



HAL
open science

Investors' attention and information losses under market stress

Dionisis Th Philippas, Catalin Dragomirescu-Gaina, Stéphane Goutte, Duc Khuong Nguyen

► **To cite this version:**

Dionisis Th Philippas, Catalin Dragomirescu-Gaina, Stéphane Goutte, Duc Khuong Nguyen. Investors' attention and information losses under market stress. *Journal of Economic Behavior and Organization*, 2021, 191, pp.1112-1127. 10.1016/j.jebo.2021.09.040 . hal-03434918

HAL Id: hal-03434918

<https://hal.science/hal-03434918>

Submitted on 5 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Investors' Attention and Information Losses under Market Stress

Dionisis Philippas^a, Catalin Dragomirescu-Gaina^b, Stéphane Goutte^{c,d,*}
and Duc Khuong Nguyen^{e,f}

^aESSCA School of Management, Paris, France

^bCenter for European Studies (CefES) – University of Milano-Bicocca, Milano, Italy

^cUniversité Paris-Saclay, CEMOTEV, France

^dParis School of Business, Paris, France

^eIPAG Business School, Paris, France

^fInternational School, Vietnam National University, Hanoi, Vietnam

Abstract

The paper proposes a novel point-wise entropy approach to measure the time-varying losses in the value of information that investors associate with market signals, financial and economic indicators, and news. We cast our approach in a Bayesian framework and assume that market agents update their beliefs to incoming signals based on a prior information set. By exploiting the distribution rather than the time-series properties of information signals, our method is able to construct univariate signal-specific, but also composite proxies of information loss, with the latter being more efficient in reducing misleading effects and interpretation errors. As an empirical illustration, we construct information loss proxies for the US equity market from several mainstream information signals and find that the majority of information loss indicators can influence investors' attention, which then intermediates the impact of information signals on market outcomes. Finally, we show that, by relying on composites rather than univariate proxies, market agents can diversify and thus reduce their information losses when interpreting signals associated with the same underlying event.

Keywords: information loss; point-wise entropy; attention; Google search volume.

JEL classifications: C32, D81, G11, G14.

* *Corresponding author:* Université Paris-Saclay, UVSQ, CEMOTEV, 47 bld Vauban 78280 Guyancourt, France. *Email:* stephane.goutte@uvsq.fr.

Email addresses: D. Philippas (dionisis.philippas@essca.fr), C. Dragomirescu-Gaina (catalingaina@gmail.com), S. Goutte (stephane.goutte@uvsq.fr), and D.K. Nguyen (duc.nguyen@ipag.fr)

1. Introduction

On 11 March 2020, the World Health Organization declared the COVID-19 outbreak to be a pandemic. With world stock markets crumbling, the Wall Street VIX fear gauge index closed the next day above 75 – a level not seen for almost 12 years. Most financial investors, who follow this index on a regular basis, understood that the economic consequences will be dire, perhaps similar if not worse than what they could remember from the 2008 global financial crisis when VIX initially traded within similar ranges. Yet, without a clear reference and a context, numbers alone might not always carry sufficient information, nor suggest a clear narrative. Investors' memory is rarely accurate, so sometimes data would just lack sufficient power to convey a message and help formulate a decision. As a matter of fact, after a terrible 2008, on 20 January 2009, VIX spiked one last time trading above 57 – or almost five sigma above its long-term average. However, this level by itself foretold investors little about the underlying causes of market stress, especially on a day marked by the festive inauguration of the U.S. president Barack Obama. In many similar situations, investors would then increase their reliance on rumours, narratives and contextual information before arriving at a decision, making therefore markets more vulnerable, less efficient and subject to sudden reversals.¹

The flow of information plays an important role for price discovery in financial markets, as it alters market agents' behaviours, and therefore market outcomes. The importance of information is perhaps greatest during market turmoil, when multiple news agencies simultaneously report the same underlying events, with varying degrees of detail. This generates an enormous increase in information signals that become difficult to process in real-time due to market agents' limited cognitive capacity and memory load (Sims, 2003; Veldkamp, 2006). Accordingly, interpretation errors and misleading effects can easily arise when analysing few, unclear or irrelevant information signals, and particularly when lacking a proper reference framework for interpreting incoming data. In this paper, we propose a novel approach that enables us to quantify market agents' information

¹ Tetlock (2011) shows that *stale news* (describing past, redundant, or irrelevant firm-specific events) generate market reactions that revert faster than it is the case for stories containing new (i.e., relevant) information.

losses when evaluating new information arrivals, and then produce a signal-specific and scalable metric derived from any type of data.

Assigning value to new information arrivals is a challenge because it is not always possible to evaluate all various types of signals (e.g., data, headlines, news, indicators) rationally, nor to accurately measure their impact on the decision-making process of different market agents. We take a Bayesian perspective and assume that, when such a signal hits the market, all agents (e.g., individual and institutional investors, portfolio managers, etc.) update their information sets, given some priors. Some might decide it is best to act sooner rather than later, thus disregarding any potential gains associated with information analysis, to instead follow the market momentum, even if this means ignoring their own private information. However, others might decide to act only on the basis of a more complete set, such that misleading effects are minimized, although waiting could worsen their market timing performances.² The quality of information processing can therefore influence investors' attention (i.e., cognitive search effort), belief heterogeneity and market volatility.³ We study these consequences in the empirical section, after developing the methodology for our new information loss proxy.

We build a mapping to quantify losses in the value of information as a time-varying point-wise entropy metric based on a tolerance window. This metric has the ability to evaluate information losses as a result of analysing market signals, media news, or various press releases related to financial markets and economic activity. It also tracks the effect of information through time, without requiring strict assumptions about the sentiment (positive or negative) embedded in the news. More importantly, it applies to any type of investor with limited cognitive abilities. We also extend our mapping to build a time-varying composite indicator that can reflect the losses in the

² Misleading effects here directly relate to incomplete and/or inaccurate information, which instead might allow faster reactions – thus generating an accuracy-speed trade-off. For a theoretical treatment of this trade-off, see Steiner et al. (2017), Fudenberg et al. (2018); for an empirical application, see Dragomirescu-Gaina et al. (2021).

³ Theoretical papers such as Veldkamp (2006), Andrei and Hasler (2015), and Kacperczyk et al. (2016) establish that the quality of investors' information processing mediate the impact of news on asset prices and market volatility.

value of information extracted from different information flows. For this, we use an aggregate portfolio framework applied to individual information losses indicators, based on the existing dynamic correlation pattern among them.

Our paper makes at least four contributions to the literature. Firstly, it proposes a novel and scalable methodology that can be used for constructing information loss indicators based on a variety of signals. These signals are retrieved from data available to all investors on a real-time basis, such as market trading data, news and announcements. Secondly, based on portfolio theory, it presents a simple way to aggregate univariate information losses into composite indicators, which are more efficient in reducing misleading effects and interpretation errors. Thirdly, in an empirical application to the U.S. data, it shows that information loss patterns overlap with well-known historical market stress periods. Finally, it analyses the importance of information losses for investors' attention and asset price dynamics, offering an empirical counterpart to theoretical models developed in Andrei and Hasler (2015) or Kacperczyk et al. (2016).

The rest of the paper is organized as follows. Section 2 describes the conceptual framework and presents our methodology for constructing an entropy metric of information loss. Section 3 develops the information loss indicator for the U.S. equity markets, thus illustrating our empirical framework together with a discussion of the main findings. Finally, Section 4 provides our concluding remarks. More detailed results can be found in the accompanying Supplement.

2. Conceptual framework

The true value of information that market agents use to validate their actions cannot be directly perceived but only be inferred indirectly, subject to a loss. This is the main purpose of the paper: to propose a proxy able to summarise (albeit from a mathematical perspective) the loss in the value of the 'information' that market agents extract from various signals (e.g., news, firm-specific announcements, macroeconomic events, and other data). Since our proxy is signal-specific, it is scalable and can be easily applied to any type of information flow data. As discussed above, this is not an easy task because not all the potential signals and indicators convey the same amount of useful 'information' to each and every investor.

Our proposed proxy is based on an entropic transformation that downplays the time-series properties of the initial signal (e.g., autocorrelation, stationarity frequency, and clustering) to focus instead on its data distribution properties. Accordingly, there is no need to differentiate market agents into different types (e.g., informed vs. uninformed; retail vs. institutional) in order to explain why some market phenomena prevail at some particular times, given some realised signal. In fact, we show in the empirical section that our concept can explain, albeit with varying degrees of confidence, how market agents' attention comes to influence the cross-sectional dispersion (i.e., belief heterogeneity) of asset returns. The only assumption we need though to support such a (mathematical) transformation of the original information signal is one about market agents using Bayesian updating – a recursive process that represents an essential ingredient in the empirical modelling of decision-making under uncertainty.

Large swaths of the early empirical literature focus on market agents' reactions to news, headlines, and financial information along with the subsequent dynamics of stock prices, volatility, trading activity and/or market liquidity (see Zhang, 2006; Mackowiak and Wiederholt, 2009; Solomon, 2012). However, many of these studies rely on market-based signals such as trading volume, share ownership, implied volatilities or fund flows, simply because they are available on a high frequency basis. The problem in such cases, as Baker and Wurgler (2007) observe, is that these metrics and the corresponding approaches suffer from being the equilibrium outcome of many factors other than just information signals and market agents' reactions. A handful of theoretical papers clarify these aspects by exposing the underlying mechanisms that intermediate the impact of optimal information set on markets and asset prices (Veldkamp, 2006; Woodford, 2014; Kacperczyk et al., 2016; Andrei and Hasler, 2015). Using Bayesian frameworks, these models employ rational inattention as a more suitable concept to explain investors' choices for risky asset holdings, which ultimately determines financial market equilibrium outcomes, like volatility and returns.

