \text{if} \quad CEI_{s+T_m-1} = 1 \\ 0 & \text{else}, \end{cases} \tag{19}$$

where again $s = T_w, T_w + 1, \dots, T - T_m$.

The forward-looking variable in Eq. (18) assumes the value of one at day s whenever the indicator CI is equal to one in one of the next T_m trading days. In the same vein, the forward-looking variable in Eq. (19) is equal to one at day s whenever the indicator CEI takes the value of one at day $s + T_m - 1$. It is also important to pay attention to the fact that by the definition of CEI, the forward-looking variable Y using this crisis indicator already includes the fact that a crisis as defined by CI might occur in one of the T_m trading days after day s.

2.3. Logit model

We suppose the probability of having financial crises can be modelled through a logit regression model (see Wooldridge (2016))

$$\mathbb{P}(Y_s = 1) = \frac{\exp(\beta_0 + \mathbf{x}_s \boldsymbol{\beta})}{1 + \exp(\beta_0 + \mathbf{x}_s \boldsymbol{\beta})}, \tag{20}$$

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

In order to analyse the effect of the explanatory variables on the probability of financial crises, we will run four types of logit regression models. The first two types consist of only the EWI $\ln RV$ and the crisis indicators CI or CEI. Therefore, only values of Y_s and $\ln RV_s$ will be considered, whereas RV_s gives the realized variance at day s. The forward-looking variables Y_s will be computed by setting $T_w = 1$ and $T_m = 22$, 66 or 132, which correspond more or less to 1 month, 1 quarter and 2 quarters. The log-realized variance is preferred over realized variance because of the studies in Barndorff-Nielsen and Shephard (2005) and Gonçalves and Meddahi (2011). The RV estimator is estimated with sparse sampling (Andersen et al., 2000 and Bandi & Russell, 2011). In order to compute these returns, we use the previous-tick interpolation method by setting the price to its most recent value. On the other side, the time series of daily returns is constructed by taking the differences between the opening log-prices of two consecutive days.

To run the third and fourth types of regressions, we introduce the following variable

$$Y_{\lambda,i} = \begin{cases} \overline{\Lambda}_{n,N,M,L} & \text{if} \quad \Lambda^i_{n,N,M,L} > \overline{\Lambda}_{n,N,M,L} \\ 0 & \text{else}, \end{cases} \tag{21}$$

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 $\Lambda^i_{n,N,M,L}$ computed over the whole temporal period. As explained in Section 2.1, large positive values of $\Lambda^i_{n,N,M,L}$ usually anticipate a significant variation in the price level, while negative values or around zero correspond to stable market directions. Our focus then falls on those values of $\Lambda^i_{n,N,M,L}$ that exceed $\overline{\Lambda}_{n,N,M,L}$ (the linear trend). $Y_{\lambda,i}$ does the same work of the indicator variable

$$I_{\lambda,i} = \begin{cases} 1 & \text{if} \quad \Lambda^i_{n,N,M,L} > \overline{\Lambda}_{n,N,M,L} \\ 0 & \text{else}. \end{cases} \tag{22}$$

The threshold $\overline{\Lambda}_{n,N,M,L}$ determines whether a given $\Lambda^i_{n,N,M,L}$ belongs to one regime (high) or the other (low). The threshold is to be considered as the point forecast of the future values of the daily feedback rate and therefore values below or above this threshold (trend) are to be considered unusual values. It is also easy to see that $I_{\lambda,i}$ acts as a dummy variable (Wooldridge, 2016) taking the value of one if $\Lambda^i_{n,N,M,L}$ is sufficiently high (indicating thus a less stable market) and zero otherwise. Before discussing our empirical findings, we anticipate already here that the value of $\overline{\Lambda}_{n,N,M,L}$ calculated using our data is positive.

These regressions use the crisis indicators *CI* or *CEI* with estimation window T_w equal to 22, 66 or 132 and regressors or EWIs given by $\lambda_{av,s} = \frac{\sum_{i=s-T_w+1}^{s} Y_{\lambda,i}}{T_w}$ and $RV_{av,s} = \frac{\sum_{i=s-T_w+1}^{s} \ln RV_i}{T_w}$, where $s = T_w, T_w+1, \ldots, T-T_m$. Thus we believe that the average of Y_λ and RV taken over the interval $s-T_w+1$ to s can predict a financial crisis in one of the next T_m trading days after the day s. Note also that using $Y_{\lambda,i}$ instead of $I_{\lambda,i}$ makes the interpretation of the regression coefficient easier.

The two risk measures VaR and ES will be computed using the normal and the t location-scale distribution. In the case of the normal distribution, $VaR_{1-\alpha}$ and $ES_{1-\alpha}$ are computed by estimating the parameters μ_{t+1} and σ_{t+1} using the historical mean and standard deviation of the daily losses. The same risk measures in the t location-scale distribution case are computed by estimating the respective parameters using the maximum likelihood estimation (MLE) method applied to the daily losses. The indicators *CI* and *Y* can then be computed by comparing the time series of our asset returns with $VaR_{1-\alpha}$. It should be noted that our thresholds VaR and ES are time independent, an assumption commonly used in the related literature. However, there is nothing in our EWSs that prevents the use of a threshold that changes with time.

We estimate A_M using the expression

$$\widehat{A}_M = \frac{\sum_{j=M-T_m+1}^{M} \tilde{r}_j \mathbf{1}_{\{-\tilde{r}_j > \widehat{VaR}_{1-\alpha}\}}}{\alpha T_m}, \tag{23}$$

where $\tilde{r}_j$ gives the realized asset return at time j and $\widehat{VaR}_{1-\alpha}$ the estimated VaR with $1-\alpha$ level of confidence. We can thus interpret $\widehat{A}_M$ as the average loss of the losses exceeding $\widehat{VaR}_{1-\alpha}$ over the interval $M-T_m+1$ to $M+1$. The computation of *CEI* and the related *Y* follows easily. As before, we let $T_m = 22$, 66 or 132 trading days.

3. Empirical application

In this section we apply the models discussed in Section 2 to data from the S&P 500 index futures recorded at the Chicago Mercantile Exchange (CME).[5]

3.1. Data description

Table 1 describes the main features of the intraday returns of the S&P 500 index futures. The sample covers the period from 3 January 2000 to 31 December 2008, a period of 2248 trading days, having 5,576,759 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; 2002-10-09: Stock Market Crash; 2005-10-08: Delphi (G.M) Bankruptcy; 2007-07-01: Subprime Crisis; 2008-09-15: Lehman Brothers Collapse. In line with Duchin et al. (2010), we define the beginning of the Subprime crisis as July 1, 2007. As is well known, the roots of the financial crisis of 2007–2008 can be found in the early 2000s, where there has been a consistent growth in mortgage lending. Our sample period therefore includes the year in which it all began and the climax year of the financial crisis.

Fig. 1 plots the futures intraday return series, its ACF and PACF. The returns exhibit a large negative spike at the occurrence of the Twin Tower Attacks. We remind that, to prevent a stock market meltdown, on September 11th the opening of the New York Stock Exchange (NYSE) was delayed after the first plane crash and trading for the day was canceled after the second one. To avoid market chaos, the NYSE remained closed until September 17th. The first day of NYSE trading after the attacks, the market fell 684 points, a 7.1% decline. That Friday, at the end of trading day, the S&P 500 index lost 11.6%. Other large negative spikes on returns are evident after the Lehman Brothers Collapse, i.e. on 24 October 2008, corresponding also to the highest negative return. The plots of the ACF and PACF are almost equal and show a significant positive first order autocorrelation of returns and persistence of autocorrelation at higher lags that reveal the presence of market microstructure effects in the high frequency time series.

Before presenting the results of our empirical application, we plot in Fig. 2 the "volatility signature plot" of the RV estimator. This kind of graph shows the behaviour of the RV as a function of the sampling frequency and was proposed in Andersen et al. (2000). The sampling frequency goes from a minimum of 1 min to a maximum of 30 min with a time step equal to one. The daily average volatility is computed as the average of the daily RV throughout the entire time horizon.

From the volatility signature plot, we note that the largest RV estimates happen at the highest sampling frequencies, a typical behaviour of liquid assets in the presence of market microstructure effects (Andersen et al., 2000). In addition, it seems that these values stabilize around the sampling frequency equal to 15 min. Consequently, to eliminate microstructure noise, we use this sampling frequency when calculating RV.

3.2. Results

This section reports and discusses the results of the four types of the logit regressions explained in Section 2.3.

3.2.1. The first and the second types of regressions

Tables 5 and 6 in Appendix report the results of the first type of logit regressions ran with the forward-looking variable Y shown in Eq. (18) and $\ln RV$. As concluded in the previous section, the RV estimator is computed using 15-minutes returns, while T_m is set equal to 22, 66 or 132. The VaR risk measure is estimated assuming a normal (Table 5) or a t location-scale distribution (Table 6) and a confidence level equal to 0.90, 0.95 or 0.99. The tables display the coefficient estimates together with the p-values of the t-statistic tests that the coefficient is equal to zero or not, McFaddens's pseudo ordinary R-squared

5 We thank Fulvio Corsi for providing us the possibility of using S&P 500 index futures data.

Table 1
Summary statistics of the intraday returns.

Mean	Minimum	Maximum	Std	Kurtosis	Skewness	JB
−9.2642e−08	−0.0866	0.0612	2.6575e−04	4.9329e+03	−6.1356	5.6474e+12

The table reports the mean, minimum, maximum, standard deviation (Std), kurtosis, skewness and the Jarque–Bera test (JB) of the intraday returns including the overnight returns of the S&P 500 index futures from 3 January 2000 to 31 December 2008 (5,576,759 trades).

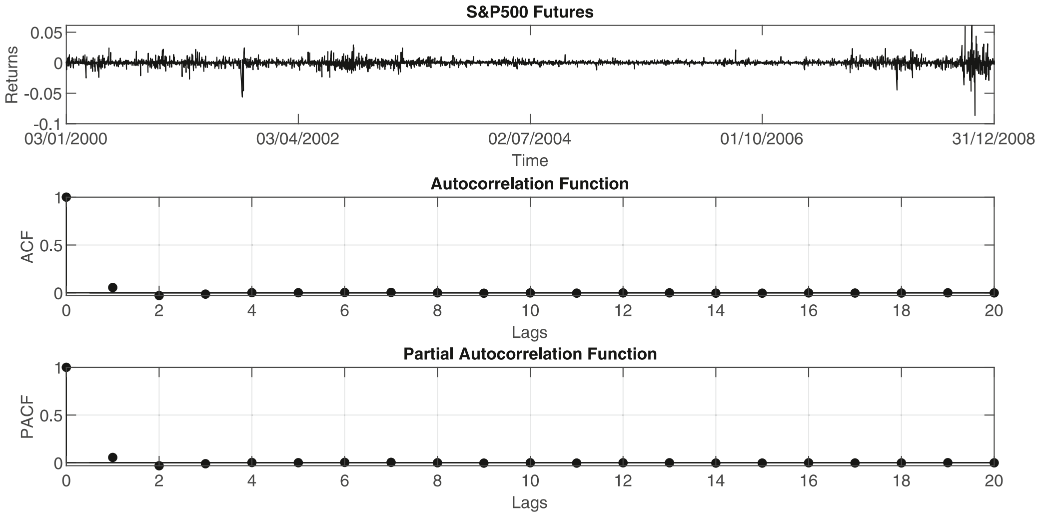


Fig. 1. Time series, ACF and PACF of the S&P 500 index futures intraday returns.

