Article

Information Entropy and Measures of Market Risk

Daniel Traian Pele [1],*, Emese Lazar [2] and Alfonso Dufour [2]

[1] Department of Statistics and Econometrics, Bucharest University of Economic Studies, Piata Romana 6, Bucharest 010371, Romania
[2] ICMA Centre, Henley Business School, University of Reading, Whiteknights, Reading RG6 6BA, UK; e.lazar@icmacentre.ac.uk (E.L.); a.dufour@icmacentre.ac.uk (A.D.)
* Correspondence: danpele@ase.ro; Tel.: +40-21-319-19-01 (ext. 383)

Abstract: In this paper we investigate the relationship between the information entropy of the distribution of intraday returns and intraday and daily measures of market risk. Using data on the EUR/JPY exchange rate, we find a negative relationship between entropy and intraday Value-at-Risk, and also between entropy and intraday Expected Shortfall. This relationship is then used to forecast daily Value-at-Risk, using the entropy of the distribution of intraday returns as a predictor.

Keywords: information entropy; entropy of a distribution function; value at risk; expected shortfall

1. Introduction

Entropy, as a measure of uncertainty of a system, is widely used in many applications, from physics to social sciences. As stated by the second law of thermodynamics, "this entropy cannot decrease in any process in which the system remains adiabatically isolated, i.e., shielded from heat exchange with its environment" [1].

From this point of view, the stock market could be regarded as a non-isolated system, subject to a constant information exchange process with the real economy. Using the terminology from the information theory (Avery, [2]), the information entropy of the system cannot increase, other way than by exchanging information with the exterior environment. The impact of the incoming information on the stock market entropy can be illustrated in the case of a collective behaviour determined by some extreme bad news: most traders will tend to sell thus reducing the overall market entropy.

There is a lot of theoretical and empirical work dealing with the relationship between entropy and financial markets. The entropy has been used as a measure of stock market efficiency in Zunino et al. [3], since high values of entropy are related to randomness in the evolution of stock prices, according to the Efficient Market Hypothesis. A variant of entropy called normalized entropy—being a dimensionless measure—can be used to assess the relative efficiency of a stock market. Risso [4], on the other hand, relates entropy to stock market crashes, his main result being that for some markets the probability of having a crash increases as the market efficiency, measured by entropy, decreases. An application in the foreign exchange markets is that of Oh et al. [5], who use the approximate entropy as a measure of the relative efficiency of the FX markets. Their results suggest that market efficiency measured by approximate entropy is correlated with the liquidity level of foreign exchange markets.

Considering the stock market a complex open system far from equilibrium, Wang et al. [6] analyse the interactions among agents based on generalized entropy. Using nonlinear evolutionary dynamic equations for the stock markets, derived from the Maximum Generalized Entropy Principle, the structural evolution of the stock market system is demonstrated.

A related use of entropy is to study the predictability of stock market returns, as in Maasoumi and Racine [7] who employ an entropy measure for the dependence between stocks' returns. They find that

the entropy is capable of detecting nonlinear dependence between the returns series. Billio et al. [8] use entropy to construct an early warning indicator for the systemic risk of the banking system. They estimate the entropy of marginal systemic risk measures like Marginal Expected Shortfall, Delta CoVaR and network connectedness. By using various definitions of entropy (Shannon, Tsallis and Rényi), they prove that entropy measures have the ability to forecast and predict banking crises.

Dionisio et al. [9] provide a comparison between the theoretical and empirical properties of the entropy and the variance, as measures of uncertainty (although volatility can be considered a measure of risk in finance, it is a measure of uncertainty in statistical terms. Here, measures that are symmetric by nature we call measures of uncertainty, and tail measures that consider certain negative outcomes we call measures of risk.). They conclude that the entropy is a more general measure of uncertainty than the variance or standard deviation, as originally suggested by Philippatos and Wilson [10] and Ebrahimi et al. [11]. An explanation is that the entropy may be related to higher-order moments of a distribution, unlike the variance, so it could be a better measure of uncertainty. Furthermore, argue Dionisio et al. [9], both measures, the entropy and the variance, reflect concentration but use different metrics; while the variance measures the concentration around the mean, the entropy measures the dispersion of the density irrespective of the location of the concentration (see also [12,13]). Moreover, as we will show in the paper, the entropy of a distribution function is strongly related to its tails and this feature is more important for distributions with heavy tails or with an infinite second-order moment (like the non-Gaussian alpha-stable distribution) for which an estimator of variance is obsolete.

Entropy-based measures have been compared to the classical coefficient of correlation as well. A measure called cross-sample entropy has been used by Liu et al. [14] to assess the degree of asynchrony between foreign exchange rates, concluding that their measure is superior to the classical correlation measure as a descriptor of the relationship between time series.

Entropy can be applied in the area of risk management as described in Bowden [15]. The authors propose a new concept called directional entropy and use it to improve the performance of classical measures like value-at-risk (VaR) in capturing regime changes. An interesting application is using measures based on the Tsallis entropy as a warning indicator of financial crises, as in Gradojevic and Gencay [16], Gencay and Gradojevic [17] and Gradojevic and Caric [18]. A further application of entropy is for option pricing, as in Stutzer [19] and Stutzer and Kitamura [20]. Also, entropy-based risk measures have been used in a decision-making model context in Yang and Qiu [21].

Besides [5,14] above, other innovative approaches involving entropy and FX markets can be found in Ishizaki and Inoue [22], showing how entropy can be a signal of turning points for exchange rates regimes. Furthermore, Bekiros and Marcellino [23] and Bekiros [24] use entropy in wavelet analysis, revealing the complex dynamics across different timescales in the FX markets.

The main objective of this paper is (1) to study the link between the entropy of the distribution function of intraday returns, and intraday and daily measures of market risk, namely VaR and Expected Shortfall (ES); and then (2) to demonstrate their VaR-forecasting ability. The entropy is considered to have more informational content than the standard measures of risk and it is also more reactive to new information. This paper uses the concept of entropy of a function (Lorentz [25]) in order to estimate the entropy of a distribution function, using a non-parametric approach, with an application to the FX market. The main advantage of this approach is that the entropy can be estimated for any distribution, without any prior knowledge about its functional form, which is especially important for distributions with no closed form for the probability distribution function.

The rest of this paper is organized as follows: in Section 2 we provide the theoretical background defining the entropy of a distribution function and measures of market risk and uncertainty. Section 3 presents the results of the empirical analysis whilst Section 4 concludes.

2. The Entropy of a Distribution Function and Measures of Market Risk and Uncertainty

2.1. The Entropy and Intraday Measures of Market Risk and Uncertainty

The entropy, as a measure of uncertainty, can be defined using different metrics (Shannon Entropy, Tsallis Entropy, Rényi Entropy etc.), based on the informational content of a discrete or continuous random variable (see Zhou et al. [26] for a comprehensive review on entropy measures used in finance). The most common entropy metric, the Shannon Information Entropy, quantifies the expected value of information contained in a discrete distribution ([27]):

Definition 1 (Shannon information entropy). *If X is a discrete random variable, with probability distribution $X : \begin{pmatrix} x_1 \ldots \ldots x_n \\ p_1 \ldots \ldots p_n \end{pmatrix}$, where $p_i = P(X = x_i)$, $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$, then the Shannon Information Entropy is defined as follows:*

$$H(X) = -\sum_i p_i \log_2 p_i. \tag{1}$$

It will reach its maximum value of $H(X) = log_2 n$ for the uniform distribution, while the minimum of 0 is attained for a distribution where one of the probabilities p_i is 1 and the rest are 0. In other words, high (low) levels of entropy are obtained for probability distributions with high (low) levels of uncertainty. If X is a continuous random variable with probability density function $f(x)$, then we can define the differential entropy as:

$$H(f) = -\int_A f(x) \log_2 f(x) dx, A = supp(X). \tag{2}$$

Unlike the Shannon entropy, the differential entropy does not possess certain desirable properties: invariance to linear transformations and non-negativity ([25,28]). However, the analogue Shannon entropy of a function can be defined through a transformation called quantization. We present this transformation, as in Lorentz [25].

Definition 2 (sampled function). *Let $f : I = [a,b] \rightarrow \mathbf{R}$ be a real valued continuous function, let $n \in \mathbf{N}^*$ be fixed and let $x_i = a + (i + 1/2)h$, for $i = 0,..,n-1$, where $h = (b-a)/n$. Then the sampled function for f is:*

$$S_n(f)(i) = f(x_i), \text{ for } i = 0,..,n-1. \tag{3}$$

If $f : I = [a,b] \rightarrow \mathbf{R}$ is essentially bounded, then the sampled function is

$$S_n(f)(i) = h^{-1} \int_{x_i - h/2}^{x_i + h/2} f(x) dx, \text{ for } i = 0,..,n-1. \tag{4}$$

The sampling defined in (3) is called point sampling, whilst the one in (4) is called mean sampling.