In a similar vein, we develop our entropic metric under a Bayesian framework through which market agents update their prior beliefs based on incoming information signals. Since market agents' attention (a by-product of this updating process) mediates the impact of the initial

information signals onto market outcomes, we will empirically estimate a system of equations to reflect this transmission mechanism. We use a market-wide attention proxy as in Andrei and Hasler (2015), a choice that is based on a large and growing literature strand that measures investors' attention using indicators based on search volume indexes from Google Trends, where data are organised according to some user-defined keywords. Such proxies are found to be significantly related to volatility patterns displayed by various financial assets and markets (see, e.g., Engelberg and Parsons, 2011; Da et al., 2011; Drake et al., 2012; Vlastakis and Markellos, 2013; Hamid and Heiden, 2015; Dimpfl and Jank, 2016; Hsieh et al., 2020). Hamid and Heiden, (2015) use *empirical similarity* to show that investors' attention (proxied by Google search volume) helps to predict market volatility.⁴ Hsieh et al. (2020) show that Google searches for specific firms (particularly small caps) are positively associated with herding behaviour and thus higher returns' dispersion. Da et al. (2011) use Google search volumes on ticker symbols as a proxy for firm-specific investor attention, showing that Google searches are positively associated with market capitalization, abnormal returns, turnover, and media attention. Finally, Drake et al. (2012), Dimpfl and Jank (2016), and Perlin et al. (2017) also use Google search volumes (and indicators built on such data) to proxy instead for information demand coming from market agents looking for additional inputs, and link it to market outcomes such as volatility, trading activity and liquidity.⁵

2.1. A mathematical formulation of information loss

Consider an information flow within a distributed system, denoted by an information signal s_i , with i being a sequence of consecutive time cells (of equal size for simplicity).⁶ Absent any significant news arriving at the i^{th} time cell, a market agent decides her strategy based on the past

⁴ *Empirical similarity* is a concept that allows one to predict a variable based on a weighted average of *all* its past observations.

⁵ Drake et al. (2012) acknowledge that investors' attention and information demand are conceptually very similar (see page 1008, footnote 5).

⁶ A distributed system is a collection of states of affairs that are able to carry information about each other. The flow of information within a distributed system is defined as the class of all its information transfers with respect to a certain process in a specified time period (Barwise and Seligman, 1997).

(aggregate) information set, available at time cell $i - 1^{th}$. Now, suppose instead the market agent receives a new significant information signal. To retain simplicity, we assume that the market agent frames her decision-making process within a *subjective expected utility* (SEU) setting (Savage, 1954). She will use a Bayesian process to update her information set and based on the prior information set, profile, preferences, and the interpretation of the new signal, she will decide her investment strategy for the next period (i.e., action or inaction). Accordingly, her problem now refers to updating prior belief distribution to a posterior information set at the i^{th} time cell, conditional to the importance of new signal. However, there can be deviations from the rational decision-making process described above, deviations arising either in relation to some features of the information signal or in relation to agent's preferences and profile.

We theorise that the parameters of interest are given via a loss function, denoted as $L(r, v; \theta_{s_i})$,⁷ where r is a vector of the characteristics related to returns, v is a vector of the characteristics related to volatility and θ_{s_i} is a parameter vector of interest related to information signals. Hence, rational agents can update θ_{s_i} conditional to information arrivals. Assume that $p(\theta_{s_i})$ represents prior beliefs about the parameters θ_{s_i} and, hence, a Bayesian posterior or updated beliefs about the parameters can be given as⁸:

$$p(\theta_{s_i}|r, v) \propto p(\theta_{s_i}) \exp(-L(r, v; \theta_{s_i})) \quad (1)$$

where $p(\theta_{s_i}|r, v)$ is the updated market agent's subjective beliefs (posterior) which is an action made under uncertainty.

With respect to the information signal, an emphasis should be placed on the sequence s_i because it may contain several (un-)observable elements that affect the posterior. These elements relate to some intrinsic features carried by any information signal (Eaton and Bawden, 1991), such as irregular/regular arrival frequency, gradual/fast intake or perception by market agents, dynamic/static nature of the evaluation process, and other qualitative aspects (economic/non-

⁷ This assumption is in line with Bissiri et al. (2016) and Dragomirescu-Gaina et al. (2021) who use loss functions to connect information in the data to functionals of interest.

⁸ See also Andrei and Hasler (2015) and Kacperczyk et al. (2016) for Bayesian updating.

economic topics, policy/politic actions, etc.). These features may generate different trade-offs between the content and the value of new information and thereby affecting the decision-making process, which makes a simple scaling (e.g., good versus bad news) for the sequence s_i insufficient. A higher precision in the evaluation of the signals is needed to improve the price discovery process. In this setting, features such as *relativity* (the case that the same state of affairs might carry different pieces of information to different agents) and *fallibility* (the case that a state of affairs does not carry the information it is supposed to carry) of the s_i sequence will affect the rational process of evaluating information (Pérez-Montoro, 2007).

With respect to market agent features, we cannot assume homogeneity in preferences and characteristics. Rational and irrational agents interact dynamically, cognitive phenomena⁹ may arise any time period, and market agents form beliefs, preferences, or make decisions given their beliefs and preferences over time, with a substantial impact on price discovery. Nevertheless, every market agent has a finite memory, limited cognitive capacity to process information, as well as limited attention to devote to new signals in a given time cell (Sims, 2003); therefore, she may need to process a high volume of new signals for the market of her interest, depending on priority and urgency.

Moreover, a market involves both retail and institutional market agents, which play different roles. Retail agents are predisposed to react emotionally on new information that have first caught their attention and might be too sensitive to a specific financial stress event and/or real economy change; therefore, they can reallocate their attention over time in order to extract relevant information from diverse and noisy sources, including market moves, asset prices, news and rumours, etc. Institutional agents have a more structural way of analysing new signals and interpreting a broader information set; therefore, they might digest new signals with an air of scepticism because they are just confident in their strategy, and, therefore, unlikely to be led astray by new market headlines and news. However, with respect to market agent's type and her information processing, the available literature focuses on two main concepts. The first relevant

⁹ There is a plethora of cognitive biases such as conservatism, representativeness, anchoring, ambiguity aversion, and so on, that play a key role in pricing anomalies and thereby on decisions and actions.

concept is the *informational cascade*, which occurs when a market agent observes the actions of others and then, despite possible contradictions in her own private information signals, engages in the same acts inferred based on other agents' actions (Easley, 2010), causing irrational herding. The second relevant, but less common concept is the *hipster effect*. According to Challet et al. (2000), upon information arrival, all market agents are interactive and have specific preferences for taking their decisions either in accordance (i.e., consensus) or against the mainstream (i.e., hipsters). However, in the end all agents look alike (*representative agent*) after incorporating the delays necessary for information to be evaluated. Touboul (2019) designs general models featuring both conformists and hipsters (or anti-conformists) who interact in an environment where information processing delays shape the collective dynamics. Even when all individuals are faced with binary choices, because of the delays in information processing these models display non-stationary solutions, with synchronizing dynamics oscillating between different states (equilibria).

Within our framework, we see the concept of *efficient market hypothesis* (EMH) to be dynamic (rather than a binary one), subject to the value of information. When new information arrives, the evaluation process has an interpretive effect on the market and will be reflected on prices with a lag, in such a way that market efficiency is the frontier of optimal information value. For example, if hipsters engage in a significant number of trades, then a market can become efficient over time, in a dynamic way, leading to higher liquidity and market turnover.

We start by assuming that the importance of the information signal itself (i.e., the *pure information*), is not as high as the value a market agent would attach on this information signal based on her preferences and objectives. Accordingly, we consider the *value of information* (*VoI*)¹⁰ as being the evaluation of the information signal in time cell i^{th} , along three important properties. First, a market agent should consider all possible outcomes along with their relative likelihood (Schepanski and Uecker, 1984). Second, a market agent should update her beliefs after receiving new information, but her decision needs to be taken based on *VoI*. Third, the process of analysing information signal is always subject to a risk that some relevant bits of information get lost or

¹⁰ The term was initially proposed by Howard (1996), according to whom *VoI* is the amount a decision maker would be willing to pay for information prior to making a decision.

disregarded in the decision-making process, a risk that we call in this paper as *loss in the value of information* or simply *information loss* and denote it as *VoL*. Therefore, *information loss* is the risk associated with analysing and incorporating an information signal in the decision-making process.

We also define the concept of *expected loss* in the value of information in the i^{th} time cell as the amount of information loss that an agent would be willing to accept, when facing a trade-off between more accurate information signals (with lower uncertainty), and less accurate signals (with higher uncertainty). This expected loss is set by the market agent at the $i - 1^{\text{th}}$ time cell and it refers to the amount of information loss the agent stands ready to lose in the next time cell i . When evaluating a signal, a market agent will disregard other aggregate signals as long as she faces an information loss lower than expected. However, if the evaluation of any new signal is too costly in terms of reasoning or cognitive capacity, and thereby triggers higher uncertainty, she would then retrieve and evaluate more signals (if possible) for the same event (i.e., thus generating a higher attention and demand for information), in line with Pérez-Montoro (2007). Hence, she will weigh potential information losses – which can be associated with different information signals – according to her explanatory power and reasoning capacity. This implies that the *VoL*, and thereby the *information loss*, might be diversified in a context of multiple higher/lower misleading effects over time. We call this concept as *diversifying information losses*.