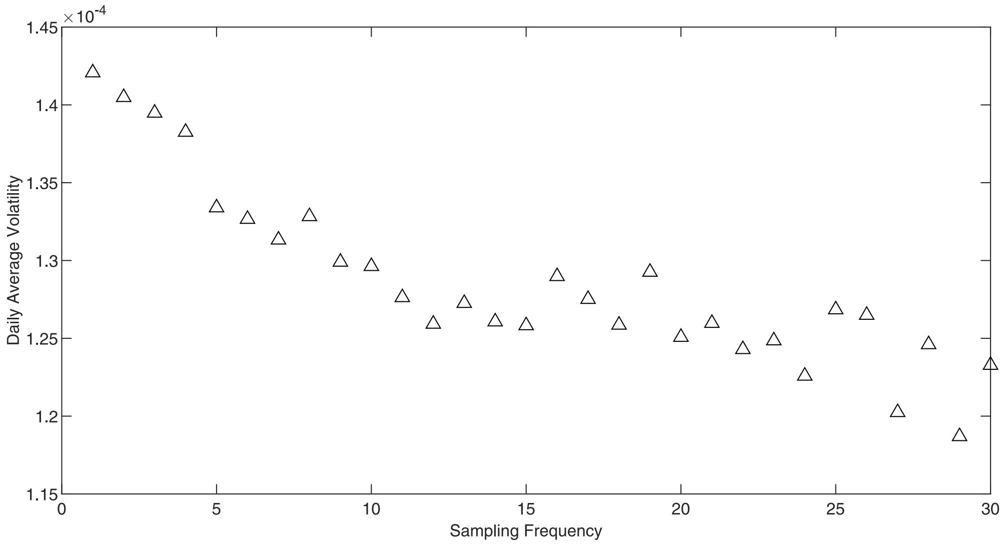


Fig. 2. Volatility signature plot.

(McFadden, 1973), hit ratio (computed with a probability threshold equal to 0.5) and the p-value of the chi-square test (likelihood ratio test) of the intercept-only model versus the intercept and regressor model (goodness). The hit ratio is a goodness-of-fit measure giving the proportion correctly predicted (Wooldridge, 2016).

Focusing on Table 5, we note that both the intercept and slope of the logit regressions are highly significant. The chi-square test is also in favour of the logit regression with the $\ln RV$ as the regressor. Taking the exponential of the slope coefficients, we can conclude that there is a positive relationship between realized variance and financial crises. Take for example the slope coefficient of the regression using $VaR_{0.90}$ and $T_m = 22$. The exponential of the slope coefficient is equal to 4.7705 which says that for one unit change in $\ln RV$, we expect about 377% increase in the odds of having a financial crisis in the next $T_m = 22$ trading days. In terms of RV this tells us that if RV increases by 10%, odds would increase by about 16%. Note also that the odds of having a financial crisis in the next $T_m = 22$ trading days are also very high when $\ln RV$ is 0. This can be easily seen by taking the exponential of the intercept estimate.

The regressions using a VaR computed with a 0.90 level of confidence have the highest values of the slope, intercept and R-squared compared with other regressions employing a VaR with a different confidence level. We can therefore assert that less extreme crises are much more likely than more extreme ones and that $\ln RV$ changes have a larger effect on the odds when predicting less extreme events. The prediction accuracy of the model is good as shown by the hit ratios. The logit regression computed with $VaR_{0.99}$ and $T_m = 22$ has the highest prediction accuracy with a hit ratio up to about 88%. Other things being equal, the increase of T_m worsens the regressions outputs for almost every output. However, the intercept and the slopes of the logit regression models are still heavily significant and have similar magnitudes confirming the ability of the RV EWI to predict crises for long horizons.

Assuming a t location-scale distribution for the losses, Table 6 shows that compared to the normal losses case, the performance of the logit regression gets worse when $VaR_{0.90}$ is used. In particular, when $T_m = 132$, the slope becomes statistically insignificant showing that the assumption of a t location-scale distribution is more conservative in terms of crises prediction. Nevertheless, the logit regression computed with $T_m = 66$ has a hit ratio equal to 0.9386 which, excluding the latter case, is the highest hit ratio among the regressions employing normal and t location-scale losses. On the other side, the performance increases when the confidence level is equal to 0.95. For example, the intercept, slope and the R-squared of the various regressions are higher than in the normal losses case. This clearly increases the odds of having a financial crisis in the next 22, 66 and 132 days. Taking $VaR_{0.99}$, the values of the slope, R-squared and hit ratio exceed those of the related logit regression derived with normal losses only in the case of $T_m = 22$. For the other cases, the logit regression under the normal losses performs better.

Table 2
Results from the Kupiec (1995) test.

Distribution	α	VaR	Failure rate	p-value
Normal	0.10	0.0181	0.0650	8.4157e−09
Normal	0.05	0.0232	0.0423	0.1697
Normal	0.01	0.0327	0.0165	0.0097
t location-scale	0.10	0.0136	0.1095	NaN
t location-scale	0.05	0.0198	0.0574	0.2300
t location-scale	0.01	0.0397	0.0107	0.5034

The table reports the failure rates and the p-values of the Kupiec (1995) test for the VaR measure computed with different confidence levels while assuming daily losses of the S&P 500 index futures follow a normal or a t location-scale distribution. When the p-value is undefined, we write NaN (Not a Number).

The reason why we include the t location-scale distribution in our analysis in addition to the normal distribution is that this fits the data particularly well. To see how, we can look at the Q-Q plots shown in Fig. 3. This figure shows that the data points lie very close to the straight line in the t location-scale distribution case. It follows that this distribution is a more reasonable approximation of the S&P 500 index futures daily losses than the normal one. Autocorrelation function plot also shows small correlations and little evidence of statistical significance between losses at different times. As a result, the calculation of VaR and ES under the t location-scale distribution is more natural.

Another way to check the performance of the VaR under the normal or the t location-scale distribution is to implement the Kupiec (1995) test. The goal of this test is to assess if a given VaR estimate is accurate by testing if the failure rate (proportion of time the VaR is exceeded) is not significantly different from the given α. Table 2 shows the p-values and failure rates of this test applied to our VaR measure. It is not difficult to note that the failure rates of the VaR under the t location-scale distribution are closer to the values of α than those of the normal distribution. We can also clearly see that the p-values are large enough when α is equal to 0.01 and 0.05 in the t location-scale case so that there is not enough statistical evidence to reject the null hypothesis. Using $\alpha = 0.10$ and t location-scale losses renders the computation of the p-value impossible even though the failure rate induces to not reject the null hypothesis. Employing normal losses, the Kupiec's test provides statistical evidence when α is set equal to 0.05.

We also investigate our regressions for possible outliers. The observations having a Pearson residual in absolute value greater than 3 are treated as outliers. But since the residuals from the logit regression are discrete, we prefer to use bin residuals as suggested by Gelman and Hill (2006). The number of bins is set equal to 30. We have found no outliers.

The performance of our proposed EWSs is also assessed on the basis of three additional performance measures, the area under the ROC curve (AUC), quadratic probability score (QPS) and log probability score (LPS). The first measure operates on the signals and is computed as the area under the receiving operating characteristic (ROC) curve (Hsieh & Turnbull, 1996). It takes values

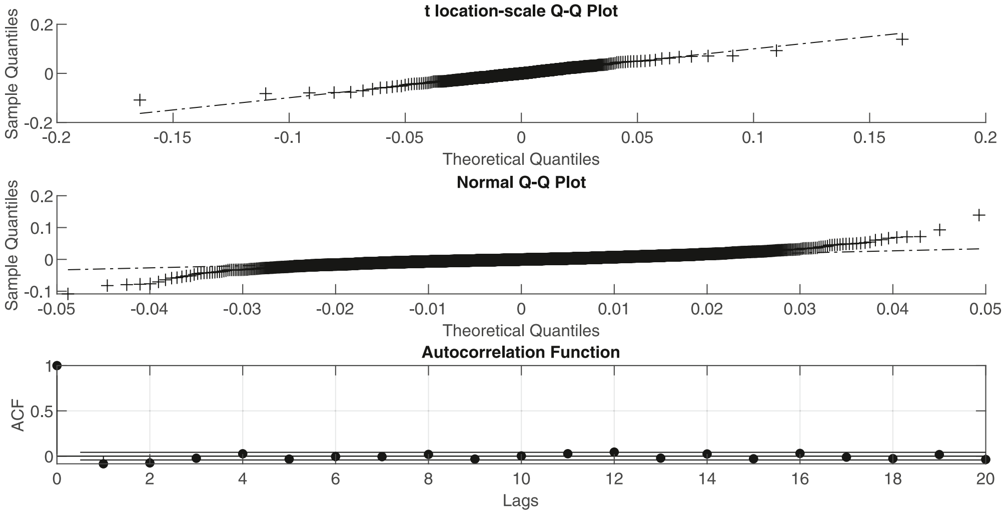


Fig. 3. t location-scale Q-Q plot, normal Q-Q plot and the autocorrelation function of the S&P 500 index futures daily losses.

between 0 and 1, with 0 indicating bad classification, 0.5 inability to distinguish between the 0 and 1 classes and 1 perfect classification. The other two measures are defined (Brier, 1950) and Good (1952)[6] as

$$QPS = \frac{2}{n} \sum_{s=T_w}^{T-T_m} (p_s - Y_s)^2, \tag{24}$$

$$LPS = -\frac{1}{n} \sum_{s=T_w}^{T-T_m} \left[Y_s \ln p_s + (1 - Y_s) \ln(1 - p_s) \right], \tag{25}$$

where $p_s = \mathbb{P}(Y_s = 1)$ and $n = T - T_m - T_w + 1$.

QPS assumes a minimum equal to 0 (perfect accuracy) and a maximum equal to 2, whereas LPS might assume values between 0 (perfect accuracy) and ∞.

Table 7 presents the values assumed by these measures for the first type of logit regressions. We also test whether *AUC* is equal to 0.5 or not using the Wald test statistics which, under the large sample theory, is assumed to have a standard normal distribution (Zhou et al., 2009). The standard error for this test is estimated through the bootstrapping method. We observe that the values of AUC for normal and t location-scale losses are higher when T_m is equal to 22 suggesting that they have a higher ability to distinguish between the two classes. More specifically, when VaR is derived with level of confidence equal to 0.99, the logit regression using t location-scale losses has the highest AUC among the regressions considered. QPS and LPS values are also at satisfactory levels with values equal to 0.0028 and 0.0104 in the $VaR_{0.90}$, $T_m = 132$ and t location-scale losses case. However, as stated before, this model performs poorly from a statistical point of view.

For limit space reasons, we cannot display the ROC curves of all the logit regressions shown in Tables 5 and 6. For this reason, we focus our attention on the logit regression model that uses $VaR_{0.99}$, $T_m = 22$ and t location-scale losses which has the highest AUC. The ROC curve of this model is shown in Fig. 4. In our work, the value 0 denotes as usual the negative class and the value 1 the positive class. On the same figure, we also plot the predicted probabilities derived from the logit regression model together with the crisis events (vertical dashed lines) outlined in the beginning of Section 3 and various probability thresholds (horizontal dashed lines). We note that the crisis events from NASDAQ Crash to Stock Market Crash are well signaled from our EWS. Indeed, we can find high predicted probabilities just before these crisis events. These probabilities are lower for the Argentine Default event. This probably might be because this event occurs outside the USA. This is also in line to what already observed by Mancino and Sanfelici (2020). The authors find that during the Argentine default the volatility remained low and no shocks occurred in the S&P index futures. The Delphi (G.M) Bankruptcy event is however not strongly signaled from the EWS proposed, whereas the Lehman Brothers Collapse is very well signaled. Moreover, the highest predicted probability is reached on 13 October 2008 being approximately equal to 0.9779.

The performance of the previous logit regression is evaluated further by looking at the following confusion matrix (see Table 3).

