A statistical test of market efficiency based on information theory

Xavier Brouty & Matthieu Garcin

Published online: 23 May 2023.

Submit your article to this journal ⬈

Article views: 51

View related articles ⬈

View Crossmark data ⬈

A statistical test of market efficiency based on information theory

XAVIER BROUTY† and MATTHIEU GARCIN ⬡*‡

†ESILV, Paris La Défense, 92916, France
‡Léonard de Vinci Pôle Universitaire, Research Center, Paris La Défense, 92916, France

We determine the amount of information contained in a time series of price returns at a given time scale, by using a widespread tool of the information theory, namely the Shannon entropy, applied to a symbolic representation of this time series. By deriving the exact and the asymptotic distribution of this market information indicator in the case where the efficient market hypothesis holds, we develop a statistical test of market efficiency. We apply it to a real dataset of stock indices, single stocks, and cryptocurrencies, for which we are able to determine at each date whether the efficient market hypothesis is to be rejected, with respect to a given confidence level.

Keywords: Market efficiency; Information theory; Shannon entropy

1. Introduction

The efficient market hypothesis (EMH) is a cornerstone of financial theory. According to Eugene Fama, '*a market in which prices always fully reflect available information is called efficient*' (Fama 1970). One can in fact distinguish different natures of market efficiency, depending on how one defines the information set. The weak form of market efficiency considers that current prices reflect all the information contained in past prices. The semi-strong form includes, in addition, public information such as announcements of annual earnings or stock splits. The strong form also includes private information for some investors.

In this work, we are interested in the weak-form efficiency. Our purpose is to determine whether this efficiency is a realistic assumption and therefore whether arbitrages may exist. In other words, we want to answer the question: can one take advantage of past prices to predict future evolutions of prices? However, one must distinguish two kinds of arbitrages. Pure arbitrage leads to a certain gain. Statistical arbitrage leads to an uncertain output which, in average only, is a gain. Since pure arbitrages are in practice almost instantaneously eliminated by the market, we are more interested here in statistical arbitrages. In this perspective, we define the market efficiency as the absence of statistical arbitrages with predictions based on past prices, that is as the characterization of prices by

martingales. Therefore, determining whether the market is efficient or not, according to this definition, is of major importance both for quantitative portfolio managers, to help them define a universe of predictable assets to trade, and for market makers, to help them adjust their prices and avoid informed trades.

The financial literature puts forward several solutions to answer this question of the relevance of the EMH. The Hurst exponent is a widespread statistic for this purpose. As a parameter of a specific model, namely the fractional Brownian motion (fBm), it is related to the covariance between various price increments (Mandelbrot and Van Ness 1968). Therefore, given a particular value of the Hurst exponent, building predictions with this model is possible (Nuzman and Poor 2000, Garcin 2017) and statistical arbitrages naturally follow (Guasoni *et al.* 2019, 2021, Garcin 2022b).

However, the fBm is not always the most realistic specification for depicting time series of log-prices. One may thus be inclined to use some extensions of the fBm, for example getting rid of the Gaussian distribution (Stoev and Taqqu 2004, Weron *et al.* 2005, Garcin 2020, Ammy-Driss and Garcin 2023), using time-varying parameters in a deterministic (Peltier and Lévy Véhel 1994, Benassi *et al.* 1997, Ayache and Lévy Véhel 2000, Garcin 2017, Bianchi and Pianese 2018) or stochastic way (Ayache and Taqqu 2005, Bianchi *et al.* 2012, Garcin 2020), or transforming the fBm in a stationary process (Flandrin *et al.* 2003, Garcin 2019, 2022a). All these extensions have

*Corresponding author. Email: matthieu.garcin@m4x.org; matthieu.garcin@devinci.fr

fractal or multifractal properties, quantified by a Hurst exponent, whose interpretation is thus different from the fBm case and may not be related to the autocorrelation of the process. In this case, the Hurst exponent is not a relevant indicator of market efficiency and one therefore needs a statistic which is less related to a specific model.

Information theory proposes some model-free alternatives to the Hurst exponent. In particular, a rich literature in econophysics applies concepts from this field to the quantification of the complexity in financial time series (Soloviev *et al.* 2020, Bielinskyi *et al.* 2021). The most widespread measure of complexity is Shannon entropy which directly derives from a probability distribution, which can for example be the distribution of prices (Stosic *et al.* 2016, Lahmiri *et al.* 2018) or of singular values of the matrix of lagged subsequences of prices (Espinosa-Paredes *et al.* 2022). One can also cite the approximate entropy which measures the complexity of dynamical systems (Pincus 1991), with applications in finance (Pincus and Kalman 2004, Kristoufek and Vosvrda 2014). But one cannot easily relate the complexity of all these distributions to the notion of market efficiency. In this perspective, the complexity of two particular probability distributions is of greater interest. First, following the method proposed by Bandt and Pompe Bandt and Pompe (2002), the permutation entropy focuses on the distribution of the ordinal patterns of a time series of prices (Zunino *et al.* 2011, 2016, Bariviera *et al.* 2019). Second, Risso's method studies the discrete distribution of the sign of successive price returns (Risso 2008, Mensi *et al.* 2014, Ducournau 2021). The two approaches are close to each other in that they both consider that information resides in the order in which price increases and decreases occur.

In this literature on the application of information theory to measuring market efficiency, it is considered that the entropy is a gradual indicator of efficiency. The underlying idea is that the market is more or less efficient. We consider instead that the question about the relevance of the EMH is summarized into the following: can one predict future evolutions of prices with a ratio of winning bets significantly higher than 50%? Therefore, the statistical significance of entropy-based market efficiency indicators is crucial for their interpretation, beyond their gradual aspect.

Focusing on Risso's method, the purpose of this work is to provide a statistical test of market efficiency. We thus compare the entropy of the market with the entropy of an ideal efficient market, the difference of entropies corresponding to the market information. We show that the estimator of this information is subject to a statistical error because of a limited number of observations. We however give the moments of the distribution of this estimator as well as a more synthetic and practical formula for its asymptotic distribution, which follows a gamma law. We study on simulations the extent to which this asymptotic distribution applies and we build a statistical test of market efficiency. An empirical application illustrates that we are able to reject the EMH for some time series at certain times. This is a stronger conclusion than the simple comparison of efficiency levels usually put forward by the traditional literature on these entropy-related tools.

The rest of the paper is organized as follows. Section 2 presents the market information, which stems from Risso's method. Section 3 presents the statistical perspectives about this market information indicator, including the estimator, its asymptotic distribution, and the statistical test for the EMH. Section 4 shows an empirical study to various financial series. Section 5 concludes.

2. Market information

We consider, for a given financial asset, a time series of $n+1$ consecutive prices $P_0, \ldots, P_n$. In the empirical study, section 4, we work with daily prices, but we can also consider other frequencies.† We transform the series of prices in a binary series $X_1, \ldots, X_n$, which indicates whether the price has increased or decreased between two consecutive observations:

$$X_i = \mathbb{1}_{\{P_i - P_{i-1} > 0\}}.$$

This symbolic representation of a series of consecutive prices is consistent with Risso's approach (Risso 2008). For example the sequence $(X_1, X_2, X_3) = (0, 1, 1)$ represents a decrease followed by two daily price increases. For a given length $L < n$, 2^L sequences are possible. We order them, for example, with Gray's binary code. If $L = 3$, the 8 possible sequences are

$$(G_1^3, \ldots, G_8^3) = ((0,0,0), (0,0,1), (0,1,1), (0,1,0),$$
$$(1,1,0), (1,0,0), (1,0,1), (1,1,1)).$$

The purpose of Shannon entropy is to determine the amount of uncertainty in a discrete probability distribution. In Risso's approach, one focuses on the distribution of the sequences of increase indicators, given a length L. We assume that the random series (X_i) is stationary. The probability to draw a particular sequence of length L is noted $p_i^L = \mathbb{P}((X_., X_{.+1}, \ldots, X_{.+L-1}) = G_i^L)$. The Shannon entropy of the discrete distribution $(p_i^L)_i$ is defined by:

$$H^L = -\sum_{i=1}^{2^L} p_i^L \log_2\left(p_i^L\right),$$

where we use the convention that $0 \log_2(0) = 0$ which is reasonable because $\lim_{x \to 0} x \log_2(x) = 0$. We note that the highest entropy corresponds to a uniform distribution: $\forall i \in [\![1, 2^L]\!], p_i^L = 2^{-L}$.

In Risso's approach, a normalized version of H^L, thus belonging to $[0, 1]$, plays the role of efficiency indicator. The larger this indicator, the more efficient the market. This leads to a gradual interpretation of market efficiency. On the contrary, the approach we put forward is based on a binary interpretation of efficiency: the market is efficient or it isn't. We therefore propose an indicator which is slightly different from that of Risso. First, we don't need to normalize Shannon entropy. Second, instead of considering the entropy, we use the amount of information contained in the distribution of sequences of increase indicators X_i, with respect to what

---

† This kind of method has for instance already been applied to intraday price series of ETFs and FX rates (Shmilovici *et al.* 2009, Calcagnile *et al.* 2020, Shternshis *et al.* 2022).

would be this distribution according to the EMH. We thus need to clearly define what are the distributions consistent with the EMH. In this perspective, we find more natural to base our analysis on conditional distributions instead of on the non-conditional distributions used in Risso's approach. Indeed, in our symbolic framework, the statistical arbitrages appearing in an inefficient market take advantage of a difference of probability between the two possible values of the future increase indicator X_{L+1}, conditionally on observed past indicators $(X_1, \ldots, X_L)$.

Following this conditional approach, we decompose a sequence $(X_1, \ldots, X_{L+1})$ observed in time L in two parts. The L first elements constitute a prefix sequence, corresponding to the last L observed increase indicators at the current time. The suffix, X_{L+1}, corresponds to the next, unobserved and random increase indicator. The discrete probability of the prefix $(X_1, \ldots, X_L) = G_i^L$ is p_i^L, as already mentioned. Conditionally on this prefix, the distribution of the suffix X_{L+1} is Bernoulli of a parameter noted $\pi_i^L = \mathbb{P}(X_{L+1} = 1 | (X_1, \ldots, X_L) = G_i^L) \in [0, 1]$. Of course, π_i^L is only defined if $p_i^L \neq 0$. The full sequence $(X_1, \ldots, X_{L+1})$ is thus equal to $(G_i^L, 1)$ with probability $p_i^L \pi_i^L$ and $(G_i^L, 0)$ with probability $p_i^L(1 - \pi_i^L)$. We then get the Shannon entropy of the discrete distribution of this full sequence of length $L + 1$:

$$H^{L+1} = -\sum_{i=1}^{2^L} \left(p_i^L \pi_i^L \log_2 \left(p_i^L \pi_i^L \right) + p_i^L (1 - \pi_i^L) \right.$$
$$\left. \times \log_2 \left(p_i^L (1 - \pi_i^L) \right) \right). \quad (1)$$

The EMH asserts that, conditionally on the prefix $(X_1, \ldots, X_L)$, the two possible values for the suffix, $X_{L+1} = 1$ and $X_{L+1} = 0$, have equal probability. The Shannon entropy $H_\star^{L+1}$ for a market following the EMH thus expresses as a particular case of equation (1), with $\pi_i^L = 1/2$:

$$H_\star^{L+1} = -\sum_{i=1}^{2^L} p_i^L \log_2 \left(\frac{p_i^L}{2} \right) = 1 + H^L. \quad (2)$$

Finally, we define our efficiency indicator, the market information, as the difference between the entropy consistent with the ideal EMH and the true entropy of the market:

$$I^{L+1} = H_\star^{L+1} - H^{L+1}. \quad (3)$$

In section 3, we use the estimated value of this indicator to build a statistical test of market efficiency. This is possible because of the important following theorem, which shows that the value of I^{L+1} discriminates efficient and inefficient markets.

THEOREM 2.1 *For $L \in \mathbb{N} \setminus \{0\}$, we have $I^{L+1} \geq 0$. Moreover,*

$$I^{L+1} = 0 \iff \forall i \in [\![1, 2^L]\!] \setminus \left\{ j | p_j^L = 0 \right\}, \pi_i^L = \frac{1}{2}.$$

The proof is postponed in appendix 1.