Note that information signals and market agents are both viewed as interactive players in our framework.¹¹ Information signals can be derived/provided from a wide range of diverse sources (e.g., databases, newspaper articles, newswires, etc.). A market agent may incorporate new information into her objectives acting upon with a conservative/momentum bias based on her evaluation, priority and urgency (Gutierrez and Kelley, 2008). However, she will use information arrival to an extent that might endogenously determine her own beliefs, since there are cases where the emergence of new information may not cause any market reaction for various reasons, e.g.,

¹¹ A separate strand in the recent literature investigates the relationship between uncertainty or market volatility on the one side, and investors' attention or behaviour with respect to information acquisition on the other side (see Andrei and Hasler, 2015; 2019; Woodford, 2014; Steiner et al., 2017; Dragomirescu-Gaina et al., 2021).

liquidity restrictions, costs, regulation restrictions, etc. Formulating this upon the same basic model structure and the previous information set, denoted as s_{i-1} , let us consider the loss function $L^{(a \text{ or } b)}$, as follows: (i) L_a is an approximation of the market agent's loss function under information signal taken to be the current period; and (ii) L_b is the approximation of market agent's loss function under *expected information loss*; therefore, updated beliefs are given as:

$$p(r_i, v_i, \theta_{s_i} | s_{i-1}) \propto p(r_{i-1}, v_{i-1}, \theta_{s_{i-1}}) \exp[L^{(a \text{ or } b)}(r_i, v_i, \theta_{s_i})] \quad (2)$$

In Eq. (2), we update information conditional on the structural parameter (θ_{s_i}), a prior $p(\theta_{s_i})$ and a subjective loss function L , with regards to the loss in the value of information.

The information loss of new signals may increase the importance of misleading effects, since market agents have a limited analytical and cognitive capacity (explanatory power).¹² This can be different for each market agent or group of similar agents, even when it is associated with the same information. If a market agent could remove all the uncertainty while evaluating an information signal, the value of such *perfect information* would provide a benchmark (i.e., *risk-free VoI*) in the evaluation of new signals. However, in real-world, such a benchmark does not exist but only incentives to retrieve more information (i.e., based on all available information supplying sources), making the market efficiency hypothesis a dynamic problem rather than a binary one (e.g., efficient or inefficient). Hence, the higher the difference between her *expected* and the *actual information loss* in the time cell i , the greater the misleading effect. As a result, she might choose to access different information flows to reduce the risk of being misled, thus contributing to a higher information demand (Moscarini and Smith, 2002).¹³ Motivated by the studies of Da et al. (2011) and Verona (2014), we assume that market agents are acquiring information signals which are already in the public domain (allowing a *hipster effect*). Since there is less clear evidence that large price swings always happen instantly after an important (financial stress) event or immediately

¹² This can occur because a market agent might face unexpected news in terms of content, topic, nature, a high frequency of information signals, or other similar challenges for her limited reasoning capacity.

¹³ Moscarini and Smith (2002) argue that information demand is a decreasing function of the informational content of a signal. When a stress event occurs, the uncertainty with respect to its consequences is greater and agents demand more information.

following a press release (Brunnermeier and Nagel, 2008; Drake et al., 2012), we can assume that the accumulation of news or results from analysing overlapping news can have a greater misleading effect over time.

2.2. Information loss indicator

Consider an information signal s_i for a *representative market agent* with s_{i-1} being her previously available information set, formulated at time cell $i - 1$, with $i = 1, 2, \dots, N$ being equally spaced in time. Based on the conceptual framework in Eqs. (1)-(2), the loss in the value of information, denoted as h_i , associated with the signal s_i , is a two-dimensional vector function, defined as:

$$h_i = [\varphi_i^0, \varphi_i^1] \quad (3)$$

where φ_i^1 is a signal concerning the loss in the value of information in the time cell i , and φ_i^0 is a signal concerning the expected information loss for the $i - 1^{\text{th}}$ time cell. Equation (3) states that the loss in the value of information is determined by the agent who chooses the precision of the signals and faces a loss in evaluation due to her limited cognitive or reasoning capacity.

We propose a point-wise entropy measurement that quantifies the loss in the value of information of a market agent in the time cell i . Entropy has been mostly used in order to represent the disorder of a system, e.g., as an optimal allocation of attention and the dynamics of prices depending on the firms' environment (Mackowiak and Wiederholt, 2009) or as a measure of rational inattention (Sims, 2003). However, accurate entropy calculation requires vast amounts of data and the results will be greatly influenced by system features; therefore, it is not practical to apply the typical entropy measure to experimental/real data but to use alternative entropy measures in order to quantify the amount of irregularity and the unpredictability of fluctuations in time-series data. Our proposed point-wise entropy measure is built in four steps, as we illustrate below.

- Step 1:

The first step follows the approximate entropy measure initially proposed by Pincus and Goldberger (2004). Form the information signals as $u(1), u(2), \dots, u(N)$, with u being a scalar.

Consider an integer m that represents the length of compared run data. Next, form a sequence of vectors $x(i) = [u(i), \dots, u(i - m + 1)]$ in \mathfrak{R}^m , i.e., in the real m -dimensional space with $i = 1, 2, \dots, T$ and $T = N - m + 1$. Also consider a positive real number r that specifies a filtering level.

- Step 2:

For each vector $x(i) \equiv \mathbf{x}_i$, calculate the distances as the maximum norm, to all the other vectors have been formed. Denote these distances as $dist(\mathbf{x}_i, \mathbf{x}) = \max|\mathbf{x}_i - \mathbf{x}|$.

- Step 3:

Let us consider all the nearest neighbours of \mathbf{x}_i which are less than r , satisfying the rule $dist(\mathbf{x}_i, \mathbf{x}) < r$. The parameter r is set by the user and it represents the scale of the nearest neighbourhood (i.e., the range of information scalars or a percentage of standard deviation, and so on). We save their indices j , as a $\{ j \}$ sequence.

- Step 4:

As last step, we calculate the point-wise entropy measure. Consider the distances $\mathbf{d}_0 = dist(\mathbf{x}_i, \mathbf{x}_j)$. The elements of \mathbf{d}_0 are the distances of the nearest neighbours of \mathbf{x}_i in the chosen range, saved in a vector \mathbf{d}_0 . Then, track the mappings of \mathbf{x}_i and the set \mathbf{x}_j to \mathbf{x}_{i+1} and \mathbf{x}_{j+1} , respectively. Calculate the distances $\mathbf{d}_1 = dist(\mathbf{x}_{i+1}, \mathbf{x}_{j+1})$. Finally, the point-wise entropy is quantified as:

$$h_i = \varphi_i^1 - \varphi_i^0 = \log(\langle \mathbf{d}_1 \rangle) - \log(\langle \mathbf{d}_0 \rangle) \quad (4)$$

where $\langle \cdot \rangle$ denoting the mean.

Equation (4) measures the information loss at the time cell i as a point-wise entropy metric in the value of information, representing the generalized error extracted from the value of information. It affects all the latent factors that determine this evaluation. In the time cell $i - 1$, the agent adjusts the precision of the historical signals so as to determine the expected information loss. In the time cell i , she receives the signal and takes the optimal decision given the evaluation of the signal that she has received. The trade-off is that a higher loss makes the price setting behaviour worst because it requires paying more attention to conditions and evaluation. Generalizing the point-wise entropy measure for all the cells over time, the univariate information loss indicator h_i shows the time-

varying measure for all sub-sequence of cells, with the subscript i being a time sequence, under the simplifying assumption that the market agent has chosen a signal precision for every time cell i , and consistent with the same information flow.

2.3. A composite indicator of information losses

In reality, market agents are able to evaluate signals from multiple information sources; hence, a *composite information loss* indicator would reflect the aggregation of multiple information loss components, when based on an efficient weighting scheme. Consider an allocation of k information signals. Generalizing Eq. (1) and (2), the allocation of information signals is given by:

$$p(r_i, v_i, \theta_{s_{i,j}} | s_{i-1,j}) \propto p(r_{i-1}, v_{i-1}, \theta_{s_{i-1,j}}) \exp[L^{(a \text{ or } b)}(r_i, v_i, \theta_{s_{i,j}})] \quad (5)$$

with $j = 1, 2, \dots, k$.

The allocation of respective individual time-varying information loss indicators h_i , for every different information signal j that can also allow different filtering levels, now takes the form:

$$h_{i,j} = (h_{i,1} \quad \dots \quad h_{i,k}) \quad (6)$$

where $h_{i,j} = [\varphi_{i,j}^0, \varphi_{i,j}^1]$

Conceptually, the composite indicator follows the same principals as the univariate case. However, the part that requires a detailed discussion is the weighting scheme of individual information loss indicators capturing their relevance to the overall loss. Different market agents have different preferences on evaluating signals from different information signals, but they all still have limited cognitive capacity. In addition, there are other determinants that may affect aggregate weighting such as the ambiguity, possible correlation or conflict between signals (even for the same event), the unexpected nature of some events (e.g., political risk), the repetition or frequency of the signals (e.g., media channels reproducing the same news) and so on. Consequently, the market agent will try to retrieve and evaluate a portfolio of information signals; however, the problem that arises is if she can weigh the relevant importance of every signal and thereby its information loss, forming a portfolio for which she can diversify the information losses in order to obtain lower misleading effects.