The probability threshold is set equal to 0.2 and follows Bussiere and Fratzscher (2006). It is well-known that there exists a trade-off between the Type I error (false positive) and Type II error (false negative) when fixing the probability threshold. The confusion matrix shows that the ACC or the proportion correctly predicted is relatively high, while the TPR, also known as the sensitivity or the proportion of actual positive correctly identified, and FNR, or the proportion of actual positive incorrectly identified, have more or less the same value. Note that ACC and hit ratio are the same thing, but we prefer to use the latter

6 See also Diebold and Rudebusch (1989) for an economic application.

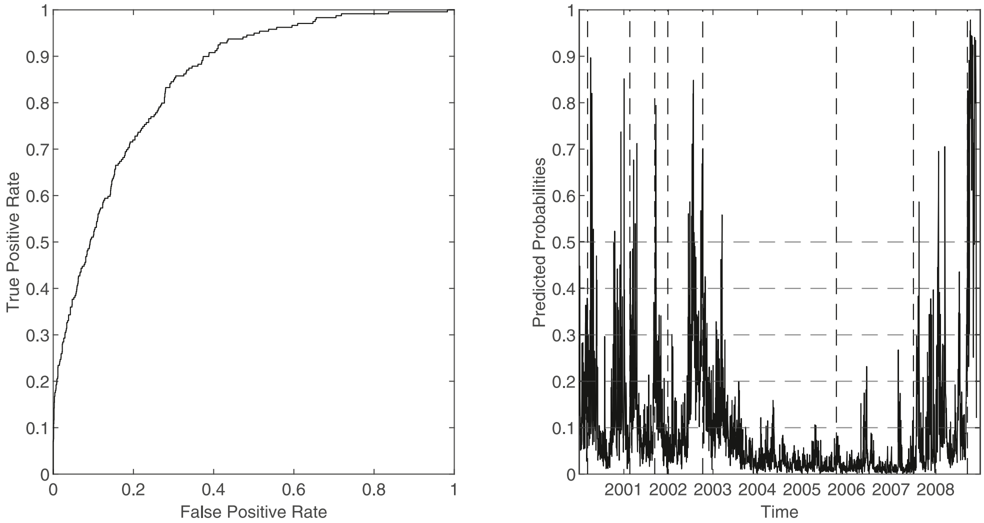


Fig. 4. ROC curve and predicted probabilities of the logit regression using $VaR_{0.99}$, $T_w = 1$, $T_m = 22$ and t location-scale losses.

Table 3
Confusion matrix.

		Actual				
		Positive	Negative	Total	FNR = 0.4728	ACC = 0.8576
Predicted	Positive	TP = 126	FP = 204	330	FPR = 0.1027	PPV = 0.3818
	Negative	FN = 113	TN = 1783	1896	TNR = 0.8973	NSR = 0.1948
	Total	239	1987	2226	TPR = 0.5272	FOR = 0.0596

The table reports the numbers of true positive (TP), false positive (FP), false negatives (FN) and true negatives (TN) of the logit regression using $T_w = 1$, $T_m = 22$, $VaR_{0.99}$ and t location-scale losses. The right part of the table shows the metrics derived from this confusion matrix, that is, the false negative rate (FNR) = FN/(FN+TP), false positive rate (FPR) = FP/(FP+TN), true negative rate (TNR) = TN/(TN+FP), true positive rate (TPR) = TP/(TP+FN), accuracy (ACC) = (TP+TN)/(TP+TN+FP+FN), positive predictive value (PPV) = TP/(TP+FP), noise-to-signal ratio (NSR) = FPR/TPR and false omission rate (FOR) = FN/(FN+TN).

term when the threshold is equal to 0.5. The TNR, also known as specificity or the proportion of actual negative correctly identified, is high and FPR, or the proportion of actual negative incorrectly identified, low. Furthermore, the PPV or the conditional crisis probability says that almost 38% of the times a positive event will be identified as such, where on the other hand, as suggested by FOR, almost 6% of the times a negative event will be identified as positive. Finally, a value of the NSR equal to 0.1984 is to be considered satisfactory if one follows Kaminsky et al. (1998) which suggests that a level of the NSR equal to or higher than unity is not helpful in predicting crises.

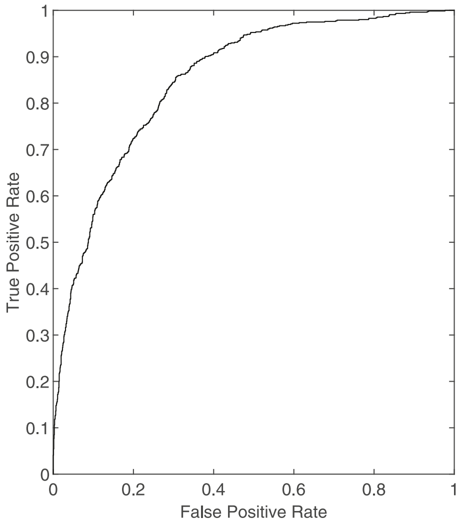
Inspired by the Q-Q plots in Fig. 3, from now on we focus on the t location-scale losses and use these to run the second type of regressions. This type of regression uses the same EWI as the first type but now employs *CEI* rather than *CI* as the indicator crisis variable. T_w still equals 1 and T_m, 22, 66 and 132. The results are reported in Table 8. We find that the regressions computed with $T_m = 22$ and $T_m = 66$ perform better than the analogous regressions using CI in terms of R-squared. The R-squared as well as the magnitude of the slope coefficients decrease as α decreases. The regression computed with $ES_{0.99}$ and $T_m = 22$ has the highest hit ratio among all regressions of second type. We also remark here that the latter regression has the same outputs of the equivalent regression derived with the crisis indicator CI. The second type of regressions using $ES_{0.90}$ produce a higher slope coefficient which, for example, in the case of $T_m = 22$, indicates that for one unit change in $\ln RV$ one would expect about 478% increase in the odds of having a crisis after $T_m = 22$ days as defined by the forward-looking variable Y using *CEI*. The hit ratios increase with the increase of the confidence level. As with the first type of regressions, possible outliers were checked but we didn't find any of these.

Table 9 in Appendix shows that the AUC of the logit regression derived with $ES_{0.90}$, equal to 0.0251, and $T_m = 22$ has an AUC equal to 0.8511, sign this of a good separability. This is also emphasized by the ROC curve in Fig. 5. The related QPS and LPS are on the other side close to zero meaning that this model guarantees a very good accuracy. The Wald test indicates that the null hypothesis $AUC = 0.5$ of all the logit regressions can be rejected.

Table 4
Summary statistics of the EWI λ_{av}.

$T_w = 22$, $T_m = 22$						
Mean	Minimum	Maximum	Std	Kurtosis	Skewness	JB
6.2250	0.0000	14.2372	2.4379	3.0962	0.2831	28.7871
$T_w = 66$, $T_m = 66$						
Mean	Minimum	Maximum	Std	Kurtosis	Skewness	JB
6.1831	1.5819	11.0734	1.6269	3.1007	0.0952	3.9620
$T_w = 132$, $T_m = 132$						
Mean	Minimum	Maximum	Std	Kurtosis	Skewness	JB
6.2039	3.4802	9.9661	1.2347	3.9701	0.8808	334.5054

The table reports the mean, minimum, maximum, standard deviation (Std), kurtosis, skewness and the Jarque–Bera test statistic (JB) of the EWI λ_{av} assuming that $T_w = T_m = 22$, $T_w = T_m = 33$ and $T_w = 132 = T_m = 132$.



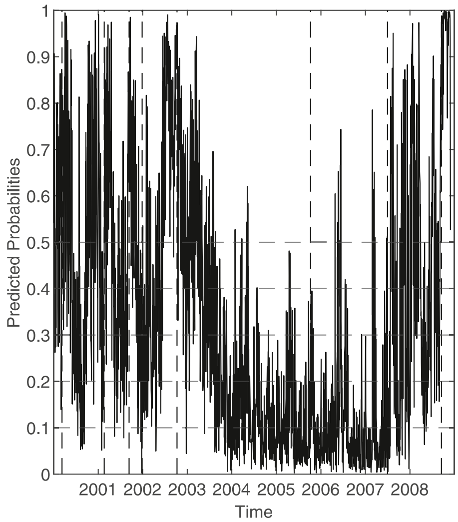


Fig. 5. ROC curve and predicted probabilities of the logit regression using $ES_{0.90}$, $T_w = 1$, $T_m = 22$ and t location-scale losses.

The predicted probabilities in the second panel of Fig. 5 increase considerably when compared to those of Fig. 4. The crisis events described in Section 3 are well predicted by this EWS. The probabilities are especially high for the first three crisis events, the Stock Market Crash and Lehman Brother Collapse crisis event. The highest predicted probability is reached on 10 October 2008 and is equal to 0.9992.

3.2.2. The third and the fourth types of regressions

Let's now turn our attention to the results obtained by running the third type of regressions. As already mentioned in the previous section, we consider only the situation in which losses follow a t location-scale distribution. We present these results in Table 10 shown in Appendix by fitting separately the full model containing the two EWIs RV_{av} and λ_{av} and the two reduced models employing separately RV_{av} or λ_{av} as their independent variable. However, the table shows only the results for which the level of confidence is equal to 0.95 and 0.99. When the level of confidence is set equal to 0.90, the coefficient of λ_{av} is statistically significant only when $T_w = 66$ and $T_m = 66$. Also, only in this occasion, the sign of this coefficient is negative and the chi-square test indicates that the use of the full model is better than the reduced models. We will try to give an explanation of this sign in the following of this section. Additionally, the coefficient of the RV_{av} variable is not statistically significant when $T_w = 132$ and $T_m = 132$. Full results, however, can be obtained from the corresponding author.

By focusing on the full model, we note that as α decreases, the magnitude of the intercept and coefficient of RV_{av} decreases. On the opposite, the coefficient of λ_{av} increases 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} goes up. We also note a decrease in the values of R-squared as we move from a high value to a low value of α. Ruling out the case $T_w = 22$ and $T_m = 22$, the result is still true for the hit ratio. The regression using $VaR_{0.95}$ and $T_w = 132$, $T_m = 132$ has the highest R-squared among all the full and reduced regressions. The coefficient attached to λ_{av} assumes the highest value when $T_w = 132$, $T_m = 132$ and $VaR_{0.99}$ and that to RV_{av} when $T_w = 132$, $T_m = 132$ and $VaR_{0.95}$. From the full model, it also emerges that the coefficient of the λ_{av} variable is negative when $T_w = 132$, $T_m = 132$ and $VaR_{0.95}$.

Interpreting the coefficient of the logit regressions is relatively easy. Take the first regression shown in Table 10. The coefficient associated to λ_{av} can be interpreted as follows - with RV_{av} held constant, incrementing λ_{av} of one unit increases the odds of having financial crises by about 13%. To interpret the coefficient of RV_{av} we can use the properties of the logarithm to write $RV_{av,s}$ as

$$RV_{av,s} = \frac{\ln \Pi_{i=s-T_w+1}^{s} RV_i}{T_w}. \tag{26}$$

Suppose that λ_{av} is held constant. It is then not difficult to see that the difference between the two log-odds computed at any two values $RV_{av,s}^{a}$ and $RV_{av,s}^{b}$ is equal to

$$\frac{\beta_{RV_{av}}}{T_w} \ln \frac{\Pi_{i=s-T_w+1}^{s} RV_i^{a}}{\Pi_{i=s-T_w+1}^{s} RV_i^{b}}. \tag{27}$$

Taking the exponential and substituting the value of the RV_{av} coefficient ($\beta_{RV_{av}}$) derived with the first regression shown in Table 10 and T_w, we can say that if $\Pi_{i=s-T_w+1}^{s} RV_i$ increases by 10%, odds would increase by about 111%.