Definition 3 (quantization). *The quantization process of a function refers to creating a simple function that approximates the original function. Let $q > 0$ be a quantum. Then the following function defines a quantization of f:*

$$Q_q(f)(x) = (i + 1/2)q, \text{ if } f(x) \in [iq, (i+1)q). \tag{5}$$

Definition 4 (entropy of a function at a quantization level q). *Let f be a measurable and essentially bounded real valued function defined on [a,b] and let q > 0. Also let $I_i = [iq, (i+1)q)$ and $B_i = f^{-1}(I_i)$. Then the entropy of f at quantization level q is*

$$H_q(f) = -\sum_i \mu(B_i) \log_2(\mu(B_i)), \tag{6}$$

where μ is the Lebesgue measure.

In light of this definition, we can calculate the entropy of any continuous function on a compact interval. The following theorem provides a conceptual framework for defining an estimator of the entropy of a continuous function.

Theorem (Lorentz, [25]). *"Let f be continuous for point sampling, measurable and essentially bounded for mean sampling. The sampling spacing is 1/n. Let $S_n(f)$ be the corresponding sampling of (3) and respectively (4). Fix q > 0 and let $Q_q S_n$ be the quantization of the samples with resolution q as given in (5). Denote the number of occurrences of the value $(i + 1/2)q$ in $Q_q S_n$ by $c_n(i) = card\{(i+1/2)q \in Q_q S_n\}$ and denote the relative probability of the occurrence of the value i by $p_n(i) = \frac{c_n(i)}{\sum_j c_n(j)} = \frac{c_n(i)}{n}$. Then we have the following result:*

$$\lim_{n\to\infty} -\sum_i p_n(i) \log_2 p_n(i) = H_q(f)." \tag{7}$$

The above theorem assures us that regardless of the sampling and quantization, we obtain a onsistent estimator of the entropy of a function. As such, we can define the entropy of a distribution function for a continuous random variable on a compact interval. In general, continuous random variables do not have a compact support, with finite Lebesgue measure. In order to meet the assumptions of Lorentz's theorem, we can define the entropy of the distribution function on a compact interval. In what follows we assume that we are dealing with a continuous random variable X, whose support is the interval [0,1]. Then its distribution function F: [0,1] → [0,1] is continuous and the conditions of Lorentz's theorem are fulfilled, so we can define the entropy of the density function $H_q(F)$ at the quantization level $q > 0$.

2.1.1. Entropy of a Distribution Function

Let $I_i = [iq, (i+1)q)$ and $B_i = F^{-1}(I_i)$. Then, according to Lorentz's theorem, the entropy of F with the quantum q is $H_q(F) = -\sum_i \mu(B_i) \log_2(\mu(B_i))$, where μ is the Lebesgue measure. In fact, $\cup_i I_i = [0,1]$, I_i are disjoint, and $\cup_i B_i = [0,1]$, with B_i not necessarily disjoint.

Note: In general, for any distribution function, defined on the set A, not necessarily of finite measure, we can consider the restriction of this function on some compact interval: $F\big|_{[a,b]} : [a,b] \to [0,1]$, $F\big|_{[a,b]}(x) = F(x)$. Then $F\big|_{[a,b]} : [a,b] \to [0,1]$ satisfies the conditions of Lorentz's theorem so we can define the entropy measure.

The framework described above can be applied to estimate the entropy of a distribution function of a continuous random variable X. The cumulative distribution function (*CDF*) is $F(x) = P(X < x)$. The distribution function is defined on the support set of X, with values on [0,1] and has the following properties:

(i) F is right continuous;
(ii) F is monotonically increasing;
(iii) $\lim_{x\to-\infty} F(x) = 0$;
(iv) $\lim_{x\to\infty} F(x) = 1$.

If in addition *F* is absolutely continuous, then there is a Lebesgue integrable function *f(x)* such as

$$F(b) - F(a) = P(a < X < b) = \int_a^b f(x)dx, \text{ with } f(x) \text{ the density function of } X. \text{ Also, } F(x) = \int_{-\infty}^x f(t)dt.$$

In practice, there are cases when the analytical form of the distribution function is unknown. When it is not known, a robust approach can be taken based on a nonparametric method.

2.1.2. Empirical Distribution Function

The cumulative density function can be estimated in a simple way by using the histogram estimator of a probability density function [29]. The basic algorithm to obtain the estimator of the *CDF* from a sample $X_1 < ... < X_n$ is as follows:

Step 1.　Let x_0 be a fixed point and let $h > 0$ be the bin width;

Step 2.　Define the bins as $I_h^m = [x_0 + mh, x_0 + (m+1)h)$, $m \in \mathbf{Z}$, obtaining a partition of the real line;

Step 3.　For $\forall x \in \mathbf{R}$, $\exists m \in \mathbf{Z}$ such as $x \in I_h^m$, let $A_x = \{X_i | X_i \in I_h^m\}$;

Step 4.　The histogram estimator of the pdf is defined as $\hat{f}(x) = \frac{1}{nh} card\ A_x$, $\forall x \in \mathbf{R}$;

Step 5.　The empirical estimator of distribution function (*CDF*) is:

$$\hat{F}_n(x) = \begin{cases} 0,\ x < X_1 \\ \sum_{X_j < x} \hat{f}(X_j)h,\ x \in [X_{i-1}, X_i). \\ 1,\ x \geq X_n \end{cases}$$

2.1.3. Kernel Density Estimator

To estimate the distribution function, we can use the Kernel Density Estimation (*KDE*) methods. If $X_1, ..., X_n$ is a sample of i.i.d. observations, then an estimator of the distribution function is:

$$\hat{F}_n(x) = \int_{-\infty}^x \hat{f}(u)du = \int_{-\infty}^x \frac{1}{nh} \sum_{i=1}^n K\left(\frac{u - X_i}{h}\right)du = \frac{1}{nh} \sum_{i=1}^n \int_{-\infty}^x K\left(\frac{u - X_i}{h}\right)du,$$

where *K* is a real function with the following properties:$K(x) \geq 0$, $\forall x \in \mathbf{R}$, $K(x) = K(-x)$, $\forall x \in \mathbf{R}$, $\int_{\mathbf{R}} K(x)dx = 1$ and $\int_{\mathbf{R}} xK(x)dx = 0$.

Such a function is called the kernel and is usually chosen from the known probability density functions. The parameter *h* is the scale parameter (also called the smoothing parameter or bandwidth), the choice of which determines the estimate. The asymptotic properties of the kernel estimator above have been studied in numerous papers [30,31], establishing the uniform convergence and convergence in probability to the theoretical distribution function, regardless of the form of the kernel used.

As a special case, the uniform distribution is considered. Given *X* a uniformly distributed random variable on interval [0,1], with distribution function:

$$F(x) = \begin{cases} 0,\ x < 0 \\ x,\ x \in [0, 1) \\ 1,\ x \geq 1 \end{cases}.$$

Then the entropy of the function $F\big|_{[0,1]}(x) = \begin{cases} x,\ x \in [0, 1) \\ 1,\ x = 1 \end{cases}$ is $H_q(F\big|_{[0,1]}) = \log_2 n$, which is the maximum value of the entropy. Next, we present the estimation methodology of the entropy of a distribution function.

2.1.4. Estimation of the Entropy of a Distribution Function

Let $X_0, \ldots, X_{n-1}$ be a sample of i.i.d. observations drawn from the distribution F. In order to ensure the comparability of results between various estimates, we assume that the observed values are normalized in the interval [0,1], through a transformation of the type $X_i \rightarrow \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$.

The following steps present the estimation of the entropy of a distribution function (a similar approach, but in a different context, can be found in [27,28,32]):

Step 1. Estimate the distribution function, obtaining values $\hat{F}_n(X_i)$ for $i = 0, .., n - 1$;

Step 2. Sample from the distribution function, using the sampled function $S_n(\hat{F}_n)(i) = \hat{F}_n(X_i)$ for $i = 0, .., n - 1$;

Step 3. Define a quantum $q > 0$; then $Q_q S_n(\hat{F}_n)(j) = (i + 1/2)q$, if $\hat{F}_n(X_j) \in [iq, (i + 1)q)$;

Step 4. Compute the probabilities $p_n(i) = \frac{c_n(i)}{\sum_j c_n(j)} = \frac{c_n(i)}{n} = \frac{card\{\hat{F}_n(X_j) \in [iq,(i+1)q)\}}{n}$;

Step 5. Estimate the entropy of the distribution function: $H_q(\hat{F}_n) = -\sum_i p_n(i) \log_2 p_n(i)$.