However, there is no one single way to form the portfolio weights; for example, a simple way would be to consider an equal weighting scheme. In this paper, we consider a straightforward method to aggregate the univariate information loss indicators. We apply the minimum variance portfolio strategy as in DeMiguel et al. (2009), who pointed out that minimum variance portfolios are not necessarily optimal portfolios, but they may be suitable for market agents wishing to load up on low-risk or ‘defensive’ assets. In this setting, every information loss indicator is treated as a unique asset aggregated within a portfolio of information losses. This portfolio is determined by the allocation $w_j h_{j,i}$ for $j = 1, 2, \dots, k$, where w_j denotes an importance weight. This allows every information loss indicator to carry either more or less significance relative to another one, thereby allowing the quantification of their relevant *holistic dependence*. The portfolio weights are determined on the basis of a *minimum variance portfolio optimization*, defined as $\min(w_j' \Omega w_j)$ where Ω is the variance-covariance matrix. The composite indicator of information loss is denoted as $h_{i,CI}$ and it is given by the following equation:

$$h_{i,CI} = (w_j \circ h_{i,j}) \Lambda_i (w_j \circ h_{i,j})' \quad (7)$$

where the term $w_j \circ h_{i,j}$ represents the Hadamard-product¹⁴ in time cell i , and Λ_i is the matrix of the time-varying cross-correlation coefficients, ρ_i , between pairs of individual information loss indicators, allowing time-varying dependence in every time cell. The matrix Λ_i takes the form:

$$\Lambda_i = \begin{bmatrix} 1 & \rho_{i,12} & \cdots & \rho_{i,1k} \\ \rho_{i,21} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{i,k1} & \cdots & \cdots & 1 \end{bmatrix} \quad (8)$$

Here, again, there are several ways to calculate the time-varying correlations ρ_i in matrix Λ_t and it is up to analyst’s intuition. In this paper, we calculate the time-varying correlations ρ_i with exponentially weighted moving average (EWMA) of respective covariances $cov_{i,j=1,\dots,k}$ and volatilities $\sigma_{i,j}$ and $\sigma_{i,\neq j}$:

$$cov_{i,j=1,\dots,k} = \lambda \sigma_{i-1,j} + (1 - \lambda) h_{i,j} h_{i,\neq j} \quad (7)$$

¹⁴ The Hadamard-product is the element-by-element multiplication of the sub-index weights vector with the vector of sub-index values in time i .

$$\sigma_{i,j}^2 = \lambda\sigma_{i-1,j}^2 + (1 - \lambda)h_{i,j}^2$$

$$\sigma_{i,\neq j}^2 = \lambda\sigma_{i-1,\neq j}^2 + (1 - \lambda)h_{i,\neq j}^2$$

$$\rho_i = \text{cov}_{i,j=1,\dots,k} / (\sigma_{i,j}\sigma_{i,\neq j})$$

This structure of the composite indicator comes with some major advantages. First, a potential repetition of information signals is a strong argument to use time-varying correlation structures. A time-varying ρ_i shows the dependency at any time cell between two information loss indicators.¹⁵ Market agents cannot always filter out media “noise” rationally when making decisions and they can easily be led astray by news reproduced by “TV experts”.¹⁶ Secondly, our composite indicator can combine signals arriving from different information sources. This is a particular advantage in empirical exercises because it can allow firm-originated press releases, macro-economic announcements and media coverage. Combining this with the overall media coverage, which also provides economic and non-economic news, a market agent can obtain a more accurate picture to reduce her misleading effects. Finally, the composite indicator captures the importance (scale) of aggregate information loss; therefore, it can be also used as an index in analysing market microstructure, market sentiment, market stress, etc. This scaling captures features such as dynamics and levels of information loss (what makes information flow A more or less valuable than B) without depending on confirmatory bias (what happens when A confirms a prior belief, while B might oppose the prior belief and be disregarded).

3. Empirical application

To illustrate the empirical relevance of our new entropy proxy, we apply the methodology presented in the previous section to U.S. equity market data. To this end, this section presents a two-step empirical framework that is based on information loss indicators extracted from various signals,

¹⁵ If all pairs of information loss indicators were perfectly correlated, this would imply a situation in which all information loss indicators stand either at historically low levels or at historically high levels at the very same time. Such situations, however, are still the exception rather than the rule.

¹⁶ This is also in line with Gentzkow and Shapiro (2010), in the sense that the biggest driver of media slant is the audience seeking like-minded news (i.e., confirmatory bias). Some news will capture all the attention (the concept of communicative complexity) of most agents, thus increasing their intensity.

which we discuss in detail in the subsection 3.1, and then graphically illustrate in subsection 3.2. As a first step, we take a simple univariate approach in subsection 3.3 and look at the heterogeneity of equity returns across four representative sectors of the S&P 500 Index. We use a quantile regression model to reveal whether this heterogeneity, which we proxy by the cross-sectional dispersion of returns, is driven by differences in investors' beliefs with respect to how they evaluate several pertinent information signals. Next, as a second step, we employ a multivariate approach in subsection 3.4 to understand the dynamics of investors' attention and how it effects the cross-sectional dispersion of equity returns. We thus see our estimated system as an empirical counterpart of the main theoretical mechanisms outlined in the rational inattention literature (e.g., Andrei and Hasler, 2015; Kacperczyk et al., 2016). Compared to the univariate step, we now expect entropy to be a relevant explanatory factor for investors' attention, which mediates the transmission of information signals to market outcomes.

3.1 Data

Our data comes from different sources and includes weekly observations for a recent period spanning from January 2017 to January 2021.¹⁷ The sample includes several economic, social, and political events that have generated important market reactions (see Appendixes A and B in the Supplement), as well as the first waves of the ongoing Covid-19 pandemic outbreak. Despite a relatively short time span, this recent sample is indicative for the ongoing change in investors' behaviour and attitude with respect to recent global trends in development and integration (see

¹⁷ As we mention in the *Introduction*, our method can be applied to any dataset, and to any information signals available at any frequency. As a matter of fact, the mathematical derivation of the entropy metric refers to "time cells" rather than time-series, in order to highlight that an equally-spaced frequency of the signal is not necessarily required when computing the entropy. There are no trade-offs from using higher/lower frequency data, if available. However, some aspects of the empirical application and the mechanisms might change, such as the proxy for attention and how this affects different market outcomes. For example, with high-frequency data, one might want to focus on news from specific sources and intraday trades that imply a faster decline in information value. Moreover, for individual/retail investors, a different attention proxy can be constructed, based on login/access data (Dierick et al., 2019).

Dragomirescu-Gaina and Philippas, 2021). As a robustness check, in Appendix D from the Supplement, we restrict our main empirical analysis to the pre-COVID-19 sample.

Our first data choice is about the information signals from which we will extract the information loss indicators that we employ in the empirical analysis. Market agents normally digest a large (and often irregular) set of data, news, press releases and financial indicators, which are almost impossible to survey in detail. Moreover, most agents will be dealing with few rather than many information signals in each time cell due to their limited cognitive capacity and memory load. We therefore use several readily available information signals (i.e., indicators) that are regularly reported in financial media and/or are popular among investors and market analysts; in addition, other signals are described in Appendix A from the Supplement, but not reported here due to space constraints.

Based on the methodology proposed in the previous section, we compute the information loss based on each of the signals considered below. To organise the discussion below, we classify them into three broad groups: (a) macroeconomic indicators, (b) market signals and (c) uncertainty indicators.

- a) Macroeconomic indicators: initial jobless claims; weekly values of the GDP tracker for the U.S. available from the OECD (Woloszko, 2020); yields for the 3-Month Treasuries, and the 10-Year Government Bonds; Moody's seasoned Baa corporate bond yields; gold and WTI oil prices (in USD).
- b) Market signals: S&P 500 Index, and its traded volume (in USD); VIX index; the St. Louis *Financial Stress Index* (FSI); the MSCI Index; and Bitcoin prices. As a robustness check, we include the U.S. Government Liquidity index in Appendix D3 from the Supplement.
- c) Uncertainty indicators derived from media sources: *Twitter Market/Economic Uncertainty* (TMU/TEU) from Baker et al. (2020); *economic policy uncertainty* (EPU) and *equity market uncertainty* (EQU) from Baker et al. (2016).

Our two-step empirical analysis centres on the cross-sectional dispersion of equity returns, focusing on 4 main sectors included in the U.S. S&P 500: *information technology*, *industrials*,

utilities, and *financials*. Data on individual stocks is from Bloomberg. *Technology* is one of the retail’s favourite recent targets, while the others are among some of the best sectors favoured by institutional investors; in particular, *Utilities* is considered a defensive sector, while *Industrials* and *Financials* are mostly considered cyclical sectors. We believe these four sectors are representative in a sample dominated by several events that altered the U.S. equity market structure and dynamics, including the substantial growth in retail investors’ market participation post-COVID-19. For each sector, the cross-sectional dispersion of returns (i.e., *cross-sectional absolute deviation* or CSAD) is computed as follows:

$$CSAD_i = \frac{\sum_1^n |r_{stock,i} - r_{m,i}|}{n - 1} \quad (8)$$

where i is a time subscript; $r_{stock,i}$ is the return of an individual stock at i ; $r_{m,i}$ is the market return at i for the corresponding sector; and n is the number of assets included in the sector at i .¹⁸ CSAD is well-known measure of investors’ heterogeneity in beliefs and evaluations that is positively associated with contemporaneous market returns: small during tranquil periods, but considerably higher when market returns exhibit extreme values and investors’ attention increases (see, Chang et al., 2000; Philippas et al., 2020; Hsieh et al., 2020). While taking heterogeneity at full value, our next focus is on what drives this heterogeneity, as measured by CSAD, and not on what would be the possible effects of heterogeneity itself on other market outcomes.

Finally, as a proxy of investors’ attention, we follow Andrei and Hasler (2015) and use the following keywords to extract the search volumes from Google Trends: “financial news”, “economic news”, “Wall Street Journal”, “Financial Times”, “CNN Money”, “Bloomberg News”, “S&P500”, “US economy”, “stock prices”, “stock market”, “NYSE”, “NASDAQ”, and “FTSE”. We then compute the cross-sectional average of all these time-series and use it as our proxy for attention. Two alternative measures yield similar results.¹⁹ Firstly, we employ the first principal component of these time-series instead of the average. Secondly, we use only “stock” as keyword in

¹⁸ Individual returns are calculated by taking the log-differences, while market returns for each sector are calculated as the simple average of the returns of all the stocks.