The reduced regressions obtained with the EWI λ_{av} show a positive association between this EWI and the log-odds. However, as seen before, when $T_w = 132$, $T_m = 132$ and $VaR_{0.95}$, the sign of the coefficient of λ_{av} is negative. Here, the good news is that this coefficient is statistically insignificant. The association becomes again stronger as α decreases with a coefficient equal to 0.3967 when $T_w = 132$, $T_m = 132$ and $VaR_{0.99}$. With RV_{av} held constant, this will increase the odds of having financial crises with about 49% when λ_{av} increases with one unit. This indicates that the ability of λ_{av} in predicting future crises is better when these are more extreme, forecasting horizon is longer and more data are used to compute it. This is the same as saying that experiencing a turbulent period (as measured by λ) for much longer can increase the probability of a crisis over the next 132 days more significantly. Note also that excluding the case where λ_{av} is statistically insignificant, the sign of the intercept for these regressions is always negative which says that a crisis event is less likely when $\lambda_{av} = 0$. This result is in strong agreement with the theory presented in Section 2.1.

What is still worth discussing from these regressions are the low values of R-squared. This is however not a problem since the p-values are fairly low indicating that a relationship exists.

The other reduced regressions using RV_{av} have significant coefficients for all levels of confidence and combinations of T_w and T_m. As already observed, the relative coefficients and p-values decrease as α decreases. Moreover, the regression ran with $T_w = 132$, $T_m = 132$ and $VaR_{0.95}$ has the highest slope coefficient and R-squared, whereas the regression using $T_w = 22$, $T_m = 22$ and $VaR_{0.99}$ the highest hit ratio. Letting again $T_w = 132$, $T_m = 132$, $VaR_{0.95}$ and bringing the results of the logit regressions computed with these parameters together, we note that λ_{av} behaves as a suppressor variable (Conger, 1974). This can be noticed by comparing the full and the reduced model employing RV_{av}. The predictive ability, as measured by the R-squared, and the coefficient of the RV_{av} variable increase their magnitude when λ_{av} is introduced in the model. Besides this, we observe that the reduced regression using λ_{av} shows no significant relationship between this EWI and the forward-looking variable, while the correlation with RV_{av} is positive which might also explain the negative sign of the λ_{av} coefficient (McNemar, 1945).

Comparing the full with the reduced models, we observe that the values of R-squared and goodness are always better than those of reduced models. As for the hit ratio, this is true only when the forecasting horizon starts to increase. Note also how in the cases $T_w = T_m = 66$, $T_w = T_m = 132$ and $VaR_{0.99}$, the reduced model using λ_{av} has the highest hit ratio. We also implement the chi-square test to determine whether the model with the EWIs RV_{av} and λ_{av} is better than the model with RV_{av} or λ_{av} alone. The test suggests that there is enough statistical evidence to reject the null hypothesis.

An interaction term was also included to our full regressions. The chi-square test shows that a full model including this new term may be appropriate. But we found high collinearity between λ_{av} and the interaction term as shown by the Pearson's correlation coefficient. The same coefficient applied to RV_{av} and λ_{av} shows moderate correlation. Our findings are also confirmed by the Belsley collinearity diagnostics (Belsley et al., 1980).

We conclude the discussion of Table 10 by reporting in Table 4 some statistics of the EWI λ_{av}. In particular, we note that when $T_w = 66$ and $T_m = 66$ the JB test indicates that λ_{av} is normally distributed.

As previously, we checked our regressions for probable outliers and find only one outlier in the case of the reduced regression using the EWI λ_{av}, $T_w = 66$, $T_m = 66$ and $VaR_{0.99}$.

The performance measures AUC, QPS and LPS of the regressions considered in Table 10 are shown in Table 11. It can be noticed that full regressions have almost always better performance measures than the reduced models and that the reduced models using RV_{av} perform much better than those using λ_{av}. Our EWS benefits greatly from the introduction of λ_{av} especially in the cases where AUC is significantly different from the AUC of the reduced regression using RV_{av} as its EWI. The test statistics used to provide evidence is obtained as follows

$$D = \frac{AUC_1 - AUC_2}{\mathrm{Std}(AUC_1 - AUC_2)}, \tag{28}$$

where D is compared to the standard normal distribution, AUC_1 and AUC_2 are the AUCs of the full and reduced model, respectively, and the standard deviation (Std) is computed using the bootstrapping technique.

In the special case where T_w and T_m are equal to 132 and VaR is computed with 0.95 level of confidence, the

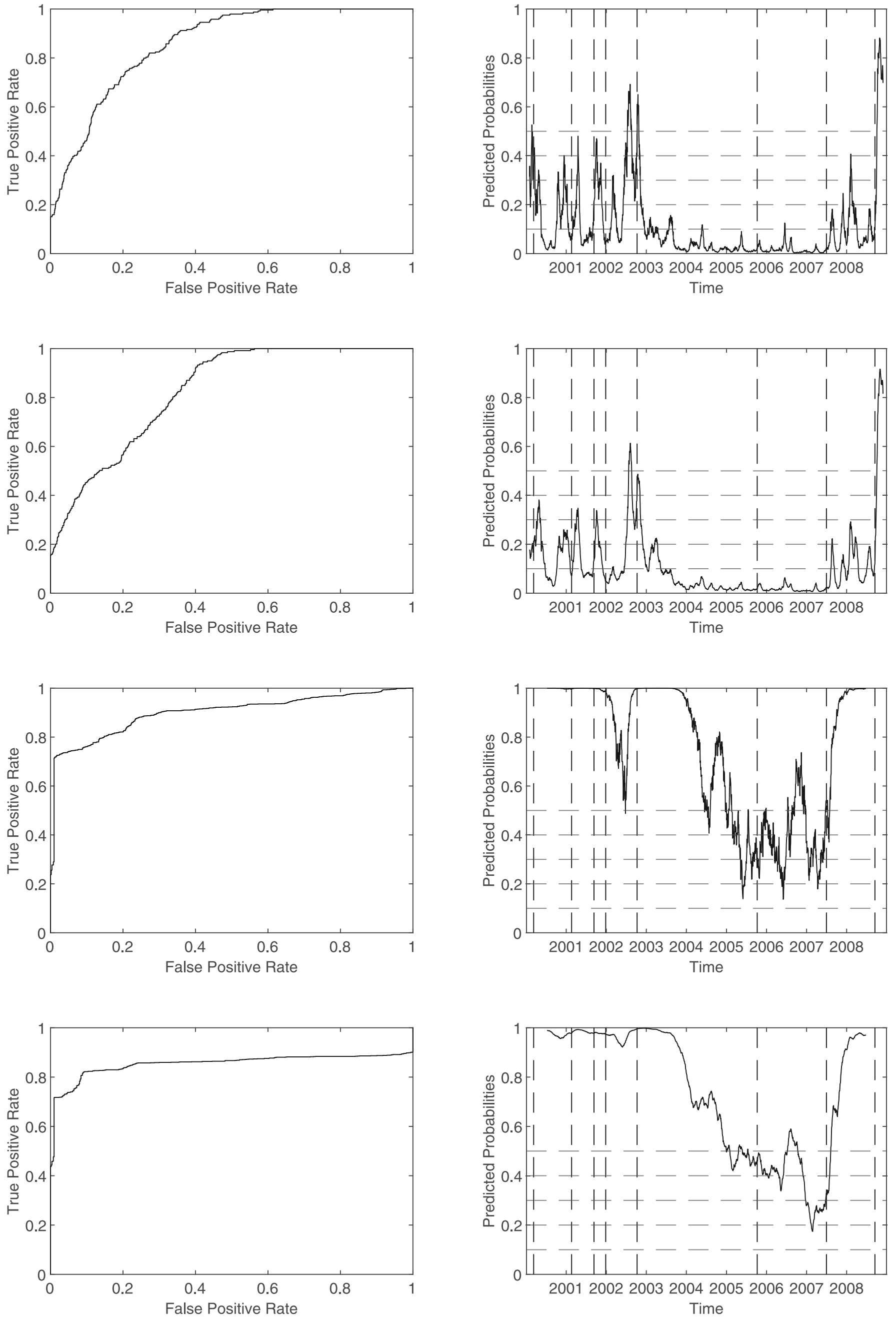


Fig. 6. From left to right: The first two panels show the ROC curve and predicted probabilities of the full logit regression using $VaR_{0.99}$, $T_w = 22$, $T_m = 22$ and t location-scale losses. The next two panels do the same, but now consider only the EWI RV_{av}. For the remaining panels, we set $VaR_{0.95}$ and $T_w = 132$, $T_m = 132$ and show the ROC curve and predicted probabilities of the full and reduced model using the RV_{av} variable.

full model has the highest AUC among all the regressions. On the other hand, when $T_w = 22$, $T_m = 22$ and $VaR_{0.99}$ the full model presents the best results in terms of QPS and LPS.

Fig. 6 shows the ROC curves and the predicted probabilities of the logit regressions mentioned above. In the same figure, we also plot the graphs for the reduced models computed with RV_{av}.

The predicted probabilities of the first full model with $VaR_{0.99}$, $T_w = 22$ and $T_m = 22$ reveal the ability of our EWS in predicting the crisis events starting with NASDAQ Crash and ending with Stock Market Crash. The predicted probabilities are especially higher just before the Twin Tower Attacks and Stock Market Crash event. However, the predicted probabilities are low during the Stock Market Crash and Subprime Crisis event. The maximum of the predicted probability equal to 0.8818 is reached on 29 October 2008. On the other side, the graph of the predicted probabilities of the related reduced regression shows a similar behaviour, but now it makes a worse prediction for the first four crisis events. The maximum of these probabilities increases to 0.9163 and is reached on 31 October 2008.

The other full regression model with $VaR_{0.95}$, $T_w = 132$ and $T_m = 132$ produces higher predicted probabilities for the Turkish Crisis, Twin Tower Attacks, Argentine Default and the Subprime Crisis event than the reduced counterpart model even though these probabilities are very high for both models. The crisis event Stock Market Crash and Delphi (G. M) Bankruptcy are predicted slightly better from the reduced model. The highest predicted probabilities equal to 1.0000 and 0.9981 for the full and reduced model are achieved on 29 April 2003 and 12 December 2002. It is important to note that in these regressions the probabilities are very high compared to the previous situation. This is probably because the information contained in the two EWIs RV_{av} and λ_{av} is higher given that they use more data when computed. Moreover, the first crisis event delineated in Section 3.1 remains outside our prediction. Indeed, in these regressions, probabilities are predicted starting from day 11 July 2000 to 24 June 2008. Look however how the Lehman Brothers Collapse is predicted very well from both models with the full model predicting slightly better.

Tables 12 and 13 illustrate the results of the confusion matrices for the two regressions discussed above. The probability threshold is set again equal to 0.2. These results basically show that the full models have better performance metrics than the reduced ones. Focusing on some of them, it can be deduced that the full models identify better true positives and negatives and have a lower NSR. Note how in the case of the second regression (Table 13) the full and reduced regression have a very high sensitivity and a very low specificity. However, the reduced model reports a specificity equal to zero suggesting that this model returns a positive result (crises) for 100% of the non-crises events that in turn can be more costly for the responsible authorities. This result can also be easily deduced by looking at the metrics FOR.

Finally, let's take a look at the results of the fourth type of regressions shown in Table 14 in Appendix. Observe that as we have often done in this paper we are focusing our calculations on the t location-scale losses case. As noted when we discussed the previous types of regressions, the magnitude of the coefficient of RV_{av} decreases with α, while the magnitude of the λ_{av} coefficient increases for at least all regressions. This is true for both full and reduced regressions. In every situation, the full regression models have a higher R-squared when compared to the reduced models. More particularly, the model employing $T_w = 132$, $T_m = 132$ and $ES_{0.90}$ has the highest R-squared. The slope coefficients have higher statistical significance in reduced regressions than in full ones. The reduced models utilizing the EWI λ_{av} have low R-squared compared to the other models. However, in some cases, they show the highest values of the hit ratio. The maximum value of this indicator is obtained for the reduced model employing RV_{av}, $T_w = 22$, $T_m = 22$ and $ES_{0.99}$.

