Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 01-2026

The Tell-Tale Clock!
Speed and Agent Composition in High Frequency Trading

Iordanis Kalaitzoglou and Stelios Arvanitis

January 2026

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

The Tell-Tale Clock!

Speed and Agent Composition in High Frequency Trading.

Iordanis Kalaitzoglou [†], Stelios Arvanitis[‡]

[†]Audencia Business School. ikalaitzoglou@audencia.com

[‡]Athens University of Economics and Business, stelios@aueb.gr

January 31, 2026

Abstract

We introduce an instantaneous, intensity-based metric for estimating the number, type, and probability of different agent-types over any time interval. It is founded on the idea that differences in trading motives are reflected in agent specific arrival rates, which are modeled using conditional hazard functions. This novel framework exhibits superior empirical performance and theoretical properties over its interval-based — and thus slower — counterparts. In addition, it enables the derivation of a local measure for HFT reflexivity that is found to be highly correlated with the presence of variance spikes and extreme events, such as flash crashes. Therefore, it can act as a propensity measure for HFT market ”stress”.

Keywords: High Frequency Trading; Agent-types; Hazard Functions.

JEL Codes: C41; C55

1 Introduction

AI [...] could lead to firms taking increasingly correlated positions and acting in a similar way during a stress, thereby amplifying shocks. Such market instability can then affect the availability and cost of funding for the real economy."

Bank of England, April 2025

The quote above stresses out that Algorithmic Trading (AT) might propagate errors in a manner and speed that can harm the real economy.[1] In principle, AT acts as a continuous market monitoring mechanism (Brogaard et al., 2014, 2019, 2025) that has the potential to improve market efficiency (Chaboud et al., 2014) by exploiting arbitrage opportunities, even at low latency (Hasbrouck and Saar, 2013), rendering it eventually "relevant for humans" (Chordia et al., 2024, Chakrabarty and Pascual, 2023). In order to do so, AT relies on trading signals (Li et al., 2021) as a proxy for fundamental information (Yadav, 2015), which gets propagated to subsequent trades; a property known as market reflexivity (Soros, 1994, 2013). At these trading speeds (measured in micro-seconds; Aquilina et al., 2022), these signals become highly entropic (Shannon, 1948), making AT conditional on a noisy signal of fundamental information.[2] In normal trading conditions, more or faster trading resolves this information faster (Kirilenko and Lo, 2013, Kirilenko et al., 2017), but in periods of stress, it becomes increasingly difficult to discern information about the informational content of other trades (Hendershott et al., 2011, Zhang and Zhang, 2025) and, thus, AT might propagate noise instead of signal.[3] In this case, reflexivity becomes toxic (Easley et al., 2014) and can

[1] The report of Bank of England on "Financial Stability in Focus: Artificial intelligence in the financial system" can be found here: https://www.bankofengland.co.uk/financial-stability-in-focus/2025/april-2025.

[2] At trading speeds (in μs; Aquilina et al., 2022) that exceed those of humans (>650ms; Johnson et al., 2013), it is increasingly unrealistic to assume AT has the time to extract, analyze, and act upon fundamental information concerning the cash flows of an asset. Instead, algorithms try to extract this information from the actions of other agents and the trading signals (e.g., trading volume, direction, etc.) they emit. This cannot be done perfectly and thus, information is being resolved only gradually, exhibiting entropic features.

[3] According to the Bank for International Settlements (BIS) in "FX execution algorithms and market functioning" (www.bis.org/publ/mktc13.pdf), AT uses the same set of data and their actions tend to be highly correlated. In normal trading periods, this can accelerate information resolution, but in periods of stress, the informational content of trades might not be that clear and, thus, AT might propagate noise, instead. This creates the fundamentals for extreme events, such as flash crashes and price spikes.

escalate to systemic risk, fairly quickly, due to High Frequency Trading (HFT).[4]

Consequently, it is essential for AT to infer the real trading motives of other market participants, from observable (trading) signals, or heuristically "who trades in the market", in real time. Previous approaches suggest interval-based metrics, which in HFT are rather slow. In this paper, we propose an empirical strategy for identifying the presence of different agent-types over any time interval, venturing the idea that their trading motives are reflected in tangible actions; in particular in the arrival rates of observable variables. Drawing on the theory of point processes, the conditional intensity (hazard function) fully describes these arrival rates and it is used to capture the presence of these agents. The new metric, referred to as the (i)ntensity-based (R)elative (P)roportion (iRP), is highly relevant in HFT and exhibits significant advantages over previous approaches.

First, iRP is an instantaneous metric and can be estimated at any point or over any time interval. Efficient markets (Fama, 1965) imply an instantaneous and unanimous information resolution, but at HFT there is a drastic shift to "trading" information as a noisy signal of "fundamental" information, which forces AT to extract it from trading (primarily speed-related; Easley and O'hara, 1992) signals. The objective is to infer the informational content (trading motives) of a trade, or else, "who trades in the market". At every point in time several agent-types co-exist and their interactions constitute the mechanism by which information is price resolved. Consequently, the agent composition is essential for conditioning AT, but this constitutes a latent piece of information, which can only be estimated in a probabilistic manner.[5] Toward this end, prior literature suggests using the aggregated, primarily over time, characteristics of trading activity as observable proxies for the actions of

[4]May 2010 and October 2014 are examples of signal misinterpretation that resulted in a sharp liquidity withdrawal, while August 2015 and October 2016 are examples of misinterpretation resulting in coordinated actions. In both cases, the inability of AT to identify the real motives behind a trading signal resulted in a market-wide shock within minutes. Smaller and faster shocks are also present in the market ((U)ltra-fast (E)xtreme (E)vents (UEE's); Johnson et al., 2013) and they are also the result of signal misinterpretation.