¹⁹ These results are available from the authors upon request.

Google Trends, since Perlin et al. (2017) find that this indicator alone is the single most important explanatory variable for returns, volatility, and traded volumes in the U.S. financial market (but also in the markets of the United Kingdom, Canada and Australia).

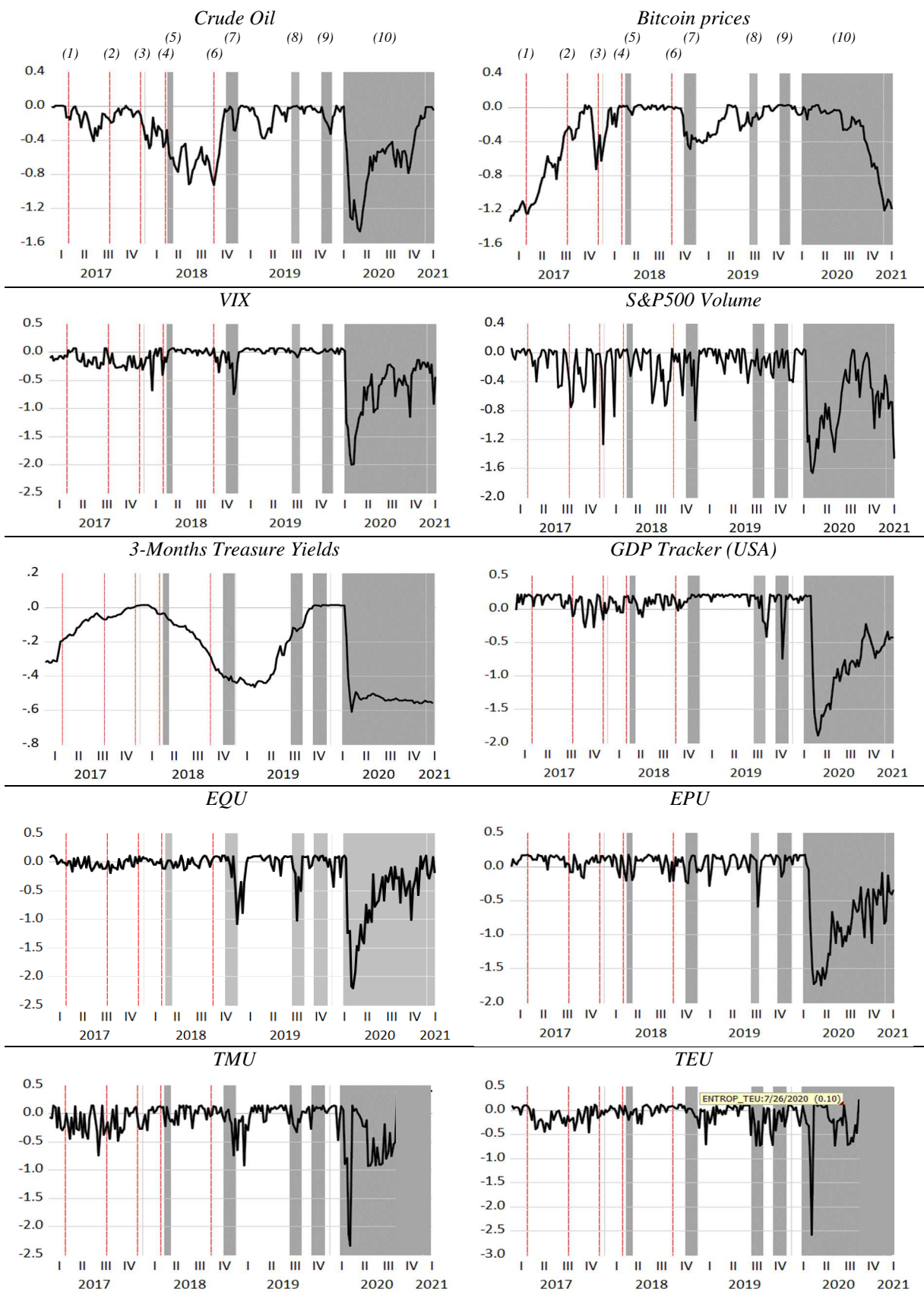
3.2 Data visualisation

An information loss indicator should increase in response to a stress event, which adds more uncertainty and complicates the analysis of its long-term consequences. However, it is not always clear whether a stress event will be associated with an increase in the loss of information value. This is because it is not known a priori whether a potential increase in the flow of information arrivals implies higher losses in the value of information (i.e., more distortions and confusion). While it is easy to identify events that disturb various segments of the financial system, it is not easy to establish which ones had a higher impact, or if there are any overlapping or nonlinear consequences. Hence, it is better to focus on the identification of those periods of risk which are associated both with a specific triggering event, as well as with a gradual build-up that cumulates several information arrivals.

One can visually evaluate the performance of our information loss indicators by simply inspecting their overlapping with some well-known periods of higher uncertainty and financial stress. Figures 1, 2 and 3 present some interesting overlapping (for some selected information loss indicators) with events that have been particularly relevant for the U.S. financial market over the last four years (see the calendar provided in Appendix B). The dynamic pattern of these indicators is broken down by red vertical lines and shadow areas that correspond to specific stress events as indicated by a number, which is then referenced in our calendar. We see a canonical pattern that connects the losses in the value of information with the associated stress observed in various information signals, ranging from macroeconomic to market indicators and uncertainty measures. What is interesting to note here is that for the same market stress event, not all the information loss indicators display the same reaction, suggesting therefore some heterogeneity in the ‘narrative’ that each of these metrics is capturing. Nevertheless, before drawing more insights, we show that these information losses correlate well with investors’ actions, trading and market outcomes.

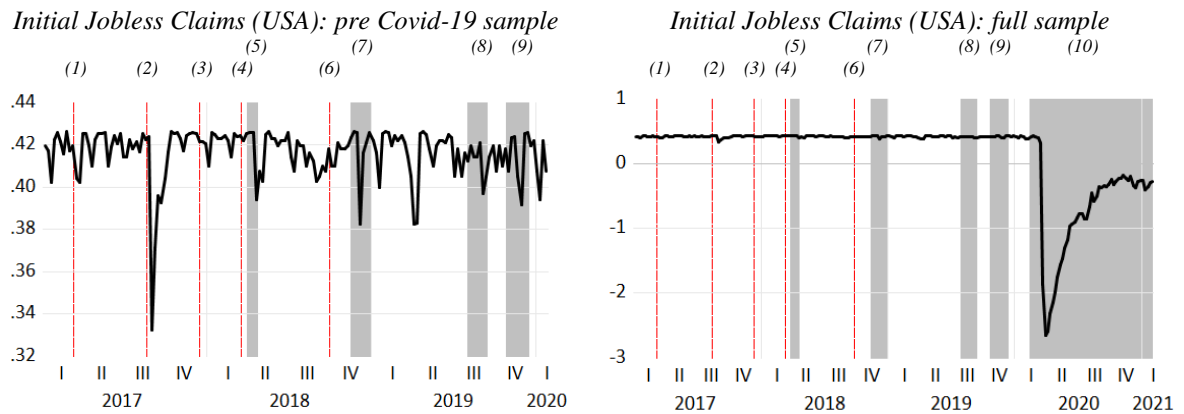
Our information loss indicators are also robust to a change in sample. Given the huge jump recorded in U.S. initial jobless claims in March 2020, we check whether our methodology is sensitive to outliers. As can be seen in Figure 2, there are important spikes in entropy associated with some specific events (described in the calendar) even before the COVID-19 pandemic. However, once a large jump in reported claims occurred, investors changed their reference point, adjusting the distribution accordingly.

Figure 1. Information loss indicators derived from specific information signals



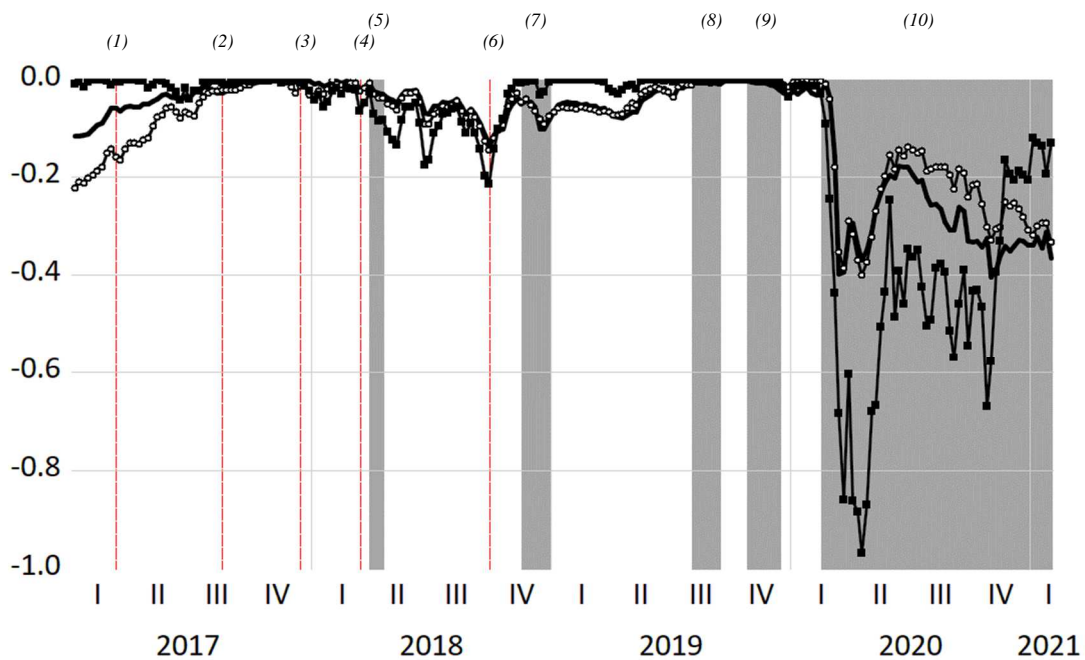
Note: The red and black vertical lines correspond to specific events mentioned in the calendar given in Appendix C of the accompanying online Supplement. Sample is January 2017 – January 2021, at a weekly frequency.

