Information Transition in Trading and its Effect on Market Efficiency: An Entropy Approach

Anqi Liu · Jing Chen · Steve Y. Yang · Alan G. Hawkes

Accepted May 2020, to appear as Chapter 5 of Proceedings of the First International Forum on Financial Mathematics and Fintech. Zhiyong Zheng (Ed.). Springer Verlag, Singapore. (To be published Dec. 2020)

Abstract The Efficient Market Hypothesis has been well explored in terms of daily responses to market movements and financial reports. However, there is lack of evidences about information efficiency after the popularization of intraday trading. We investigate the time series properties of information adopted in the intraday market, in particular the causality effects. We use 30-min market price and news data to represent the past market data and the public information respectively, so that our analysis is in line with the EMH framework. Traders' responses to such information are associated with the financial crisis. There was strong overreaction to market data right before the 2008 crisis and traders tend to rely more on news data during the crisis. We confirm that, in terms of the intraday information efficiency, it is worthwhile to adopt both types of information. Furthermore, there is still room for improving the price discovery process to reveal such information more effectively.

Keywords news sentiment · information entropy · market efficiency · intraday equity market

Anqi Liu
School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK
E-mail: liua5@cardiff.ac.uk

Jing Chen
School of Mathematics, Cardiff University, Senghennydd Road, Cardiff, CF24 4AG, UK
E-mail: chenj60@cardiff.ac.uk

Steve Y. Yang
School of Business, Stevens Institute of Technology, 1 Castle Point on Hudson, Hoboken, NJ 07030
E-mail: syang14@stevens.edu

Alan G. Hawkes
School of Management, Swansea University, Bay Campus, Fabian Way, Swansea, SA1 8EN, UK
E-mail: a.g.hawkes@swansea.ac.uk

1 Introduction

In the recent 20 years, the financial market entered the electronic age. Trading is highly relying on advanced computing techniques, such as big data, text mining. In particular, the increasing popularity of intraday trading requires quick reactions to business and market information. In this paper, we explore the information that is widely adopted in intraday trading activities, in particular the causality relationships. They indicates the channel(s) of information transition that forms the market movements; moreover, the effects on market efficiency.

The motivation of this study hinges on the Efficient Market Hypothesis [6]. It involves three information efficiency stages that reflect investors' responses to past market data, public information (i.e. companies' financial reports), and non-public information. It has been widely accepted that most developed markets achieve semi-strong form market efficiency [2, 7, 8]. However, most studies that examined this argument were in 1980s and only focus on daily market responses. Apparently, such data frequency is too slow to fully represent the market condition now. Increasing high frequency trading volume sharply reduced the responding time of information on the markets [4]. Moreover, development of information technology has led to an "information revolution" of investment industry. Traders tend to explore broader categories of information to seek profits through the fast moving electronic markets. Apart from the traditional fundamental analysis of financial statements, investors also attempt to "decode" textual information into trading signals [3, 14]. For example, traders in hedge funds track breaking news to respond to shocks quickly. Business news vendors, such as Bloomberg and Thomson Reuters, start providing news analysis solutions that allow investors to develop new strategies. Obviously, business news has replaced the fundamental information as "public information" in the intraday markets.

In this study, we bridge the gap of the literature by breaking the limited types of information considered in price discovery and market efficiency. We investigate the two types of information that are the main leads to fast reactions in the modern financial markets: 1) the "instant" market price changes, to represent the past market data; and 2) the business news, to represent the public information. Our objective is to analyze the dynamics of information transitions such that we would bring some new insights of information efficiency on financial markets. Intuitively, the two types of information should show causal relationships to each other. The price changes caused by news information can be explained by traders' responses to business news. On the other hand, the cause of market information to news means the public is informed about the market conditions through the news media in a timely manner; such effect is expected to be more significant when there are shocks to the markets. Furthermore, both types of information should also have "self-causality effect". Apparently, the "habit" of following-up stories in the news industry would naturally result in a feedback pattern of news information. The rationale of causality in price information is that the technical trading highly involves the use of past market data. Similar to the causality from news to market, this indicates traders' responses to market information. We believe that, for the entire market, there is an optimal level of reactions to each type of information such that the rest of the market would not be able to discover a profitable strategy using the same information. In other words, causality from news to market and from market to market would be stabilized at certain values which indicate sufficient responses; too high and too low causalities mean overreaction and underrection respectively. Our results of causality relationships and regressions to market efficiency confirm this point.

In traditional financial studies, the causality relationship is usually validated by the Granger causality test. However, this approach has a few limitations that do not fit our context. First, the Granger test is based on linear regression. Price movements caused by diversified trading strategies would be barely simplified into a linear model. For instance, technical analysis involves both momentum trading and reversal trading activities, such that a price increase (or a decrease) may be driven by either direction of price movements. This is obviously a non-linear causal relationship that cannot be identified by linear regression. Second, the Granger test is built under the assumption of Gaussian distributed variables. Its validity for the non-Gaussian financial data is questionable. Furthermore, the Granger test is designed for cross-sectional causality analysis. It is not appropriate to apply it for the self-causality. Therefore, in this study, we introduce the use of entropy measures to identify the causality relationships. This approach has been adopted in a few financial market studies, in particular to solve non-linear causality relationships [5, 10]. Despite overcoming the limitations of Granger test mentioned above, the rationale of using the entropy as an alternative approach is two-fold. From the perspective of multivariate stochastic modeling, it can be interpreted as a causality problem measured by conditional distributions. The entropy measures indicate whether the distribution of one variable is conditional on the process itself or other time series(es). Also, as a well-developed approach in the field of information theory, it provides valuable insights of quantifying information in financial markets.