[5]The agent-type composition cannot be known either ex-post or ex-ante. Arthur (2013) argues that identifying individual agent-type strategies is rather unattainable using only public information. Consequently, mean field theory (Jovanovic and Rosenthal, 1988) cannot address the complexity arising from the emergent properties of agent interactions. Thus, the identification of agent-types cannot rely on observing their individual actions but must instead be approached probabilistically, as a sample property.

different agent-types.[6] They are mostly based on fluctuations in the aggregated magnitude of a variable over an interval, which albeit intuitive, suffers from a trade-off between sampling bias and frequency.[7] In HFT, where the timing of possessing information is crucial, because it can turn time priority into an information advantage ("*to be uninformed is to be slow*"; Haldane, 2012) a faster identification, ideally as a point estimate, would be more appropriate.

This is the primary concern of this study, which suggests matching differentials in the conditional intensities (hazard functions) of observable events (e.g., trades or volume) to different trading motives, as they are reflected on these events, in order to identify the presence of different agent-types and provide a point estimate for their probabilities. Unlike previous approaches that focus on the 'aggregated properties' of the accumulated actions of each agent-type, iRP focuses on the properties of their 'accumulation rates'. The starting point is the presence of the agent-type herself, who is assumed to act upon stimuli in a distinct time-invariant manner. Consequently, her actions exhibit a distinct time-invariant arrival rate. However, her presence is conditional on market conditions and thus,time-variant. The market as a whole is seen as an infinite mixture of their presence or absence, considering that all agents (as experts with time invariant characteristics) contribute to shaping the overall market activity, with time varying weights that account for market conditions. This is a flexible framework to estimate the instantaneous probability that a particular event is initi-

[6] Since its early stages, the microstructure literature has associated the arrival rate of trades (Easley and O'hara, 1992) and volume (Easley et al., 1996; Easley et al., 1997b; Easley et al., 1997a) with the existence of informed agents. Other agent-types have also been identified based on observable characteristics, such as trade sidedness (Sarkar and Schwartz, 2009), which is linked to asymmetric information; trade initiation runs (Patterson and Sharma, 2010), associated with information cascades and herding; and trade/order imbalances (Easley et al., 2011; Easley et al., 2014), associated with order flow toxicity.

[7] Longer time intervals reduce noise, but result in slower pace in information extraction. Arthur (2013) claims that a top-down approach captures better the emergent properties arising from agent interactions. Interval-based measures are more likely to capture these aggregated properties in a manner that reduces the impact of noisy signals (Pöppe et al., 2016). However, there is no clear consensus on what constitutes an optimal sampling interval. Empirical evidence (e.g., Easley et al., 2011; Easley et al., 2014; Andersen and Bondarenko, 2014) reports significant sampling bias, as these metrics face a trade-off: longer intervals better address noisy signals (e.g., trade initiation, Easley et al., 2014) but are less sensitive due to mean reversion, while shorter intervals are more sensitive but also more prone to noise (Andersen and Bondarenko, 2014). Consequently, higher sampling frequencies, which are more relevant to HFT, may render interval-based metrics ineffective for capturing the finer properties of the data due to noise.

ated by a specific agent-type, which can be extended over any interval without introducing sampling bias (Andersen and Bondarenko, 2014).

Second, iRP is a significant generalization, where any number or type of agents can be identified, because their presence is linked to a precise statistical measure (hazard function) rather than to a specific variable, such as volume to informed trading. The trading characteristics of an agent-type can be described by any set of variables, but the detection of the agent-type is done by how they affect the conditional intensity, without assuming any prior link. This introduces a flexible framework that can be used either i) to estimate the probability of a specific agent-type, with a well defined trading characteristics or ii) in an agnostic manner, to estimate, the number, type and probability of the agent-types present in the market with no prior assumption about their number or the observable factors their actions might affect. We illustrate the former with an application on private information in Section 4 and the latter in a simulation in Section 5.

Finally, our framework enables an explicit assessment of the degree of reflexivity, derived from the per-trade informational entropy (informational gain). We consider, in the spirit of Jacobs et al. (1991), a (M)ixture of (E)xperts approach, where the different agent-types affect the trading activity in a collective manner, both with their presence (conditional intensity), as well as with their absence (survival function). Their actions and interactions are manifested into the overall market activity, as a highly non-linear mixture that exhibits time variant characteristics due to varying market conditions, as well as by their arrival rates. This enables the estimate of (sample-wide) prior probabilities and local revisions (posterior probabilities), depending on assessment errors. This separates the impact of market conditions from the impact of arrival rates and can be used in order to assess how much trading is affected by the former, which constitutes a signal, or the latter, which indicates endogenous trading; thus, a higher degree of reflexivity. We develop a new metric to capture this and we show that it is highly correlated with higher variance and the presence of extreme events, such as flash-crashes and price spikes.