The chi-square test suggests that a model with an additional interaction term is preferred for six out of nine logit regressions shown in Table 14. However, the problem of collinearity is present similarly as for the third type of regressions. The full model is, on the contrary, preferred over the reduced models in all cases. We found no outliers in our regressions.

The results in Table 15 show that the measures AUC, QPS and LPS are always better for the full model. This remains true for all the combinations of T_w, T_m and VaR. However, we decide to report only the results for which there is a significant difference between the AUC of the full model and that of the reduced model employing RV_{av}. As stated before, results are available upon request from the corresponding author. Here, we are excluding the case $T_w = 22$, $T_m = 22$ and $ES_{0.99}$ since this logit regression has the same outputs as the equivalent logit regression displayed in Table 10. Consequently, it also shares the same performance measures. Also worth noting that the difference between the two AUCs is always significant when $\alpha = 0.01$ demonstrating the benefits of using λ_{av} in these particular situations. As a special case, we show in Fig. 7 the predicted probabilities of the full logit regression obtained with $T_w = 132$, $T_m = 132$ and $ES_{0.90}$ together with the related reduced regression using RV_{av}. This also corresponds to the case where the full model has the highest AUC. Excluding the Delphi (G.M) Bankruptcy and the Subprime Crisis event, both models show high predicted probabilities for the other crisis events. On the other hand, TPR is equal to 0.8973 and 0.9109 for the full and reduced model, whereas TNR equal to 0.6542 and 0.6916 respectively.

4. Further considerations

To conclude our empirical application, we report in Table 16 in Appendix the fractions of ones of the forward-looking variable Y for each of the four types of regressions discussed in the previous section. The fractions are shown for t location-scale losses and for each combination of T_w, T_m and α. What emerges from Table 16 is the fact that when $\alpha = 0.10$ and T_m is equal to 66 or 132, the fractions of ones for the first and third types of regressions are very high. This is particularly evident when $T_m = 132$. It is no coincidence that in this situation Y has little variation and we did not find significant results.

As natural, moving from a higher to a lower level of α results in a decrease in the fractions of ones making the sample more unbalanced. We also notice from Table 16 that the samples used by many regressions are well-balanced.

Different studies report problems of the logit regression with rare events data. However, this is particularly a problem when the sample size is large (in thousands or millions) and the event rate (the fraction of ones in our study) very small (King & Zeng, 2001 and Fithian & Hastie, 2014).

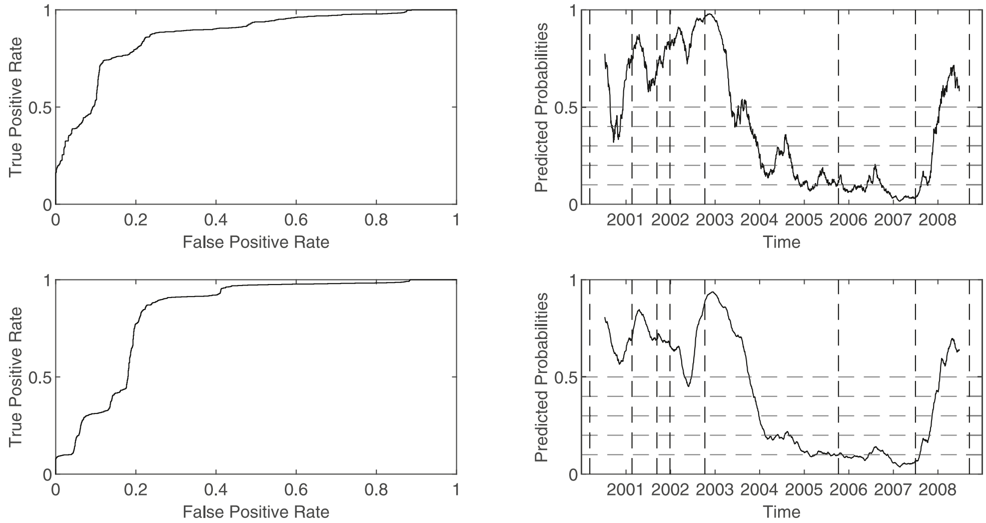


Fig. 7. From left to right: The first two panels show the ROC curve and predicted probabilities of the full logit regression using $ES_{0.90}$, $T_w = 132$, $T_m = 132$ and t location-scale losses. The next two panels do the same, but now consider only the EWI RV_{av}.

We will evaluate the out-of-sample performance of our EWS models using 10-fold cross-validation. This procedure randomly divides the original dataset into 10 non-overlapping folds having approximately equal size. While the first fold is treated as the validation set, the various logit regressions are ran on the remaining 9 folds. We compute on these folds the measures AUC, QPS, LPS and hit ratio. Finally, the procedure is repeated 10 times with each of the 10 folds used precisely once as validation set. The 10-fold AUC, QPS, LPS and hit ratio estimates are then obtained by averaging the values we found previously. The choice of the 10-fold cross-validation follows the suggestions of James et al. (2013).

An excerpt from our 10-fold cross validation is presented in Table 17. Due to space reasons, we show only the results for the third type of regressions. Those that we would conclude for this regression however remain true for all the regressions considered in this paper. When looking at the results in Table 17, we note that the four performance measures used to validate our EWS are almost similar to the same measures obtained from the original regression models. This indicates that the performance of our EWSs do not suffer in cross-validation.

5. Conclusions

We propose and analyze different logit regression Early Warning Systems (EWSs) where Early Warning Indicators (EWIs) based on the daily realized variance and on the integrated feedback rate are considered. The forward-looking variable is defined in terms of the Value-at-risk (VaR) or Expected shortfall (ES). We test the proposed EWSs by using tick-by-tick prices of the S&P 500 index futures covering the period from 3 January 2000 to 31 December 2008.

The EWSs suggest that the EWIs using realized variance (RV) can help in predicting financial crises. Indeed, we find positive significant regression coefficients for these EWIs. We also often find that in many regressions the magnitude of the coefficient of RV decreases as the significance level decreases. On the opposite, in nearly all cases, the estimated regression coefficients of the EWI employing the price-volatility feedback rate increase their magnitude when the significance level decreases. This suggests a stronger capacity of this EWI to predict more extreme financial crises. The logit regression models ran with this EWI and the EWI using the average of the logarithms of RVs are preferred to the single reduced models.

The results are also robust to changes in the estimation and prediction intervals. The prediction ability of our EWSs is also confirmed by the 10-fold cross validation analysis and it is not affected by the class imbalance problem and the outliers presence. As a consequence, we can assert that the EWSs proposed can be a useful tool in predicting financial instabilities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was partially supported by the INdDAM-GNCS project MZG1VARI01.

Appendix

See Tables 5–17.

Table 5
Logit output results based on the first type of regressions.

$T_w = 1, T_m = 22$ - $VaR_{0.90}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
15.0413	9.8344e−103	1.5625	6.1474e−105	0.2548	0.7475	1.3191e−172
$T_w = 1, T_m = 22$ - $VaR_{0.95}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
13.4594	6.4660e−88	1.4810	1.2442e−95	0.2459	0.7722	3.3509e−154
$T_w = 1, T_m = 22$ - $VaR_{0.99}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
11.1373	3.2734e−50	1.4042	2.4566e−64	0.2488	0.8814	1.8812e−97
$T_w = 1, T_m = 66$ - $VaR_{0.90}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
14.5245	4.8927e−92	1.3975	1.6507e−86	0.2021	0.7475	3.225e−125
$T_w = 1, T_m = 66$ - $VaR_{0.95}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
12.5196	5.8323e−87	1.2718	7.9362e−87	0.1819	0.7204	1.5954e−121
$T_w = 1, T_m = 66$ - $VaR_{0.99}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
9.8160	1.0361e−54	1.1448	1.6242e−65	0.1565	0.7663	1.8746e−85
$T_w = 1, T_m = 132$ - $VaR_{0.90}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
14.2413	3.2438e−75	1.3050	1.3049e−66	0.1755	0.7495	2.2283e−91
$T_w = 1, T_m = 132$ - $VaR_{0.95}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
9.9795	4.2902e−62	0.9512	9.3414e−57	0.1132	0.6730	4.2169e−70
$T_w = 1, T_m = 132$ - $VaR_{0.99}$ - Normal case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
10.3626	6.2100e−66	1.1237	6.0279e−72	0.1522	0.6834	1.3383e−94

The table reports the outputs of the logit regressions applied to the forward-looking variable defined in Eq. (18) based on the daily returns of the S&P 500 index futures. T_w is set equal to 1 and T_m equal to 22, 66 and 132. VaR is computed using the normal distribution. The sample contains respectively 2226, 2182 and 2116 observations.

Table 6
Logit output results based on the first type of regressions.

$T_w = 1, T_m = 22$ - $VaR_{0.90}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
12.3882	3.6161e−78	1.1625	1.9468e−70	0.1608	0.7192	1.0719e−96
$T_w = 1, T_m = 22$ - $VaR_{0.95}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
14.7084	4.3805e−100	1.5497	9.4716e−104	0.2542	0.7597	1.4829e−170
$T_w = 1, T_m = 22$ - $VaR_{0.99}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
11.0670	1.7213e−44	1.4353	1.0065e−58	0.2620	0.9066	1.7412e−88
$T_w = 1, T_m = 66$ - $VaR_{0.90}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
11.1983	1.2888e−24	0.8411	1.2973e−15	0.0705	0.9386	3.5075e−17
$T_w = 1, T_m = 66$ - $VaR_{0.95}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
13.8173	1.3779e−89	1.3354	1.8939e−84	0.1905	0.7360	9.1837e−120
$T_w = 1, T_m = 66$ - $VaR_{0.99}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
8.9275	2.3902e−46	1.0670	7.5983e−58	0.1397	0.7782	9.6456e−73
$T_w = 1, T_m = 132$ - $VaR_{0.90}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
14.2883	0.0369	0.7646	0.2426	0.0318	0.9986	0.22963
$T_w = 1, T_m = 132$ - $VaR_{0.95}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
14.2413	3.2438e−75	1.3050	1.3049e−66	0.1755	0.7495	2.2283e−91
$T_w = 1, T_m = 132$ - $VaR_{0.99}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
9.3204	1.5409e−57	1.0234	3.6855e−64	0.1318	0.6829	3.796e−81

The table reports the outputs of the logit regressions applied to the forward-looking variable defined in Eq. (18) based on the daily returns of the S&P 500 index futures. T_w is set equal to 1 and T_m equal to 22, 66 and 132. VaR is computed using the t location-scale distribution. The sample contains respectively 2226, 2182 and 2116 observations.

Table 7
Performance measures of the first type of regressions.