Figure 2: Before and after COVID-19 pandemic



Note: This figure presents the dynamic patterns of information loss indicators built based on the initial jobless claims reported data, for before COVID-19 and for the full sample. The red and black vertical lines correspond to specific events mentioned in the calendar given in Appendix C of the online Supplement. The effective sample is January 2017 – January 2021, at a weekly frequency.

Figure 3: Composite information loss indicators



Note: The red vertical lines and grey shadow areas correspond to specific events mentioned in the calendar given in Appendix C of the online Supplement. Sample is January 2017 – January 2021, at a weekly frequency. The following composites are illustrated. Composite 1 is illustrated with a black line and aggregates individual entropy indicators for 3-Month Treasury yields, S&P 500 Index and WTI oil prices (in USD). Composite 2 is illustrated by the black line with circles and aggregates individual entropy indicators for 3-Month Treasury yields, S&P 500 Index, WTI oil prices and Bitcoin prices. Finally, Composite 3 is illustrated by the black line with black squares and aggregates individual entropy indicators for WTI oil prices, gold prices, and VIX.

Figure 3 illustrates three composite indicators of information losses computed by aggregating the following market signals: 3-Month Treasury yields, VIX, S&P 500 index, Bitcoin prices, gold, and

oil prices but in various combinations, with optimal weights derived from portfolio theory. More technical details on the construction of these composite indicators are provided in Appendix A from the Supplement. As can be seen in Figure 3, all these composites are sensitive to several market stress events described in our calendar.

3.3. Univariate approach: Quantile regression analysis

In the first step of our empirical framework, we use several information loss indicators to explain the dispersion in stock returns across four representative sectors of the US equity market, i.e., *information technology*, *industrials*, *utilities*, and *financials*. We model the dynamics of (log) CSAD in each sector using a univariate quantile regression model, which is more robust to potential outliers such as those generated by investors' heterogeneous expectations with respect to particular stocks within a sector. Our quantile regression model is specified as:

$$CSAD_{i,\tau}^{(s)} = c + b_{\tau,AR(1)} * CSAD_{i-1,\tau}^{(s)} + b_{\tau,h} * h_i + u_i \quad (9)$$

where $CSAD_{i,\tau}^{(s)}$ is the cross-sectional absolute deviation of sectoral returns, c is the constant term, h_i is the information loss indicator, and $b_{\tau,h}$, $b_{\tau,AR(1)}$ are the coefficients associated with a given quantile τ . For expositional purposes, we consider only the 25th, 50th and 75th quantiles.²⁰

Conceptually, low CSAD levels are indicative of a greater propensity to follow the market consensus, whereas high CSAD levels suggest that market agents tend to price stocks farther away from the market consensus. Hence, our intention is to examine how information loss is endogenized in market agents' decisions, thus amplifying or dampening their propensity to trade along with the market consensus, or alternatively, to gain exposure towards other factors that are unrelated to that market. Consider for example a portfolio manager using stocks from a specific industry to implement a long-short strategy designed to exploit differences in stocks' sensitivities to some fundamental factors, e.g., long-term yields or oil prices. An increase in returns' dispersion would directly affect the volatility of her portfolio returns, and therefore lower her performance. While there are standard tools to hedge this increase in portfolio risk when driven by changes in the

²⁰ Detailed results, along with methodological notes, are relegated to Appendix C from the Supplement.

underlying fundamental factors, there are fewer options available when facing market (information) uncertainty that only leads to higher dispersion, which directly affects the profitability of a long-short strategy. The practical aspect of our approach is that when investors know that the entropy of various information signals correlates with dispersion, they can design and implement better hedging strategies to protect against spikes in risk driven by market information uncertainty.²¹

In our quantile regression analysis (which we present in detail in Appendix C, from the Supplement) we find that information loss proxies do a better job at explaining *low* and *high* CSAD quantiles rather than median values. The main implication of this result is that our entropy can be a relevant proxy for the quality of information processing performed by investors on various signals, although it might not yet fully reflect the underlying transmission mechanism; for a more complete view, in the second step of our framework we will include investors' attention, which encompasses the main transmission mechanisms outlined in both empirical (e.g., Da et al., 2011; Dimpfl and Jank, 2016; Perlin et al., 2017) and theoretical literature strands (e.g. Andrei and Hasler, 2015). Being a scarce cognitive resource, attention plays the key role in investors' decision-making, which for example in Andrei and Hasler (2015) is akin to a constrained optimization from where one can derive solutions for both financial (portfolio) allocations as well as cognitive search effort allocations.

As a measure of returns' dispersion, the dynamics of CSAD mostly reflects differences in the evaluations of each company's prospects, whereas these same prospects are 'averaged' to converge to a main narrative, which is the one that drives the returns of the sectoral index (i.e. market consensus). Since different risk types (e.g., on fundamentals, market liquidity, macroeconomic factors) can affect different stocks to varying extents, losses in the informational content of different market-wide (aggregate) signals may mislead market agents to reconsider their strategies in relation to the market consensus. Misleading evaluations correspond to high information losses stemming from good/bad news, announcements with positive/negative publicity, market stress, governmental

²¹ Turtle and Wang (2017) confirm the key role of uncertainty in influencing investors' reactions to public information. In particular, investors underreact when uncertainty is high, leading in the medium-to-long run to an over-performance of stocks with strong fundamentals.

decisions, etc. that can increase or reduce individual stock returns (with respect to market returns). As financial markets become more interconnected, and hedging opportunities become harder to find, investors' portfolios become exposed to more complex risks, which are amplified by the misevaluation of aggregate information signals. To a large extent, our results here confirm this interpretation regarding the empirical relevance of our entropy metrics for *high* and *low* CSAD levels, which are more likely to be the consequences of misevaluation.

3.4 Multivariate approach

In this second step, we use *seemingly unrelated regression* (SUR) methods to estimate a system of equations for stock returns' dispersion and investors' attention.²² In this setup, information loss enters as a determinant that now explains the dynamics of investors' attention, which then affects the dispersion in returns (i.e., a measure of the heterogeneous consequences of investors' beliefs) proxied by CSAD.²³ Since we group all sectors together, the estimated five-equation system is the following:

$$\begin{cases} ATT_i = a + b * h_{i,X_i} + c * ATT_{i-1} + e_i \\ CSAD_i^{(sect)} = a^{(sect)} + \beta^{(sect)} * ATT_i + \gamma^{(sect)} * CSAD_{i-1}^{(sect)} + u_i^{(sect)} \end{cases} \quad (10)$$

where $sect = \{tech, indu, util, fina\}$ is a sector-specific superscript; i is a time subscript; $CSAD^{(sect)}$ denotes the (log) cross-sectional absolute deviation of the weekly returns displayed by the constituencies of a given sector, $sect$; X_i denotes the information signal, such that $h_i(X_i)$

²² The SUR method estimates the parameters of the system, accounting for heteroskedasticity and contemporaneous correlations in the errors across equations (Zellner, 1962). Interesting extensions of our empirical system can be considered; for example, one may be willing to frame the problem in a non-linear setting (e.g., considering regime changes, etc.). The methodology proposed in the paper is not affected by the adopted empirical modelling approach, the dataset, nor by the market of interest.

²³ Investors' attention acts as a function composition operation, where the system of CSAD equations is applied simultaneously to the result of applying the function ATT_i to h_{i,X_i} . That is; $ATT_i: X \rightarrow Y$ or $ATT_i = f(h_{i,X_i})$ and $g: Y \rightarrow Z$ or $CSAD_i^{(sect)} = g(ATT_i)$ are composed to yield one entangled bivariate function that is an epimorphism in the category theory. The benefit of using this morphism category is, in fact, what we are trying to show in our framework: information loss enters as a determinant that explains the dynamics of investors' attention, which then affects the dispersion in returns, i.e., a measure of the heterogeneous consequences of investors' beliefs) proxied by CSAD.

denotes the information loss indicator of that signal; $u_i^{(sect)}$ and e_i are idiosyncratic shocks. Our focus is on ATT_i which denotes investors' attention, proxied using data extracted from Google Trends, as in Andrei and Hasler (2015). Attention represents, therefore, the core of the transmission mechanisms that we borrow from the theory we survey in section 2, running from information signals, through entropy, and then reflected onto market outcomes (i.e., return dispersion in our case).

Table 1 summarises the main results from estimating system (10). We are mostly interested in the statistical significance of β and especially b coefficients, which reflect the impact of information losses associated with several relevant information signals (which are displayed on the first column of Table 1).