In this study, we adopt two entropy measures, the conditional block entropy and the transfer entropy, to tackle the self-causality and the cross-sectional causality of the two types of information respectively. We use the S&P 500 index as an representation of market information and the Thomson Reuters news analytics data to reflect news information. We track the changes of causality relationships using a 3-year rolling window and find that all these causalities are closely associated with the market conditions. As mentioned above, we highlight the causalities from market to market and from news to market because they reflect the predictability of price movements through different types of information. We observe a sharp increase of market self-causality before the 2008 crisis which is due to the price chasing activity; also indicate an overuse of market information. At the same time, news has not yet been widely adopted in trading. It was until the financial crisis that news information started to be involved in trading strategies and partially replaced the impacts of market information. We find that both causalities are stabilized after the market recovery from the double crisis. This confirm our thought that the market eventually evolves itself to figure out the efficient use of different information.

Another contribution of this study is to examine how the information entropy affect market efficiency. We find that, apart from the causality from market to news, all other entropy measures have significant linear relationships with market efficiency. The results indicate that both technical analysis and trading through news are closely associated with market efficiency. Based on the regression coefficients and the observation of more peaceful responses to information, we think it is worthwhile to adopt these two types of information in price discovery process.

2 Information entropy

The concept of entropy was initially introduced by Claude E. Shannon in 1948 to quantify how much information is contained in a signal process. Through decades of research, several entropy measures have been proposed to analyze more complex information systems. This approach has also been applied to examine lagged impacts among financial time series. For example, . We also show the equivalence of transfer entropy and the Granger test subject to the Gaussian distributed variables.

2.1 Entropy measures

In information theory, entropy measures uncertainty of a process. If we view from the trading perspective, higher uncertainty means less predictability. Therefore, we use entropy to analyze whether new information would help traders to "predict" price changes. Three entropy measures are applied in this study.

– Shannon entropy: It measures the amount of information in a random process. A larger entropy denotes less informative, also a higher uncertainty [12].

$$H(X) = -\sum p(x) \log_2 p(x)$$

– Conditional block entropy: It measures the amount of information in a subseries that affects the subsequent observation [10].

$$h_X(k) = -\sum p(x_{t+1}, x_t^{(k)}) \log_2 p(x_{t+1}|x_t^{(k)}) \tag{1}$$

– Transfer entropy: It measures the amount of information in one process that affects the subsequent observation in another process. It examines the asymmetric dynamics of two processes [11].

$$T_{Y\to X}(k,l) = \sum_{x,y} p(x_{t+1}, x_t^{(k)}, y_t^{(l)}) \log_2 \frac{p(x_{t+1}|x_t^{(k)}, y_t^{(l)})}{p(x_{t+1}|x_t^{(k)})} \tag{2}$$

2.2 Propositions of causal relationships in transfer entropy

In the area of finance, conditional block entropy and transfer entropy can be used to explain causal type of relationships. We provide a few propositions to establish the relationship, in particular between Granger causal relationships and the transfer entropy measure.

Proposition 1 *If X is a sequence of i.i.d. random variables, then there is no self information flow within the series X i.e. the conditional block entropy shall be equal to the Shannon entropy.*

Proof For an i.i.d. sequence X, we have

$$p(x_{t+1}|x_t^{(k)}) = p(x_{t+1}) \tag{3}$$

$$\text{and } p(x_{t+1}, x_t^{(k)}) = p(x_t^{(k)})p(x_{t+1}|x_t^{(k)}) = p(x_t^{(k)})p(x_{t+1}) \tag{4}$$

Then from Equation (1)

$$
\begin{aligned}
h_X(k) &= -\sum p(x_{t+1}, x_t^{(k)}) \log_2 p(x_{t+1}|x_t^{(k)}) \\
&= \sum p(x_t^{(k)})\{-\sum p(x_{t+1}) \log_2 p(x_{t+1}))\} \\
&= \sum p(x_t^{(k)}) H_X \\
&= H_X
\end{aligned} \tag{5}
$$

Since $\sum p(X_t^{(k)}) = 1$.

Thus, the conditional block entropy is the same as the unconditional entropy or, the past provides no information about the future.

Proposition 2 *For two independent series X and Y, the transfer entropy between them will be zero (i.e. no causal relationships between X and Y).*

Proof For the two series X, Y, the transfer entropy satisfies Equation (2).

If the two series are independent, we have $p(x_{t+1}|x_t^{(k)}, y_t^{(l)}) = p(x_{t+1}|x_t^{(k)})$. Then for all possible series values the logarithmic term in the above expression becomes $log_2(1) = 0$.

So $T_{Y\to X} = 0$ for any positive integers k and l.

Similarly, $T_{X\to Y} = 0$ as well.

This proposition shows that non-zero transfer entropy between two time series indicates the existence of a dependent relationship. We can safely assume there is a causal relationship from the source series (Y) to the target series (X) if the transfer entropy $T_{Y\to X} > 0$.