$T_w = 1, T_m = 22$ - Normal case	AUC	QPS	LPS
$VaR_{0.90}$	0.8228***	0.3401	0.5153
$VaR_{0.95}$	0.8206***	0.3130	0.4821
$VaR_{0.99}$	0.8380***	0.1808	0.2976
$T_w = 1, T_m = 66$ - Normal case	AUC	QPS	LPS
$VaR_{0.90}$	0.8041***	0.3419	0.5125
$VaR_{0.95}$	0.7829***	0.3810	0.5665
$VaR_{0.99}$	0.7714***	0.3142	0.4741
$T_w = 1, T_m = 132$ - Normal case	AUC	QPS	LPS
$VaR_{0.90}$	0.7999***	0.3076	0.4562
$VaR_{0.95}$	0.7321***	0.4000	0.5800
$VaR_{0.99}$	0.7623***	0.3828	0.5603
$T_w = 1, T_m = 22$ - t location-scale case	AUC	QPS	LPS
$VaR_{0.90}$	0.7727***	0.3451	0.5104
$VaR_{0.95}$	0.8230***	0.3343	0.5107
$VaR_{0.99}$	0.8498***	0.1481	0.2516
$T_w = 1, T_m = 66$ - t location-scale case	AUC	QPS	LPS
$VaR_{0.90}$	0.7278***	0.1130	0.2146
$VaR_{0.95}$	0.7945***	0.3545	0.5273
$VaR_{0.99}$	0.7628***	0.3018	0.4592
$T_w = 1, T_m = 132$ - t location-scale case	AUC	QPS	LPS
$VaR_{0.90}$	0.7306^{LTCF}	0.0028	0.0104
$VaR_{0.95}$	0.7999***	0.3076	0.4562
$VaR_{0.99}$	0.7482***	0.3884	0.5667

The table displays the performance measures AUC, QPS and LPS of the logit regressions obtained using the forward-looking variable defined in Eq. (18) and the daily returns of the S&P 500 index futures. T_w is set equal to 1 and T_m equal to 22, 66 and 132. VaR is computed using the normal and the t location-scale distribution. The sample contains respectively 2226, 2182 and 2116 observations. *** indicates rejection of null hypothesis test $AUC = 0.5$ at 0.01% level of significance. LTCF stands for "less than two classes are found" and therefore the p-value cannot be defined.

Table 8
Logit output results based on the second type of regressions.

$T_w = 1, T_m = 22$ - $ES_{0.90}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
16.0567	3.1889e−98	1.7546	5.3506e−105	0.2997	0.7848	7.7089e−188
$T_w = 1, T_m = 22$ - $ES_{0.95}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
15.2104	2.3851e−90	1.6977	8.2251e−99	0.2921	0.7934	1.8629e−173
$T_w = 1, T_m = 22$ - $ES_{0.99}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
11.0670	1.7213e−44	1.4353	1.0065e−58	0.2620	0.9066	1.7412e−88
$T_w = 1, T_m = 66$ - $ES_{0.90}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
15.7642	2.4653e−98	1.6933	1.0184e−103	0.2681	0.7566	4.0465e−38
$T_w = 1, T_m = 66$ - $ES_{0.95}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
14.8973	1.2699e−89	1.6386	5.3356e−97	0.2579	0.7727	6.3058e−155
$T_w = 1, T_m = 66$ - $ES_{0.99}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
9.6579	3.3009e−47	1.1692	2.2691e−59	0.1600	0.8103	7.4972e−77
$T_w = 1, T_m = 132$ - $ES_{0.90}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
12.3024	2.1455e−80	1.3065	6.4725e−85	0.1895	0.7108	7.0519e−120
$T_w = 1, T_m = 132$ - $ES_{0.95}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
10.9148	1.1513e−69	1.1822	1.0043e−75	0.1642	0.6971	9.2605e−102
$T_w = 1, T_m = 132$ - $ES_{0.99}$ - t location-scale case						
β_0	$p-$ value	$\ln RV$	$p-$ value	R-squared	Hit ratio	Goodness
3.9508	2.6489e−12	0.5514	1.0981e−20	0.0430	0.7949	8.6121e−22

The table reports the outputs of the logit regressions applied to the forward-looking variable defined in Eq. (19) based on the daily returns of the S&P 500 index futures. T_w is set equal to 1 and T_m equal to 22, 66 and 132. VaR and ES are computed using the t location-scale distribution. The sample contains respectively 2226, 2182 and 2116 observations.

Table 9
Performance measures of the second type of regressions.

$T_w = 1, T_m = 22$ - t location-scale case	AUC	QPS	LPS
$ES_{0.90}$	0.8511***	0.2930	0.4486
$ES_{0.95}$	0.8463***	0.2776	0.4292
$ES_{0.99}$	0.8498***	0.1481	0.2516
$T_w = 1, T_m = 66$ - t location-scale case	AUC	QPS	LPS
$ES_{0.90}$	0.8372***	0.3194	0.4827
$ES_{0.95}$	0.8342***	0.3050	0.4636
$ES_{0.99}$	0.7843***	0.2680	0.4142
$T_w = 1, T_m = 132$ - t location-scale case	AUC	QPS	LPS
$ES_{0.90}$	0.7902***	0.3696	0.5478
$ES_{0.95}$	0.7711***	0.3746	0.5516
$ES_{0.99}$	0.6555***	0.3134	0.4837

The table displays the performance measures AUC, QPS and LPS of the logit regressions obtained using the forward-looking variable defined in Eq. (19) and the daily returns of the S&P 500 index futures. T_w is set equal to 1 and T_m equal to 22, 66 and 132. VaR and ES are computed using the t location-scale distribution. The sample contains respectively 2226, 2182 and 2116 observations. *** indicates rejection of null hypothesis test $AUC = 0.5$ at 0.01% level of significance.

Table 10
Logit output results based on the third type of regressions.

$T_w = 22$, $T_m = 22$ - $VaR_{0.95}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	18.0473	9.5475e−101	18.9383	1.5512e−115	−1.2347	3.8326e−23
RV_{av}	1.9740	3.5979e−114	1.9859	7.6320e−119		
λ_{av}	0.1242	1.1679e−08			0.1531	2.0474e−16
R-squared	0.3104		0.2993		0.0347	
Hit ratio	0.7578		0.7714		0.5941	
Goodness		1.4947e−198		8.2788e−193		4.1564e−17
$T_w = 22$, $T_m = 22$ - $VaR_{0.99}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	11.3467	2.8515e−33	12.7697	1.5429e−41	−3.5829	1.5942e−63
RV_{av}	1.6024	7.8047e−54	1.5980	5.8441e−53		
λ_{av}	0.2119	2.3621e−12			0.2212	3.7409e−15
R-squared	0.2679		0.2345		0.0445	
Hit ratio	0.9011		0.9034		0.8916	
Goodness		5.5803e−88		2.2914e−78		2.3756e−15
$T_w = 66$, $T_m = 66$ - $VaR_{0.95}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	18.8842	1.8680e−89	19.7039	8.5115e−104	−0.4254	0.0180
RV_{av}	1.9162	1.1066e−98	1.9334	8.5008e−101		
λ_{av}	0.1064	0.0033			0.1587	3.4529e−08
R-squared	0.2493		0.2463		0.0315	
Hit ratio	0.8035		0.8063		0.6297	
Goodness		5.8742e−142		1.4218e−141		2.2781e−08
$T_w = 66$, $T_m = 66$ - $VaR_{0.99}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	7.3775	5.9416e−18	8.4846	7.3142e−27	−2.5510	5.1846e−30
RV_{av}	0.9825	1.8223e−31	1.0214	9.2599e−35		
λ_{av}	0.1167	3.5094e−04			0.1898	1.9179e−08
R-squared	0.1561		0.1505		0.0917	
Hit ratio	0.7940		0.7851		0.7950	
Goodness		3.0366e−40		1.2010e−38		1.4998e−08
$T_w = 132$, $T_m = 132$ - $VaR_{0.95}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	62.6174	1.5113e−53	32.2423	1.7006e−80	1.4549	2.5311e−08
RV_{av}	5.2447	7.1335e−56	3.0822	9.2769e−79		
λ_{av}	−1.2789	9.7240e−31			−0.0639	0.1186
R-squared	0.4355		0.3565		0.0332	
Hit ratio	0.8484		0.8055		0.7421	
Goodness		7.8809e−206		3.5723e−167		0.1202
$T_w = 132$, $T_m = 132$ - $VaR_{0.99}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	9.9899	1.2760e−29	12.3911	1.2731e−53	−3.2677	3.2265e−36
RV_{av}	1.2924	2.0850e−51	1.3550	2.5440e−59		
λ_{av}	0.2863	1.0933e−12			0.3967	4.8490e−23
R-squared	0.2348		0.2157		0.1402	
Hit ratio	0.7164		0.6650		0.7421	
Goodness		7.9391e−80		1.1217e−69		3.3892e−24

The table reports the outputs of the third type of logit regressions applied to the forward-looking variable defined in Eq. (18) based on the daily returns of the S&P 500 index futures. T_w is assumed to be equal to T_m and is set equal to 22, 66 and 132. VaR is computed using the t location-scale distribution. The sample contains respectively 2205, 2117 and 1985 observations.

Table 11
Performance measures on the third type of regressions.

$T_w = 22, T_m = 22$ - t location-scale case	AUC	QPS	LPS
$VaR_{0.95}$ - Full model	0.8461	0.3117	0.4767
$VaR_{0.95}$ - Reduced model RV_{av}	0.8420	0.3180	0.4844
$VaR_{0.95}$ - Reduced model λ_{av}	0.6067	0.4748	0.6673
$VaR_{0.99}$ - Full model	0.8576*	0.1521	0.2520
$VaR_{0.99}$ - Reduced model RV_{av}	0.8247	0.1561	0.2635
$VaR_{0.99}$ - Reduced model λ_{av}	0.6402	0.1873	0.3289
$T_w = 66, T_m = 66$ - t location-scale case	AUC	QPS	LPS
$VaR_{0.95}$ - Full model	0.8094	0.3251	0.5040
$VaR_{0.95}$ - Reduced model RV_{av}	0.8055	0.3228	0.5061
$VaR_{0.95}$ - Reduced model λ_{av}	0.5487	0.4592	0.6502
$VaR_{0.99}$ - Full model	0.7095	0.3066	0.4643
$VaR_{0.99}$ - Reduced model RV_{av}	0.6917	0.3115	0.4673
$VaR_{0.99}$ - Reduced model λ_{av}	0.5477	0.3179	0.4997
$T_w = 132, T_m = 132$ - t location-scale case	AUC	QPS	LPS
$VaR_{0.95}$ - Full model	0.9003***	0.2199	0.3330
$VaR_{0.95}$ - Reduced model RV_{av}	0.8542	0.2634	0.3796
$VaR_{0.95}$ - Reduced model λ_{av}	0.6275	0.3826	0.5703
$VaR_{0.99}$ - Full model	0.7751**	0.3589	0.5323
$VaR_{0.99}$ - Reduced model RV_{av}	0.7337	0.3767	0.5457
$VaR_{0.99}$ - Reduced model λ_{av}	0.5569	0.4056	0.5982

The table displays the performance measures AUC, QPS and LPS of the logit regressions obtained using the forward-looking variable defined in Eq. (18) and the daily returns of the S&P 500 index futures. T_w is assumed to be equal to T_m and is set equal to 22, 66 and 132. VaR is computed using the t location-scale distribution. The sample contains respectively 2205, 2117 and 1985 observations. *, ** and *** indicate rejection of null hypothesis test $AUC_{\text{Full model}} = AUC_{\text{Reduced model}}$ at 5%, 1% and 0.01% level of significance. The reduced model is derived using the RV_{av} variable.

Table 12
Confusion matrix.