As previously shown, the entropy of the distribution function reaches its maximum value for the uniform distribution. One can define a dimensionless measure of uncertainty, the normalized entropy, defined as the ratio between the entropy of the distribution function and the entropy of the uniform distribution:

$$NH_q(\hat{F}_n) = \frac{-\sum_i p_n(i) \log_2 p_n(i)}{\log_2 n} \in [0, 1]. \tag{8}$$

In the following sections, we will refer to the entropy of the distribution function as the normalized entropy of the distribution function: $H(F) \equiv NH_q(F) \in [0, 1]$.

2.1.5. Properties and Asymptotic Behaviour of the Entropy of a Distribution Function

To illustrate the properties of the entropy estimator, we performed a Monte Carlo experiment, estimating the entropy for simulated distributions, using a sample of 400 observations and replicating the experiment 1000 times. We have simulated several α-stable distributions, allowing for higher probabilities in the tails. Stable distributions have some important properties: they allow for heavy tails and more, any linear combination of independent stable variable follows a stable distribution, up to a scale and location parameter [33]; the Gaussian distribution is a particular case of a stable distribution.

In the literature there are several parameterizations of α-stable distributions. In this paper we use the *S1* parameterization [33]: a random variable X follows a α-stable distribution $S(\alpha, \beta, \gamma, \delta; 1)$ if its characteristic function is:

$$\phi(t) = \mathbf{E}[e^{itX}] = \begin{cases} \exp(-\gamma^\alpha |t|^\alpha [1 - i\beta \tan(\frac{\pi\alpha}{2})sign(t)] + i\delta t), \alpha \neq 1 \\ \exp(-\gamma |t| [1 + i\beta t \frac{2}{\pi} sign(t)(\ln(|t|)] + i\delta t), \alpha = 1 \end{cases}.$$

In the above notation $\alpha \in (0.2]$ is the characteristic parameter (for a normal distribution $\alpha = 2$), $\beta \in [-1, 1]$ is the skewness parameter, $\gamma \in (0, \infty)$ is the scale parameter and $\delta \in \mathbf{R}$ is the location parameter. To simulate a α-stable distribution $S(\alpha, \beta, \gamma, \delta; 1)$ we have used the algorithm from [34] (see Appendix A). The results of the simulations are presented in Table 1. The entropy reaches its maximum value for the uniform distribution and as the α parameter decreases, the entropy of the stable distribution decreases, too. As expected, low entropy values are associated with heavy-tailed distributions; as the tail probability increases, the expected value of the entropy goes down. Figure 1 presents the relationship between the α parameter of a stable distribution, and entropy. Before turning our attention to the link between entropy and measures of risk and uncertainty, we consider the optimal sampling frequency to efficiently compute returns (ignoring noise).

Table 1. The entropy of simulated distributions.

Distribution	Average Value of the Entropy of the Distribution Function	Standard Deviation of the Entropy of the Distribution Function
Uniform (0,1)	0.9982	0.0011
Normal (0,1)	0.8933	0.0319
Stable ($\alpha = 1.9$)	0.6725	0.1600
Stable ($\alpha = 1.5$)	0.4788	0.1186
Stable ($\alpha = 1$)	0.3979	0.1125
Stable ($\alpha = 0.5$)	0.2858	0.0984
Stable ($\alpha = 0.1$)	0.1872	0.0610

Note: This table presents the average value of the entropy and its standard deviation, estimated by simulating a sample of 400 cases and repeating the experiment 1000 times.
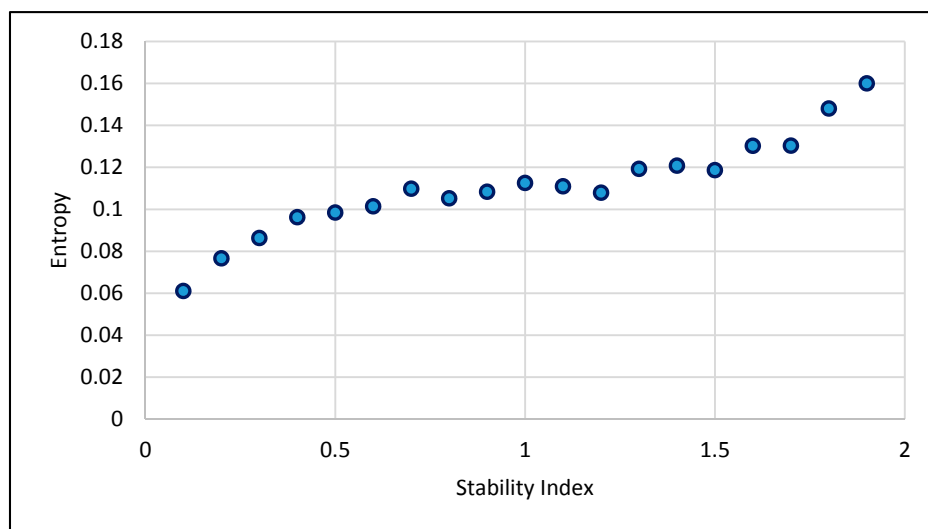
Figure 1. The entropy of distribution functions of simulated alpha-stable distributions, as a function of α.

2.1.6. Optimal Sampling Frequency

When dealing with intraday data, one problem is to separate the fundamental dynamics from market noise. Assuming that the trading price can be decomposed into an efficient component and a noise component, reflecting market microstructure frictions, one way to distinguish between the informational content of these components is to choose an optimal sampling frequency for the intraday data. Following Bandi and Russell [35], we assume that the observed logprice is given by:

$$\widetilde{p}_i = p_i + \eta_i, i = 1...n,$$

where n is the number of trading days, p_i is the efficient log-price and η_i is the microstructure noise. Now we divide the trading day into M subperiods and define the observed intraday logreturns as:

$$\widetilde{r}_{ij} = \widetilde{p}_{i-1+j\delta} - \widetilde{p}_{i-1+(j-1)\delta}, j = 1...M,$$

where $\delta = 1/M$ is the sampling frequency of the intraday returns, used to estimate the daily entropy. Then the intraday returns can be decomposed into an unobserved efficient return and a market microstructure disturbance as:

$$\widetilde{r}_{ij} = r_{ij} + \varepsilon_{ij}, \text{ with } r_{ij} = p_{i-1+j\delta} - p_{i-1+(j-1)\delta} \; \varepsilon_{ij} = \eta_{i-1+j\delta} - \eta_{i-1+(j-1)\delta}.$$

In terms of the probability density function, if the unobserved efficient return and the market microstructure disturbance are independent, then the probability density function of the observed returns is the convolution of the probability density functions of unobserved returns and microstructure noise:

$$f_{\tilde{r}_{ij}}(x) = f_{r_{ij}+\varepsilon_{ij}}(x) = \int f_{r_{ij}}(y)f_{\varepsilon_{ij}}(x-y)dy = (f_{r_{ij}} * f_{\varepsilon_{ij}})(x).$$

Assuming that the entropy of the distribution function is estimated using intraday data with a quantum $q = 0.05$, while the distribution function is estimated using the empirical distribution function, the following measures of risk and uncertainty can be defined, for an $\alpha \in (0,1)$:

(1)　Intraday VaR at significance level α computed from observations at frequency ν, being the α-quantile of the distribution of intraday returns, so that the following is satisfied:

$$P(\tilde{r}_{ij,\nu} < -IVaR_{\alpha,\nu}) = \alpha;$$

(2)　Intraday ES at significance level α computed from observations at frequency ν, defined as:

$$IES_{\alpha,\nu} = \frac{1}{\alpha}\int_0^\alpha IVaR_{\gamma,\nu}d\gamma;$$

(3)　Intraday Realized Volatility computed from intraday returns at frequency ν, computed as:

$$IRV_\nu = \sqrt{\sum_{i=1}^n \tilde{r}_{ij,\nu}^2}.$$

If $[1, T]$ is the time-horizon of daily data, then we can compute daily estimates of the following 4 measures: entropy $H_{\delta;t}$, intraday Value at Risk $IVaR_{\alpha,\nu;t}$, intraday Expected Shortfall $IES_{\alpha,\nu;t}$ and Intraday Realized Volatility $IRV_{\nu;t}$.