2 An Intensity-Based Estimator of Agent-Types

2.1 The Market as a Collection of Different Agents

The market is a collection of $k = 1, 2, ..., K$ agent-types that interact and formulate the overall trading activity. Each agent-type is driven by a stationary, time-invariant motivation for trading, e.g., information, technical rule, etc., which is characterized by a set of parameters $\boldsymbol{\tau}^k \in \mathbb{R}^Q$; $Q_{\in\mathbb{Z}}$ is the number of parameters that describe differentials in trading behavior. These parameters govern the shape, scale, and entropy behavior of her arrival rate, which, are assumed to be time invariant, for traceability reasons. Then, the conditional intensity $\lambda^k(t|\mathcal{F}_s) \equiv \lambda^k\left(t, \boldsymbol{\tau}^k|\mathcal{F}_s\right)$ describes fully the arrival rate of each agent-type k.[8]

These intensities are latent and only the intensity of the market, $\lambda(t|\mathcal{F}_s) \equiv \lambda(t, \boldsymbol{\tau}_s|\mathcal{F}_s) = \frac{f(t,\boldsymbol{\tau}_s|\mathcal{F}_s)}{S(t,\boldsymbol{\tau}_s|\mathcal{F}_s)}$, can be observed at any point in time $t = s$. $f(t, \boldsymbol{\tau}_s|\mathcal{F}_s)$ and $S(t, \boldsymbol{\tau}_s|\mathcal{F}_s)$ are the distribution and the survival function respectively. Our objective is to infer the presence of all agent-types from the observable $\lambda(t, \boldsymbol{\tau}_s|\mathcal{F}_s)$. We consider that each agent-type contributes to the overall market activity in a latent non-linear manner, because emergence properties (Johnson et al., 2003) that are not necessarily the sum of the actions of all agent-types. Consequently, we employ a "Mixture of Experts" (MoE) approach (Jacobs et al., 1991), rather than a "Mixture of Densities" (MoD) like in Bowsher (2007), which embeds the contributions of all agents/"experts" $\boldsymbol{\tau}^k$ as a convex combination in the parameters that govern the overall market activity, $\boldsymbol{\tau}_s$.

$$\lambda(t, \boldsymbol{\tau}_s|\mathcal{F}_s) = \lambda\left(t, \sum_{k=1}^{K}(p_t^k|\mathcal{F}_s)\boldsymbol{\tau}^k|\mathcal{F}_s\right) \tag{1}$$

[8] $\lambda^k\left(t, \boldsymbol{\tau}^k|\mathcal{F}_s\right)$ is the conditional intensity of a simple point process and describes mathematically the entire trading behavior of agent-type k. $\{T_i\}_{i\in\mathbb{Z}}$ denotes a simple point process on $[0, \infty)$, defined as a sequence of non-negative random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such as $0 < T_i < T_{i+1} \forall i$. $N(t)$ is the counting process of $\{T_i\}_{i\in\mathbb{Z}}$ defined as $N(t) = \sum_{i\geq 1} 1(T_i < t)$ that counts the events up to time t. Then $\lambda(t|\mathcal{F}_s)$ is defined as the intensity of $N(t)$, given some filtration $\mathcal{F}_s$, if $\mathbb{E}(N(t) - N(s)|\mathcal{F}_s) = \mathbb{E}\left(\int_s^t \lambda(u)\, du\Big|\mathcal{F}_s\right)$, for $0 < s < t$, and fully describes $\{T_i\}_{i\in\mathbb{Z}}$. The counting function can be defined as $N^k(t) = \sum_{i\geq 1} 1(T_i < t)(Z_i = k)$, with a cumulative function $\mathbb{E}(N^k(t) - N^k(s)|\mathcal{F}_s) = \mathbb{E}\left(\int_s^t \lambda^k(u)\, du\Big|\mathcal{F}_s\right)$.

Each agent k contributes – $(p_t^k|\mathcal{F}_s)\boldsymbol{\tau}^k$ – to the market activity either with their presence – $\lambda\left(t, \sum_{k=1}^{K}(p_t^k|\mathcal{F}_s)\boldsymbol{\tau}^k|\mathcal{F}_s\right)$ – or their absence – $S\left(t, \sum_{k=1}^{K}(p_t^k|\mathcal{F}_s)\boldsymbol{\tau}^k|\mathcal{F}_s\right)$ – albeit in an agnostic manner. This makes $p_t^k|\mathcal{F}_s$ a measure of the probability of agent k entering the market, prior to its arrival; thus, a prior probability, which is inferred from the trading characteristics $\boldsymbol{\tau}^k$.

Collectively, the market can be seen as a time-varying (infinite) mixture, $p_t^k|\mathcal{F}_s$, of the contribution of different agent-types, $\boldsymbol{\tau}^k$, to the overall market activity, $\boldsymbol{\tau}_s$. Unlike the intensities of each agent-type that are latent, the intensity of the market, $\lambda(t|\mathcal{F}_s) \equiv \lambda(t, \boldsymbol{\tau}_s|\mathcal{F}_s)$, and its counting function, $N(t) = \sum_{i\geq 1} 1(T_i < t)$, are observable. They are assumed to be the combined outcome of the interactions of all the agent-types present in the market, reflected on the parameters $\boldsymbol{\tau}_t$, which define the conditional intensity of the market:

$$\mathbb{E}(N(t) - N(s)|\mathcal{F}_s) = \mathbb{E}\left(\int_s^t \lambda(u)\, du\middle|\mathcal{F}_s\right) \equiv \mathbb{E}\left(\int_s^t \lambda(u, \boldsymbol{\tau}_u)\, du\middle|\mathcal{F}_s\right)$$

$$= \mathbb{E}\left(\int_s^t \lambda(u, \sum_{k=1}^{K}(p_u^k|\mathcal{F}_s)\boldsymbol{\tau}^k)\, du\middle|\mathcal{F}_s\right) \tag{2}$$