Proposition 3 *Granger causality and transfer entropy are equivalent if all variables involved are distributed as multivariate normal distributions.*

Proof This is a more succinct proof of a result of [1]. For any random vector Z with probability density $f(Z)$ the entropy is defined as

$$
H(Z) = -\int f(z) \ln f(z) dz = -E[\ln f(Z)]. \tag{6}
$$

Note that we are using "Natural" logarithms rather than base 2 logs that are common in information theory. If Z has multi-Normal distribution $Z \sim MN(\mu, \Sigma(Z))$ the probability density is

$$
\begin{aligned}
f(z) =& (2\pi)^{-\frac{1}{2}d_Z} |\Sigma(Z)|^{-\frac{1}{2}} \\
& \exp\left\{-\frac{1}{2}(z-\mu)'\Sigma(Z)^{-1}(z-\mu)\right\},
\end{aligned} \tag{7}
$$

where d_Z is the dimension of Z. Then

$$
\begin{aligned}
H(Z) = & \frac{1}{2} d_Z \ln(2\pi) + \frac{1}{2} \ln |\Sigma(Z)| \\
& + E[\frac{1}{2}(Z-\mu)'\Sigma(Z)^{-1}(Z-\mu)].
\end{aligned} \tag{8}
$$

But the quadratic form in the final term has a chi-squared distribution with d_Z degrees of freedom, and so has expectation d_Z. Therefore

$$\begin{aligned} H(Z) &= \frac{1}{2}\ln|\Sigma(Z)| + \frac{1}{2}d_Z ln(2\pi) + \frac{1}{2}d_Z \\ &= \frac{1}{2}\ln|\Sigma(Z)| + \frac{1}{2}d_Z ln(2\pi e). \end{aligned} \tag{9}$$

Now let $Z = \begin{pmatrix} X \\ W \end{pmatrix}$, then Equation (7) can be written as

$$f(z) = f(w)f(x|w), \tag{10}$$

where $f(w)$, similar to Equation (7),

$$\begin{aligned} f(w) =&(2\pi)^{-\frac{1}{2}d_W}|\Sigma(W)|^{-\frac{1}{2}} \\ &\exp\left\{-\frac{1}{2}(w-\mu_W)'\Sigma(W)^{-1}(w-\mu_W)\right\}, \end{aligned} \tag{11}$$

The conditional density is

$$\begin{aligned} f(x|w) =&(2\pi)^{-\frac{1}{2}d_X}|\Sigma(X|W)|^{-\frac{1}{2}} \\ &\exp\left\{-\frac{1}{2}(x-\mu_{X|W})'\Sigma(X|W)^{-1}(x-\mu_{X|W})\right\}, \end{aligned} \tag{12}$$

where the conditional dispersion matrix is

$$\Sigma(X|W) = \Sigma(X) - \Sigma(X,W)\Sigma(W)^{-1}\Sigma(W,X) \tag{13}$$

with

$$\Sigma(Z) = \Sigma\begin{pmatrix} X \\ W \end{pmatrix} = \begin{pmatrix} \Sigma(X) & \Sigma(X,W) \\ \Sigma(W,X) & \Sigma(W) \end{pmatrix}. \tag{14}$$

Note that, from Equations (7) and (10) to (12)

$$|\Sigma(Z)| = |\Sigma(W)||\Sigma(X|W)|. \tag{15}$$

Let $x_{t+1}, x_l^{(k)}, y_t^{(l)}$ have a multivariate Normal distribution. Then transfer entropy is

$$\begin{aligned} T_{Y\to X}(k,l) &= H(x_{t+1}|x_t^{(k)}) - H(x_{t+1}|x_t^{(k)}, y_t^{(l)}) \\ &= \frac{1}{2}\ln|\Sigma(x_{t+1}|x_t^{(k)})| + \frac{1}{2}\ln(2\pi e) \\ &- \frac{1}{2}\ln|\Sigma(x_{t+1}|x_t^{(k)}, y_t^{(l)})| - \frac{1}{2}\ln(2\pi e) \\ &= \frac{1}{2}\ln\left\{\frac{\Sigma(x_{t+1}|x_t^{(k)})}{\Sigma(x_{t+1}|x_t^{(k)}, y_t^{(l)})}\right\} \end{aligned} \tag{16}$$

The argument of the logarithm is just the ratio of the variance of x_{t+1} conditional on $x_t^{(k)}$ and the variance of x_{t+1} conditional on both $x_t^{(k)}$ and $y_t^{(l)}$. As we are dealing with multivariate Normal, these are calculated by appropriate forms of

Equation (13), which is a standard result for linear regression (whether or not distributions are Normal). This is therefore exactly the criterion that is used to determine whether Y Granger causes X, and so Granger causality and transfer entropy are equivalent if all variables involved are distributed as multivariate Normal.

From the above three propositions, we conclude that the entropy method is sufficient in identifying causal relationships that are relevant and important for many fundamental financial problems.

3 Methodology

In this section, we outline the use of entropy measures to explore empirical features of intraday market price information and news information. The purpose is to show the dynamics of information transition on "high-frequency" level, in particularly consider news as the main source of public information for intraday trading.

3.1 Measuring causality using entropy

We investigate self-causality and cross-sectional causality of price and news information.