In order to asses the relationship between entropy of the distribution of the intraday returns and intraday measures of market risk, we estimate static and dynamic linear regression models using entropy as (one of) the explanatory variable(s):

2.1.7. Static Models

The first class of models study the explanatory power of the (daily) entropy to explain different measures of market risk and uncertainty: daily observations of Intraday ES, Intraday VaR and Intraday Realized Volatility estimated at different time scales ν, running the following regressions:

$$\begin{cases} IES_{\alpha,\nu;t} = \beta_0 + \beta_1 H_{\delta;t} + \varepsilon_t \\ IVaR_{\alpha,\nu;t} = \alpha_0 + \alpha_1 H_{\delta;t} + \upsilon_t \\ IRV_{\nu;t} = \gamma_0 + \gamma_1 H_{\delta;t} + \zeta_t \end{cases}. \tag{9}$$

2.1.8. Dynamic Models

The second class of models aims to check whether the entropy of the distribution of intraday returns can provide additional information to that contained in the latest observation of the risk measures by estimating the following regressions:

$$\begin{cases} IES_{\alpha,\nu;t} = \beta_0 + \beta_1 H_{\delta;t-1} + \beta_2 IES_{\alpha,\nu;t-1} + \varepsilon_t \\ IVaR_{\alpha,\nu;t} = \alpha_0 + \alpha_1 H_{\delta;t-1} + \alpha_2 IVaR_{\alpha,\nu;t-1} + \upsilon_t \\ IRV_{\nu;t} = \gamma_0 + \gamma_1 H_{\delta;t-1} + \gamma_2 IRV_{\nu;t-1} + \zeta_t \end{cases}. \tag{10}$$

2.2. *Quantile Regressions*

Classical linear regression is used to estimate the conditional mean of a dependent variable, given the values of an explanatory variable. However, the presence of outliers and/or heteroskedasticity can affect the results. Also, in many situations not just the conditional mean of a variable is required, but its entire conditional distribution, in particular the conditional quantiles. For a random variable Y with distribution function F, the τth quantile is defined as the inverse of distribution function $Q(\tau) = \inf\{y, F(y) \geq \tau\}$, where $\tau \in (0, 1)$.

The τth sample quantile $\xi(\tau)$ is the minimizer of the expression below:

$$\xi(\tau) = \text{argmin}_\xi \sum_{i=1}^n \rho_\tau(y_i - \xi),$$

where $\rho_\tau(z) = z(\tau - I(z < 0))$, $\tau \in (0, 1)$, and $I(.)$ is the indicator function.

For a given $\tau \in (0, 1)$, quantile regressions estimate the linear conditional quantile function $Q(\tau | X = x) = x'\beta(\tau)$ by solving:

$$\hat{\beta}(\tau) = \text{argmin}_\beta \sum_{i=1}^n \rho_\tau(y_i - x_i'\beta).$$

The quantity $\hat{\beta}(\tau)$ is the estimate of the τth regression quantile.

2.2.1. Quantile Regressions for Intraday VaR

To better understand the effect of the entropy on IVaR, we estimate a quantile regression model using the entropy of the distribution of intraday returns as explanatory variable. The model is:

$$Q_{IVaR}(\tau | H_{\delta;t} = x) = \beta_0 + \beta_1 H_{\delta;t}, \tag{11}$$

where, for a given time scale v and significance level α:

$$Q_{IVaR}(\tau) = inf\{IVaR_{\alpha,v;t}, F(IVaR_{\alpha,v;t}) \geq \tau\}.$$

2.2.2. Quantile Regressions for Daily Returns

Quantile regressions can be used to assess the relationship between extreme values of daily returns and the entropy of the distribution of intraday returns. We estimate the following model:

$$Q_R(\tau | H_{\delta;t} = x) = \beta_0 + \beta_1 H_{\delta;t}, \tag{12}$$

where, denoting the daily log-returns by R_t, the quantile of the returns is denoted by:

$$Q_R(\tau) = inf\{R_t, F(R_t) \geq \tau\}.$$

2.3. *Forecasting Daily VaR Using Entropy*

The daily VaR at probability level α can be defined by the following equation:

$$Pr(R_t < -VaR_{\alpha;t}) = \alpha.$$

VaR can be forecasted using various methods. Furthermore, many VaR measures and forecasts fail to react fast enough to new information (market shocks) so often underestimate risk. Entropy, on the other hand, is very sensitive to new information so can be used to update and improve VaR forecasts. In order to forecast the daily entropy-based VaR, with $t \in \{k+1, ..., k+w\}$, $k \in \{0, ..., T-w+1\}$ and T the number of daily returns, Equation (13) below is estimated (we tried adding extra lags of the

entropy in the regression, but the extra lags were not significant, the optimal lag length was found to be one) using a rolling window of length w:

$$Q_{R,t}(\alpha) = \beta_0 + \beta_1 H_{\delta;t-1}. \tag{13}$$

Estimating this on the time interval $[k+1, k+w]$, the parameter estimates β_0^k and β_1^k are obtained. Then the forecast of VaR for the next trading day is given by the following:

$$\hat{VaR}_{\alpha;k+w+1} = -\beta_0^k - \beta_1^k H_{\delta;k+w}. \tag{14}$$

Equation (13) can be extended to include an autoregressive term for the VaR as below (it is not required to add extra lags of the quantile in the regression as the quantile is highly autocorrelated):

$$Q_{R,t}(\alpha) = \beta_0 + \beta_1 H_{\delta;t-1} + \beta_2 Q_{R,t-1}(\alpha). \tag{15}$$

Then the forecast of VaR for the next trading day can be computed as:

$$\hat{VaR}_{\alpha;k+w+1} = -\beta_0^k - \beta_1^k H_{\delta;k+w} + \beta_2^k VaR_{\alpha;k+w}. \tag{16}$$

The results obtained based on the models (13) & (15) and forecasting formulae (14) & (16), respectively, are compared with the VaR forecasting results obtained from a historical VaR forecasting model, and a VaR forecasting model based on the GARCH(1,1) model below:

$$\begin{cases} R_t = \mu + \varepsilon_t \\ \varepsilon_t = \sigma_t z_t \\ \sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \beta_1 \varepsilon_{t-1}^2 \end{cases}. \tag{17}$$

Model (17) is estimated using the same rolling windows as above, with the error term z_t being a standard normal variable or following a Student's t distribution, and the forecast of the VaR for the next trading day is given by the formula below:

$$\hat{VaR}_{k+w+1} = -(q_\alpha \hat{\sigma}_{k+w+1} + \mu_{k+w+1}), \tag{18}$$

where q_α is the Gaussian quantile or the Student-t quantile with the degrees of freedom estimated.

Thus, we compare the VaR forecasting ability of the following five models:

1. Historical VaR forecasts, estimated using a rolling window of length w;
2. Normal GARCH(1,1) VaR forecasts, with (17), $z_t \sim N(0,1)$ and forecasting Equation (18);
3. Student's t-GARCH(1,1) VaR forecasts, with (17), $z_t \sim$ Student's t and Equation (18);
4. Entropy-based VaR forecasts, given by (13) and forecasting Equation (14);
5. Entropy-based autoregressive VaR forecasts, given in (15) and (16).

In order to test the forecasting ability of the above models, Christoffersen's [36] tests are used: the LR test of Unconditional Coverage, the LR test of Independence and the LR test of Independence and Conditional Coverage. Also, we employed the forecast performance tests of Diebold and Mariano [37] and West [38], using the loss function of Giacomini and White [39] (see Appendixes B and C):

$$L\left(R_{t+1}, \hat{VaR}_{t+1|t}\right) = \left(R_{t+1} + \hat{VaR}_{t+1|t}\right)\left[\alpha - I_{R_{t+1} < -\hat{VaR}_{t+1|t}}\right]. \tag{19}$$

3. Empirical Analysis

In order to illustrate the application of the entropy of the distribution of intraday returns in financial risk management, we consider the EUR/JPY exchange rate (sourced from Disk Trading). FX rates are mostly symmetric, so the two tails of the distributions are similar, and so the entropy is closer related to the VaR estimate. For highly asymmetrical distributions, like stocks and commodities, the link between entropy and VaR could be weaker. The time period considered is 1999–2005. The database used for estimation has two components: (1) intraday prices (2025 transaction days and 2,340,624 minute-by-minute intraday observations); and (2) daily prices (2025 daily observations). Using the methodology from Bandi and Russell [35], the optimal sampling frequency for intraday data was estimated at 10 minutes on average ($\delta = \delta^* = 10$); this frequency is used to compute the entropy of distribution of intraday returns. We use the frequencies of $v \in \{1, 10, 15\}$ (minutes) to compute intraday measures of risk and uncertainty, and compute the risk measures at 1% significance level.