In Eq.(2) the overall market activity, $\mathbb{E}\left(\int_s^t \lambda(u, \boldsymbol{\tau}_u)\, du\middle|\mathcal{F}_s\right)$, is the combined outcome of the interactions of all agent-types, $\mathbb{E}\left(\int_s^t \lambda(u, \sum_{k=1}^{K}(p_u^k|\mathcal{F}_s)\boldsymbol{\tau}^k)\, du\middle|\mathcal{F}_s\right)$. Each agent-type is characterized by a uniquely identifiable conditional intensity, $\lambda^k\left(t, \boldsymbol{\tau}^k|\mathcal{F}_s\right)$, that is defined by a set of time-invariant parameters, $\boldsymbol{\tau}^k$. This describes their trading pattern that captures their intrinsic motivation and can be thought of as their baseline intensity. Their presence in the market is defined jointly by their intrinsic motivation and their interactions, $p_t^k|\mathcal{F}_s$, with the presence $f^k\left(t, \boldsymbol{\tau}^k|\mathcal{F}_s\right)$, or absence $S^k\left(t, \boldsymbol{\tau}^k|\mathcal{F}_s\right)$ of all other agents. This results in $\sum_{k=1}^{K}(p_u^k|\mathcal{F}_s)\boldsymbol{\tau}^k$, where conditionality is reflected on "how" each "expert" contributes to the shape of the overall trading activity $\boldsymbol{\tau}_s$.

2.2 Toward an Empirical Specification

To use Eq.(2) empirically, $\lambda^k(t, \boldsymbol{\tau}^k|\mathcal{F}_s)$ and $p_t^k|\mathcal{F}_s$, must be extracted from the observed $\lambda(t, \boldsymbol{\tau}_s|\mathcal{F}_s)$. We use the conditional density of arrival times, $f(t, \boldsymbol{\tau}_s|\mathcal{F}_s)$ (Kalaitzoglou and

Ibrahim, 2013), and focus on durations , $x_i = t_i - t_{i-1}$, i.e., the time between two consecutive events, which are modeled with an ACD specification (Engle and Russell, 1998):

$$x_i = \psi_i \epsilon_i \tag{3}$$

$\psi_i = \mathbb{E}(x_i|\mathcal{F}_{i-1})$ is the expected duration, conditional on past information $\mathcal{F}_{i-1}$. $\epsilon_i = \frac{\chi_i}{\psi_i}$ is the standardized duration, the distribution of which, $f(\epsilon_i|\mathcal{F}_{i-1})$, is used to derive the distribution of χ_i, $f(\chi_i|\mathcal{F}_{i-1}) = f(\epsilon_i|\mathcal{F}_{i-1})\psi^{-1}$, which in turn is used to derive the conditional intensity for the whole market, $\lambda(\chi_i|\mathcal{F}_{i-1})$. Following Eq.(2), $f(\chi_i|\mathcal{F}_{i-1})$ is modeled as:

$$f(\chi_i, \boldsymbol{\tau}_i|\mathcal{F}_{i-1}) = f(\chi_i, \sum_{k=1}^{K}(p_i^k|\mathcal{F}_{i-1})\boldsymbol{\tau}^k|\mathcal{F}_{i-1}) \tag{4}$$

Eq.(4) defines $f(\chi_i, \boldsymbol{\tau}_i|\mathcal{F}_{i-1})$, the shape of which is defined collectively by all agent-types as a weighted average $\sum_{k=1}^{K}(p_i^k|\mathcal{F}_{i-1})\boldsymbol{\tau}^k$ of their distinct characteristics $\boldsymbol{\tau}^k$ and the prevailing market conditions $p_i^k|\mathcal{F}_{i-1}$. $p_i^k|\mathcal{F}_{i-1} \equiv p_t^k|\mathcal{F}_s$ is the prior probability in event time, i, rather than in continuous time t. We estimate $p_i^k|\mathcal{F}_{i-1}$ and $f^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$ by linking the distinctive characteristics of each agent-type to a set of $q = 1, ..., Q$ parameters, different values of which would indicate a differently shaped distribution. Eq.(4) can be rewritten as (conditionality is denoted by i):

$$f(\chi_i, \boldsymbol{\tau}_i|\mathcal{F}_{i-1}) = f(\chi_i, \boldsymbol{\tau}_i(\boldsymbol{W}_i(:), \boldsymbol{\tau}^m|\mathcal{F}_{i-1}) = f(\chi_i, \sum_{k=1}^{K} L_i^k(\boldsymbol{W}_i)\boldsymbol{\tau}^k|\mathcal{F}_{i-1}) \tag{5}$$

$\boldsymbol{\tau}_i = (\tau_i^q)_{q=1}^{Q}$ is a vector of $q = 1, ..., Q$ parameters, dissected into $m = 1, ..., M$ regimes, that determine the shape/scale of $f(\chi_i, \boldsymbol{\tau}_i|\mathcal{F}_{i-1})$, which is expressed as a function, $\boldsymbol{\tau}_i(:)$, of a $[Q \times M]$ matrix of weighting functions, $\boldsymbol{W}_i(:)$, and a $[Q \times M]$ vector of shape/scale parameters, $\boldsymbol{\tau}^m = \left(\left(\tau_m^Q\right)_{q=1}^{Q}\right)_{m=1}^{M}$. Different combinations of τ_m^q lead to differently shaped conditional intensities $\lambda^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$ that can match the actions of an agent-type k. The prior probability, $p_i^k|\mathcal{F}_{i-1}$, is expressed as a function $L_i^k(:)$ of $\boldsymbol{W}_i$. Differentiation in Eq.(5) is reduced to estimable parameters τ_m^q, which lead to vectors $\boldsymbol{\tau}^k = \left(\tau_{m:k}^{q=1}, ..., \tau_{m:k}^{q=Q}\right)$, with $m : k$ indicating the regimes m that lead to an intensity $\lambda^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$, which fully describes the baseline intensity of agent-type k. Accordingly, each parameter, $\tau_i^q, \forall q$, is defined as:

$$\tau_i^q = \sum_{m=1}^{M} W_{m,i}^q \tau^m \tag{6}$$

which is a data driven way to identify different regimes, m, of observable variables, $\boldsymbol{J}_i$ that are associated with the parameters τ^m. The k's are latent and they have to be inferred from the data. This is done by associating different values of τ^m's with specific trading characteristics. Then, we employ the smooth transition functions $\boldsymbol{G}_i$ (Teräsvirta, 1994) to assess $\boldsymbol{W}_i$. $\forall q$:

$$W_{m,i}^q = \left(G_{m,i}^q - G_{m,i+1}^q\right), \; with \; G_{1,i}^q = 1, \; G_{M+1,i}^q = 0 \; and \; \sum_{m=1}^{M} W_{m,i}^q = 1 \tag{7}$$

$\boldsymbol{G}_i = \left(\left(G_{m,i}^q\right)_{q=1}^{Q}\right)_{m=1}^{M}$ is a matrix of smooth transition functions across regimes, which estimate the weighting of each regime m in τ_i^q. Then, depending on the complexity of $f(\chi_i, \boldsymbol{\tau}_i|\mathcal{F}_{i-1})$, more than one parameters ($q \geq 1$) might define the shape/scale of the hazard function. Consequently, the prior probability of a trade initiated by an agent-type k, $p_i^k|\mathcal{F}_{i-1}$ can be estimated via combinatorial intersections ($L_i^k(:)$) of $\boldsymbol{W}_i = \left(\left(W_{m,i}^q\right)_{q=1}^{Q}\right)_{m=1}^{M}$, as:

$$\mathbb{P}(Z_i = k|\mathcal{F}_{i-1} = L_i^k(\boldsymbol{W}_i) = \sum_{Q \bigotimes m:k} \prod_{Q \bigotimes M} W_{m,i}^q \tag{8}$$

$\prod_{Q \bigotimes M} W_{m,i}^q$ is the cross-multiplication product of $W_{m,i}^q$ that defines the probability of the intersection between parameters, q, and regimes, m, with dimension M^Q. $\sum_{Q \bigotimes m:k}(:)$ is the intersections $Q \bigotimes m : k$ where the parameters result in a distribution that matches the characteristics of agent-type k. Then, $p_i^k|\mathcal{F}_{i-1}$ is the sum of all $Q \bigotimes m : k$. Conditionality is summarized into $\boldsymbol{G}_i$'s, which are modeled as smooth transition functions (Teräsvirta, 1994):

$$G^q_{m,i} = G^q_{m,i}(\boldsymbol{J}_i : g^q_m, j^q_m) = \left(1 + e^{-g^q_m(\boldsymbol{J}_i - j^q_m)}\right)^{-1} \tag{9}$$

where, $\boldsymbol{J}_i = \left(\left(J^q_{v,i}\right)^Q_{q=1}\right)^V_{v=1}$ is a vector of $v = 1, ..., V$ threshold variables, measurable with respect to $\mathcal{F}_{i-1}$. The magnitude of $\boldsymbol{J}_i$ relative to the threshold values, j^q_m, determines the allocation of each event into a regime (Lof, 2012). Then, each τ^q_i can be defined as:

$$\begin{aligned}\tau^q_i = &\underbrace{\left(\overbrace{G^q_{m=1,i}}^{1} - G^q_{m=2,i}\right)}_{W^q_{m=1,i}} \tau^q_{m=1} + \underbrace{\left(G^q_{m=2,i} - G^q_{m=3,i}\right)}_{W^q_{m=2,i}} \tau^q_{m=2} + ... + \\ &\underbrace{\left(G^q_{m=M-1,i} - G^q_{m=M,i}\right)}_{W^q_{m=M-1,i}} \tau^q_{m=M-1} + \underbrace{\left(G^q_{m=M,i} - \overbrace{G^q_{m=M+1,i}}^{0}\right)}_{W^q_{m=M,i}} \tau^q_{m=M}\end{aligned} \tag{10}$$

Eq.(10) is a specific derivation of Eq.(6), which links the value of each shape or scale parameter τ^q_i to observable variables $\boldsymbol{J}_i$ through the weighting functions $G^q_{m,i}$, which eventually lead to the prior probabilities $\mathbb{P}(Z_i = k|\mathcal{F}_{i-1})$, through the combinatorial intersections $L^k_i(:)$. This is a data driven way to estimate $\mathbb{P}(Z_i = k|\mathcal{F}_{i-1})$ and can be used in two ways:

- *Estimate the probability of an agent-type.* Should the trading characteristics of an agent-type be reflected on a set of observable variables and a uniquely shaped hazard function; Eq.(10) can be used to estimate the probability of this agent-type entering the market.

- *Estimate the number of different agent-types.* Eq.(10) can be used to identify the number of regimes and intersections that fit the data best and thus; estimate how many different agents exist in the market and their probabilities.