Theorem 2.1 states that if the market follows the EMH, then I^{L+1} is equal to 0. If the market does not follow the EMH by leading to some $\pi_i^L \neq 1/2$, then $I^{L+1} > 0$. Nevertheless, we

are aware that some time series of prices may have all their π_i^L equal to 1/2, while they are not consistent with the EMH. The reason for this is that we have summarized a price increment in X_i, which can take only two values, 0 or 1. Some information contained in the time series is lost with this two-state symbolic approach. One could thus be tempted to generalize Risso's approach to a greater number of states or even to a continuum of states (Stosic *et al.* 2016, Lahmiri *et al.* 2018). We leave this possible extension to further research and focus on the classical two-state approach, which we believe to be enlightening. We however underline below some challenges regarding the extension to more than two states.

First, we recall that the aim of our work is to build a statistical test of market efficiency, in which we determine whether we can reject the null hypothesis corresponding to the EMH. If we reject the EMH in our simplified two-state approach with a given confidence (typically 99%, e.g.), we know that we are able to reject the EMH in a more realistic framework with an even greater confidence. In other words, using only two states tends to diminish the power of the statistical test but should not alter its significance level. Second, if we include more states than only two, it is more difficult to link entropy to the notion of market efficiency. Indeed, with only two states, it is clear that the efficient market corresponds to the uniform conditional distribution of the suffix ($\pi_i^L = 1/2$). With a greater number of states, the definition of the set of distributions consistent with the EMH is not as trivial and Shannon entropy may not be an appropriate quantity for discriminating efficient and inefficient markets. In particular, outside the two-state framework, a same level of entropy may correspond to several distributions, among which some but not all are consistent with the EMH. Third, given a prefix sequence, increasing the number of states will decrease the number of observations of each suffix and thus increase the variance of the estimator of the probability of each state. In other words, the statistical significance of the estimated market information may be reduced.

3. Statistical perspective

We now focus on the statistical properties of the concept of market information defined in section 2. We begin by introducing a simple estimator along with some of its properties. Then we determine the asymptotic distribution of this estimator and we use these elements to build a statistical test of market efficiency.

3.1. Estimator of the market information

The definition of the market information, as displayed in equation (3), relies on unobserved probabilities p_i^L and π_i^L. Replacing these theoretical probabilities by their empirical version leads to an estimator $\widehat{I}^{L+1}$ of the market information I^{L+1}. This plug-in estimator corresponds to the maximum-likelihood estimator. It converges almost surely, as $n \to +\infty$, toward the true market information (Verdú 2019). We note that it is the value of π_i^L, not the value of p_i^L, which makes it possible to conclude about the relevance of the EMH. We will

thus be particularly interested in the behaviour of the market information with respect to the estimated conditional probabilities $\widehat{\pi}_i^L$. Consequently, most of the results below will be stated conditionally on the probabilities p_i^L.

A first remark regarding $\widehat{I}^{L+1}$ is that this estimate also follows theorem 2.1 in which probabilities are replaced by empirical probabilities. In other words, we still have $\widehat{I}^{L+1} \geq 0$. But if the EMH holds, we may have $\widehat{I}^{L+1}$ *slightly* higher than 0 because, due to the statistical error, the estimated probabilities $\widehat{\pi}_i^L$ may be different from π_i^L, which are 1/2 in this case. Building a statistical test is thus mandatory for answering the question of the efficiency of the market. It will be the purpose of section 3.3.

Considering that the theoretical value of the market information is 0, according to the EMH, and that the estimator $\widehat{I}^{L+1}$ is nonnegative and may be different from zero even when the EMH holds, we conclude that $\widehat{I}^{L+1}$ is biased. We will see however in section 3.2.2 that the bias tends to zero when n tends to infinity, at least in the EMH case. We note that more advanced estimators of the entropy have been proposed in the literature (Gao *et al.* 2008, Verdú 2019), such as estimators based on the Lempel–Ziv compression algorithm (Kontoyiannis *et al.* 1998), on the context-tree weighting algorithm (Gao *et al.* 2008), or on Bayesian context trees (Papageorgiou and Kontoyiannis 2022). They are intended to bypass mainly two drawbacks of the plug-in estimator: the undersampling of the method when L is large and the bias.

We provide some insight on the distribution of the estimator $\widehat{I}^{L+1}$ with its exact moment-generating function, displayed in the following proposition. We focus on the case where the EMH holds, that is for $\pi_i^L = 1/2$ for all i, but for an empirical market information based on $\widehat{\pi}_i^L$ instead of π_i^L. The provided expression also assumes that we have access to the true probabilities of the prefixes: $\widehat{p}_i^L = p_i^L$.

PROPOSITION 3.1 *For $L \in \mathbb{N} \setminus \{0\}$, the moment-generating function of $\widehat{I}^{L+1}$, conditionally on the event $\mathcal{E} = \{\forall i \in [\![1, 2^L]\!], \widehat{p}_i^L = p_i^L, \pi_i^L = 1/2\}$, is:*

$$M_{\widehat{I}^{L+1}} : t \mapsto \mathbb{E}\left[e^{t\widehat{I}^{L+1}} \middle| \mathcal{E}\right] = e^t \prod_{i=1}^{2^L} \sum_{j=0}^{n_i} C_{i,j}^L(t),$$

for values of t for which this quantity is defined and where

$$C_{i,j}^L(t) = \binom{n_i}{j} \frac{1}{2^{n_i}} \left(\frac{j}{n_i}\right)^{t p_i^L j/n_i \ln(2)} \left(1 - \frac{j}{n_i}\right)^{t p_i^L (1-j/n_i)/\ln(2)}.$$

The proof of this proposition is postponed in appendix 2.

We can then deduce from proposition 3.1 the moments of the estimated market information, like in the following proposition, whose proof is in appendix 3.

PROPOSITION 3.2 *For $L \in \mathbb{N} \setminus \{0\}$ and $r \in \mathbb{N}$, the conditional moment of order r of the estimator of market information is*

$$\mathbb{E}\left[\left(\widehat{I}^{L+1}\right)^r \middle| \mathcal{E}\right] = \sum_{m=0}^r \binom{r}{m} \sum_{j_1=0}^{n_1} \cdots \sum_{j_{2^L}=0}^{n_{2^L}} \alpha_{j_1,\dots,j_{2^L}}^m \prod_{i=1}^{2^L} \binom{n_i}{j_i} \frac{1}{2^{n_i}},$$

where

$$\alpha_{j_1,\dots,j_{2^L}} = \sum_{k=1}^{2^L} p_k^L \left(\frac{j_k}{n_k} \log_2\left(\frac{j_k}{n_k}\right) + \left(1 - \frac{j_k}{n_k}\right) \right.$$
$$\left. \times \log_2\left(1 - \frac{j_k}{n_k}\right)\right).$$

The formula with the nested sum in proposition 3.2 leads to a slow computation of moments, in particular when L is big. We can however also provide a simpler expression of the moments of low order, for example

$$\mathbb{E}\left[\widehat{I}^{L+1} \middle| \mathcal{E}\right] = \sum_{i=1}^{2^L} \left[-p_i^L \log_2\left(\frac{p_i^L}{2}\right)\right.$$
$$+ p_i^L \sum_{j_i=0}^{n_i} \binom{n_i}{j_i} \frac{1}{2^{n_i}} \left(\frac{j_i}{n_i} \log_2\left(p_i^L \frac{j_i}{n_i}\right)\right.$$
$$\left.\left. + \left(1 - \frac{j_i}{n_i}\right) \log_2\left(p_i^L \left(1 - \frac{j_i}{n_i}\right)\right)\right)\right]$$
$$= 1 + \sum_{i=1}^{2^L} p_i^L 2^{1-n_i} \sum_{j_i=0}^{n_i} \binom{n_i}{j_i} \frac{j_i}{n_i} \log_2\left(\frac{j_i}{n_i}\right).$$

This expression relies on the direct calculation of the first moment, without using the moment-generating function. Nevertheless, the moment-generating function or any similar transform is useful to fully describe the probability distribution of the estimator of market efficiency. A more condensed asymptotic expression of such a transform would thus be helpful in the perspective of a practical application. It is the purpose of the next section.

3.2. Asymptotic analysis

We decompose the study of the asymptotic distribution of the empirical market information in two steps: first we focus on the summands, then on the market information itself.

3.2.1. Summands of the market information. We introduce the function g_j, defined, for $(t, x) \in \mathbb{R} \times (0, 1)$ and a given $j \in [\![1, 2^L]\!]$, by:

$$g_j(t, x) = \exp\left(it\left[p_j^L x \log_2(p_j^L x) + p_j^L(1-x)\log_2(p_j^L(1-x))\right.\right.$$
$$\left.\left. - p_j^L \log_2\left(\frac{p_j^L}{2}\right)\right]\right), \tag{4}$$

where i is the imaginary unit. As related to a summand of $\widehat{I}^{L+1}$, this functions appears in the characteristic function of the market information, if we assume the independence of $\widehat{\pi}_j$ with $\widehat{\pi}_i$, for $i \neq j$:

$$\varphi_{\widehat{I}^{L+1}} : t \in \mathbb{R} \mapsto \mathbb{E}\left[e^{it\widehat{I}^{L+1}} \middle| \mathcal{E}\right] = \prod_{j=1}^{2^L} \mathbb{E}\left[g_j(t, \widehat{\pi}_j)\right]. \tag{5}$$

Before giving an asymptotic expression for the characteristic function, we provide some useful properties on the function g_j, beginning by an expression of its derivatives.

PROPOSITION 3.3 *For $k \geq 1$ and $j \in [\![1, 2^L]\!]$, the kth derivative with respect to $x \in (0, 1)$ of the function g_j defined in equation* (4), *is*

$$\partial_x^k g_j(t, x) = g_j(t, x) \sum_{l=1}^{k} \left(\frac{itp_j^L}{\ln(2)} \right)^l B_{k,l} \left(\lambda(x), \frac{\mathrm{d}}{\mathrm{d}x} \lambda(x), \ldots, \right.$$
$$\left. \frac{\mathrm{d}^{k-l}}{\mathrm{d}x^{k-l}} \lambda(x) \right), \tag{6}$$

where $B_{k,l}$ is a Bell polynomial and $\lambda(x) = \ln(x) - \ln(1 - x)$. In particular, when $x = 1/2$, we have

$$\partial_x^k g_j \left(t, \frac{1}{2} \right) = 2^k \sum_{l=1}^{k} \left(\frac{itp_j^L}{\ln(2)} \right)^l B_{k,l} (0, 0!, 0, 2!, 0, 4!, 0, \ldots)$$
$$\tag{7}$$

and, in the particular case where k is odd, $\partial_x^k g_j(t, 1/2) = 0$.

The proof of this proposition is postponed in appendix 4.

It is also easy to see, from equation (4), that for all $y \in [0, 1/2)$, $g_j(t, y + 1/2) = g_j(t, -y + 1/2)$. As a consequence,

$$\partial_x^k g_j \left(t, y + \frac{1}{2} \right) = (-1)^k \partial_x^k g_j \left(t, -y + \frac{1}{2} \right). \tag{8}$$

So the value of the function $x \mapsto |\partial_x^k g_j(t, x)|$ evolves symmetrically with respect to $x = 1/2$. The next proposition gives some insight on the amplitude, understood as the L^q norm of the derivatives of g_j restricted to an interval whose left bound is 1/2. It can be easily extended to integrals on an interval centred in 1/2 thanks to equation (8).

PROPOSITION 3.4 *For $k \geq 1$, $j \in [\![1, 2^L]\!]$, $q \geq 1$, $z \in [1/2, 1)$, and $t \in \mathbb{R}$, the $L^q([1/2, z])$ norm of the function g_j, defined in equation* (4), *admits the following bound:*

$$\left(\int_{1/2}^{z} \left| \partial_x^k g_j(t, x) \right|^q \mathrm{d}x \right)^{1/q}$$
$$\leq \sum_{l=1}^{k} \left| \frac{tp_j^L}{\ln(2)} \right|^l \mathcal{L}(k, l) \left(\frac{2}{(r(k-l) + 1)r(k-l)} \right)^l$$
$$\times \frac{(1 - z)^{-lr(k-l) + 1/q}}{(qlr(k-l) - 1)^{1/q}},$$

where $\mathcal{L}(k, l)$ is a Lah number, equal to $\binom{k-1}{l-1} \frac{k!}{l!}$, and $r : \mathbb{N} \to \mathbb{N}$ is defined by:

$$r(k) = \max \left(5, 2 \left\lceil \frac{k-1}{2} \right\rceil + 1 \right). \tag{9}$$

The proof of this proposition is postponed in appendix 5.