		Actual				
		Positive	Negative	Total	FNR=0.4142	ACC=0.9496
Predicted	Positive	TP = 140	FP = 239	379	FPR=0.1216	PPV=0.3694
	Negative	FN = 99	TN = 1727	1826	TNR=0.8784	NSR=0.2076
	Total	239	1966	2205	TPR=0.5858	FOR=0.0542

		Actual				
		Positive	Negative	Total	FNR=0.5146	ACC=0.8331
Predicted	Positive	TP = 116	FP = 245	361	FPR=0.1246	PPV=0.3213
	Negative	FN = 123	TN = 1721	1844	TNR=0.8754	NSR=0.2567
	Total	239	1966	2205	TPR=0.4854	FOR=0.0667

The first panel reports the numbers of true positive (TP), false positive (FP), false negatives (FN) and true negatives (TN) of the third type of full logit regressions using $T_w = 22$, $T_m = 22$, $VaR_{0.99}$ and t location-scale losses. The right part of the table shows the metrics derived from this confusion matrix, that is, the false negative rate (FNR) = FN/(FN+TP), false positive rate (FPR) = FP/(FP+TN), true negative rate (TNR) = TN/(TN+FP), true positive rate (TPR) = TP/(TP+FN), accuracy (ACC) = (TP+TN)/(TP+TN+FP+FN), positive predictive value (PPV) = TP/(TP+FP), noise-to-signal ratio (NSR) = FPR/TPR and false omission rate (FOR) = FN/(FN+TN). The second panel repeats the same computations to the reduced model derived with the RV_{av} variable.

Table 13
Confusion matrix.

		Actual				
		Positive	Negative	Total	FNR=0.0014	ACC=0.7526
Predicted	Positive	TP = 1471	FP = 489	1960	FPR=0.9551	PPV=0.7505
	Negative	FN = 2	TN = 23	25	TNR=0.0449	NSR=0.9564
	Total	1473	512	1985	TPR=0.9986	FOR=0.0800

		Actual				
		Positive	Negative	Total	FNR=0.0075	ACC=0.7365
Predicted	Positive	TP = 1462	FP = 512	1974	FPR=1.0000	PPV=0.7406
	Negative	FN = 11	TN = 0	11	TNR=0.0000	NSR=1.0076
	Total	1473	512	1985	TPR=0.9925	FOR=1.0000

The first panel reports the numbers of true positive (TP), false positive (FP), false negatives (FN) and true negatives (TN) of the third type of full logit regressions using $T_w = 132$, $T_m = 132$, $VaR_{0.95}$ and t location-scale losses. The right part of the table shows the metrics derived from this confusion matrix, that is, the false negative rate (FNR) = FN/(FN+TP), false positive rate (FPR) = FP/(FP+TN), true negative rate (TNR) = TN/(TN+FP), true positive rate (TPR) = TP/(TP+FN), accuracy (ACC) = (TP+TN)/(TP+TN+FP+FN), positive predictive value (PPV) = TP/(TP+FP), noise-to-signal ratio (NSR) = FPR/TPR and false omission rate (FOR) = FN/(FN+TN). The second panel repeats the same computations to the reduced model derived with the RV_{av} variable.

Table 14
Logit output results based on the fourth type of regressions.

$T_w = 22$, $T_m = 22$ - $ES_{0.90}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	18.5659	2.3701e−94	19.3098	2.9188e−106	−1.5513	2.2623e−32
RV_{av}	2.0799	2.3552e−108	2.0923	5.6354e−112		
λ_{av}	0.0985	5.9800e−06			0.1360	1.2162e−12
R-squared	0.3199		0.3125		0.0185	
Hit ratio	0.7755		0.7642		0.6694	
Goodness		1.2388e−195		1.1516e−192		5.9834e−13
$T_w = 22$, $T_m = 22$ - $ES_{0.95}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	17.7429	9.8552e−81	18.9403	1.9592e−97	−2.1961	5.1838e−53
RV_{av}	2.0841	1.4966e−98	2.0870	3.5603e−104		
λ_{av}	0.1814	4.2941e−15			0.2024	4.1190e−23
R-squared	0.3362		0.3115		0.0162	
Hit ratio	0.8159		0.7814		0.7134	
Goodness		1.0123e−59		1.3142e−181		2.0493e−24
$T_w = 22$, $T_m = 22$ - $ES_{0.99}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	11.3467	2.8515e−33	12.7697	1.5429e−41	−3.5829	1.5942e−63
RV_{av}	1.6024	7.8047e−54	1.5980	5.8441e−53		
λ_{av}	0.2119	2.3621e−12			0.2212	3.7409e−15
R-squared	0.2655		0.2321		0.0415	
Hit ratio	0.9011		0.9034		0.8916	
Goodness		5.5803e−88		2.2914e−78		2.3756e−15

(*continued on next page*)

Table 14 (*continued*).

$T_w = 66$, $T_m = 66$ - $ES_{0.90}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	21.7576	1.4688e−99	22.4589	1.0206e−111	−1.8326	2.6035e−22
RV_{av}	2.4047	6.0039e−112	2.3944	2.1282e−115		
λ_{av}	0.1291	4.8665e−05			0.1962	1.4554e−11
R-squared	0.3207		0.3147		0.0171	
Hit ratio	0.7350		0.7227		0.6821	
Goodness		5.3817e−192		3.4774e−10		7.1338e−12
$T_w = 66$, $T_m = 66$ - $ES_{0.95}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	14.9751	2.6389e−61	16.3233	7.2768e−80	−2.5756	2.5228e−36
RV_{av}	1.7707	1.2075e−80	1.7871	2.8067e−86		
λ_{av}	0.1908	1.0259e−09			0.2718	1.7664e−18
R-squared	0.2230		0.2081		0.0315	
Hit ratio	0.7289		0.6826		0.7303	
Goodness		2.5460e−125		1.8724e−118		2.2904e−19
$T_w = 66$, $T_m = 66$ - $ES_{0.99}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	5.7183	2.1290e−09	9.0230	3.2894e−26	−4.3593	1.2574e−59
RV_{av}	0.9822	6.6998e−26	1.1021	1.8007e−34		
λ_{av}	0.3337	3.4612e−19			0.4305	3.3667e−29
R-squared	0.1322		0.0874		0.0702	
Hit ratio	0.8281		0.8219		0.8299	
Goodness		8.7265e−57		5.2784e−39		1.2123e−31
$T_w = 132$, $T_m = 132$ - $ES_{0.90}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	19.7071	5.7170e−81	22.4716	3.0111e−117	−3.8114	1.5580e−44
RV_{av}	2.3946	1.4451e−110	2.3469	1.3450e−119		
λ_{av}	0.5248	1.6536e−24			0.5506	1.3942e−37
R-squared	0.3489		0.3018		0.0722	
Hit ratio	0.8156		0.7914		0.6610	
Goodness		5.4557e−204		4.3623e−178		5.0553e−44
$T_w = 132$, $T_m = 132$ - $ES_{0.95}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	14.4298	9.4606e−56	16.6683	1.1588e−84	−3.0903	1.3624e−33
RV_{av}	1.7330	3.6317e−81	1.7690	7.9166e−89		
λ_{av}	0.3035	4.8653e−13			0.4080	1.6400e−24
R-squared	0.2193		0.1980		0.0431	
Hit ratio	0.7798		0.7290		0.7139	
Goodness		3.1420e−125		1.3610e−114		2.4383e−26
$T_w = 132$, $T_m = 132$ - $ES_{0.99}$ - t location-scale case						
		$p-$ value		$p-$ value		$p-$ value
β_0	6.8771	3.8982e−12	10.2528	1.1247e−31	−4.3769	2.9187e−51
RV_{av}	1.0816	6.8287e−30	1.1941	8.6669e−39		
λ_{av}	0.3595	3.3694e−17			0.4821	1.5668e−28
R-squared	0.1303		0.0944		0.0615	
Hit ratio	0.7456		0.7557		0.8035	
Goodness		3.1906e−59		2.3764e−44		1.7874e−29

The table reports the outputs of the fourth type of logit regressions applied to the forward-looking variable defined in Eq. (19) based on the daily returns of the S&P 500 index futures. T_w is assumed to be equal to T_m and is set equal to 22, 66 and 132. VaR and ES are computed using the t location-scale distribution. The sample contains respectively 2205, 2117 and 1985 observations.

Table 15
Performance measures on the fourth type of regressions.

$T_w = 66, T_m = 66$ - t location-scale case	AUC	QPS	LPS
$ES_{0.99}$ - Full model	0.7646**	0.2544	0.4002
$ES_{0.99}$ - Reduced model RV_{av}	0.7066	0.2739	0.4209
$ES_{0.99}$ - Reduced model λ_{av}	0.6291	0.2607	0.4288
$T_w = 132, T_m = 132$ - t location-scale case	AUC	QPS	LPS
$ES_{0.90}$ - Full model	0.8671**	0.2808	0.4400
$ES_{0.90}$ - Reduced model RV_{av}	0.8297	0.3077	0.4718
$ES_{0.90}$ - Reduced model λ_{av}	0.6137	0.4355	0.6270
$ES_{0.95}$ - Full model	0.8160*	0.3374	0.5141
$ES_{0.95}$ - Reduced model RV_{av}	0.7806	0.3558	0.5282
$ES_{0.95}$ - Reduced model λ_{av}	0.5677	0.4366	0.6302
$ES_{0.99}$ - Full model	0.7787***	0.3001	0.4528
$ES_{0.99}$ - Reduced model RV_{av}	0.7117	0.3167	0.4715
$ES_{0.99}$ - Reduced model λ_{av}	0.6078	0.3113	0.4886

The table displays the performance measures AUC, QPS and LPS of the logit regressions obtained using the forward-looking variable defined in Eq. (19) and the daily returns of the S&P 500 index futures. T_w is assumed to be equal to T_m and is set equal to 66 and 132. VaR and ES are computed using the t location-scale distribution. The sample contains respectively 2117 and 1985 observations. *, ** and *** indicate rejection of null hypothesis test $AUC_{\text{Full model}} = AUC_{\text{Reduced model}}$ at 5%, 1% and 0.01% level of significance. The reduced model is derived using the RV_{av} variable.

Table 16
The fractions of ones of the forward-looking variable Y.

First type of regressions - $T_w = 1, T_m = 22$; $T_w = 1, T_m = 66$; $T_w = 1, T_m = 132$				
Distribution	α	Fraction	Fraction	Fraction
t location-scale	0.10	0.7031	0.9386	0.9986
t location-scale	0.05	0.4353	0.6434	0.7580
t location-scale	0.01	0.1074	0.2255	0.3587
Second type of regressions - $T_w = 1, T_m = 22$; $T_w = 1, T_m = 66$; $T_w = 1, T_m = 132$				
Distribution	α	Fraction	Fraction	Fraction
t location-scale	0.10	0.3392	0.3712	0.4074
t location-scale	0.05	0.2947	0.3171	0.3719
t location-scale	0.01	0.1074	0.1948	0.2037
Third type of regressions - $T_w = 22, T_m = 22$; $T_w = 66, T_m = 66$; $T_w = 132, T_m = 132$				
Distribution	α	Fraction	Fraction	Fraction
t location-scale	0.10	0.7002	0.9386	0.9986
t location-scale	0.05	0.4299	0.6400	0.7556
t location-scale	0.01	0.1084	0.2212	0.3523
Fourth type of regressions - $T_w = 22, T_m = 22$; $T_w = 66, T_m = 66$; $T_w = 132, T_m = 132$				
Distribution	α	Fraction	Fraction	Fraction
t location-scale	0.10	0.3329	0.3519	0.4071
t location-scale	0.05	0.2889	0.2962	0.3693
t location-scale	0.01	0.1084	0.1734	0.2151

The table reports the fractions of ones of the forward-looking variable Y for the four types of regressions shown in Section 2.3. From left to right: Each panel shows the distribution of the losses, level of significance and the fractions of ones for $T_w = 22, T_m = 22$; $T_w = 66, T_m = 66$ and $T_w = 132, T_m = 132$.

Table 17
10-fold performance measures on the third type of regressions.