2.3 Estimation

Eqq. (5)-(10) set an optimization problem that corresponds to a non-linear, parameter-based generalization of the MoE framework of Jacobs et al. (1991). Each agent-type ("expert") contributes to the market activity, through her presence or absence, but not directly, since the type of this agent cannot be known with certainty. Instead, all agents do so, in a combined manner, the aggregated output of which, is observable as the overall market activity. To simulate this, we consider that each agent-type ($\boldsymbol{\tau}^k$) contributes to the overall market activity $\lambda(\chi_i, \boldsymbol{\tau}_i|\mathcal{F}_{i-1})$ through a "gating" mechanism, which is captured by the impact of functions $\boldsymbol{G}_i$ or in $L_i^k(:)$'s on the shape/scale of the overall market activity $\boldsymbol{\tau}_i$, rather than directly. This way all parameters can be estimated based on the distributional assumptions (Engle and Russell, 1998) for the market durations (Eq.(5)) and the likelihood function $\sum_i \log f(\chi_i; \boldsymbol{\tau}_i = \sum_k L_i^k(\boldsymbol{W}_i; \theta))\boldsymbol{\tau}^k|\mathcal{F}_{i-1}))$, with $\theta = \{g_m^q, j_m^q, \boldsymbol{\tau}^k\}$.

The objective of this work is to capture the presence of a multitude of agents and this depends on the flexibility of the distribution in producing non-monotonic hazard functions. Distributions employed in previous literature (e.g., Weibull, Burr, (generalized) Gamma), primarily to capture the dipole of informed/uninformed, can only generate monotonic hazard functions. Instead, another distribution is 'indicatively' proposed here, the q-Weibull distribution, for its flexibility in generating non-monotonic hazard functions and its link to "information entropy". [9] The "q-Weibull" distribution can be defined for $(\chi_i|\mathcal{F}_{i-1}) \sim W_q\left(A_i, \tau_i^{q=1}, \tau_i^{q=2}\right)$:

$$f(\chi_i|\mathcal{F}_{i-1}; \boldsymbol{\tau}_i) = (2 - \tau_i^{q=1})\frac{\tau_i^{q=2}}{\chi_i}\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}} e_q\left(-\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}}\right) \tag{11}$$

$A_i = \left[{}^{\Gamma\left(1+1/\tau_i^{q=2}\right)^{-\tau_i^{q=2}}}/_{\psi_i}\right]$ is the scale parameter, "q":= $\tau_i^{q=1}$ is the entropy parameter,

[9]The selection of the q-form is based on "information entropy" (Tsallis, 1988) or the "information rate" of a DGP (Shannon, 1948), which is introduced in the "q"-form by measuring the impact of the "surprise" through the Box-Cox transformed parameter (1 − "q"), which captures the degree of extensivity of the stochastic DGP; or else the impact of an observation in changing the moments of the overall distribution.

$\tau_i^{q=2}$ is a shape parameter and e_q is the q-Exponential distribution (Box and Cox, 1964) that collapses to the exponential when $\tau_i^{q=1} = 1$. It generates non-monotonic hazard functions:

	$0 < \tau^{q=2} < 1$	$\tau^{q=2} \to 1$	$\tau^{q=2} > 1$
$0 < \tau^{q=1} < 1$	Bath-tub	Increasing	Increasing
$\tau^{q=1} \to 1$	Decreasing	Flat	Increasing
$1 < \tau^{q=1} < 2$	Decreasing	Decreasing	Unimodal

The cummulative hazard function, $H^k(t|\mathcal{F}_s)$ (Eq.(14)), is (Nadarajah and Kotz, 2007):

$$H(t|\mathcal{F}_s) = \begin{cases} \left(2-\tau^{q=1}\right)\left(A_i(t-s)\right)^{\tau^{q=2}} \sum_{j=0}^{\infty} \frac{\left[\left(\left(1-\tau^{q=1}\right)A_i\right]^j}{j+1} & when\, \tau^{q=1} < 1 \\ \left(A_i(t-s)\right)^{\tau^{q=2}} & when\, \tau^{q=1} \to 1 \\ \frac{2-\tau^{q=1}}{\tau^{q=1}-1} ln\left[1 + \left(\tau^{q=1}-1\right)\left(A_i(t-s)\right)^{\tau^{q=2}}\right] & when\, 1 < \tau^{q=1} < 2 \end{cases} \tag{12}$$

2.4 Hazard Functions and Agent Types

Our framework can cover a wide range of different agent-types, as long as their actions can be linked to a distinct pattern in the arrival rate of observable factors. In particular, the underlying concept is that the actions of a particular agent type are reflected on a set of observable factors (variables, $\boldsymbol{J}_i$), such as trading volume (Easley and O'hara, 1992), trading imbalance (Sarkar and Schwartz, 2009), order flow (Easley et al., 2014), etc, which, ultimately determine the arrival rate (shape of the hazard function $\lambda^k(t, \boldsymbol{\tau}^k|\mathcal{F}_s)$) of a variable ($\chi_i$) that is inextricably linked to their presence. This is done by using $\boldsymbol{J}_i$ to split the sample into sub-groups, each one of which exhibits a hazard function with a different shape (values of $\boldsymbol{\tau}^k$). Consequently, the building blocks for our modelling are:

Table 1 summarizes the necessary inputs for identifying the presence of an agent-type. For illustration, we introduce this to identify the presence of informed agents in Section 3, but other applications are also discussed in the Appendix B.3. In section 3, we follow relevant literature (Easley and O'hara, 1992, Easley et al., 2014, Hujer and Vuletić, 2007) and we assume that the actions of informed agents are reflected on trading intensity ($\boldsymbol{J}_i$ is defined in Section 3). Then, Eqq.(9)-(10) link different values of trading intensity to the