The following theorem provides an approximation for $\mathbb{E}[g_j(t, X/n_j)]$, where X is a binomial variable of parameters n_j and 1/2, with an upper bound for the error. This approximation will then be useful for deriving the characteristic function of the market information, as defined in equation (5).
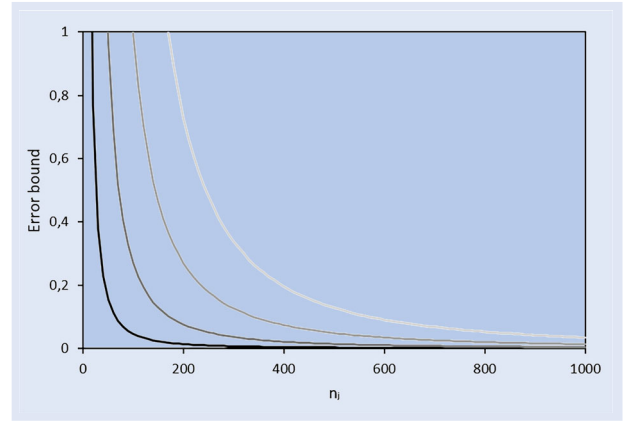

Figure 1. Assymptotic upper bound of $|R(t, n_j)|$ according to theorem 3.1, for several values of q, namely, from the bottom to the top: 2, 3, 4, 5. We have considered $p_j^L = 1/2$ and $t = 1$.

THEOREM 3.1 *For $j \in [\![1, 2^L]\!]$, $t \in \mathbb{R}$, g_j defined in equation* (4), *and $X \sim \mathcal{B}(n_j, 1/2)$, we have*

$$\mathbb{E} \left[g_j \left(t, \frac{X}{n_j} \right) \right] = 1 + \frac{itp_j^L}{2 \ln(2) n_j} + R(t, n_j),$$

where, for all $q \in \mathbb{N} \setminus \{0, 1\}$, we have for all $\varepsilon > 1$, the existence of $v \in \mathbb{N}$ such that, for all $n_j \geq v$,

$$|R(t, n_j)| \leq \frac{\varepsilon}{96} (q - 1)^{1-1/q} (4q - 1)^3$$
$$\times \left(\sum_{l=1}^{4} \left| \frac{2^5 tp_j^L}{15 \ln(2)} \right|^l \frac{\mathcal{L}(4, l)}{(5ql - 1)^{1/q}} \right) n_j^{-2+1/2q}.$$
$$\tag{10}$$

The proof of this theorem is postponed in appendix 6.

The upper bound of the error term in theorem 3.1 depends on two free parameters, ε and q. It is obvious that the lower possible value for ε, the tighter the bound. When $n_j \to \infty$, replacing ε by 1 in equation (10) provides an asymptotic expression of the upper bound of $|R(t, n_j)|$. Regarding the selection of q, it is directly related to the convergence speed with respect to n_j because of the term $n_j^{-2+1/2q}$. On the one hand, a higher q increases the convergence rate of the error term with respect to n_j, on the other hand it also increases the constant term in the formula of the upper bound. Figure 1 displays a numerical evaluation of the asymptotic bound of theorem 3.1 for several possible values for q. It finally suggests that the lower q, the lower the bound, regardless of n_j. With $q = 2$, the error term is asymptotically $\mathcal{O}(n_j^{-1.75})$ which is to be compared to the n_j^{-1} appearing in the approximation of $\mathbb{E}[g_j(t, X/n_j)]$ in theorem 3.1.

3.2.2. Asymptotic distribution of the market information estimator. We are now interested in the characteristic function of the empirical market information, conditionally on $\mathcal{E}$. If we assume again the independence of $\hat{\pi}_j$ with $\hat{\pi}_i$, for $i \neq j$, according to equation (5), and using theorem 3.1, we have,
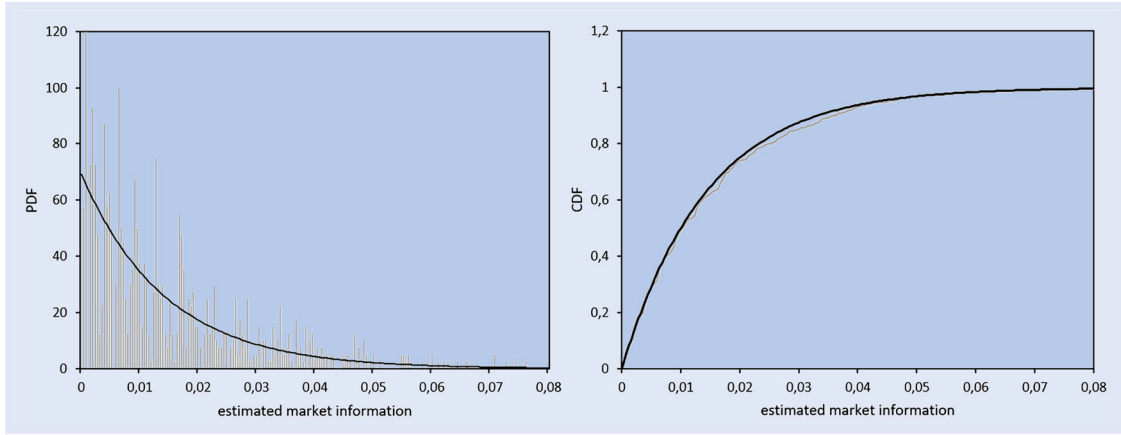
Figure 2. Asymptotic $\Gamma(2^{L-1}, 1/\ln(2)n)$ (black) and simulated (on 1000 trajectories, grey) distributions of the estimated market information $\widehat{I}^2$ for $n = 100$.
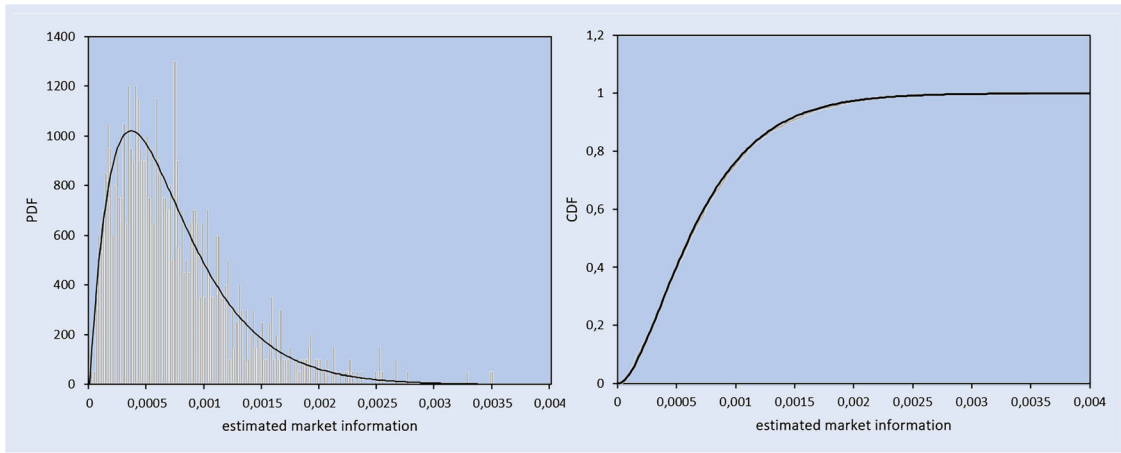

Figure 3. Asymptotic $\Gamma(2^{L-1}, 1/\ln(2)n)$ (black) and simulated (on 1000 trajectories, grey) distributions of the estimated market information $\widehat{I}^3$ for $n = 4000$.

for $t \in \mathbb{R}$ and $p_j^L = n_j/n > 0$ whatever j,

$$\varphi_{\widehat{I}^{L+1}}(t) = \prod_{j=1}^{2^L} \mathbb{E}\left[g_j(t, \widehat{\pi}_j)\right]$$

$$\overset{n\to\infty}{\sim} \prod_{j=1}^{2^L}\left(1 + \frac{it}{2\ln(2)n}\right)$$

$$\overset{n\to\infty}{\sim} \left(1 - \frac{it}{\ln(2)n}\right)^{-2^{L-1}}. \tag{11}$$

We recognize the characteristic function of the gamma distribution $\Gamma(k, \theta)$ of shape parameter $k = 2^{L-1}$ and scale parameter $\theta = 1/\ln(2)n$. We note that the result does not depend on the specific value of each p_j^L, but only on n and L.

We underline that a related literature on transfer entropy concludes that the information, defined as a difference of conditional entropies, follows asymptotically a chi-square distribution, using Wilks' theorem (Barnett and Bossomaier 2012, Kontoyiannis and Skoularidou 2016, Bongiorno and Challet 2022). While the framework and definition of information are slightly different from ours, we note that we would also obtain such an asymptotic chi-square distribution if we changed the scale of our market information, in

particular by replacing the base 2 of the logarithm by a base e in equations (1) and (2). Moreover, beyond the asymptotic distribution, we have also contributed to give some insight into the way the error of the approximation behaves, thanks to theorem 3.1. The chi-square asymptotic result echoes a more ancient literature about the asymptotic distribution of Shannon entropy, limited to non-conditional distributions, after which the limit distribution is Gaussian if the probability of the states is not uniform and chi-square if it is uniform (Zubkov 1974, Harris 1977, Regnault 2011).

Figures 2 and 3 confirm with the help of simulations the relevance of the gamma distribution, with the parameters $k = 2^{L-1}$ and $\theta = 1/\ln(2)n$, as the asymptotic distribution of the empirical market information under the hypothesis of an efficient market, for two values of L. However, this confirmation is only visual. In order to study more thoroughly the accuracy of our gamma approximation, we conduct a statistical test to assess whether the very slight difference one can see between the simulated and asymptotic cumulative distribution functions in figures 2 and 3 is significant or not, depending on the value of n. We answer this question with a Kolmogorov–Smirnov test, whose conclusion, for $L = 1$, is that the asymptotic distribution is valid for $n \gtrsim 100$, as displayed in figure 4.
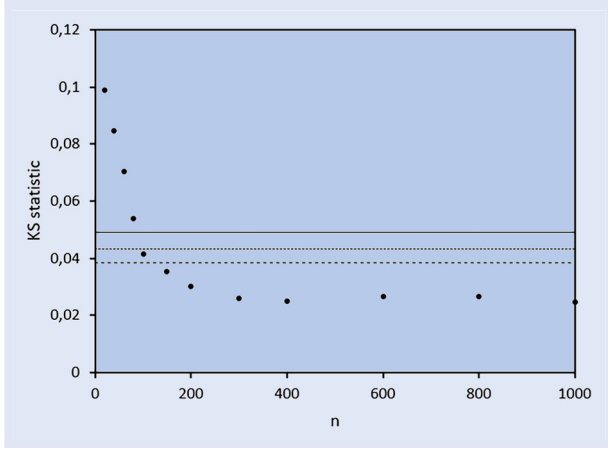
Figure 4. Kolmogorov–Smirnov statistic between the asymptotic and the simulated (on 1000 trajectories) distributions of the estimated market information $\widehat{I}^2$ for various values of n. The three lines indicate the values of the statistics leading to a p-value of 5%, 1%, and 0.1% (from bottom to top).

Figure 5. Information corresponding to a *p*-value of 5%, 1%, and 0.1% (from bottom to top) in the test with a market information $I^2 = 0$ as a null hypothesis, for various values of n.

3.3. Statistical test of market efficiency

The purpose of our work is to statistically test the existence of market efficiency using the empirical market information as a statistic. More precisely, given $L \in \mathbb{N}$, we define the null hypothesis by $\forall i \in [\![1, 2^L]\!] \setminus \{j | p_j^L = 0\}, \pi_i^L = 1/2$. The alternative hypothesis thus corresponds to the existence of an i such that $\pi_i^L \neq 1/2$. One could indeed use this imbalance to do statistical arbitrages. As already exposed, the true definition of market efficiency is more restrictive than the formalism we chose in our null hypothesis. Nevertheless, we think that this approach is interesting insofar as rejecting our null hypothesis for a given confidence leads to rejecting the market efficiency with an even higher confidence.