3.1. Entropy and Intraday Measures of Market Risk and Uncertainty

Figure 2 presents a comparison of the entropy and the intraday ES (estimated at 1% level). The two series show a strong negative correlation; the entropy has a similar relationship with IVaR and IRV as well. Next, the models given in (9) and (10) are estimated, using the statistical software SAS 9.3.
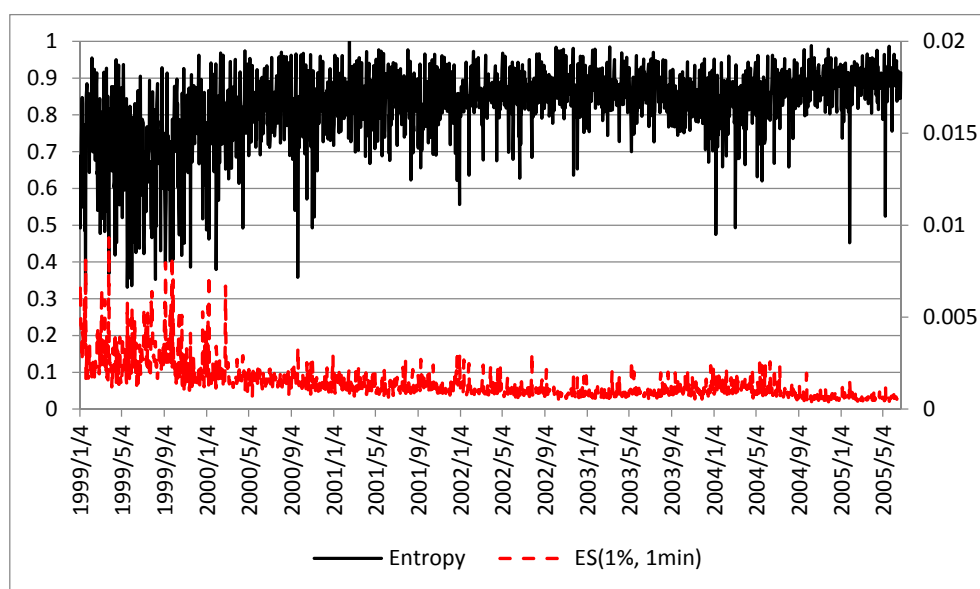

Figure 2. Intraday ES and the entropy of the distribution of intraday returns.

Panel A in Table 2 presents the results of the regressions specified in (9), using level $\alpha = 1\%$ for the risk measures. The R^2 estimates show that the entropy is strongly linked with intraday VaR, intraday ES and intraday Realized Volatility and, as expected, the coefficients are all negative and significant. Panel B in Table 2 presents the results of the dynamic regression models (10) using significance level $\alpha = 1\%$ for the risk measures. The coefficients of the entropy remain in all cases negative and significant, showing that the entropy has forecasting power for intraday measures of risk and uncertainty, even after taking past values of intraday measures of risk and uncertainty into account.

Table 2. The relationship between intraday measures of risk and uncertainty and entropy.

Sampling Frequency	v = 1 min			v = 10 min			v = 15 min		
	Dependent Variable			Dependent Variable			Dependent Variable		
	$IVaR_{\alpha,v;t}$	$IES_{\alpha,v;t}$	$IRV_{v;t}$	$IVaR_{\alpha,v;t}$	$IES_{\alpha,v;t}$	$IRV_{v;t}$	$IVaR_{\alpha,v;t}$	$IES_{\alpha,v;t}$	$IRV_{v;t}$
Panel A. Static Models									
$H_{\delta;t}$	−0.0050 ***	−0.0070 ***	−0.049 ***	−0.004 ***	−0.7565 ***	−0.044 ***	−0.0089 ***	−1.0669 ***	−0.889 ***
	[0.0003]	[0.0003]	[0.0015]	[0.0002]	[0.0205]	[0.0009]	[0.0002]	[0.0234]	[0.0018]
R^2_{adj}	0.46	0.58	0.66	0.29	0.51	0.63	0.48	0.51	0.53
Panel B. Dynamic Models									
$IVaR_{\alpha,v;t-1}$	0.5355 ***	-	-	0.2920 ***	-	-	0.3687 ***	-	-
	[0.0234]	-	-	[0.0351]	-	-	[0.0271]	-	-
$IES_{\alpha,v;t-1}$	-	0.6024 ***	-	-	0.1991 ***	-	-	0.3599 ***	-
	-	[0.0262]	-	-	[0.0428]	-	-	[0.0280]	-
$IRV_{v;t-1}$	-	-	0.629 ***	-	-	0.569 ***	-	-	0.513 ***
	-	-	[0.0300]	-	-	[0.0433]	-	-	[0.0280]
$H_{\delta;t-1}$	−0.0009 ***	−0.0420 ***	−0.001 ***	−0.0011 ***	−0.1607 ***	−0.0003 ***	−0.0017 ***	−0.1835 ***	−0.0002 ***
	[0.0002]	[0.0246]	[0.0018]	[0.0003]	[0.0455]	[0.0024]	[0.0004]	[0.0420]	[0.0034]
R^2_{adj}	0.40	0.41	0.38	0.16	0.11	0.33	0.22	0.21	0.26

Note: Estimation results for regressions (9) and (10) in Panel A and B, respectively. Risk measures IVaR and IES are at level α = 1%. White's heteroscedasticity consistent standard errors are given in brackets. *** signify significance at 1%.

3.2. Quantile Regression Results

We use quantile regressions to see the effect of the entropy on the quantiles of VaR, estimating Equation (11) for frequencies of 1, 10 and 15 min and α = 1%. Panel A of Table 3 presents the estimation results; as expected, there is a positive correlation between the entropy and the upper tails of the distribution of IVaR. Also, IVaR is more sensitive to entropy in the upper tail of its distribution, meaning that high values of IVaR have a stronger relationship with entropy. Furthermore, the relationship is strongest for v = 15 min frequency. As an example, Figure 3 gives a visual presentation of the dependence of the 99% quantile on entropy.

Table 3. Quantile regression results.

	Panel A. for Intraday VaR						Panel B. for Daily Returns	
	τ = 1% Q for IVaR			τ = 5% Q for IVaR			τ = 1% Q for Returns	τ = 5% Q for Returns
Sampling Frequency	v = 1 min	v = 10 min	v = 15 min	v = 1 min	v = 10 min	v = 15 min	Daily	Daily
$H_{\delta;t}$	0.0111 ***	0.0092 ***	0.0154 ***	0.0078 ***	0.0073 ***	0.0140 ***	0.0728 ***	0.0368 ***
	[0.0008]	[0.0012]	[0.0012]	[0.0005]	[0.0005]	[0.0005]	[0.0155]	[0.0075]
Confidence Interval (95%)	0.0126	0.0116	0.0177	0.0088	0.0083	0.0150	0.1031	0.0516
	0.0095	0.0068	0.013	0.0068	0.0063	0.0131	0.0424	0.0220
t-Value	13.77	7.58	12.76	15.53	14.39	28.92	4.71	4.88

Note: Quantile regressions for IVAR (Panel A) and daily log-returns (Panel B) illustrating the effect of entropy on the quantiles of IVaR and daily log-returns; the models estimated are (11) and (12). *** signify significance at 1%.
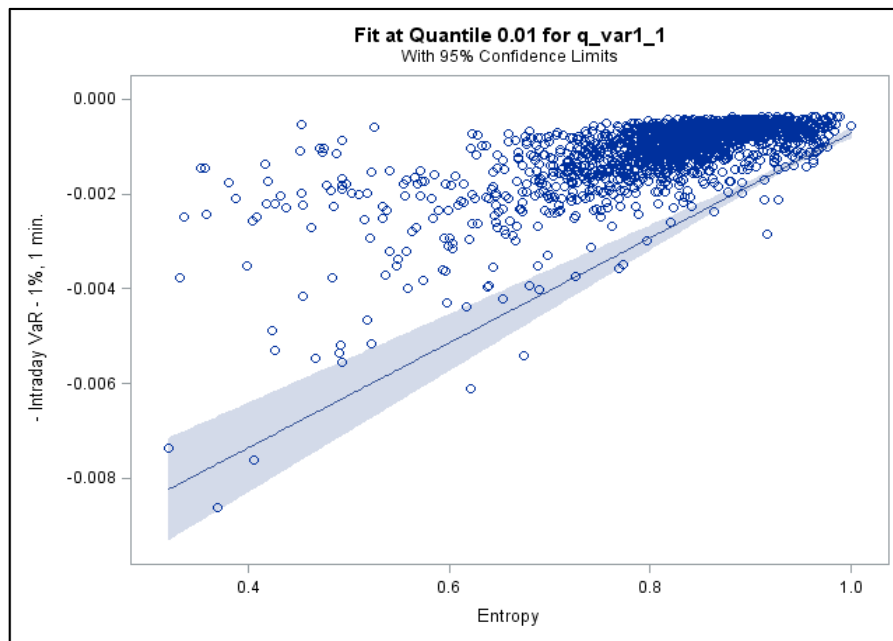
Figure 3. Quantile regression results for the 1% quantile of IVaR as a function of the entropy, for ν = 1 min.