Table 1. SUR estimates, by information signal indicator

	Investors' attention		CSAD Technology	CSAD Industrials	CSAD Utilities	CSAD Financials
Information Signal, X						
a) Macroeconomic Indicators						
Initial jobless claims						
b	0.015	β	1.199***	1.561***	1.056***	1.557***
GDP tracker, USA						
b	0.027	β	1.191***	1.530***	1.039***	1.526***
3M Treasury yields						
b	-0.013	β	1.216***	1.601***	1.092***	1.597***
10Y Bond yields						
b	-0.119***	β	1.246***	1.713***	1.236***	1.708***
Moody Baa corp. y.						
b	-0.102***	β	1.254***	1.667***	1.231***	1.662***
Crude oil prices						
b	-0.049*	β	1.179***	1.607***	1.093***	1.602***
Gold prices						
b	-0.060***	β	1.254***	1.658***	1.205***	1.652***
b) Market signals						
S&P500 volume						
b	-0.056**	β	1.178***	1.616***	1.137***	1.611***
S&P500						
b	-0.036	β	1.221***	1.593***	1.137***	1.588***
VIX						
b	-0.248***	β	1.026***	1.495***	1.159***	1.492***
FSI						
b	-0.109***	β	1.138***	1.605***	1.071***	1.600***
MSCI						
b	-0.036*	β	1.220***	1.584***	1.129***	1.579***
Bitcoin prices						
b	0.032	β	1.216***	1.633***	1.086***	1.629***
c) Uncertainty metrics						
TMU ¹⁾						
b	-0.120***	β	0.961***	1.397***	0.952***	1.397***
TEU ¹⁾						
b	-0.103***	β	1.026***	1.408***	0.926***	1.406***
EQU						
b	-0.125***	β	1.179***	1.621***	1.182***	1.614***
EPU						
b	0.006	β	1.206***	1.583***	1.076***	1.578***

Note: The effective estimation sample is February 2017 – January 2021, at a weekly frequency; ⁽¹⁾ for TEU and TMU the sample ends in September 2020. Due to space constraints, only the relevant coefficients and their statistical significance (* for 10%, ** for 5%, *** for 1%) are displayed. The Adjusted R² for the *ATT* equation is in the range [0.54, 0.67]; for CSAD^{tech} the range is [0.18, 0.24]; for CSAD^{indu} the range is [0.28, 0.36]; for CSAD^{util} the range is [0.42, 0.46]; for CSAD ^{fina} the range is [0.28, 0.36]. In Appendix B of the online Supplement, we show that similar results obtain when adding to this five-equation system (10) a sixth equation describing the dynamics of CSAD for the overall S&P 500 index.

The main insights from Table 1 are the followings. Firstly, all β 's are positive and statistically significant, meaning that higher investors' attention determines higher returns' dispersion for all

sectors – an idea that aligns with the existing wealth of empirical evidence on the topic (Da et al., 2011; Hsieh et al, 2020). Secondly, most of the estimated b coefficients are negative and statistically significant, meaning that larger losses (i.e., more negative entropy levels) in the value of information would lead to higher investors’ attention. Given the observed overlaps in subsection 3.2, a negative b suggests that during market turmoil the ‘message’ provided by many of these information signals gets distorted or is not sufficiently clear, such that higher attention levels (i.e., more Google searches) are required.

In fact, VIX – which is particularly visible to many investors – or the number of Tweets discussing about *market/economic uncertainty* (i.e., TMU and TEU) can point to different (sometimes even contradicting) ‘narratives’ that are hard to distil one from other. At best, they can be interpreted as ‘symptoms’ rather than ‘diagnostics’, perhaps because they do not contain enough information about what specific ‘narrative’ is behind the recent market moves.²⁴ In contrast, news about initial jobless claims, the level of 3-Month Treasury yields or Bitcoin prices do not seem to significantly affect investors’ attention, perhaps because, even during market turmoil, their informational value is not too much distorted, and so it converges to the main market ‘narrative’. In a nutshell, while some information signals can provide a reference framework for understanding deviations from the market consensus (or from the main narrative), other signals are less able to do so, attracting higher attention and more information demand from investors. It is worth emphasizing that all the estimated β ’s are highly statistically significant in all of our robustness checks, even when different attention proxies are used.

We get similar insights (see Appendix D in the Supplement) where we redo the entropy transformation for a shorter, pre-COVID-19 sample that ends in January 2020. In Appendix D2 we find that only some information loss indicators that we derive from interest rates, VIX and FSI signals are statistically significant for investors’ attention (i.e., trigger a search for more information when distortions arise), suggesting that narratives converge faster during calm periods than during market turmoil (see also Bertsch et al., 2021).

²⁴ Bertsch et al. (2021) document a relative convergence in available narratives during business cycle expansions, and a divergence during contractions.

A single market stress event might generate (similar or opposite) changes in some or in all information signals that are regularly followed by market agents. Accordingly, a composite indicator of information losses might be able to reduce the misleading effects associated with that particular event. Table 2 illustrates our SUR estimates when using composite indicators of information losses, with optimal weights derived from the correlation patterns existing between different information signals. We use the following indicative combinations that include information signal indicators from different groups:: Composite 1 is based on 3-Month Treasury yields, oil prices, and S&P 500 Index values; Composite 2 is based on the same components as Composite 1, to which we add Bitcoin prices; Composite 3 is based on oil and gold prices, and VIX.

Table 2: SUR estimates, by composite indicator

	Investors' attention		CSAD Technology	CSAD Industrials	CSAD Utilities	CSAD Financials
Composite information signals						
Composite 1						
<i>b</i>	-0.223***	β	1.246***	1.663***	1.209***	1.658***
Composite 2						
<i>b</i>	-0.164*	β	1.214***	1.606***	1.129***	1.602***
Composite 3						
<i>b</i>	-0.189***	β	1.199***	1.676***	1.200***	1.671***

Note: The estimation sample is February 2017 - January 2021, at a weekly frequency. Due to space constraints, only the relevant coefficients and their statistical significance (* for 10%, ** for 5%, *** for 1%) are displayed. The Adjusted R^2 for the *ATT* equation is in the range [0.54, 0.55]; for $CSAD^{tech}$ the range is [0.18, 0.19]; for $CSAD^{indu}$ the range is [0.29, 0.31]; for $CSAD^{util}$ the range is [0.42, 0.43]; for $CSAD^{fina}$ the range is [0.30, 0.32]. Weights for each composite indicator are as follows. Composite 1: 3-M treasury yields, $w_1 = 0.50$; Crude oil, $w_2 = 0.25$; S&P500, $w_3 = 0.25$; Composite 2: Weights: 3-M treasury yields, $w_1 = 0.40$; Crude oil, $w_2 = 0.30$; S&P500, $w_3 = 0.10$; Bitcoin prices, $w_4 = 0.20$; Composite 3: Weights: Gold, $w_1 = 0.42$; VIX, $w_2 = 0.14$; Crude oil, $w_3 = 0.44$.

The main difference between Table 1 and 2 is that all composite indicators are now statistically significant in the equation that describes the dynamics of investors' attention (i.e., all the *b* coefficients are negative and statistically significant), even when we include signals with insignificant coefficients according to Table 1. This clearly suggests that misleading effects can be reduced when aggregating information loss indicators (extracted from different signals, each with its own relevance and 'narrative') based on weights derived from optimal portfolio theory. In summary, Table 2 shows that several information loss indicators can be aggregated to provide a

clearer reference framework for investors, who would thus be able to reduce cognitive costs (and search effort) or diversify information losses in order to concentrate their attention on the best available signals. This would instead reduce the cross-sectional dispersion of returns, along with herding propensity and market momentum chasing, bringing thus improvements for price discovery and market efficiency.

Our metric has important implications for market agents, as well as for policymakers. We show how sensitive our information loss indicators can be to exogenous market stress events, and therefore the implicit “costs” they capture in the value of information associated with those events. We also find that, depending on the nature of the information signal, the dynamic losses in the information value can have a transitory (and interactive) impact on investors’ beliefs during turbulent periods. Our metric can thus serve for *benchmarking* the current level of risk during market stress events, as long as one can specify the size of loss that counts as risk and the way it affects a specific financial portfolio. Benchmarking risk levels can help setting up trading rules or better monitoring, even if they remain conditional on information flows. In today’s fluid data environment, where news and indicators are made available on a real-time basis, an entropic transformation performed along the lines described in this paper can therefore allow investors to build more efficient strategies (e.g., hedging options) and policymakers to better monitor markets.

The metric we are proposing here can also be employed to explore other interesting avenues. Two promising directions for future research might examine how different markets or investor types react to information losses derived from the same signal or investigate the impact of information losses on specific portfolios while relying on manager-defined information signals and indicators. As outlined before, in other situations more complex and/or non-linear empirical approaches might better fit the dataset one is focusing on when employing our proposed entropy metrics.

4. Conclusion

This paper proposes a new methodology able to produce signal-specific, scalable metrics which we associate with losses in the value of information that market agents derive from different types of data and indicators. We frame our approach in a Bayesian setting to specify the updating rule that

agents use to process incoming information signals. We then build a time-varying point-wise entropy metric based on a tolerance window, and apply this transformation to different popular information signals, ranging from macroeconomic and stock market indicators to uncertainty measures.

Bayesian updating features in many theoretical models explaining investors' choices with respect to information acquisition and their investment in risky assets (Andrei and Hasler, 2015; Kacperczyk et al 2016). Most of these models rely on rational inattention (developed in Sims 2003) to explain the impact of information on several market outcomes, such as volatility, herding, trading activity and returns' dispersion (i.e., a proxy for investors' belief heterogeneity). To empirically validate our methodology, we estimate a system of equations that reflect these transmission mechanisms. We find that many different information loss indicators, which we construct from various signals, can affect investors' attention, and then spill over into higher cross-sectional dispersion in four representative sectors of the U.S. equity market. We then apply portfolio theory to derive the optimal weights needed to aggregate several information loss indicators into a composite that is more statistically relevant and, thus, can reduce misleading effects and interpretation errors. These positive effects demonstrate that market agents can diversify their information losses when interpreting signals associated with the same underlying event, a result that can have positive consequences on their behaviour, as well as on price discovery and market efficiency.