Thanks to theorem 2.1, the null hypothesis is equivalent to having the market information I^{L+1} equal to zero. We evaluate this quantity with the statistic $\widehat{I}^{L+1}$. Unfortunately, under the null hypothesis, there is a bias in $\widehat{I}^{L+1}$, which asymptotically disappears according to equation (11). We have also provided in section 3.2.2 an asymptotic gamma distribution for $\widehat{I}^{L+1}$ under the null hypothesis, which we can use to quantify the p-value corresponding to the obtained test statistic and determine whether we are able or not to reject the EMH. Figure 5 reports the link between the value of the test statistic, its p-value, and n. We note that an alternative to our approach has previously been published, in which the confidence intervals were estimated by simulations instead of our straightforward asymptotic distribution (Shternshis *et al.* 2022).

In order to get an insight into the power of the test, we consider, in the case where $L = 1$, the alternative hypothesis $(\pi_1^L, \pi_2^L) = (1/2 - \zeta, 1/2 + \zeta)$. The quantity $\zeta \in (0, 1/2]$ is the excess conditional probability. This model makes it possible to define the time series X_i as a stationary Markov chain with $p_1^L = p_2^L = 1/2$.

Figure 6 illustrates the power of the test. It is obtained by the mean of 10 000 simulated Markov chains. If $n = 400$, the statistical test of confidence level 95% (respectively 99% and 99.9%) reaches a power of 95% as soon as ζ is higher than 0.10 (resp. 0.11 and 0.13). In other words, for the statistical
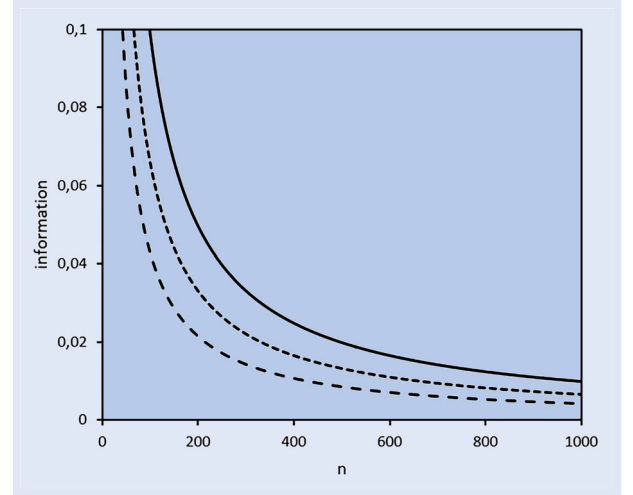
test of confidence 95% and $n = 400$, if $\pi_2^L = 0.60$, the null hypothesis of EMH is properly rejected with a probability of 95%; but if $\pi_2^L = 0.55$, which is a situation much closer to the EMH, the null hypothesis is properly rejected with a probability of only 40%. Of course, if one increases the size of the sample, then one can reach the same power with a lower excess conditional probability ζ. For instance, when $n = 1000$, the test of confidence level 95% reaches the power of 95% for every $\zeta > 0.06$.

4. Empirical study

We now apply the method introduced above to real financial data. We focus on several years of daily data for time series of prices of three kinds of assets: three stock indices, eight standalone stocks, and four cryptocurrencies. All the time series were imported from Yahoo Finance. The starting date in our study depends on the asset considered and on the availability of the data, but the end date is the same for all the assets: end December 2022.

Using the estimator $\widehat{I}^{L+1}$ introduced in section 3.1, where the empirical estimator $\widehat{\pi}_i^L$ is based on all overlapping blocks of $L + 1$ consecutive symbols, we evaluate the market information for each of these series, using a rolling window of n consecutive daily price returns. We focus on a particular setting with $L = 1$ and various values of n, depending of the asset class.

First, we consider three stock indices: the CAC 40 index, representing the 40 most significant stocks on the French market, the Russell 3000 index, a wide index of 3000 stocks and thus representing almost the entire U.S. market, and the Russell 2000 index, a small-cap index comprising the 2000 smallest constituents of the Russell 3000 index. We address the sensitiveness of the results to the size of the rolling window by selecting various values of n, corresponding to 6 months, 1 year, 2 years, and 3 years of data. We display the estimated market information between July 1991 and end December 2022 in figure 7.
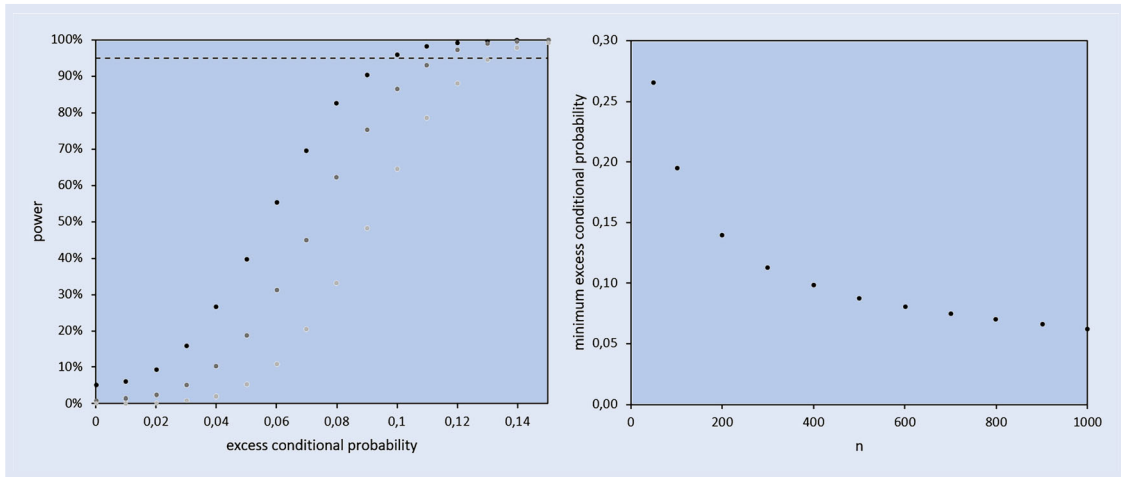
Figure 6. Left: Power of the test of confidence 95%, 99%, and 99.9% (from the darker to the lighter) with respect to the excess conditional probability ζ, for $n = 400$. Right: Minimal value of ζ such that the test of confidence level 95% reaches a power of 95%, with respect to n.
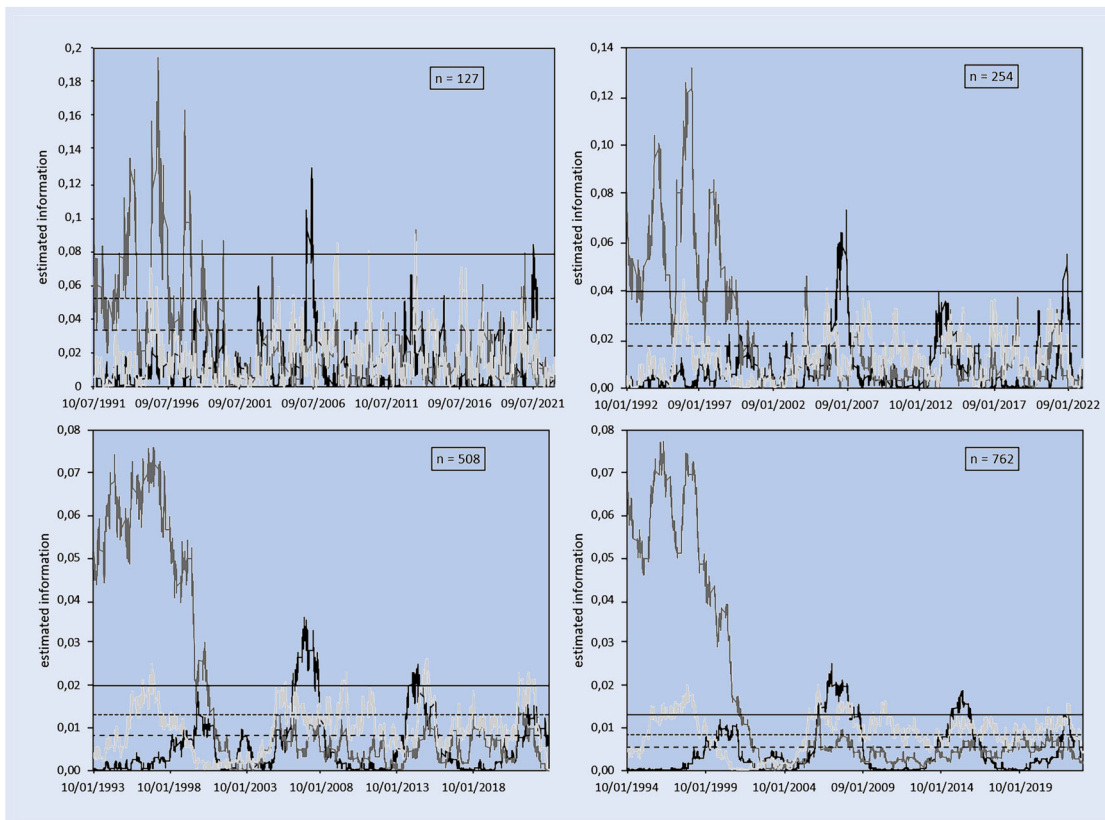

Figure 7. Estimated information $\widehat{I}^2$ in a rolling window of $n \in \{127, 254, 508, 762\}$ business days, for the CAC 40 index (black), Russell 2000 index (dark grey), Russell 3000 index (light grey). The horizontal lines are the confidence intervals of the statistical test of absence of information, with probabilities of 95%, 99%, and 99.9%.

We observe in figure 7 different behaviours for the three stock indices. Before 2002, Russell 2000 index, which is made up of less liquid stocks, has largely the highest information among the three indices. Between 2000 and 2004, Russell 3000 index has a very low market information but, since 2005, it has a market information which is very often higher than the one of Russell 2000. Besides the gradual interpretation of the market information, we are interested in applying the statistical test of market efficiency exposed in section 3.3 and which states that the information of Russell 3000 index is almost always significantly different from 0 since 2005, with a 95% confidence level. Regarding the CAC 40 index, its market information is fluctuating, alternating periods of efficiency with periods of inefficiency (1999–2000, 2006–2008, 2014–2015, 2021–2022). This economic cycle is particularly apparent for $n = 508$ or $n = 762$. As a complement to figure 7, we display in table 1 the proportion of days for which we reject the null hypothesis. In this multiple testing of the EMH, we also implement the Bonferroni correction, which provides a more conservative result. Though the null hypothesis of market efficiency is quite often rejected for the three stock indices whatever the value of n, only the Russell 2000 index

Table 1. Proportion of days for which the null hypothesis of market efficiency is rejected for stock indices, with confidence levels of 95%, 99%, and 99.9%, with and without the Bonferroni correction.

n	Stock index	Without correction and a confidence of			Bonferroni correction and a confidence of		
		95%	99%	99.9%	95%	99%	99.9%
127	CAC 40	8.3%	3.1%	1.0%	0.0%	0.0%	0.0%
	Russell 3000	10.7%	2.9%	0.3%	0.0%	0.0%	0.0%
	Russell 2000	25.9%	14.1%	6.9%	1.5%	0.9%	0.1%
254	CAC 40	13.5%	7.7%	2.7%	0.1%	0.0%	0.0%
	Russell 3000	26.9%	7.6%	0.2%	0.0%	0.0%	0.0%
	Russell 2000	35.5%	26.2%	18.4%	7.7%	4.9%	3.0%
508	CAC 40	23.6%	13.8%	7.3%	0.1%	0.0%	0.0%
	Russell 3000	57.4%	26.8%	4.4%	0.0%	0.0%	0.0%
	Russell 2000	39.3%	26.9%	24.9%	22.2%	21.9%	20.3%
762	CAC 40	31.6%	19.9%	12.0%	0.3%	0.0%	0.0%
	Russell 3000	75.5%	53.2%	15.9%	0.0%	0.0%	0.0%
	Russell 2000	45.7%	29.0%	25.0%	23.1%	22.7%	22.6%

has a significant proportion of days for which the market is inefficient if we consider the Bonferroni correction.