Table 1: Hazard Functions and Agent Types

Economic Concept	**Statistical Variable**	**Notation**
Observable Factors	Set of threshold variables	$\boldsymbol{J}_i$
The speed of events	The modelled variable	χ_i
The arrival rate of the events	The Hazard Function	$\lambda^k(t, \boldsymbol{\tau}^k \vert \mathcal{F}_s)$
Variations in observable factors	Regimes of the threshold variables	$J_i^q < j_{m=2}^q$ ... $j_{m=\mathcal{M}-1}^q \leq J_i^q < j_{m=\mathcal{M}}^q$ ...$J_i^q \geq j_{m=M}^q$
Variations in the arrival rates	Different distributions defined by shape/scale parameters	τ_m^q

shape parameter $\tau_m^{q=2}$ of the q-Weibull distribution in Eq.(11). The value of this parameter is associated with differently shaped arrival rates of trades (in the case of PIN in Section 3.1) or volume (in VPIN in Section 3.2). According to previous studies (Easley et al., 2014), uninformed agents arrive randomly, following a flat hazard function. This is observed when $\tau_m^{q=2} \to 1$. So the presence of uninformed agents can be identified by the values of $\boldsymbol{J}_i$ that are associated with $\tau_m^{q=2} \to 1$. No prior assumption is required concerning the values of $\boldsymbol{J}_i$ and their link to the presence of this agent-type. The data will identify this through the shape of the hazard function. Furthermore, considering that informed agents have no incentive to trade when there is no price unresolved information, their probability of entering the market should be higher when new information arrives and should decay to zero, when information becomes public. So, in our setup, informed trading would be identified by the values of $\boldsymbol{J}_i$ that are associated with $\tau_m^{q=2} < 1$ and, thus, with a decreasing hazard function. In consistency with theory and previous studies, but without any prior assumption, we find that higher trading intensity is associated with an increased probability of informed trading.

2.5 A Decision-Theoretic Interpretation

But how is the temporal footprint of each agent linked to their demand functions, in a unique way, in order to become identifiable? To address this point we model the behavior of a representative agent within each agent-type as an optimization problem, which ultimately determines their demand and, consequently, their trading pattern. Representative agents choose a timing distribution over $t \in [0, +\infty)$ (rather than a single point), trying to i) maximize their utility with respect to the timing of their actions and ii) minimize the adverse impact of their actions. They choose a conditional density $f^k(t_i \mid \mathcal{F}_{s(i)})$, where $\mathcal{F}_{s(i)}$ denotes the filtration available at the decision epoch $s(i)$ for unit i and solve an entropy-regularized utility maximization problem, with $\mathcal{D}$ being the set of admissible densities on $[0, T]$:

$$\max_{f^k \in \mathcal{D}} \mathbb{E}_{f^k} \left[U^k(t_i \mid \mathcal{F}_{s(i)}) \right] + \eta^k \cdot H(f^k), \tag{13}$$

This formulation captures the agent's trade-off between expected utility and strategic placements. $U^k(t \mid \mathcal{F}_{s(i)})$ is the utility from acting at time t and, $H(f) := -\int_0^T f(t) \log f(t)\, dt$ is the Shannon entropy, which captures the "surprise" implied by acting at time t_i and thus, of the impact of this action. The first part, accounts for the direct utility gained from acting according to this trading pattern and it is directly derived from their intrinsic motivation. The second part is a regularization of this pattern according to the subsequent trading impact. This accounts for potential strategic action implied by these intrinsic characteristics. [10] Heterogeneity within and across agent types could be accounted for with $\eta^k \geq 0$, which controls the strength of the regularization.

To illustrate this, consider an uninformed agent, with a constant utility with respect to time (Easley et al., 2014), $U^{\text{uninf}}(t \mid \mathcal{F}_{s(i)}) = \bar{U}$. Her optimization reduces to pure entropy maximization, such as minimizing trading costs, by minimizing the impact of her trades, $\max_{f \in \mathcal{D}} H(f)$. This yields a unique solution; the exponential distribution, $f(t) = \lambda e^{-\lambda t}$, with a flat hazard $\lambda(t) = \lambda$. This binds the characteristics of uninformed agents to a unique shape (flat) of the hazard function.

Along the same lines, informed agents (Section 3) have an incentive to act before private information decays in value and thus, their utility decreases in t. For instance, $U^{\text{inf}}(t \mid$

[10]For example, a liquidity trader would try to minimize the liquidity impact of her trades (Hasbrouck and Schwartz, 1988), an informed agent would try to minimize the price impact (Engle, 2000), while an impatient agent would try to minimize the time impact Foucault et al. (2005), Engle and Russell (1998) of her actions.

$\mathcal{F}_{s(i)}) = -\alpha t^q$, with $\alpha > 0$ and $q > 0$. In this case, the optimal density from (13) takes the Gibbs-form $f(t) \propto \exp\left(-\alpha t^q/\eta\right)$, which is a member of the q-Weibull family: $f(t) = {}^a\!/_b\,({}^t\!/_b)^{a-1}\exp_q\left(-\left({}^t\!/_b\right)^a\right)$, where $a, b > 0$ and $\exp_q(-x)$ denotes the q-exponential function. This formulation directly links the shape parameter a and the entropic index q to the agent's underlying preferences and strategic environment. Steeply decreasing utilities result in decreasing hazard functions ($a < 1$).

Consequently, Eq. (13) is the underlying mechanism that links the intrinsic qualitative characteristics of each agent-type to a uniquely shaped hazard function. As long as, the characteristics of an agent type can be mapped into a unique utility function and entropy sensitivity, then her presence can be linked to a unique hazard function. Conditionality and heterogeneity can also be further addressed in the parameters of Eqq. (3)-(11).