Regarding the relationship between extreme values of daily returns and the entropy of the distribution of intraday returns, Panel B of Table 3 presents the results of regression (12), whilst Figure 4 presents the scatter plot and the regression line of the estimation. As expected, the relationship between the quantile (equal with minus VaR) and entropy is positive and significant. Low (high) values of the entropy of the distribution of intraday returns generally correspond to high (low) absolute VaR estimates for daily returns; in line with the results in the previous section.
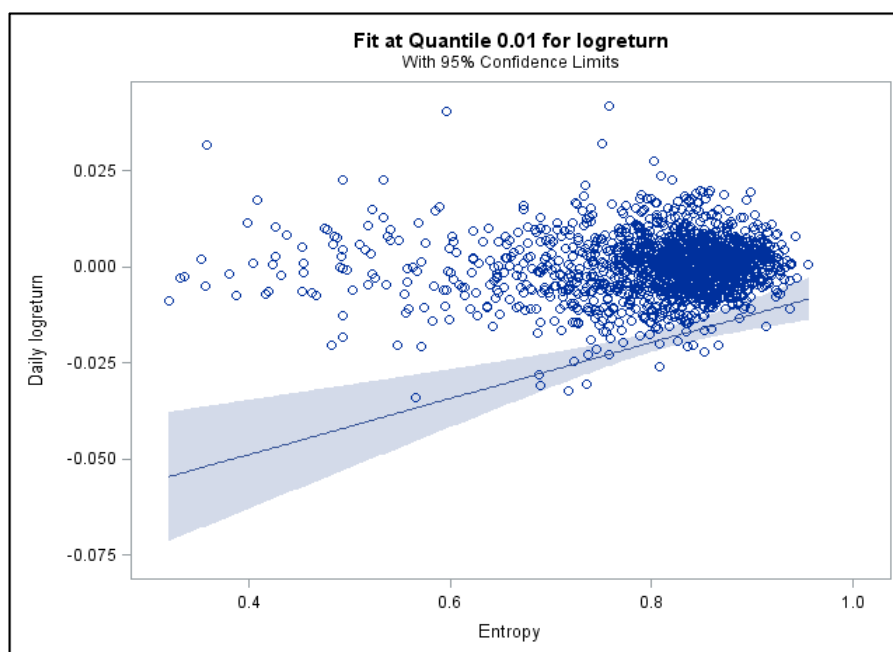

Figure 4. Quantile regression results for the 1% quantile of the returns as a function of the entropy.

3.3. Forecasting Daily VaR Using Entropy

Entropy can be used to forecast VaR, and for this Equation (13) is estimated on a rolling basis. We use a length of $w = 1000$ days for the estimation windows. The time series for daily returns are shown and plotted against the entropy-based 1% VaR forecasts and the GARCH-based 1% VaR forecasts in Figures 5 and 6. As expected, the GARCH-based VaR is more stable whilst the entropy-based VaR forecast reacts faster to new information.
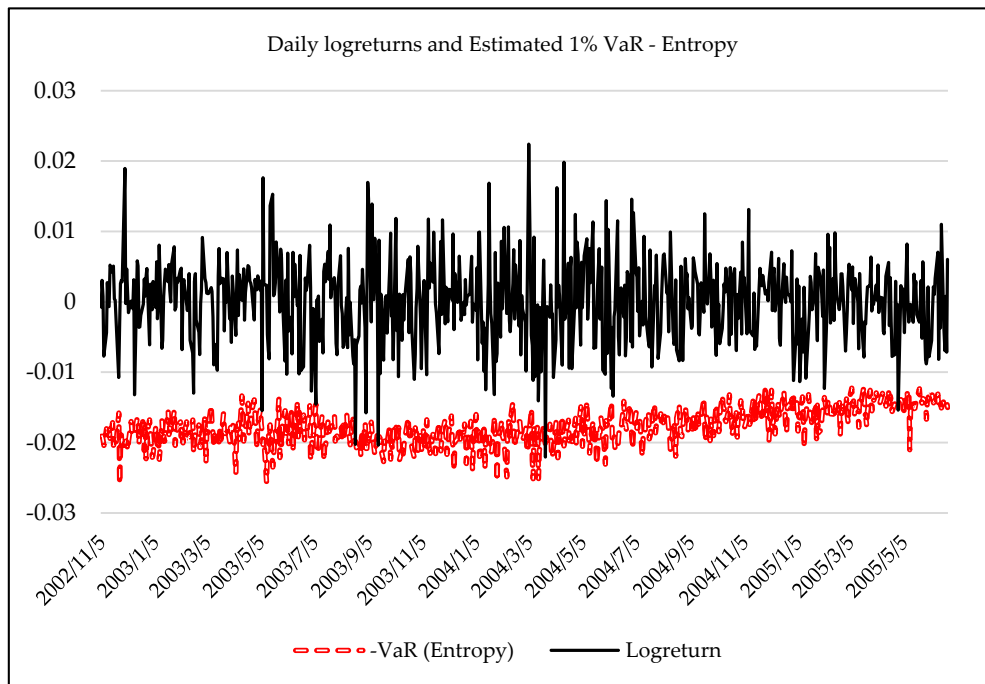

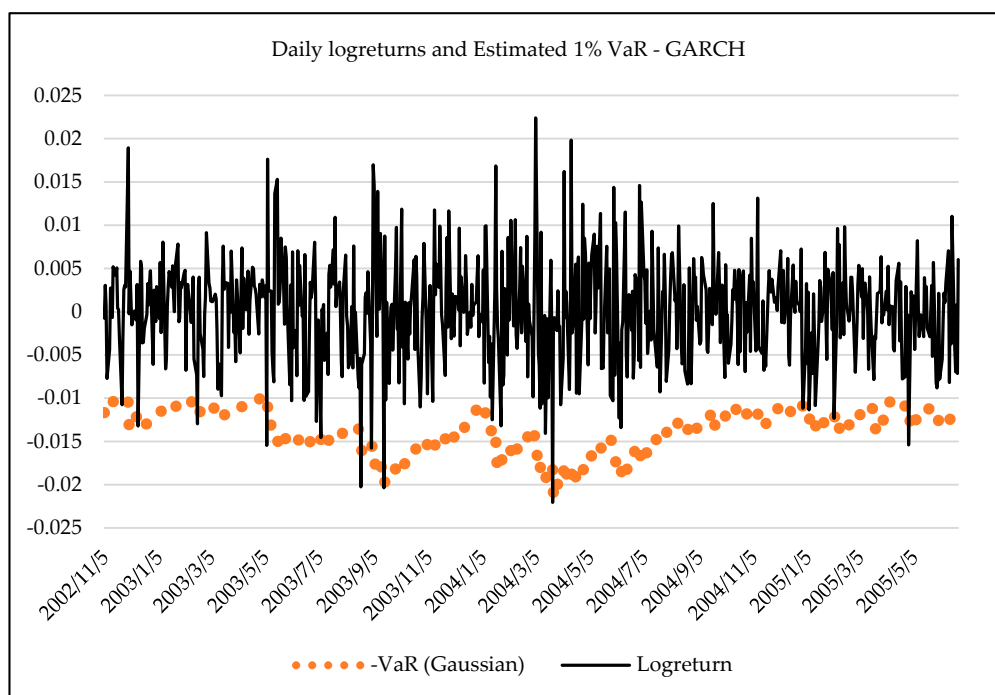**Figure 5.** Forecasting daily 1% VaR using entropy.

Figure 6. Forecasting daily 1% VaR using a normal GARCH(1,1) model.