Our results confirm and supplement previous studies, which show that large-cap U.S. stock indices are very often efficient, except during financial crisis periods (Risso 2008, Giglio *et al.* 2008, Alvarez-Ramirez and Rodriguez 2021), and that the CAC 40 index, and even more so the small-cap U.S. index Russell 2000, are less efficient than large-cap U.S. stock indices (Giglio *et al.* 2008, Castura *et al.* 2010).

We now consider eight individual stocks, both in the French and in the U.S. markets, and a value of n corresponding to a one-year window. The dataset consists of stocks of various sectors and market capitalizations. We analyse four French stocks with a time series starting in 2000: L'Air liquide (AI.PA), Société Générale (GLE.PA), LVMH (MC.PA), and Michelin (ML.PA). We analyse four U.S. stocks: Apple (AAPL) since 1981, JPMorgan Chase & Co. (JPM) since 1984, Pfizer (PFE) since 1973, and Perficient (PRFT) since 2000. The last two stocks are in the same pharmaceutical sector but Perficient belongs to Russell 2000 whereas Pfizer has a much larger capitalization. Figure 8 shows the evolution of market information for these eight stocks.

We observe very high levels of market information at the beginning of the sample for Pfizer and Perficient, rejecting the EMH with a very high confidence. This is confirmed by table 2. For other stocks, we obtain an illustration of the economic cycle, with phases of efficiency and phases of inefficiency which are not synchronized between the stocks.

Last, we study four cryptocurrencies: Bitcoin (BTC-USD), Ethereum (ETH-USD), Ripple (XRP-USD), and Tezos (XTZ-USD). We have data since 2015 for Bitcoin and since 2018 for the three other cryptocurrencies. We display in figure 9 the evolution of their information through time for three values of n corresponding to 6 months, 1 year, and 2 years. Since the cryptocurrencies are traded every day, n is bigger for cryptocurrencies than for stocks for a same window duration of say 1 year.

Bitcoin and Ethereum have a reputation for behaving closely against each other. In particular, a recent study shows that they belong to a same cluster of market efficiency (El

Table 2. Proportion of days for which the null hypothesis of market efficiency is rejected for stocks, with confidence levels of 95%, 99%, and 99.9%, $n = 254$, with and without the Bonferroni correction.

Stock	Without correction and a confidence of			Bonferroni correction and a confidence of		
	95%	99%	99.9%	95%	99%	99.9%
AI.PA	19.4%	6.3%	0.8%	0.0%	0.0%	0.0%
GLE.PA	9.1%	1.9%	0.0%	0.0%	0.0%	0.0%
MC.PA	6.5%	1.7%	0.2%	0.0%	0.0%	0.0%
ML.PA	2.4%	0.1%	0.0%	0.0%	0.0%	0.0%
AAPL	26.1%	10.5%	2.0%	0.0%	0.0%	0.0%
JPM	19.2%	8.2%	1.7%	0.2%	0.0%	0.0%
PFE	25.6%	20.9%	18.3%	14.2%	12.3%	10.0%
PRFT	22.7%	17.2%	12.6%	10.7%	9.4%	7.7%

Montasser *et al.* 2022). Figure 9 confirms this behaviour with highly synchronized time series of information between Bitcoin and Ethereum. We also observe a very large information for Bitcoin in 2017 and 2018, outside confidence intervals of the EMH. Since then, we observe an alternation of periods of efficiency with periods of inefficiency. The information time series of Ripple and Tezos seem incommensurate with the ones of Bitcoin and Ethereum. Figure 9 and table 3 also stress the very low information of Tezos, for which we cannot reject the EMH.

Our results are to be compared to previous studies. It is difficult to find a consensus in the literature about the market efficiency of cryptocurrencies. For instance, in the period 2010–2016, Urquhart (2016) concludes that Bitcoin is inefficient whereas Nadarajah and Chu (2017) obtain the opposite conclusion by considering a power transformation of the price returns. Regarding more recent study periods up to 2018, several articles have stressed that, despite a certain market inefficiency, several cryptocurrencies became progressively more efficient shortly before 2018 (Drozdz *et al.* 2018, Kristoufek and Vosvrda 2019). We also observe a decrease of the market information of Bitcoin at the beginning of 2018 in figure 9, confirming the findings of these papers.
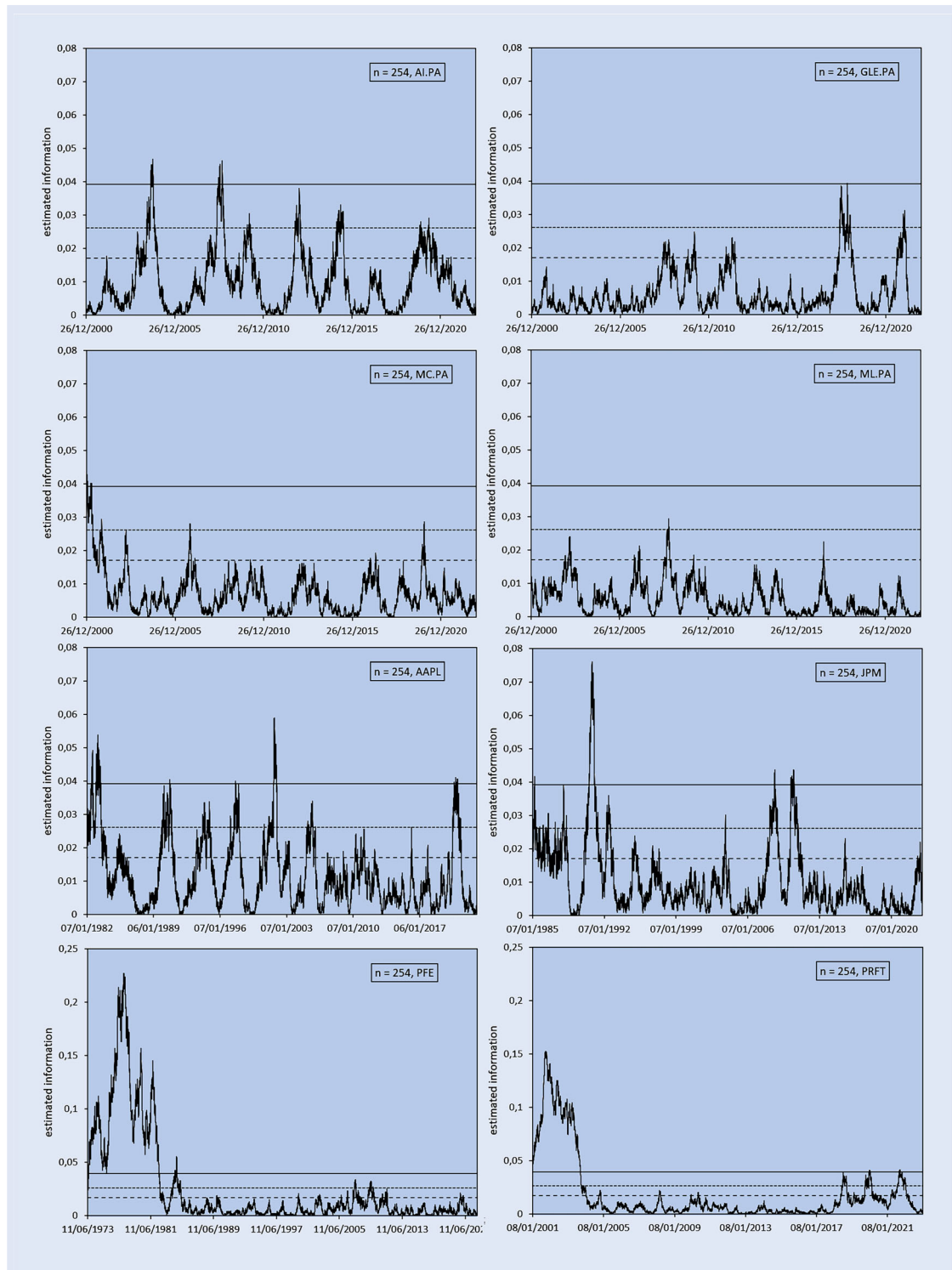
Figure 8. Estimated information $\widehat{I}^2$ in a rolling window of 254 business days, for eight stocks. The horizontal lines are the confidence intervals of the statistical test of absence of information, with probabilities of 95%, 99%, and 99.9%.

5. Conclusion

Determining whether the EMH holds or not for a particular financial asset is of major importance in the asset management industry. It can indeed help market agents in their investment decisions. In this article, we have proposed a statistical test of market efficiency which may be used in this practical perspective. This statistical test is based on a market information estimator relying on Shannon entropy and a symbolic representation of a series of successive price returns. Applying this tool to real financial data shows a diversity among financial markets regarding their efficiency. Future research in this field could focus on extending the statistical test introduced above to other information statistics, such as the permutation entropy (Bandt and Pompe 2002).
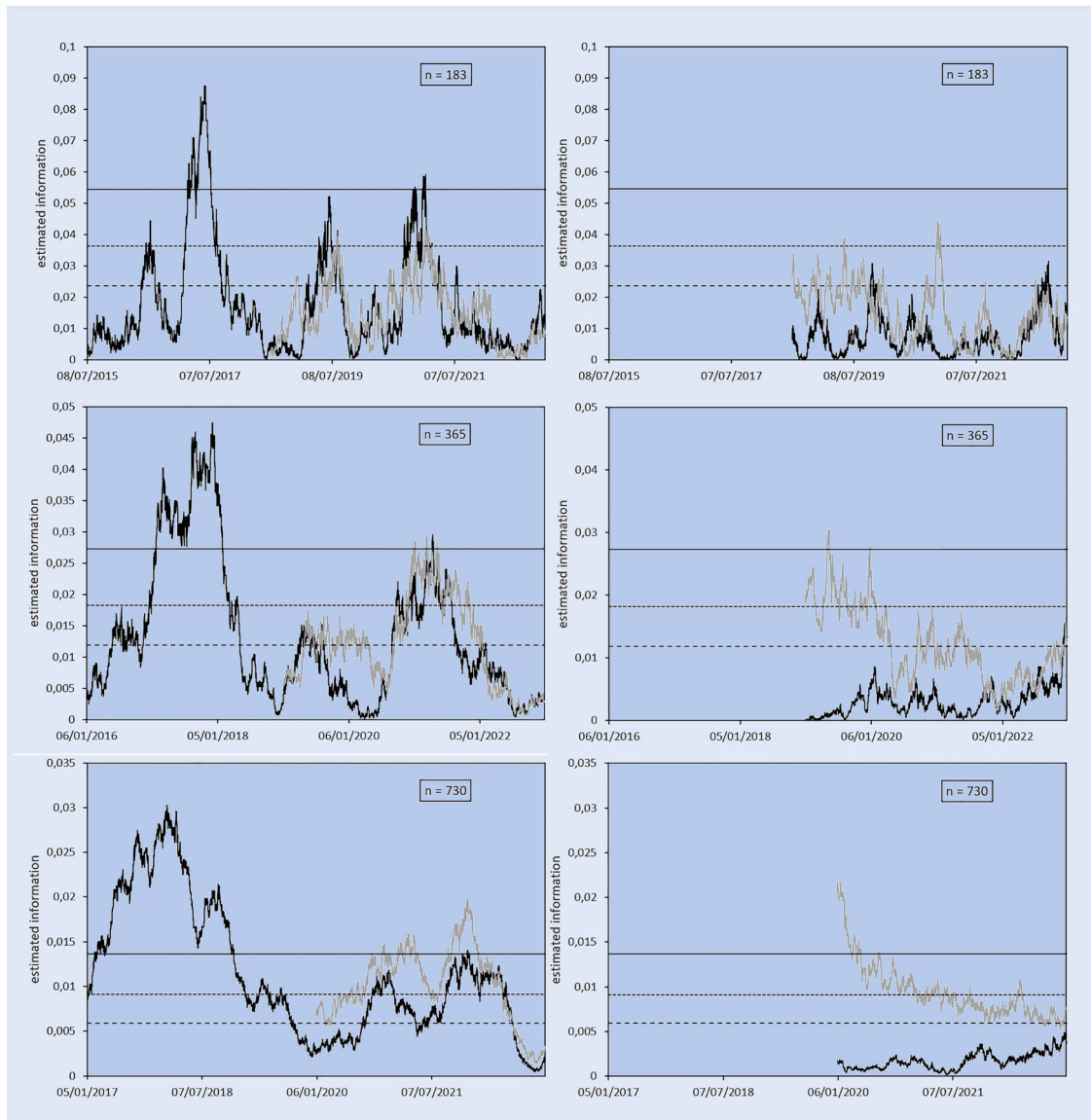
Figure 9. Estimated information $\widehat{I}^2$ in a rolling window of $n \in \{183, 365, 730\}$ days, for Bitcoin (left, black), Ethereum (left, grey), Ripple (right, grey), and Tezos (right, black). The horizontal lines are the confidence intervals of the statistical test of absence of information, with probabilities of 95%, 99%, and 99.9%.