3.1.1 Self-causality of information

The first property we examine is the self-causality effect. We think both types of information should carry some memory. The impact of price to itself is due to the use of past market information in trading, which is consistent with the understanding of technical analysis. The existence of technical trading is also crucial to ensure, at least, the weak form market efficiency. The memory of news information is often called "news of news". For instance, one important Fed's announcement could be followed by hundreds of news articles. In addition, it often hinges on the more complicated nature of news, namely the speed of news publication, contents of news, news sources and their validity, etc.

The self-causality can be interpreted as the statistical property that the probability distribution on time t is conditional on the filtration of previous observation(s) $\mathcal{F}_{t-1}$. This is slightly different from the serial correlation as it may lead to non-linear relationships. As suggested by the definition of conditional block entropy, it quantifies the uncertainty of signal(s) based on known information. This is a perfect match to solve the identification of self-causality.

We denote $\Delta_X(k)$ as the contribution of memory $x_t^{(k)}$ (see Equation 17). The larger block size k, the longer memory is available to estimate x_{t+1} and the larger $\Delta_X(k)$.

$$\Delta_X(k) = H_X - h_X(k) \tag{17}$$

In other words, $\Delta_X(k)$ increases until k reaches the memory length k_X. This is demonstrated in Figure **??**, and it is clear that the conditional block entropy $h_X(k)$

is reduced by the increase in the contribution of the memory $\Delta_X(k)$. Obviously, the $\Delta_X(k)$ is bounded by the entropy,

$$0 \leq \Delta_X(k) \leq H_X.$$

We standardize this measure through the Equation 18 that maps the value to $[0, 1]$, in order to keep the consistency for comparison. It can also be interpreted as "the contribution of previous information in percentage."

$$\frac{\Delta_X(k)}{H_X} = 1 - \frac{h_X(k)}{H_X} \tag{18}$$

3.1.2 Cross-sectional causality of information

The causality from price to news, in reality, would happen. For instance, if a sharp price increase or decrease occurs on the market index, this information is usually spread out to the public immediately as a breaking news. While intuitively, such causality effect would be weaker than the opposite direction, i.e. from news information to price movements. As news is considered as a reliable input for price discovery nowadays, this causality is directly relating to trading dynamics. Furthermore, we think this is a supplementary of the EMH which failed to consider the potential of using varied types of public information, apart from the companies' reports, in rational trading practice.

According to the definition in [11], transfer entropy $T_{Y \to X}$ measures the amount of information in Y to "forecast" X excluding the information of X itself used in the self-causality process. We adopt transfer entropy to investigate the causality between different types of information. It is different from the Granger causality in terms of the non-linearity and non-Normality. We also verify that these two measures are equivalent for Gaussian distributed variables. Similar to the self-causality, we also standardize the transfer entropy $T_{Y \to X}$ based on its upper boundary $h_X(k)$, i.e.

$$\frac{T_{Y \to X}}{h_X(k)}.$$

3.2 Entropy calibration

The entropy calibration involves modeling the probability distribution(s) of the observed processes. Even though our data of both price and news information is continuous, we use discrete distributions in this study. This is firstly because the computing complexity is too high for continuous distributions. More importantly, we think discrete states of information are associated with investors' trading philosophy. As trading decisions are usually based on optimistic / pessimistic prospect, we define three states for each type of information (see Equation 19).

$$L(t) = \begin{cases} -1, & x(t) < \mu - d \\ 0, & \mu - d \leq x(t) \leq \mu + d \\ 1, & x(t) > \mu + d \end{cases} \tag{19}$$

in which $x(t)$ is the observation at time t. The threshold d is determined based on the "even" distribution criteria

$$Pr\,(-1) \approx Pr\,(0) \approx Pr\,(+1) \approx \frac{1}{3}.$$

Obviously, μ is the mean. This also provides the largest Shannon entropy among all three-state probability distributions.

To measure the self-causality, it is crucial to choose an appropriate block size for the conditional block entropy. The principle is to make the block size as large as possible such that the useful information in previously observed X is fully extracted. Ideally, $\Delta_X(k)$ reaches 0 if k is large enough, This means the process X does not hold a memory longer than k time intervals. However, due to the small sample bias in real practice, we may not be able to get an accurate estimation of $\Delta_X(k)$ after k increases to a certain point. This issue is also mentioned by [10]. They introduced a solution called "effective transfer entropy", which is to remove the noise by shuffling process X. We adopt this modification in our study. The block size determined for $\Delta_X(k)$ is applied in the estimation of cross-sectional causality $T_{X\to Y}$; and the block size for Y in this transfer entropy is set to 1.

4 Data

In this study, the market information and news information are represented by market index price and news sentiment respectively. We use 30-minute time intervals to observe the intraday dynamics. The dataset is from January 1, 2003 to December 31, 2014, excluding non-trading hours.

4.1 Market index price

Stock market indexes are proxies of equity market performance. In this study we use S & P 500 (.SPX), a capitalization-weighted market index, to best represent the U.S. stock market. We collect 30-minute intraday prices of the market index from Thomson Reuters Tick History^{TM}(TRTH). We use 30-minute return for the entropy calculation as return is stationary.

$$r_t = \log \frac{P_t}{P_{t-\Delta t}}$$

in which P_t denotes the index price and Δt is 30 minutes.