Acknowledgements

We would like to thank the Editor(s) and the three anonymous reviewers for helpful comments and suggestions that streamlined the discussions and the analysis presented in the paper. We are also grateful to the participants at the 2021 Financial Economics Meeting in Paris, and at the 2019 Financial Engineering & Banking Society conference in Athens, as well as to Konstantinos (Costas) Siriopoulos and Dimitris Asteriou for their many comments and suggestions. All remaining errors are our own.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. *Review of Financial Studies* 28, 33–72. <https://doi.org/10.1093/rfs/hhu059>
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129–151. [10.1257/jep.21.2.129](https://doi.org/10.1257/jep.21.2.129)
- Baker, S. B., Bloom, N., Davis, S. J., Renault, T., 2020. Economic uncertainty measures derived from twitter. Working paper, June.
- Baker, S. R., Bloom, N., Davis, S. J. 2016. Measuring economic policy uncertainty. *The Quarterly Journal of Economics* 131, 1593–1636. <https://doi.org/10.1093/qje/qjw024>
- Barber, B. M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21, 785–818. <https://doi.org/10.1093/rfs/hhm079>
- Barwise, J., Seligman, J., 1997. *Information flow: The logic of distributed systems*. Cambridge: Cambridge University Press.
- Bissiri, P. G., Holmes, C. C., Walker, S. G., 2016. A general framework for updating belief distributions. *Journal of Royal Statistical Society – Series B* 78, 1103–1130. <https://doi.org/10.1111/rssb.12158>
- Bertsch, C., Hull, I., Zhang, X., 2021. Narrative fragmentation and the business cycle. *Economics Letters* 201, 109783. <https://doi.org/10.1016/j.econlet.2021.109783>
- Brunnermeier, M., Nagel, S., 2008. Do wealth fluctuations generate time-varying risk aversion? Micro-evidence on individuals. *American Economic Review* 98, 713–736. DOI: [10.1257/aer.98.3.713](https://doi.org/10.1257/aer.98.3.713)
- Challet, D., Marsili, M., Zhang, Y.C., 2000. Modeling market mechanism with minority game. *Physica A: Statistical Mechanics and its Applications* 276, 284–315. [https://doi.org/10.1016/S0378-4371\(99\)00446-X](https://doi.org/10.1016/S0378-4371(99)00446-X)
- Chang, E., Cheng, J., Khorana, A. 2000. An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance* 24, 1651–1679. [https://doi.org/10.1016/S0378-4266\(99\)00096-5](https://doi.org/10.1016/S0378-4266(99)00096-5)
- Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. *Journal of Finance* 66, 1461–1499. <https://doi.org/10.1111/j.1540-6261.2011.01679.x>
- DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22, 1915–1953. <https://doi.org/10.1093/rfs/hhm075>
- Dierick, N., Heyman, D., Inghelbrecht, K., Stieperaere, H., 2019. Financial attention and the disposition effect. *Journal of Economic Behavior & Organization* 163, 190–217. <https://doi.org/10.1016/j.jebo.2019.04.019>

- Dragomirescu-Gaina, C., Philippas, D., Tsionas M.G., 2021. Trading off accuracy for speed: Hedge funds' decision-making under uncertainty. *International Review of Financial Analysis* 75, 101728. <https://doi.org/10.1016/j.irfa.2021.101728>
- Dragomirescu-Gaina, C., Philippas, D., 2021. Local versus global factors weighing on stock market returns during the COVID-19 pandemic. *Finance Research Letters*, 102270. <https://doi.org/10.1016/j.frl.2021.102270>
- Drake, M. S., Roulstone, D. T., Thornock, J. R., 2012. Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research* 50, 1001–1040. <https://doi.org/10.1111/j.1475-679X.2012.00443.x>
- Easley, D., 2010. *Networks, crowds and markets: Reasoning about a highly connected world*. Cambridge University Press. 483–506.
- Eaton, J. J., Bawden, D., 1991. What kind of resource is information? *International Journal of Information Management* 11, 156–165. [https://doi.org/10.1016/0268-4012\(91\)90006-X](https://doi.org/10.1016/0268-4012(91)90006-X)
- Engelberg, J., Parsons, C. A., 2011. The causal impact of media in financial markets. *Journal of Finance* 66, 67–97. <https://doi.org/10.1111/j.1540-6261.2010.01626.x>
- Fudenberg, D., Strack, P., Strzalecki, T., 2018. Speed, accuracy, and the optimal timing of choices. *American Economic Review* 108, 3651–3684. DOI: 10.1257/aer.20150742
- Gentzkow, M., Shapiro, J. M., 2010. What drives media slant? Evidence from U.S. daily newspapers. *Econometrica* 78, 35–71. <https://doi.org/10.3982/ECTA7195>
- Gutierrez, R. C., Kelley, E. K., 2008. The long-lasting momentum in weekly returns. *Journal of Finance* 63, 415–447. <https://www.jstor.org/stable/25094444>
- Hamid, A., Heiden, M., 2015. Forecasting volatility with empirical similarity and Google Trends. *Journal of Economic Behavior & Organization* 117, 62–81. <https://doi.org/10.1016/j.jebo.2015.06.005>
- Howard, R. A. 1966. Information value theory. *IEEE Transactions on Systems Science and Cybernetics (SSC-2)*, 22–26. DOI: 10.1109/TSSC.1966.300074
- Hsieh, S. F., Chan, C. Y., Wang, M. C., 2020. Retail investor attention and herding behavior. *Journal of Empirical Finance* 59, 109–132. <https://doi.org/10.1016/j.jempfin.2020.09.005>
- Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L., 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84, 571–626. <https://doi.org/10.3982/ECTA11412>
- Koenker, R., 2005. *Quantile Regression*. New York: Cambridge University Press.
- Mackowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. *American Economic Review* 99, 769–803. DOI: 10.1257/aer.99.3.769
- Mitchel, M. L., Mulherin, J. H., 1994. The impact of public information on the stock market. *Journal of Finance* 49, 923–950. <https://doi.org/10.2307/2329211>
- Moscarini, G., Smith, L., 2002. The law of large demand for information. *Econometrica* 70, 2351–2366. <https://www.jstor.org/stable/3081989>

- Pérez-Montoro, M., 2007. *The phenomenon of information: a conceptual approach to information flow*. Medford, NJ: The Scarecrow Press, Inc.
- Perlin, M. S., Caldeira, J. F., Santos, A. A., Pontuschka, M. 2017. Can we predict the financial markets based on Google's search queries? *Journal of Forecasting* 36, 454–467. <https://doi.org/10.1002/for.2446>
- Philippas, D., Philippas, N., Tziogkidis, P., Rjibad, H., 2020. Signal-herding in cryptocurrencies. *Journal of International Financial Markets, Institutions and Money* 65, 101191. <https://doi.org/10.1016/j.intfin.2020.101191>
- Pincus, S. M., Kalman, E. K., 2004. Irregularity, volatility, risk and financial market time series. *Proceedings of the National Academy of Sciences* 101(38), 13709–13714.
- Savage, L.J., 1954. *The Foundations of Statistics*. New York, Wiley.
- Schepanski, A., Uecker, W. C., 1984. The value of information in decision making. *Journal of Economic Psychology* 5, 177–194. [https://doi.org/10.1016/0167-4870\(84\)90005-9](https://doi.org/10.1016/0167-4870(84)90005-9)
- Sims, C. A., 2003. Implications of rational inattention. *Journal of Monetary Economics* 50, 665–690. [https://doi.org/10.1016/S0304-3932\(03\)00029-1](https://doi.org/10.1016/S0304-3932(03)00029-1)
- Solomon, D., 2012. Selective publicity and stock prices. *Journal of Finance* 67, 599–638. <https://doi.org/10.1111/j.1540-6261.2012.01726.x>
- Steiner, J., Stewart, C., Matjka, F., 2017. Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica* 85, 521–553. <https://doi.org/10.3982/ECTA13636>
- Tetlock, P. C., 2011. All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies* 24, 1481–1512. <https://doi.org/10.1093/rfs/hhq141>
- Touboul, J. D., 2019. The hipster effect: When anti-conformists all look the same. *American Institute of Mathematical Science* 24, 4379–4415. DOI: 10.3934/dcddb.2019124
- Turtle, H. J., Wang, K., 2017. The value in fundamental accounting information. *Journal of Financial Research* 40, 113–140. <https://doi.org/10.1111/jfir.12119>
- Veldkamp, L., 2006. Media frenzies in markets for financial information. *American Economic Review* 96, 577–601. DOI: 10.1257/aer.96.3.577
- Verona, F., 2014. Investment dynamics with information costs. *Journal of Banking and Finance* 46,1627–1656. <https://doi.org/10.1111/jmcb.12161>
- Vlastakis, N., Markellos, R.N., 2013. Information demand and stock volatility. *Journal of Banking and Finance* 36, 1808–1821. <https://doi.org/10.1016/j.jbankfin.2012.02.007>
- Woloszko, N., 2020. A weekly tracker of activity based on machine learning and Google Trends. *OECD Economics Department Working Papers* 1634, Paris: OECD Publishing.
- Woodford, M., 2014. Stochastic choice: An optimizing neuroeconomic model. *American Economic Review* 104, 495–500. DOI: 10.1257/aer.104.5.495

- Zellner, A., 1962. An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *Journal of the American Statistical Association* 57, 348–368.
<https://doi.org/10.2307/2281644>
- Zhang, F.X., 2006. Information uncertainty and stock returns. *Journal of Finance* 61, 105–137.
<https://doi.org/10.1111/j.1540-6261.2006.00831.x>