Table 3. Proportion of days for which the null hypothesis of market efficiency is rejected for cryptocurrencies, with confidence levels of 95%, 99%, and 99.9%, with and without the Bonferroni correction.

n	Cryptocurrency	Without correction and a confidence of			Bonferroni correction and a confidence of		
		95%	99%	99.9%	95%	99%	99.9%
183	BTC-USD	26.6%	13.0%	5.1%	0.1%	0.0%	0.0%
	ETH-USD	19.0%	1.6%	0.0%	0.0%	0.0%	0.0%
	XRP-USD	15.0%	1.3%	0.0%	0.0%	0.0%	0.0%
	XTZ-USD	2.6%	0.0%	0.0%	0.0%	0.0%	0.0%
365	BTC-USD	44.3%	26.1%	14.7%	1.6%	0.0%	0.0%
	ETH-USD	51.5%	20.5%	0.5%	0.0%	0.0%	0.0%
	XRP-USD	43.9%	18.6%	0.9%	0.0%	0.0%	0.0%
	XTZ-USD	1.6%	0.0%	0.0%	0.0%	0.0%	0.0%
730	BTC-USD	75.3%	53.9%	30.1%	15.1%	8.6%	0.4%
	ETH-USD	85.1%	61.4%	19.1%	0.1%	0.0%	0.0%
	XRP-USD	97.3%	44.1%	9.4%	0.0%	0.0%	0.0%
	XTZ-USD	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Acknowledgments

Disclosure statement

ORCID

Matthieu Garcin http://orcid.org/0000-0003-3296-6486

References

Alvarez-Ramirez, J. and Rodriguez, E., A singular value decomposition entropy approach for testing stock market efficiency. *Physica A*, 2021, **583**, 126337.

Ammy-Driss, A. and Garcin, M., Efficiency of the financial markets during the COVID-19 crisis: Time-varying parameters of fractional stable dynamics. *Physica A*, 2023, **609**, 128335.

Arias López, R. and Garrido, J., Some properties and inequalities related to the *k*th inverse moment of a positive binomial variate. *Rev. Mat.: Teoría Apl.*, 2001, **8**(2), 1–18.

Ayache, A. and Lévy Véhel, J., The generalized multifractional Brownian motion. *Stat. Inference Stoch. Process.*, 2000, **3**(1–2), 7–18.

Ayache, A. and Taqqu, M.S., Multifractional processes with random exponent. *Pub. Mat.*, 2005, **49**(2), 459–486.

Bandt, C. and Pompe, B., Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.*, 2002, **88**, 174102.

Bariviera, A.F., Font-Ferrer, A., Sorrosal-Forradellas, M.T. and Rosso, O.A., An information theory perspective on the informational efficiency of gold price. *N. Am. J. Econ. Finance*, 2019, **50**, 101018.

Barnett, L. and Bossomaier, T., Transfer entropy as a log-likelihood ratio. *Phys. Rev. Lett.*, 2012, **109**(13), 138105.

Benassi, A., Jaffard, S. and Roux, D., Elliptic Gaussian random processes. *Rev. Mat. Iberoam.*, 1997, **13**(1), 19–90.

Bianchi, S. and Pianese, A., Time-varying Hurst–Hölder exponents and the dynamics of (in)efficiency in stock markets. *Chaos Solitons Fract.*, 2018, **109**, 64–75.

Bianchi, S., Pantanella, A. and Pianese, A., Modeling and simulation of currency exchange rates using multifractional process with random exponent. *Int. J. Model. Optim.*, 2012, **2**(3), 309–314.

Bielinskyi, A., Serdyuk, O., Semerikov, S. and Soloviev, V., Econophysics of cryptocurrency crashes: An overview. *SHS Web Conf.*, 2021, **107**, 03001.

Bongiorno, C. and Challet, D., Statistical inference of lead-lag at various timescales between asynchronous time series from *p*-values of transfer entropy. Preprint, 2022.

Calcagnile, L.M., Corsi, F. and Marmi, S., Entropy and efficiency of the ETF market. *Comput. Econ.*, 2020, **55**(1), 143–184.

Castura, J., Litzenberger, R., Gorelick, R. and Dwivedi, Y., Market efficiency and microstructure evolution in U.S. equity markets: A high-frequency perspective. Working paper, 2010.

Drozdz, S., Gebarowski, R., Minati, L., Oswiecimka, P. and Watorek, M., Bitcoin market route to maturity? Evidence from return fluctuations, temporal correlations and multiscaling effects. *Chaos Interdiscip. J. Nonlinear Sci.*, 2018, **28**(7), 071101.

Ducournau, G., Symbol dynamics, information theory and complexity of economic time series. Preprint, 2021.

El Montasser, G., Charfeddine, L. and Benhamed, A., COVID-19, cryptocurrencies bubbles and digital market efficiency: Sensitivity and similarity analysis. *Finance Res. Lett.*, 2022, **46**, 102362.

Espinosa-Paredes, G., Rodriguez, E. and Alvarez-Ramirez, J., A singular value decomposition entropy approach to assess the impact of Covid-19 on the informational efficiency of the WTI crude oil market. *Chaos Solitons Fract.*, 2022, **160**, 112238.

Fama, E.F., Efficient capital markets: A review of theory and empirical work. *J. Finance*, 1970, **25**(2), 383–417.

Flandrin, P., Borgnat, P. and Amblard, P.-O., From stationarity to self-similarity, and back: Variations on the Lamperti transformation. In *Processes with Long-Range Correlations*, pp. 88–117, 2003 (Springer: Berlin-Heidelberg).

Gao, Y., Kontoyiannis, I. and Bienenstock, E., Estimating the entropy of binary time series: Methodology, some theory and a simulation study. *Entropy*, 2008, **10**(2), 71–99.

Garcin, M., Estimation of time-dependent Hurst exponents with variational smoothing and application to forecasting foreign exchange rates. *Physica A*, 2017, **483**, 462–479.

Garcin, M., Hurst exponents and delampertized fractional Brownian motions. *Int. J. Theor. Appl. Finance*, 2019, **22**(5), 1950024.

Garcin, M., Fractal analysis of the multifractality of foreign exchange rates. *Math. Methods Econ. Finance*, 2020, **13-14**(1), 49–73.

Garcin, M., A comparison of maximum likelihood and absolute moments for the estimation of Hurst exponents in a stationary framework. *Commun. Nonlinear Sci. Numer. Simul.*, 2022a, **114**, 106610.

Garcin, M., Forecasting with fractional Brownian motion: A financial perspective. *Quant. Finance*, 2022b, **22**(8), 1495–1512.

Giglio, R., Matsushita, R., Figueiredo, A., Gleria, I. and Da Silva, S., Algorithmic complexity theory and the relative efficiency of financial markets. *Europhys. Lett.*, 2008, **84**(4), 48005.

Guasoni, P., Nika, Z. and Rásonyi, M., Trading fractional Brownian motion. *SIAM J. Financ. Math.*, 2019, **10**(3), 769–789.

Guasoni, P., Mishura, Y. and Rásonyi, M., High-frequency trading with fractional Brownian motion. *Finance Stoch.*, 2021, **25**, 277–310.

Harris, B., The statistical estimation of entropy in the non-parametric case. *Colloq. Math. Soc. János Bolyai*, 1977, **16**, 323–355.

Kontoyiannis, I. and Skoularidou, M., Estimating the directed information and testing for causality. *IEEE Trans. Inf. Theory*, 2016, **62**(11), 6053–6067.

Kontoyiannis, I., Algoet, P.H., Suhov, Y.M. and Wyner, A.J., Non-parametric entropy estimation for stationary processes and random fields, with applications to English text. *IEEE Trans. Inf. Theory*, 1998, **44**(3), 1319–1327.

Kristoufek, L. and Vosvrda, M., Commodity futures and market efficiency. *Energy Econ.*, 2014, **42**, 50–57.

Kristoufek, L. and Vosvrda, M., Cryptocurrencies market efficiency ranking: Not so straightforward. *Physica A*, 2019, **531**, 120853.

Lahmiri, S., Bekiros, S. and Avdoulas, C., Time-dependent complexity measurement of causality in international equity markets: A spatial approach. *Chaos Solitons Fract.*, 2018, **116**, 215–219.

Mandelbrot, B. and Van Ness, J., Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 1968, **10**(4), 422–437.

Mensi, W., Beljid, M. and Managi, S., Structural breaks and the time-varying levels of weak-form efficiency in crude oil markets: Evidence from the Hurst exponent and Shannon entropy methods. *Int. Econ.*, 2014, **140**, 89–106.

Nadarajah, S. and Chu, J., On the inefficiency of Bitcoin. *Econ. Lett.*, 2017, **150**, 6–9.

Nuzman, C.J. and Poor, H.V., Linear estimation of self-similar processes via Lamperti's transformation. *J. Appl. Probab.*, 2000, **37**(2), 429–452.

Papageorgiou, I. and Kontoyiannis, I., Posterior representations for Bayesian context trees: Sampling, estimation and convergence. Preprint, 2022.

Peltier, R.F. and Lévy Véhel, J., A new method for estimating the parameter of fractional Brownian motion. Technical report 2396, INRIA, 1994.

Pincus, S. and Kalman, R.E., Irregularity, volatility, risk, and financial market time series. *Proc. Natl. Acad. Sci.*, 2004, **101**(38), 13709–13714.

Pincus, S.M., Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci.*, 1991, **88**(6), 2297–2301.

Regnault, P., Estimation using plug-in of the stationary distribution and Shannon entropy of continuous time Markov processes. *J. Stat. Plan. Inference*, 2011, **141**(8), 2711–2725.

Risso, W.A., The informational efficiency and the financial crashes. *Res. Int. Bus. Finance*, 2008, **22**(3), 396–408.

Shmilovici, A., Kahiri, Y., Ben-Gal, I. and Hauser, S., Measuring the efficiency of the intraday forex market with a universal data compression algorithm. *Comput. Econ.*, 2009, **33**(2), 131–154.

Shternshis, A., Mazzarisi, P. and Marmi, S., Measuring market efficiency: The Shannon entropy of high-frequency financial time series. *Chaos Solitons Fract.*, 2022, **162**, 112403.

Skorski, M., Handy formulas for binomial moments. Preprint, 2020.

Soloviev, V., Semerikov, S. and Solovieva, V., Lempel–Ziv complexity and crises of cryptocurrency market. In *III International Scientific Congress Society of Ambient Intelligence 2020*, pp. 299–306, March 2020 (Atlantis Press).

Stoev, S. and Taqqu, M.S., Stochastic properties of the linear multifractional stable motion. *Adv. Appl. Probab.*, 2004, **36**(4), 1085–1115.

Stosic, D., Stosic, D., Ludermir, T., de Oliveira, W. and Stosic, T., Foreign exchange rate entropy evolution during financial crisis. *Physica A*, 2016, **449**, 233–239.

Urquhart, A., The inefficiency of Bitcoin. *Econ. Lett.*, 2016, **148**, 80–82.

Verdú, S., Empirical estimation of information measures: A literature guide. *Entropy*, 2019, **21**(8), 720.

Weron, A., Burnecki, K., Mercik, S. and Weron, K., Complete description of all self-similar models driven by Lévy stable noise. *Phys. Rev. E.*, 2005, **71**(1), 016113.