4.2 News sentiment data

The news data is provided by the Thomson Reuters News Analytics^{TM}(TRNA). It is a professional news sentiment database that has been adopted by previous studies [15]. The database contains over 80 metadata fields about financial news. News sentiment is the tone of news articles, i.e. good news or bad news. In the context of financial news, it tells prospect of bull or bear markets. The advantage of the TRNA database is that, apart from the sentiment of each piece of news, it provides a "relevance score" for each company mentioned in the news article to show relevance of the news to individual stocks.

The metadata fields we used for sentiment calibration in this paper are listed below.

- *datetime*: The date and time of a news article.
- *ric*: Reuters Instrument Code (RIC) of a stock for which the sentiment scores apply.
- *pos*, *obj*, *neg*: Positive, neutral, and negative sentiment probabilities (i.e., $pos + obj + neg = 1$).
- *relevance*: A real-valued number between 0 and 1 indicating the relevance of a piece of news to a stock. One news article may refer to multiple stocks. A stock with more mentions will be assigned a higher relevance.

To evaluate the sentiment score of each record in the database, we calculate the standardized expectation of sentiment probabilities adjusted by relevance value (see Equation 20). Then we calculate 30-minute time-weighted-average news sentiment. The news published in non-trading hours are counted into the first 30 minutes of the following trading day.

$$Sentiment = relevance \times (pos - neg) \times (1 - obj) \tag{20}$$

5 Results

The objective of this study is to investigate the non-linear statistical relationships of the two types of information. More precisely, we would like to show how the self- and cross-sectional causality change over time. Based on this, we will then explore some insights of information efficiency issue in intraday markets.

In general, the calibration of entropy needs large samples. We use a 3-year daily rolling window in this study. Every sample contains almost $10,000$ observations.

5.1 The distributions and the Shannon entropy

We use the full dataset to calibrate threshold defined in Equation 19 (see Table 1).

Table 1 The distributions of market and news information

	Mean μ	Threshold d
S&P500 return	0.00	0.0006
News sentiment	0.05	0.0291

As mentioned in Section 3.2, the criterion of data partition is to achieve equivalent probabilities for the 3 states, of which the entropy value is 1.585. This is a benchmark and an upper boundary for any 3-state probability spaces. The further the entropy is below this value, the more observations are biased from equiprobability. However, it does not tell whether the values are biased to the positive or negative side.

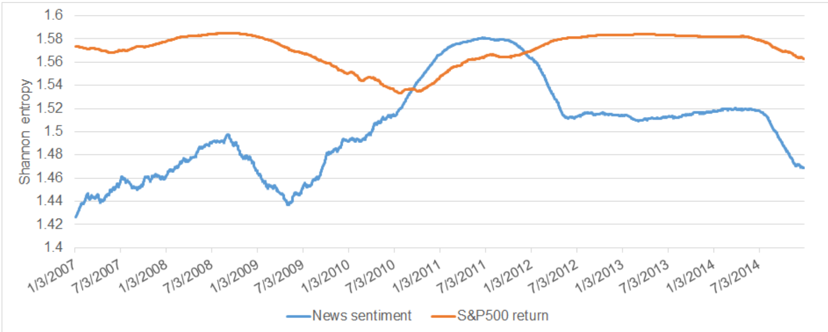

Fig. 1 Shannon entropy

In the S&P500 returns, we observe a "shock" in Shannon entropy (see Figure 1). We conjecture these entropy changes are highly associated with the formation of a price bubble that burst in 2008, thereby, detonating the crisis. The full period of the recent liquidity crisis has been featured in low entropy during late 2009. The Shannon entropy of news sentiment is less stable apart from the years after the 2008 financial crisis. For most of the time, the news sentiment entropy appears far from the benchmark 1.585, which may suggest that, in a 3-year time window, news articles usually indicates strong prospect in positive or negative market conditions. News information in the early ages is a bit "shaky" such that the entropy is volatile, showing a lack of consistent indication throughout short time periods. We observe a drop-then-rise pattern before and after the financial crisis. We think this is because good information transfer to bad information and eventually news becomes neutral after the shock in market. In recent years, recovery of financial markets brings some good news and this explains the decrease of news sentiment entropy after 2012. Even though we observe unevenly distributed news sentiment again, we believe the news information is not biased too much as the entropy is still much higher than that before the financial crisis.

5.2 Results of self-causality

We introduce the rationale of self-causality in both market and news information in Section 3.1.1. In general, it can be regarded as the "memory" of a time series. The memory length (i.e. the optimized block size) and strength (i.e. conditional block entropy) are meaningful features of both types of information.