$T_w = 22, T_m = 22$ - t location-scale case	AUC	QPS	LPS	Hit ratio
$VaR_{0.95}$ - Full model	0.8457	0.3126	0.4780	0.7569
$VaR_{0.95}$ - Reduced model RV_{av}	0.8412	0.3190	0.4857	0.7696
$VaR_{0.95}$ - Reduced model λ_{av}	0.6055	0.4757	0.6682	0.5941
$VaR_{0.99}$ - Full model	0.8569	0.1530	0.2535	0.8998
$VaR_{0.99}$ - Reduced model RV_{av}	0.8246	0.1566	0.2643	0.9038
$VaR_{0.99}$ - Reduced model λ_{av}	0.6389	0.1877	0.3297	0.8916
$T_w = 66, T_m = 66$ - t location-scale case	AUC	QPS	LPS	Hit ratio
$VaR_{0.95}$ - Full model	0.8093	0.3262	0.5058	0.8044
$VaR_{0.95}$ - Reduced model RV_{av}	0.8069	0.3237	0.5075	0.8063
$VaR_{0.95}$ - Reduced model λ_{av}	0.5488	0.4602	0.6513	0.6278
$VaR_{0.99}$ - Full model	0.7057	0.3076	0.4656	0.7926
$VaR_{0.99}$ - Reduced model RV_{av}	0.6924	0.3124	0.4686	0.7846
$VaR_{0.99}$ - Reduced model λ_{av}	0.5507	0.3189	0.5013	0.7950
$T_w = 132, T_m = 132$ - t location-scale case	AUC	QPS	LPS	Hit ratio
$VaR_{0.95}$ - Full model	0.8987	0.2207	0.3345	0.8463
$VaR_{0.95}$ - Reduced model RV_{av}	0.8547	0.2641	0.3804	0.8046
$VaR_{0.95}$ - Reduced model λ_{av}	0.6250	0.3830	0.5707	0.7421
$VaR_{0.99}$ - Full model	0.7760	0.3603	0.5344	0.7154
$VaR_{0.99}$ - Reduced model RV_{av}	0.7343	0.3776	0.5468	0.6635
$VaR_{0.99}$ - Reduced model λ_{av}	0.5565	0.4065	0.5993	0.7420

The table displays the 10-fold AUC, QPS, LPS and hit ratio estimates of the logit regressions obtained using the forward-looking variable defined in Eq. (18) and the daily returns of the S&P 500 index futures. T_w is assumed to be equal to T_m and is set equal to 22, 66 and 132. VaR is computed using the t location-scale distribution. The sample contains respectively 2205, 2117 and 1985 observations.

References

Abiad, A. G. (2003). *Early warning systems: A survey and a regime-switching approach*: *IMF working paper 03/32*.

Acerbi, C., & Szekey, B. (2014). *Backtesting expected shortfall-introducing three model-independent, non-parametric backtest methodologies for expected shortfall*: *Working paper MSCI Inc*.

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, *26*(7), 1487–1503.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2000). Great realizations. *Risk*, *13*, 105–108.

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

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, *73*(2), 251–277.

Antunes, A., Bonfim, D., Monteiro, N., & Rodrigues, P. M. (2018). Forecasting banking crises with dynamic panel probit models. *International Journal of Forecasting*, *34*(2), 249–275.

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203–228.

Babecký, J., Havránek, T., Matějů, J., Rusnák, M., Šmídková, K., & Vašíček, B. (2014). Banking, debt, and currency crises in developed countries: Stylized facts and early warning indicators. *Journal of Financial Stability*, *15*, 1–17.

Bandi, F. M., & Russell, J. R. (2011). Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. *Journal of Econometrics*, *160*(1), 145–159.

Barndorff-Nielsen, O. E., & Shephard, N. (2005). How accurate is the asymptotic approximation to the distribution of realized variance. Identification and inference for econometric models. *A Festschrift in Honour of TJ Rothenberg*, 306–311.

Barrell, R., Davis, E. P., Karim, D., & Liadze, I. (2010). Bank regulation, property prices and early warning systems for banking crises in OECD countries. *Journal of Banking & Finance*, *34*(9), 2255–2264.

Barucci, E., Malliavin, P., Mancino, M. E., Renò, R., & Thalmaier, A. (2003). The price-volatility feedback rate: an implementable mathematical indicator of market stability. *Mathematical Finance*, *13*(1), 17–35.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics*. New York: John Wiley & Sons.

Berg, A., & Pattillo, C. (1999). Predicting currency crises:: The indicators approach and an alternative. *Journal of International Money and Finance*, *18*(4), 561–586.

Bertsimas, D., Lauprete, G. J., & Samarov, A. (2004). Shortfall as a risk measure: properties, optimization and applications. *Journal of Economic Dynamics & Control*, *28*(7), 1353–1381.

Beutel, J., List, S., & Schweinitz, G. Von. (2018). *An evaluation of early warning models for systemic banking crises: Does machine learning improve predictions?*: *IWH discussion papers, No. 2/2019*.

Billio, M., Casarin, R., Costola, M., & Pasqualini, A. (2016). An entropy-based early warning indicator for systemic risk. *Journal of International Financial Markets, Institutions and Money*, *45*, 42–59.

Borio, C. E., & Drehmann, M. (2009). Assessing the risk of banking crises–revisited. *BIS Quarterly Review*.

Breden, D. (2008). Monitoring the operational risk environment effectively. *Journal of Risk Management in Financial Institutions*, *1*, 156–164.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*(1), 1–3.

Bussiere, M., & Fratzscher, M. (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance*, *25*(6), 953–973.

Caggiano, G., Calice, P., & Leonida, L. (2014). Early warning systems and systemic banking crises in low income countries: A multinomial logit approach. *Journal of Banking & Finance*, *47*, 258–269.

Caggiano, G., Calice, P., Leonida, L., & Kapetanios, G. (2016). Comparing logit-based early warning systems: Does the duration of systemic banking crises matter? *Journal of Empirical Finance*, *37*, 104–116.

Candelon, B., Dumitrescu, E. I., & Hurlin, C. (2014). Currency crisis early warning systems: Why they should be dynamic. *International Journal of Forecasting*, *30*(4), 1016–1029.

Christensen, K., Oomen, R. C. A., & Podolskij, M. (2014). Fact or friction: jumps at ultra high-frequency. *Journal of Financial Economics*, *114*, 576–599.

Conger, A. J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational and Psychological Measurement*, *34*(1), 35–46.

Curato, I. V. (2019). Estimation of the stochastic leverage effect using the Fourier transform method. *Stochastic Processes and their Applications*, *129*(9), 3207–3238.

Curato, I. V., & Sanfelici, S. (2015). Measuring the leverage effect in a high frequency trading framework. *Handbook of High Frequency Trading*, *42*, 5–446.

Dawood, M., Horsewood, N., & Strobel, F. (2017). Predicting sovereign debt crises: an early warning system approach. *Journal of Financial Stability*, *28*, 16–28.

Diebold, F. X., & Rudebusch, G. D. (1989). Scoring the leading indicators. *Journal of Business*, *62*, 369–391.

Douady, R., & Kornprobst, A. (2018). An empirical approach to financial crisis indicators based on random matrices. *International Journal of Theoretical and Applied Finance*, *21*(03), Article 1850022.

Drehmann, M., & Juselius, M. (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting*, *30*(3), 759–780.

Duchin, R., Ozbas, O., & Sensoy, B. A. (2010). Costly external finance, corporate investment, and the subprime mortgage credit crisis. *Journal of Financial Economics*, *97*(3), 418–435.

Fithian, W., & Hastie, T. (2014). Local case-control sampling: Efficient subsampling in imbalanced data sets. *The Annals of Statistics*, *42*(5), 1693.

Fuertes, A. M., & Kalotychou, E. (2007). Optimal design of early warning systems for sovereign debt crises. *International Journal of Forecasting*, *23*(1), 85–100.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Ghosh, S. R., & Ghosh, A. R. (2003). Structural vulnerabilities and currency crises. *IMF Staff Papers*, *50*(3), 481–506.

Gonçalves, S., & Meddahi, N. (2011). Box–Cox transforms for realized volatility. *Journal of Econometrics*, *160*(1), 129–144.

Gonzalez, M. Rodriguez., Basse, T., Saft, D., & Kunze, F. (2021). Leading indicators for US house prices: New evidence and implications for EU financial risk managers. *European Financial Management*.

Good, I. (1952). Rational decisions. *Journal of the Royal Statistical Society. Series B.*, *14*(1), 107–114.

Hsieh, F., & Turnbull, B. W. (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *The Annals of Statistics*, *24*(1), 25–40.

Inkaya, A., & Okur, Y. Y. (2014). Analysis of volatility feedback and leverage effects on the ISE30 index using high frequency data. *Journal of Computational and Applied Mathematics*, *259*, 377–384.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning (Vol. 112)* (p. 18). New York: Springer.

Jorion, P. (2009). Risk management lessons from the credit crisis. *European Financial Management*, *15*(5), 923–933.

Kaminsky, G., Lizondo, S., & Reinhart, C. M. (1998). The leading indicators of currency crises. *IMF Staff Papers*, *45*(1), 1–48.

Kauppi, H., & Saikkonen, P. (2008). Predicting US recessions with dynamic binary response models. *The Review of Economics and Statistics*, *90*(4), 777–791.

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, *9*(2), 137–163.

Kumar, M., Moorthy, U., & Perraudin, W. (2003). Predicting emerging market currency crashes. *Journal of Empirical Finance*, *10*(4), 427–454.

Kunze, F., Basse, T., Gonzalez, M. R., & Vornholz, G. (2020). Forward-looking financial risk management and the housing market in the United Kingdom: is there a role for sentiment indicators? *The Journal of Risk Finance*.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, *3*(2).

Li, W. X., Chen, C. C. S., & French, J. J. (2015). Toward an early warning system of financial crises: What can index futures and options tell us? *The Quarterly Review of Economics and Finance*, *55*, 87–99.

Malliavin, P., & Mancino, M. E. (2002). Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, *6*(1), 49–61.

Malliavin, P., & Mancino, M. E. (2002). Instantaneous liquidity rate, its econometric measurement by volatility feedback. *Comptes Rendus Mathematique*, *334*(6), 505–508.

Malliavin, P., & Mancino, M. E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. *The Annals of Statistics*, *37*(4), 1983–2010.

Malliavin, P., & Thalmaier, A. (2006). *Stochastic calculus of variations in mathematical finance*. Springer Science & Business Media.

Manasse, P., Roubini, N., & Schimmelpfennig, A. (2003). *Predicting sovereign debt crises*: *IMF working paper 03/221*, Washington: International Monetary Fund.

Mancino, M. E., Recchioni, M. C., & Sanfelici, S. (2017). *Fourier-malliavin volatility estimation: Theory and practice*. Springer International Publishing.

Mancino, M. E., & Sanfelici, S. (2008). Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. *Computational Statistics & Data Analysis*, *52*(6), 2966–2989.

Mancino, M. E., & Sanfelici, S. (2020). Identifying financial instability conditions using high frequency data. *Journal of Economic Interaction and Coordination*, *15*(1), 221–242.

McCarthy, J., & Orlov, A. G. (2012). Time-frequency analysis of crude oil and S & P500 futures contracts. *Quantitative Finance*, *12*(12), 1893–1908.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*. New York: Academic Press.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools-revised edition*. Princeton University Press.

McNemar, Q. (1945). The mode of operation of suppressant variables. *The American Journal of Psychology*.

Rochette, M. (2009). From risk management to ERM. *Journal of Risk Management in Financial Institutions*, *2*(4), 394–408.

Schnatz, B. (1998). *Macroeconomic determinants of currency turbulences in emerging markets (No. 1998, 03e)*: *Discussion paper*.

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach*. Nelson Education.

Zhou, X. H., McClish, D. K., & Obuchowski, N. A. (2009). *Statistical methods in diagnostic medicine (Vol. 569)*. John Wiley & Sons.