Zubkov, A.M., Limit distributions for a statistical estimate of the entropy. *Theory Prob. Its Appl.*, 1974, **18**(3), 611–618.

Zunino, L., Tabak, B., Serinaldi, F., Zanin, M., Pérez, D.G. and Rosso, O.A., Commodity predictability analysis with a permutation information theory approach. *Physica A*, 2011, **390**, 876–890.

Zunino, L., Bariviera, A.F., Belén Guercio, M., Martinez, L.B. and Rosso, O.A., Monitoring the information efficiency of European coporate bond markets with dynamical permutation min-entropy. *Physica A*, 2016, **456**, 1–9.

Appendices

Appendix 1. Proof of theorem 2.1

Proof Let the function f be defined by $f : (x, y) \in (0, 1] \times (0, 1) \mapsto xy \log_2(xy) + x(1 - y) \log_2(x(1 - y)) - x \log_2(x/2)$. We extend f by continuity for $y \in [0, 1]$ noting that $f(x, 0) = f(x, 1) = x$. Obtaining some of its derivatives for $x \in (0, 1]$ and $y \in (0, 1)$ is straightforward:

$$\begin{cases} \dfrac{\partial f}{\partial x}(x, y) = y \log_2(xy) + (1 - y) \log_2(x(1 - y)) - \log_2(x/2) \\[2mm] \dfrac{\partial^2 f}{\partial y \partial x}(x, y) = \log_2\left(\dfrac{y}{1 - y}\right). \end{cases}$$
(A1)

From these expressions, we note that f, $\partial f / \partial x$, and $\partial^2 f / \partial y \partial x$ are all equal to zero when evaluated in $y = 1/2$, regardless of the value of $x \in (0, 1]$. Moreover, when $y > 1/2$ (respectively $< 1/2$), we have $y/(1 - y) > 1$ (resp. < 1) and thus $\partial^2 f / \partial y \partial x$ is positive

for $y \in (1/2, 1)$ and negative for $y \in (0, 1/2)$. Its primitive $y \mapsto \partial f / \partial x(x, y)$ is thus strictly decreasing in $(0, 1/2)$ and strictly increasing in $(1/2, 1)$, with a value of 0 in $1/2$. Consequently, $\partial f / \partial x$ is nonnegative for $y \in (0, 1)$, and $1/2$ is the only value of y for which it is equal to zero. From equation (A1), we observe that, for all $(x, y) \in (0, 1] \times (0, 1)$,

$$f(x, y) = x \frac{\partial f}{\partial x}(x, y).$$

Using in addition the extension of f to $y \in [0, 1]$, we can thus conclude that f is nonnegative for $y \in [0, 1]$, and $1/2$ is the only value of y for which it is equal to zero, regardless of $x \in (0, 1]$. Going back to the market information formula, we note that

$$I^{L+1} = \sum_{i=1}^{2^L} f\left(p_i^L, \pi_i^L\right).$$

Each element of this sum is positive, as soon as $(p_i^L, \pi_i^L) \in (0, 1] \times [0, 1]$, with a value of zero if and only if $\pi_i^L = 1/2$. If one probability p_i^L is equal to zero, the ith element of the sum is equal to zero, because of the convention used, namely $0 \log_2(0) = 0$. This leads to the conclusion of theorem 2.1. ∎

Appendix 2. Proof of proposition 3.1

Proof By definition

$$M_{\widehat{I}^{L+1}}(t) = e^{tH_\star^{L+1}} \prod_{i=1}^{2^L} \mathbb{E}\left[\exp\left(t\left[p_i^L \widehat{\pi}_i^L \log_2\left(p_i^L \widehat{\pi}_i^L\right)\right.\right.\right.$$
$$\left.\left.\left. + p_i^L(1 - \widehat{\pi}_i^L) \log_2\left(p_i^L(1 - \widehat{\pi}_i^L)\right)\right]\right)\Big| \pi_i^L = \frac{1}{2}\right]. \quad (A2)$$

Under the condition $\pi_i^L = 1/2$, the variable $n_i \widehat{\pi}_i^L$ follows a binomial distribution of parameters n_i and $1/2$. We thus have $\mathbb{P}(\widehat{\pi}_i^L = j/n_i) = \binom{n_i}{j} 2^{-n_i}$ and:

$$\mathbb{E}\left[\exp\left(t\left[p_i^L \widehat{\pi}_i^L \log_2\left(p_i^L \widehat{\pi}_i^L\right)\right.\right.\right.$$
$$\left.\left.\left. + p_i^L(1 - \widehat{\pi}_i^L) \log_2\left(p_i^L(1 - \widehat{\pi}_i^L)\right)\right]\right)\Big| \pi_i^L = \frac{1}{2}\right]$$
$$= \sum_{j=0}^{n_i} \exp\left(t\left[p_i^L \frac{j}{n_i} \log_2\left(p_i^L \frac{j}{n_i}\right) + p_i^L\left(1 - \frac{j}{n_i}\right)\right.\right.$$
$$\left.\left. \times \log_2\left(p_i^L\left(1 - \frac{j}{n_i}\right)\right)\right]\right) \binom{n_i}{j} 2^{-n_i}$$
$$= e^{tp_i^L \log_2(p_i^L)} \sum_{j=0}^{n_i} C_{i,j}^L(t). \quad (A3)$$

Combining this equation with equation (A2), we get:

$$M_{\widehat{I}^{L+1}}(t) = e^{tH_\star^{L+1}} \prod_{i=1}^{2^L} e^{tp_i^L \log_2(p_i^L)} \sum_{j=0}^{n_i} C_{i,j}^L(t)$$
$$= e^{tH_\star^{L+1}} e^{-tH^L} \prod_{i=1}^{2^L} \sum_{j=0}^{n_i} C_{i,j}^L(t)$$
$$= e^t \prod_{i=1}^{2^L} \sum_{j=0}^{n_i} C_{i,j}^L(t),$$

because $H_\star^{L+1} = 1 + H^L$, according to equation (2). ∎

Appendix 3. Proof of proposition 3.2

Proof By definition, the rth moment appearing in proposition 3.2 is obtained by the rth derivative of the moment-generating function, $M_{\hat{l}^{L+1}}^{(r)}(0)$. By developing the product in the formula provided in proposition 3.1, we write

$$M_{\hat{l}^{L+1}}(t) = e^t \sum_{j_1=0}^{n_1} \cdots \sum_{j_{2^L}=0}^{n_{2^L}} \prod_{i=1}^{2^L} C_{i,j_i}^L(t).$$

Then, Leibniz rule leads to

$$M_{\hat{l}^{L+1}}^{(r)}(t) = \sum_{m=0}^r \binom{r}{m} e^t \sum_{j_1=0}^{n_1} \cdots \sum_{j_{2^L}=0}^{n_{2^L}} \left(\prod_{i=1}^{2^L} C_{i,j_i}^L(t) \right)^{(m)}. \quad (A4)$$

A straightforward calculation gives

$$\left(\prod_{i=1}^{2^L} C_{i,j_i}^L(t) \right)^{(m)} = \alpha_{j_1,\dots,j_{2^L}}^m \prod_{i=1}^{2^L} C_{i,j_i}^L(t). \quad (A5)$$

We also note that

$$C_{i,j_i}^L(0) = \binom{n_i}{j_i} \frac{1}{2^{n_i}}. \quad (A6)$$

Combining equations (A4), (A5), and (A6) together gives

$$M_{\hat{l}^{L+1}}^{(r)}(0) = \sum_{m=0}^r \binom{r}{m} \sum_{j_1=0}^{n_1} \cdots \sum_{j_{2^L}=0}^{n_{2^L}} \alpha_{j_1,\dots,j_{2^L}}^m \prod_{i=1}^{2^L} C_{i,j_i}^L(0)$$

$$= \sum_{m=0}^r \binom{r}{m} \sum_{j_1=0}^{n_1} \cdots \sum_{j_{2^L}=0}^{n_{2^L}} \alpha_{j_1,\dots,j_{2^L}}^m \prod_{i=1}^{2^L} \binom{n_i}{j_i} \frac{1}{2^{n_i}},$$

which is the result stated in proposition 3.2. ∎

Appendix 4. Proof of proposition 3.3

Proof Equation (6) is a direct consequence of Faà di Bruno's formula to the function $x \mapsto \exp([itp_j^L/ln(2)]\Lambda(x)) = g_j(x)$, where Λ is a primitive of λ: $\Lambda(x) = x \ln(p_j^L x) + (1-x) \ln(p_j^L(1-x)) - \ln(p_j^L/2)$.

For proving equation (7), we simply observe by recurrence that, for $k \geq 1$:

$$\frac{d^k}{dx^k} \lambda(x) = \frac{(k-1)!(-1)^{k+1}}{x^k} + \frac{(k-1)!}{(1-x)^k}. \quad (A7)$$

As a consequence, we have:

$$\frac{d^k}{dx^k} \lambda\left(\frac{1}{2}\right) = \begin{cases} 0 & \text{if } k \in 2\mathbb{N} \\ (k-1)! 2^{k+1} & \text{if } k \in 2\mathbb{N}+1. \end{cases} \quad (A8)$$

The Bell polynomial writes:

$$B_{k,l}\left(\lambda(x), \frac{d}{dx}\lambda(x),\dots\right) = \sum \frac{k!}{j_0! j_1! \cdots j_{k-l}!}$$

$$\times \prod_{m=0}^{k-l} \left(\frac{1}{(m+1)!} \frac{d^m}{dx^m} \lambda(x) \right)^{j_m}, \quad (A9)$$

where the sum is taken over the set of $j_0,\dots,j_{k-l}$ submitted to the traditional Bell conditions, namely $\sum_m j_m = l$ and $\sum_m (m+1)j_m =$

k. So, for $x = 1/2$, we get:

$$B_{k,l}\left(\lambda\left(\frac{1}{2}\right), \frac{d}{dx}\lambda\left(\frac{1}{2}\right),\dots\right) = \sum \frac{k!}{j_0! j_1! \cdots j_{k-l}!} \mathbb{1}_{\{j_0=j_2=j_4=\cdots=0\}}$$

$$\times \prod_{m=1}^{k-l} \left(\frac{(m-1)!}{(m+1)!} 2^{m+1} \right)^{j_m}$$

$$= \sum \frac{k!}{j_0! j_1! \cdots j_{k-l}!} \mathbb{1}_{\{j_0=j_2=j_4=\cdots=0\}}$$

$$\times 2^{\sum_m(m+1)j_m} \prod_{m=1}^{k-l} \left(\frac{(m-1)!}{(m+1)!} \right)^{j_m}$$

$$= 2^k B_{k,l}(0, 0!, 0, 2!, 0, 4!, 0, \dots).$$

Combining this last equation and $g_j(t, 1/2) = 1$ in equation (6), we obtain equation (7).

The reason why $\partial_x^k g_j(t, 1/2) = 0$ when k is odd lies in the properties of Bell polynomial. Indeed, when we have $j_m = 0$, for all m even numbers, we have

$$\sum_{m=0}^{k-l} (m+1)j_m = \sum_{m=0}^{\lfloor \frac{k-l-1}{2} \rfloor} (2m+2)j_{2m+1} \in 2\mathbb{N}$$

and so this quantity cannot be equal to an odd k, as required by the second Bell condition. Therefore, at least one j_m of even index m must be different from 0 in each product of equation (A9). But, as the corresponding mth derivative of λ is equal to zero in $x = 1/2$, the Bell polynomial (and the kth derivative of g_j) is equal to zero. ∎

Appendix 5. Proof of proposition 3.4

In order to prove proposition 3.4, we first introduce and prove the following lemma.