Recall that we consider a 3-year rolling window to incorporate sufficient data to obtain the optimal memory length reflecting the impact on the market. In this case, the information flow of each point at time t actually represents an accumulative effect of the past 3 years prior to time t. According to Figure 2, the memory lengths of both types of information cluster into three time periods: pre-crisis (before 2008), crisis (2008-2013, covering both 2008 liquidity crisis and EuroDebt crisis) and post-crisis (after 2013). This pattern clearer for news information. News tends to update slower than returns and we think this will also be related to the memory length and strength. We observe that the memory of news increase from 2 blocks (i.e. 1 hour) to 6 blocks (i.e. 3 hours). As discussed in Section 3.1.1,

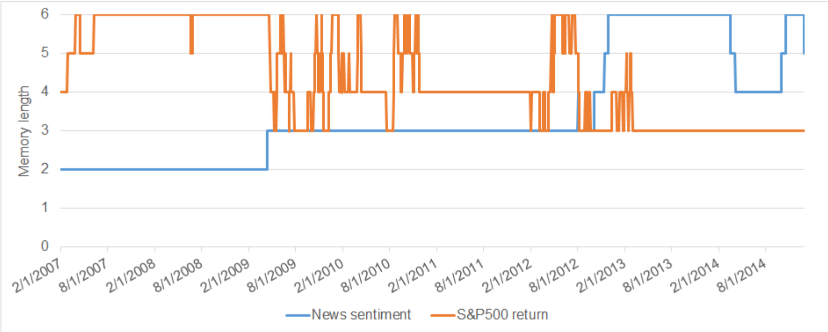

Fig. 2 Memory length of market and news information

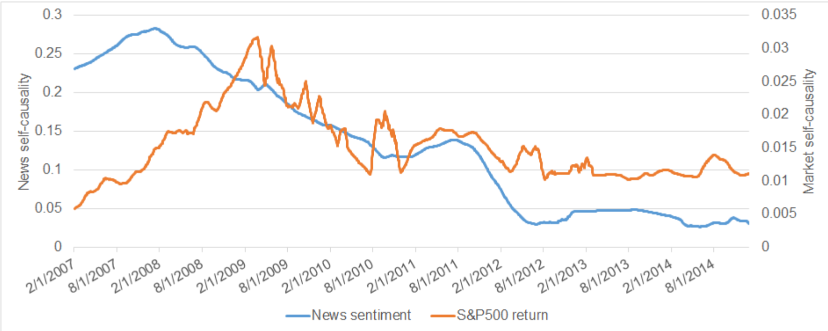

Fig. 3 Self-causality of market and news information

a piece of news often has many follow-up stories. This could be due to varied broadcasting speed for different news vendors. Generally, big business news vendors tend to gather information and broadcast them faster than others and they also can follow up the development of the news that could last for hours. Sometimes even for the same piece of news, the subjective opinions and reporting styles of different vendors may cause differences in prospect of information. There is no doubt that the number of business news vendors have increased during the past decades. The increasing memory length indicates that, as a consequence of "news of news", the timeline of follow-up news became longer. Furthermore, we find that increasing news memory length does not result in stronger memory strength. Instead, self-causality of news information decreased. This indeed confirms the diverted opinions from or subjects covered by different news vendors, which result in less repetition of news information in recent years.

On the contrary to the increasing news memory length, the memory length of market information decreases from 3 hours to a volatile stage of 1.5-3 hours during the crisis, then stabilized at 1.5 hours after 2013. We think this indicates more efficient technical trading in the intraday market which leads to a faster price discovery process. Obviously, the self-causality of market information is much weaker than that of the news information. This supports the argument of weak form market efficiency. The market movements should in general hard to be predicted by using past market information, which results in a very low self-causality. However, we also observe the sharp increase of memory before the financial crisis. We believe this is an indication of pricing chasing trading activities; a lot of investors was seeking profits in the bull market without investigating fundamental values and raised the price. Eventually, the overreaction t0 good market information ended up with a crash. Therefore, after the financial crisis, the market returned back to a rational environment in which the technical trading revealed the past market information into price discovery immediately. This can be supported by the stabilized but low self-causality of market information. In other words, the market reaches the weak form efficiency such that the price movements is not predictable only through technical analysis.

5.3 Results of cross-sectional causality

Causality between the two types of information is measured by transfer entropy. It tells how much an extra source of information would contribute to the predictability. Similar to the self-causality, we also observe the 3 ages for cross-sectional

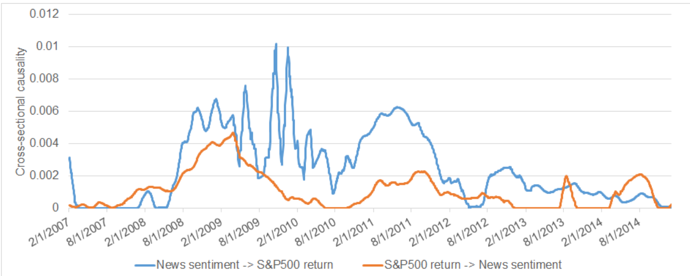

Fig. 4 Cross-sectional causality of market and news information

causality in both directions. The causality from market information to news information tells the development of the financial news industry. Back to 2000s the news industry was under a transformation from paper-based to an online media. We can see that before 2008 there was almost no news tells market conditions in a timely manner, at least within 1 block time (i.e. 30 minutes). However, things began to change during the financial crisis. The popularizing online news media was responding to the shocks in the market. This naturally leads to sharply increasing causality of market to news. It also explains why such causality effect died down in recent years; “no news is good news” when the market is stable.