The backtesting results of the VaR forecasts for the five models given in Section 2.3, for $\alpha = 1\%$ are presented in Table 4. The second column gives the probabilities that the returns are below the negative of the VaR, for different models; it can be seen that the Historical VaR model (with a probability of 0.29) provides the highest VaR forecasts on average. Looking at the unconditional test results, the entropy-based AR VaR has the smallest test statistic and the Historical VaR model marginally fails the test. Considering the independence and full results, it can be seen that the Historical VaR model performs the worst failing the tests, whilst the normal GARCH model passes all three tests. We conclude that the best results overall are obtained by the entropy-based AR VaR forecast model.

Table 4. VaR forecast backtest results.

Model	$\Pr(R_t < -\hat{VaR_t})$	LR_{uc} Test	*p*-Value	LR_i Test	*p*-Value	LR_{full} Test	*p*-Value
Historical VaR	0.290%	4.867 **	0.027	7.443 ***	0.006	12.310 ***	0.002
n.GARCH(1,1) VaR	1.597%	2.097	0.148	1.658	0.198	3.755	0.153
t-GARCH(1,1) VaR	0.581%	1.442	0.230	5.108 **	0.024	6.550 **	0.038
Entropy VaR	0.726%	0.579	0.447	4.329 **	0.037	4.908 *	0.086
Entropy AR VaR	1.016%	0.002	0.966	3.156 *	0.076	3.158	0.206

Note: Backtest results for daily 1% VaR forecasts based on the following five models specified in Section 2.3: (1) Historical VaR forecasts; (2) Normal GARCH(1,1) VaR forecasts; (3) Student's t-GARCH(1,1) VaR forecasts; (4) Entropy-based VaR forecasts and (5) Entropy-based AR VaR forecasts. Christoffersen's backtests were used (see Appendix B) with $N = 689$. *, ** and *** signify significance at 10%, 5% and 1%, respectively.

Additionally, we employed the VaR forecast comparison tests of Diebold and Mariano [37] and West [38]. The test statistic only considers the unconditional forecasting ability of the VaR models, and is highly asymmetric, the loss function in (19) favouring models which overestimate VaR and strongly penalizes models which, even very mildly, underestimate it. Our results in Table 5 show that the Historical VaR is the best performer, whilst the t-GARCH(1,1) based VaR and the entropy-based VaR models are favoured over the normal GARCH(1-1) VaR. Furthermore, the difference between the performance of the entropy-based AR VaR model and the entropy-based VaR is not statistically significant. However, these results depend very strongly on the loss function. Looking at the overall picture, we conclude that the entropy has good forecasting power for VaR.

Table 5. The Diebold and Mariano test results for VaR forecasts.

Model	n.GARCH(1,1) VaR	t-GARCH(1,1) VaR	Entropy VaR	Entropy AR VaR
Historical VaR	−3.146 ***	−1.004	−1.755 **	−2.294 **
n.GARCH(1,1) VaR	-	2.744 ***	2.167 **	1.428
t-GARCH(1,1) VaR	-	-	−0.448	−1.141
Entropy VaR	-	-	-	−1.428

Note: 1% VaR forecast comparison tests of Diebold and Mariano [37] and West [38], with loss function given in [19]. The DM test statistics reported are for comparisons of the models on the left side against models in the column titles. ** and *** signify significance at 5% and 1%, respectively. $N = 689$.

4. Conclusions

This paper investigates the link between entropy and various measures of market risk such as Value-at-Risk or Expected Shortfall. Based on the result of Lorentz [25], we developed the concept of entropy of a distribution function and we applied this concept to estimate the entropy of the distribution of intraday returns. Using Monte-Carlo simulations, we showed that there is an inverse relationship between entropy and the probability in the tails of a distribution, high levels of entropy being characteristic of a distribution with light tails like the normal distribution, and low entropy values being associated with heavy-tailed distributions.

Furthermore, we investigated the relationship between risk measures and the entropy of the distribution of intraday returns, in a static and dynamic setting. The entropy of a distribution function has more informational content than the classical measures of market risk, as it takes into account

the entire distribution. We found evidence of a strong, negative relationship between entropy and intraday Value-at-Risk, intraday Expected Shortfall and intraday Realized Volatility. From a dynamic point of view, the entropy proves to be a strong predictor for IVaR, IES and IRV, with R^2 values up to 41%. Our quantile results confirm that the entropy has a strong explanatory power for the quantiles of the intraday VaR as well as the quantiles of the daily returns. The final part of our empirical study compares entropy-based VaR estimates, which are mostly more reactive to new information than standard VaR models, with competing VaR forecasts. Whilst the Historical VaR model is the preferred model based on the Diebold-Mariano (unconditional) test results, when Cristoffersen's unconditional, conditional and joint test results are considered, it comes out as the worst performer, failing the tests. Cristoffersen's three tests favour the entropy-based AR VaR model, and we conclude that the entropy is a strong predictor of daily VaR, performing better than the competing VaR models. As it takes into account the extreme events that happen at an intraday level, it proves to be able to provide reliable VaR forecasts.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Algorithm for Simulation of Stable Distributions (Weron [34])

Step 1. Generate the following two random variables: $U \sim Unif\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ uniformly distributed and $E \sim Exp(1)$ exponentially distributed;

Step 2. Compute:

$$
X = \begin{cases}
\frac{2}{\pi}\left(\frac{\pi}{2} + \beta U\right)tanU - \beta ln \frac{\frac{\pi}{2} E cosU}{\frac{\pi}{2} + \beta U} & if\ \alpha = 1 \\
\left(1 + \beta^2 tan^2 \frac{\pi\alpha}{2}\right)^2 \frac{sin(U+B(\alpha,\beta))}{(cosU)^{1/\alpha}} \left(\frac{U - \alpha(U+B(\alpha,\beta))}{E}\right)^{\frac{1-\alpha}{\alpha}} & otherwise
\end{cases}
,
$$

where $B(\alpha, \beta) = \frac{\arctan\left(\beta \tan \frac{\pi\alpha}{2}\right)}{\alpha}$;

Step 3. Compute $Y = \begin{cases} \gamma X + \frac{2}{\pi}\beta\gamma ln\gamma + \delta & if\ \alpha = 1 \\ \gamma X + \delta & otherwise \end{cases}$, which will follow a stable distribution $S(\alpha, \beta, \gamma, \delta; 1)$.

Appendix B. VaR Forecasting Tests (Christoffersen [36])

Appendix B.1. The LR Test of Unconditional Coverage

Let $\hat{VaR}_t$ denote the forecasted VaR and let R_t be the observed logreturn;

Define $I_t = \begin{cases} 1, if\ R_t < -\hat{VaR}_t \\ 0, otherwise \end{cases}$; The hypothesis being tested is: $\begin{cases} H_0 : E(I_t) = \alpha \\ H_A : E(I_t) \neq \alpha \end{cases}$;

The test statistic is defined as:

$$
LR_{uc} = -2\log\frac{L(\alpha)}{L(\hat{\alpha})} = -2\log\frac{\alpha^{n_0}(1-\alpha)^{n-n_0}}{\hat{\alpha}^{n_0}(1-\hat{\alpha})^{n-n_0}} \approx \chi^2(1), where
$$

$$
\hat{\alpha} = \frac{n_0}{n} = \Pr(I_t = 0)
$$

Appendix B.2. The LR Test of Independence

Consider I_t a first-order Markov chain with transition probability matrix:

$$\Pi_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}, \text{where } \boldsymbol{\pi}_{ij} = \Pr(\boldsymbol{I}_t = \boldsymbol{i} | \boldsymbol{I}_{t-1} = \boldsymbol{j}).$$

Then the likelihood function is defined as $L(\Pi_1) = (1 - \boldsymbol{\pi}_{01})^{n_{00}} \boldsymbol{\pi}_{01}{}^{n_{01}} (1 - \boldsymbol{\pi}_{11})^{n_{10}} \boldsymbol{\pi}_{11}{}^{n_{11}}$.