LEMMA A.1 *Let $x \in [1/2, 1)$, $k \in \mathbb{N}$, the function r be defined by equation (9), and $u_k(x)$ by $(1/(k+1)!)d^k\lambda(x)/dx^k$, where λ is defined like in proposition 3.3. Then, for all $k' \in [\![0, k]\!]$, we have $0 \leq \min(u_0(x), u_1(x)) \leq u_{k'}(x) \leq u_{r(k)}(x)$.*

Proof The power series expansion of $x \mapsto u_k(x)$ leads to:

$$u_k(x) = \sum_{j=0}^{\infty} \frac{1}{j!(k+1)!} \frac{d^{k+j}\lambda(1/2)}{dx^{k+j}} \left(x - \frac{1}{2} \right)^j.$$

Noting thanks to equation (A8) that $d^{k+2}\lambda(1/2)/dx^{k+2} = 4k(k+1)d^k\lambda(1/2)/dx^k$, we also have:

$$u_{k+2}(x) = \sum_{j=0}^{\infty} \frac{1}{j!(k+3)!} \frac{d^{k+2+j}\lambda(1/2)}{dx^{k+2+j}} \left(x - \frac{1}{2} \right)^j$$

$$= \sum_{j=0}^{\infty} \frac{4(k+j)(k+j+1)}{(k+2)(k+3)} \frac{1}{j!(k+1)!}$$

$$\times \frac{d^{k+j}\lambda(1/2)}{dx^{k+j}} \left(x - \frac{1}{2} \right)^j.$$

Let $a_{k,j} = 4(k+j)(k+j+1)/(k+2)(k+3)$. When $j \geq 2$, $a_{k,j}$ is trivially larger than 1. When $j = 1$, we have $a_{k,j} = (2k+2)(2k+4)/(k+2)(k+3) > 1$. Finally, when $j = 0$, solving a simple binomial equation in k, we obtain that the only $k \in \mathbb{N} \setminus \{0\}$ such that $a_{k,0} < 1$ is $k = 1$. So, we have

$$\forall k \geq 2, u_{k+2}(x) \geq u_k(x). \quad (A10)$$

With a similar reasoning and solving a new binomial equation in k, we show that $\forall k \in \mathbb{N} \setminus \{0\}$, $16k(k+1)/(k+5)(k+4) > 0$ and

finally that

$$\forall k \geq 1, u_{k+4}(x) \geq u_k(x). \tag{A11}$$

Moreover, if $k > 0$ is an even number, we have from equation (A7) that $u_k(x) \leq 1/k(k+1)(1-x)^k$. The same equation also leads to $u_{k+1}(x) \geq 1/(k+1)(k+2)(1-x)^{k+1} \geq 2/(k+1)(k+2)(1-x)^k$. Noting that for $k \geq 2$, we have $2/(k+2) \geq 1/k$, this simply proves that

$$\forall k \geq 1, u_{2k}(x) \leq u_{2k+1}(x). \tag{A12}$$

Last, noting that $u_0(1/2) = 0 = u_2(1/2)$ and that

$$\frac{d}{dx}(u_0(x) - u_2(x)) = \frac{1}{x}\left(1 - \frac{1}{3x^2}\right) + \frac{1}{1-x}\left(1 - \frac{1}{3(1-x)^2}\right)$$

$$\leq \frac{1}{x}\frac{2}{3} - \frac{1}{1-x}\frac{4}{3} \leq -\frac{4}{3},$$

we get $u_0(x) \leq u_2(x)$. Combining this last equation with equations (A10), (A11), and (A12), and noting that for $k \geq 5$, $r(k) = k$ (respectively $r(k) = k+1$) when k is odd (resp. even), we prove lemma A.1. ∎

We can now prove proposition 3.4.

Proof Combining equation (6) and Minkowski inequality, we get

$$\left(\int_{1/2}^{z}\left|\partial_x^k g_j(t,x)\right|^q dx\right)^{1/q}$$

$$\leq \sum_{l=1}^{k}\left(\int_{1/2}^{z}\left|\frac{tp_j^L}{\ln(2)}\right|^{ql}\right.$$

$$\left.\times \left|B_{k,l}\left(\lambda(x), \frac{d}{dx}\lambda(x), \ldots, \frac{d^{k-l}}{dx^{k-l}}\lambda(x)\right)\right|^q dx\right)^{1/q}. \tag{A13}$$

Using lemma A.1, we can bound the Bell polynomial, in which we recall that $\sum_m j_m = l$ and $\sum_m(m+1)j_m = k$:

$$\left|B_{k,l}\left(\lambda(x), \frac{d}{dx}\lambda(x), \ldots, \frac{d^{k-l}}{dx^{k-l}}\lambda(x)\right)\right|$$

$$= \sum \frac{k!}{j_0!j_1!\cdots j_{k-l}!}\prod_{m=0}^{k-l}\left(\frac{1}{(m+1)!}\frac{d^m}{dx^m}\lambda(x)\right)^{j_m}$$

$$\leq \sum \frac{k!}{j_0!j_1!\cdots j_{k-l}!}\left(\frac{1}{(r(k-l)+1)!}\frac{d^{r(k-l)}}{dx^{r(k-l)}}\lambda(x)\right)^{\sum_m j_m}$$

$$= B_{k,l}(1!, 2!, \ldots, (k-l+1)!)$$

$$\times \left(\frac{(r(k-l)-1)!}{(r(k-l)+1)!}\left(\frac{1}{x^{r(k-l)}} + \frac{1}{(1-x)^{r(k-l)}}\right)\right)^l$$

$$\leq \mathcal{L}(k,l)\left(\frac{2}{(r(k-l)+1)r(k-l)}\right)^l(1-x)^{-lr(k-l)},$$

because $B_{k,l}(1!, 2!, \ldots, (k-l+1)!) = \mathcal{L}(k,l)$. We then use this bound in equation (A13):

$$\left(\int_{1/2}^{z}\left|\partial_x^k g_j(t,x)\right|^q dx\right)^{1/q}$$

$$\leq \sum_{l=1}^{k}\left(\left|\frac{tp_j^L}{\ln(2)}\right|^{ql}\mathcal{L}(k,l)^q\left(\frac{2}{(r(k-l)+1)r(k-l)}\right)^{ql}\right.$$

$$\left.\times \int_{1/2}^{z}(1-x)^{-qlr(k-l)}dx\right)^{1/q}$$

$$= \sum_{l=1}^{k}\left|\frac{tp_j^L}{\ln(2)}\right|^{l}\mathcal{L}(k,l)\left(\frac{2}{(r(k-l)+1)r(k-l)}\right)^{l}$$

$$\times \left(\frac{(1-z)^{-qlr(k-l)+1} - 2^{qlr(k-l)-1}}{qlr(k-l)-1}\right)^{1/q}$$

$$\leq \sum_{l=1}^{k}\left|\frac{tp_j^L}{\ln(2)}\right|^{l}\mathcal{L}(k,l)\left(\frac{2}{(r(k-l)+1)r(k-l)}\right)^{l}$$

$$\times \frac{(1-z)^{-lr(k-l)+1/q}}{(qlr(k-l)-1)^{1/q}}.$$

This concludes the proof. ∎

Appendix 6. Proof of theorem 3.1

Proof Noting that the derivatives are provided by proposition 3.3 and more precisely by equation (7) when $x = 1/2$, we can do a Taylor expansion of g_j around $x = 1/2$:

$$g_j(t,x) = 1 + \frac{2itp_j^L}{\ln(2)}\left(x - \frac{1}{2}\right)^2 + \mathcal{R}_j(t,x), \tag{A14}$$

where

$$\mathcal{R}_j(t,x) = \frac{1}{6}\int_{1/2}^{x}\partial_x^4 g_j(t,y)(x-y)^3 dy.$$

Let $q > 1$. According to Hölder's inequality, we have:

$$|\mathcal{R}_j(t,x)| \leq \frac{1}{6}\int_{1/2}^{x}\left|\partial_x^4 g_j(t,y)(x-y)^3\right|dy$$

$$\leq \frac{1}{6}\left(\int_{1/2}^{x}\left|\partial_x^4 g_j(t,y)\right|^q dy\right)^{1/q}$$

$$\times \left(\int_{1/2}^{x}\left|(x-y)^3\right|^{q/(q-1)}dy\right)^{1-1/q}.$$

The second integral is easily calculated and its value is $((q-1)/(4q-1))(x-1/2)^{(4q-1)/(q-1)}$. For the first integral, we use proposition 3.4, noting that the $r(k-l)$ are all equal to 5 since $k = 4$. Therefore,

$$|\mathcal{R}_j(t,x)| \leq \frac{1}{6}\left(\sum_{l=1}^{4}\left|\frac{tp_j^L}{15\ln(2)}\right|^{l}\mathcal{L}(4,l)\frac{(1-x)^{-5l+1/q}}{(5ql-1)^{1/q}}\right)$$

$$\times \left(\frac{q-1}{4q-1}\right)^{1-1/q}\left(x - \frac{1}{2}\right)^{4-1/q}.$$

Then, replacing x by X/n_j leads to

$$\mathbb{E}\left[\left|\mathcal{R}_j\left(t, \frac{X}{n_j}\right)\right|\right]$$

$$\leq \frac{1}{6}\left(\frac{q-1}{4q-1}\right)^{1-1/q}\sum_{l=1}^{4}\left|\frac{tp_j^L}{15\ln(2)}\right|^l$$

$$\times \mathcal{L}(4,l)\mathbb{E}\left[\frac{(1-X/n_j)^{-5l+1/q}}{(5ql-1)^{1/q}}\left(\frac{X}{n_j}-\frac{1}{2}\right)^{4-1/q}\right],$$

in which we can transform the expectation, which we note ξ, in the right-hand side of the inequality, thanks to the Cauchy-Schwarz inequality:

$$\xi^2 \leq \mathbb{E}\left[\frac{(1-X/n_j)^{-10l+2/q}}{(5ql-1)^{2/q}}\right]\mathbb{E}\left[\left(\frac{X}{n_j}-\frac{1}{2}\right)^{8-2/q}\right].$$

Noting that $1-X/n_j$ is distributed like X/n_j, we can apply an asymptotic result on the negative moments of a positive binomial variable (Arias López and Garrido 2001, Corollary 3.4): $\lim_{n_j\to\infty}\mathbb{E}[(1-X/n_j)^{-10l+2/q}]=2^{10l-2/q}$. Regarding the second equation, we first use Jensen's inequality, in which we impose to have $q\in\mathbb{N}$, and we conclude by noting that, for $d\in\mathbb{N}$, $\mathbb{E}[(X-n_j/2)^{2d}]\leq d^{2d}(n_j/4)^d$ (Skorski 2020):

$$\mathbb{E}\left[\left(\frac{X}{n_j}-\frac{1}{2}\right)^{8-2/q}\right]\leq\left(\mathbb{E}\left[\left(\frac{X}{n_j}-\frac{1}{2}\right)^{8q-2}\right]\right)^{1/q}$$

$$\leq (4q-1)^{8-2/q}(4n_j)^{-4+1/q}.$$

As a consequence, for all $\varepsilon>1$, there exists $\nu\in\mathbb{N}$ such that for all $n_j\geq\nu$ and all q and l, we have

$$\xi\leq\varepsilon\frac{2^{5l-4}(4q-1)^{4-1/q}}{(5ql-1)^{1/q}}n_j^{-2+1/2q}$$

and finally

$$\mathbb{E}\left[\left|\mathcal{R}_j\left(t,\frac{X}{n_j}\right)\right|\right]\leq\frac{\varepsilon}{96}(q-1)^{1-1/q}(4q-1)^3$$

$$\times\left(\sum_{l=1}^{4}\left|\frac{2^5tp_j^L}{15\ln(2)}\right|^l\frac{\mathcal{L}(4,l)}{(5ql-1)^{1/q}}\right)n_j^{-2+1/2q}.$$

(A15)

Going back to the Taylor expansion in equation (A14), and knowing that the variance of the binomial variable X is $n_j/4$, we then have:

$$\mathbb{E}\left[g_j\left(t,\frac{X}{n_j}\right)\right]=1+\frac{2itp_j^L}{\ln(2)}\mathbb{E}\left[\left(\frac{X}{n_j}-\frac{1}{2}\right)^2\right]+\mathbb{E}\left[\mathcal{R}_j\left(t,\frac{X}{n_j}\right)\right]$$

$$=1+\frac{itp_j^L}{2\ln(2)n_j}+\mathbb{E}\left[\mathcal{R}_j\left(t,\frac{X}{n_j}\right)\right].$$

We define $R(t,n_j)=\mathbb{E}[\mathcal{R}_j(t,X/n_j)]$ and Jensen's inequality provides us with $|R(t,n_j)|\leq\mathbb{E}[|\mathcal{R}_j(t,X/n_j)|]$, for which we know an upper bound thanks to equation (A15). This concludes the proof. ∎