On the other hand, the causality from news information to market information tells the story of news-based trading strategies. This direction of causality is more insightful as it reflects the efficiency of public information in intraday market. We observe very clear indication of the use of news information during the double crisis periods. In particular, the causality during the 2008 crisis is stronger than the Eurodollar crisis. This makes sense as the market we examine is the U.S. market, which experienced the most severer crash in 2008. Recall that we also observed stronger self-causality of market information in the same period. This means, albeit traders’ reaction to bear market helped to bring back the rational price, it was not sufficient so that trading based on news information also contributed. The predictability of news also dropped in recent years, some time even to zero. We think this shows a semi-strong form of market efficiency. In general, the intraday traders have explored news information well to reveal new public information in the price movements and killed the potential profitability.

5.4 Market inefficiency and information entropy

One ultimate goal of this study is to understand, after investigating the statistical properties of both types of information, how the market efficiency can be affected. We still scope the problem in the intraday market. We want to know when the market changes to be more efficient/inefficient and whether the causalities of information appear to be systematic. These are very important questions as overuse of information would result in overreactions and then harm the information efficiency in financial markets. In this study, we explore the relation of information entropy and market efficiency through regression analysis.

We adopt the efficiency index (EI) proposed by [9] to estimate financial market efficiency (see Equation 21).

$$EI = \sqrt{\sum_{i}^{n} (\frac{\widehat{M_i} - M_i^*}{R_i})^2} \tag{21}$$

where $\widehat{M_i}$ is the ith efficiency measure, M_i^* is the expected value of ith measure for the efficient market, and R_i is the range of ith measure. Obviously, $EI = 0$ for an efficient market, and the higher the EI the stronger the inefficiency. According to [9], the Hurst exponent, the fractal dimension, and the first order autocorrelation are selected as three efficiency measures in this efficiency index. For the efficient market, market return follows Brownian motion so that expected values of the three measures are 0.5, 1.5 and 0.0, respectively; the ranges of the three measures are 1.0, 1.0 and 2.0, respectively.

To be consistent with the evaluations of information entropy, we also use a 3-year rolling window for the efficiency index. According to the definition, this index actuality measures the level of market inefficiency. In other words, the smaller index value, the higher market efficiency. Overall, the efficiency did not change much during the sample period; the value ranged between 0.08 and 0.11 (see Figure 5). An interesting observation was the increasing efficiency before the financial crisis, when the market was undoubtedly in an irrational age with a lot of positive feedback trading [13]. We know that, according to the EMH, rational investment leads to market efficiency. Meanwhile, according to the efficiency index, the inverse (i.e. irrational investment leads to market inefficiency) does not hold.

We examine the following two linear models for the information entropy and market efficiency. In terms of the notation, EI is the efficiency index, R denotes the market return and S denotes the news sentiment. Recall that we use entropy to measure causalities, for example, $I_{S \to R}$ is the causality from news to market.

Model I: Market efficiency vs. All information entropy.

$$EI(t) = \beta_0 + \beta_1 EI(t-1) + \beta_2 I_{R \to R}(t) + \beta_3 I_{S \to S}(t) + \beta_4 I_{S \to R}(t) + \beta_5 I_{S \to S}(t)$$

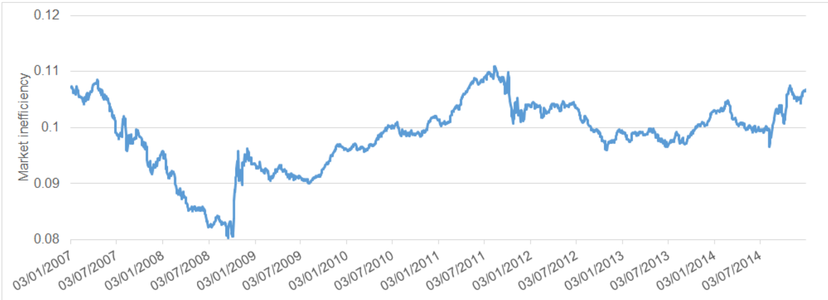

Fig. 5 Market (in)efficiency index

Model II: Market efficiency vs. Trading related causalities.

$$EI(t) = \beta_0 + \beta_1 EI(t-1) + \beta_2 I_{R\to R}(t) + \beta_3 I_{S\to R}(t)$$

In both models, we use the lag-1 efficiency index as a control variable. This is to ensure the validity of the linear regressions as the market efficiency is highly autocorrelated. The Model I involves all information entropy even though some are not directly associated with market movements. The result of this model is validate our understanding of causalities among market and news information we interpreted throughout this paper. The two trading related causalities (i.e. $I_{R\to R}$ and $I_{S\to R}$) show strong and significant linear relationship with the market efficiency (see Table 2). The coefficient $I_{S\to S}$ is also significant, however, the relation is much weaker. It indicates the "news of news" somehow triggers trading reactions in the market, but not as strong as the "first-hand" information. Apparently, the $I_{R\to S}$ is the most irrelevant to trading activities and, as expected, the coefficient is not significant.

Table 2 Market inefficiency vs. Information entropy

Dependent variable	Model I $EI(t)$	Model II $EI(t)$
Const.	0.0223*** (21.955)	0.0202*** (21.317)
$EI(t-1)$	0.8009*** (87.906)	0.8182*** (95.131)
$I_{R\to R}(t)$	−0.1779*** (−10.463)	−0.1793*** (−13.318)
$I_{S\to S}(t)$	−0.0035*** (−5.469)	
$I_{S\to R}(t)$	0.1945*** (7.556)	0.1969*** (7.613)
$I_{R\to S}(t)$	0.0608 (0.923)	
F-statistic	2994.***	4910.***
Adj. R-squared	0.881	0.880
Number of observations	2014	2014
Residual Degrees of Freedom	2009	2011

Significance level code: 0 '***' 0.001 '**' 0.01 '*' 0.05.