The likelihood under the null hypothesis of independence is

$$L(\Pi_2) = (1 - \boldsymbol{\pi}_2)^{n_{00}+n_{10}} \boldsymbol{\pi}_2{}^{n_{01}+n_{11}}, \text{where } \boldsymbol{\pi}_2 = \frac{n_{01} + n_{11}}{n_{01} + n_{00} + n_{10} + n_{11}};$$

The test statistic is defined as:

$$LR_i = -2 \log \frac{L(\hat{\Pi}_2)}{L(\hat{\Pi}_1)} \approx \chi^2(1).$$

Appendix B.3. The Joint Test of Coverage and Independence

The test statistic is given by:

$$LR_{full} = -2 \log \frac{L(\alpha)}{L(\hat{\Pi}_1)} \approx \chi^2(2).$$

Appendix C. The Diebold-Mariano Test for VaR Forecast Comparisons (Diebold and Mariano [37] and West [38])

Let $\hat{VaR}_t^A$ and $\hat{VaR}_t^B$ denote two competing VaR forecasts given by models A and B, respectively and let R_t be the observed logreturn. We denote by $L(R_{t+1}, \hat{VaR}_{t+1|t}^X)$ the loss function of the VaR forcast of model $X \in \{A, B\}$ (based on the distance between the returns and the VaR forecast of model X) and let d_t below be the difference between the two loss functions:

$$d_t = L(R_{t+1}, \hat{VaR}_{t+1|t}^A) - L(R_{t+1}, \hat{VaR}_{t+1|t}^B).$$

The hypothesis being tested is: $\begin{cases} H_0 : E(d_t) = 0 \\ H_A : E(d_t) \neq 0 \end{cases}$.

We compute the average of these differences as (assuming there are T 1-day ahead forecasts):

$$\bar{d} = \frac{1}{T} \sum_{t=k+1}^{k+T} d_t$$

and let $\hat{V}(d)$ be a HAC-consistent variance estimator of the true variance of d_t (computed, for example, using the Newey-West estimator).

The Diebold-Mariano test statistic is computed as:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(d)/T}}.$$

This will follow an asymptotical normal distribution. Large negative (positive) values indicate that model A (B) provides superior forecasts.

References

1. Uffink, J. Bluff your way in the second law of thermodynamics. *Stud. Hist. Philos. Mod. Phys.* **2001**, *32*, 305–394. [CrossRef]
2. Avery, J. *Information Theory and Evolution*, 2nd ed.; World Scientific Publishing: Singapore, 2012.
3. Zunino, L.; Zanin, M.; Tabak, B.M.; Pérez, D.G.; Rosso, O.A. Forbidden patterns, permutation entropy and stock market inefficiency. *Phys. A Stat. Mech. Appl.* **2009**, *388*, 2854–2864. [CrossRef]
4. Risso, A. The informational efficiency and the financial crashes. *Res. Int. Bus. Financ.* **2008**, *22*, 396–408. [CrossRef]
5. Oh, G.; Kim, S.; Eom, C. Market efficiency in foreign exchange markets. *Phys. A Stat. Mech. Appl.* **2007**, *382*, 209–212. [CrossRef]
6. Wang, Y.; Feng, Q.; Chai, L. Structural evolutions of stock markets controlled by generalized entropy principles of complex systems. *Int. J. Mod. Phys. B* **2010**, *24*, 5949–5971. [CrossRef]
7. Maasoumi, E.; Racine, J. Entropy and predictability of stock market returns. *J. Econ.* **2002**, *107*, 291–312. [CrossRef]
8. Billio, M.; Casarin, R.; Costola, M.; Pasqualini, A. An entropy-based early warning indicator for systemic risk. *J. Int. Financ. Mark. Inst. Money* **2016**, *45*, 42–59. [CrossRef]
9. Dionisio, A.; Menezes, R.; Mendes, D.A. An econophysics approach to analyse uncertainty in financial markets: An application to the Portuguese stock market. *Eur. Phys. J. B* **2006**, *50*, 161–164. [CrossRef]
10. Philippatos, G.C.; Wilson, C. Entropy, market risk and the selection of efficient portfolios. *Appl. Econ.* **1972**, *4*, 209–220. [CrossRef]
11. Ebrahimi, N.; Maasoumi, E.; Soofi, E.S. Ordering univariate distributions by entropy and variance. *J. Econ.* **1999**, *90*, 317–336. [CrossRef]
12. Ebrahimi, N.; Maasoumi, E.; Soofi, E.S. Measuring Informativeness of Data by Entropy and Variance. In *Advances in Econometrics: Income Distribution and Methodolgy of Science, Essays in Honor of Camilo Dagum*; Springer: Heidelberg, Germany, 1999.
13. Allen, D.E.; McAleer, M.; Powell, R.; Singh, A.K. A non-parametric and entropy based analysis of the relationship between the VIX and S&P 500. *J. Risk Financ. Manag.* **2013**, *6*, 6–30.
14. Liu, L.Z.; Qian, X.Y.; Lu, H.Y. Cross-sample entropy of foreign exchange time series. *Phys. A Stat. Mech. Appl.* **2010**, *389*, 4785–4792. [CrossRef]
15. Bowden, R.J. Directional entropy and tail uncertainty, with applications to financial hazard. *Quant. Financ.* **2011**, *11*, 437–446. [CrossRef]
16. Gradojevic, N.; Gencay, R. Overnight interest rates and aggregate market expectations. *Econ. Lett.* **2008**, *100*, 27–30. [CrossRef]
17. Gencay, R.; Gradojevic, N. Crash of '87—Was it expected? Aggregate market fears and long range dependence. *J. Empir. Financ.* **2010**, *17*, 270–282. [CrossRef]
18. Gradojevic, N.; Caric, M. Predicting systemic risk with entropic indicators. *J. Forecast.* **2017**, *36*, 16–25. [CrossRef]
19. Stutzer, M.J. Simple entropic derivation of a generalized Black-Scholes option pricing model. *Entropy* **2000**, *2*, 70–77. [CrossRef]
20. Stutzer, M.J.; Kitamura, Y. Connections between entropic and linear projections in asset pricing estimation. *J. Econ.* **2002**, *107*, 159–174. [CrossRef]
21. Yang, J.; Qiu, W. A measure of risk and a decision-making model based on expected utility and entropy. *Eur. J. Oper. Res.* **2005**, *164*, 792–799. [CrossRef]
22. Ishizaki, R.; Inoue, M. Time-series analysis of foreign exchange rates using time-dependent pattern entropy. *Phys. A Stat. Mech. Appl.* **2013**, *392*, 3344–3350. [CrossRef]
23. Bekiros, S. Timescale analysis with an entropy-based shift-invariant discrete wavelet transform. *Comput. Econ.* **2014**, *44*, 231–251. [CrossRef]
24. Bekiros, S.; Marcellino, M. The multiscale causal dynamics of foreign exchange markets. *J. Int. Money Financ.* **2013**, *33*, 282–305. [CrossRef]
25. Lorentz, R. On the entropy of a function. *J. Approx. Theor.* **2009**, *158*, 145–150. [CrossRef]
26. Zhou, R.; Cai, R.; Tong, G. Applications of entropy in finance: A review. *Entropy* **2013**, *15*, 4909–4931. [CrossRef]

27. Pele, D.T.; Mazurencu-Marinescu, M. Uncertainty in EU stock markets before and during the financial crisis. *Econophys. Sociophys. Multidiscip. Sci. J.* **2012**, *2*, 33–37.

28. Pele, D.T. Information entropy and occurrence of extreme negative returns. *J. Appl. Quant. Methods* **2011**, *6*, 23–32.

29. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman and Hall: London, UK, 1986.

30. Yamato, H. Uniform convergence of an estimator of a distribution function. *Bull. Math. Stat.* **1973**, *15*, 69–78.

31. Chacón, J.E.; Rodríguez-Casal, A. A note on the universal consistency of the kernel distribution function estimator. *Stat. Probab. Lett.* **2009**, *80*, 1414–1419. [CrossRef]

32. Pele, D.T. Uncertainty and Heavy Tails in EU Stock Markets before and during the Financial Crisis. In Proceedings of the 13th International Conference on Finance and Banking, Lessons Learned from the Financial Crisis, Ostrava, Czech Republic, 12–13 October 2011; pp. 501–512.

33. Nolan, J.P. *Stable Distributions—Models for Heavy Tailed Data*; Birkhauser: Boston, MA, USA, 2011.

34. Weron, R. On the Chambers-Mallows-Stuck method for simulating skewed stable random variables. *Stat. Probab. Lett.* **1996**, *28*, 165–171. [CrossRef]

35. Bandi, F.; Russell, J. Separating microstructure noise from volatility. *J. Financ. Econ.* **2006**, *79*, 655–692. [CrossRef]

36. Christoffersen, P. Evaluating interval forecasts. *Int. Econ. Rev.* **1998**, *39*, 841–862. [CrossRef]

37. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263. [CrossRef]

38. West, K.D. Asymptotic inference about predictive ability. *Econometrica* **1996**, *64*, 1067–1084. [CrossRef]

39. Giacomini, R.; White, H. Tests of conditional predictive ability. *Econometrica* **2006**, *74*, 1545–1578. [CrossRef]