We further examines the two trading related information entropy in Model II. The coefficients are similar to those in the Model I. It is clear that the impacts of these two trading strategies are opposite. The increasing $I_{R\to R}$ is associated with higher efficiency. On the contrary, the increasing $I_{S\to R}$ is paired to lower efficiency. We think this could be explained by different ways to use information. We have mentioned that technical trading would lead to non-linear self-causality in market information. However, the reactions to news information are usually linear; positive news should cause price increase and vise versa. Therefore, $I_{R\to R}$ does not necessarily mean predictability of price movements as we cannot tell which direction it is. Consider the fact that technical analysis has been broadly explored for decades, we think this result shows technical trading would facilitate the price discovery using market information. This argument is consistent with the EMH in terms of weak form efficiency. On the other hand, $I_{S\to R}$ is associated with predictability of news to market and indicates the news information is not fully revealed. Hence, the coefficient of $I_{S\to R}$ is positive. It also tells sufficient use of news information would kill the predictability of this causality and improve the market efficiency. This is another argument that is consistent with the EMH in terms of semi-strong form efficiency.

6 Conclusions

The objective of our study is to find the information causality in the intraday market. We follow the EMH so that market data and news data (i.e. public information) are selected in this analysis. We find that the self-causality of both types of information declined during the entire sample period. In terms of news data, this is an indication of more diversified news publishers. On the other hand, the self-causality of market data tend to converging after 2012. We think this shows there is no overreaction in the market. The changing of cross-sectional causalities match the financial crisis. In both the 2008 global crisis and the Euro Dollar crisis, we observe increasing news followed up financial market condition and more market movements caused by updating news. Both causalities diminished after the crisis. Firstly, traders are less sensitive to breaking news under good market condition. Second, and more importantly, the news data is not fully investigated by traders.

To examine the EMH in the modern, fast moving financial markets, we also highlight two regression models of the market (in)efficiency index. In the first model, we apply all causality measures as dependent variables and find out that their associations with the market efficiency are all linearly significant apart from the causality from market to news. To confirm traders' use of information in price discovery would improve market efficiency, we run a second linear model that only involve the self-causality of market data (i.e. traders' use of past market data) and causality from news to market (i.e. traders' use of public information). The coefficients are consistent with that in the first model and are extremely larger than the coefficient of self-causality of news.

In summary, we verify that the good use of information can improve market efficiency. More precisely, sufficient use of market data based on technical analysis would improve the market efficiency. We think the market is a bit conservative according to the relative low level of market to market causality. News has been overused during the financial crisis as it was an alternative to market data. We think that was a special case caused by the unreliable price movements. In general, news has not been well explored and there is no sign of overreaction after the financial crisis.

References

1. Barnett, L., Barrett, A.B., Seth, A.K.: Granger causality and transfer entropy are equivalent for Gaussian variables. Physical review letters **103**, 238701 (2009)
2. Basu, S.: The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence. Journal of financial economics **12**(1), 129–156 (1983)
3. Boulland, R., Degeorge, F., Ginglinger, E.: News dissemination and investor attention. Review of Finance **21**(2), 761–791 (2017)
4. Brogaard, J., Hendershott, T., Riordan, R.: High-frequency trading and price discovery. The Review of Financial Studies **27**(8), 2267–2306 (2014)
5. Dimpfl, T., Peter, F.J.: Using transfer entropy to measure information flows between financial markets. Studies in Nonlinear Dynamics and Econometrics **17**(1), 85–102 (2013)
6. Eugene F. Fama: Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance **25**(2), 28–30 (1970)
7. Goss, B.A.: The semi-strong form efficiency of the london metal exchange. Applied Economics **15**(5), 681–698 (1983)
8. Groenewold, N., Kang, K.C.: The semi-strong efficiency of the australian share market. Economic Record **69**(4), 405–410 (1993)
9. Kristoufek, L., Vosvrda, M.: Measuring capital market efficiency : Global and local correlations structure. Physica A: Statistical Mechanics and its Applications **392**(1), 184–193 (2013)
10. Marschinski, R., Kantz, H.: Analysing the information flow between financial time series. The European Physical Journal B **30**(2), 275–281 (2002)
11. Schreiber, T.: Measuring Information Transfer. Physical review letters **85**(2), 461 (2000)
12. Shannon, C.E.: The Shannon information entropy of protein sequences. The Bell System Technical Journal **27**, 379–423, 623–656 (1948)
13. Shleifer, A.: Inefficient markets: An introduction to behavioral finance. Oxford university press (2000)
14. Sutton, S.G., Arnold, V., Bedard, J.C., Phillips, J.R.: Enhancing and structuring the md&a to aid investors when using interactive data. Journal of Information Systems **26**(2), 167–188 (2012)
15. Yang, S.Y., Song, Q., Mo, S.Y.K., Datta, K., Deane, A.: The Impact of Abnormal News Sentiment on Financial Markets. Journal of Business and Economics **6**(10), 1682–1694 (2015)