2.6 The (i)ntensity-based (R)elative (P)roportion

Eq.(8), defines the probability of an agent-type entering the market, based on past information. The prediction is formed by aggregating the outputs of multiple specialized components (experts) through a gating mechanism ($\boldsymbol{G}_i$ and L_i^k's), which map covariates to regime probabilities and, ultimately, to prior weights over type-specific parameters $\boldsymbol{\tau}^k$. The resulting mixture, assuming a flexible distribution, enables the model to adaptively capture complex temporal trading dynamics through a structured yet data-driven prior. The prior L_i^k is then a full-sample summary of how often each expert is activated under optimal parameterization.

This estimate functions as a belief distribution over latent agent-types, conditional on market features $\boldsymbol{J}_i$ that determine $\boldsymbol{G}_i$ and $L_i^k(:)$'s, but the presence or the absence of an agent-type, affects these probabilities only as a sample-wide property. In HFT, though, a continuously updated probability that takes into consideration the arrival of agents would make more sense, because it could provide a probability at any point in time. We introduce a continuous time revision of the prior probabilites L_i^k, based on the cumulative hazard functions $H^k(\chi_i; \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$. We name this the (i)ntensity-based (R)elative (P)roportion (iRP):

$$\mathrm{iRP}_t^k = \frac{\mathbb{E}\left(N^k(t) - N^k(s)\middle|\mathcal{F}_s\right)}{\sum_{k=1}^K \mathbb{E}\left(N^k(t) - N^k(s)|\mathcal{F}_s\right)} = \frac{\mathbb{E}\left(\left(p_t^k|\mathcal{F}_s\right)\int_s^t \lambda^k(u)\,du\middle|\mathcal{F}_s\right)}{\sum_{k=1}^K \mathbb{E}\left(\left(p_t^k|\mathcal{F}_s\right)\int_s^t \lambda^k(u)\,du\middle|\mathcal{F}_s\right)} \Rightarrow$$

$$\mathrm{iRP}_i^k = \frac{L_i^K(:)H^k(\chi_i;\boldsymbol{\tau}^k|\mathcal{F}_{i-1})}{\sum_{k=1}^K L_i^K(:)H^k(\chi_i;\boldsymbol{\tau}^k|\mathcal{F}_{i-1})} \tag{14}$$

$\mathrm{iRP}_i^k \equiv \mathbb{P}_i^k$ is a posterior estimate of $p_i^k|\mathcal{F}_{i-1} = L_i^k(:)$ that is continuously updated by the presence $\lambda^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$ or the absence $S^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$ of a specific agent-type. After revealing $\boldsymbol{J}_i$, the prior probability $p_i^k|\mathcal{F}_{i-1}$, is the equivalent of a sample wide weighting for each agent. However, depending on whether the arrival time of the trade is "as expected" by the prior $\lambda^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$ and as time lapses $H^k(\chi_i; \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$, the prior is revised accordingly. This, continuous revision, enables the estimate of a posterior probability over any interval.

According to Jacobs et al. (1991), the posterior distribution $\mathbb{P}_i = \sum_{k=1}^K \mathbb{P}_i^k$ is not directly obtained via Bayes' rule but emerges as the solution to a variational optimization problem, motivated by expert aggregation based on a regularized partial likelihood loss. Given the prior belief $L_i = \sum_{k=1}^K L_i^k(:)$ and the observed duration χ_i, $\mathbb{P}_i$ can be expressed as the minimizer of the following objective function, with Δ^K being the standard $K-1$ simplex:

$$\mathbb{P}_i = \arg\min_{\mu\in\Delta^K}\left\{\sum_k \mu(k)(-\log H^k(\chi_i;\boldsymbol{\tau}^k|\mathcal{F}_{i-1})) + \mathrm{KL}(\mu\|L_i)\right\}, \tag{15}$$

where, $H^k(\chi_i; \boldsymbol{\tau}^k|\mathcal{F}_{i-1})$ is the hazard rate implied by type k at the observed duration χ_i and $\mathrm{KL}(\mu\|L_i) = \sum_k \mu(k)\log(\mu(k)/L_i^k(:))$ is the Kullback-Leibler divergence penalizing deviation from the prior. μ is the estimated quantity (i.e., distribution of weights across agent-types k), which, according to the minimization criterion above is a balance between a data dependent loss ($\sum_k \mu(k)(-\log H^k(\chi_i; \boldsymbol{\tau}^k|\mathcal{F}_{i-1}))$) and overall model consistency ($\mathrm{KL}(\mu\|L_i)$).

The first element measures a "loss", $\ell_k = -\log H(\chi_i; \boldsymbol{\tau}_k)$, at every observation that penalizes deviation from prior. At time i, when the duration χ_i is observed, ℓ_k gets lower (higher) values when the inferred agent-type k is (not) the one observed and, thus, the inferred and actual hazard functions do (not) match. This makes ℓ_k an explicit measure of "surprise" at observation level: How different from expectation is the probability that the trade at time i is initiated by agent-type k, based on information summarized in χ_i. In parallel, the second element, $\mathrm{KL}(\mu \| L_i)$, controls for overall (sample) model consistency, by penalizing μ for larger deviations from the prior.

2.6.1 iRP and Market Reflexivity

$\mathbb{P}_i$ is a continuous time update of prior beliefs based on the realization of actual events. When (the duration of) a trade is as expected the revision is small, indicating that the new trade does not introduce any new information. In the opposite case, a larger deviation from the prior suggests that the new trade carries some kind of new information, not incorporated in the prior. This makes the difference between the posterior $\mathbb{P}_i$ and the prior L_i an observable "surprise statistic" $S_i := \mathrm{KL}(\mathbb{P}_i \| L_i)$ that captures the information gain at trade i:

$$S_i := \sum_{k=1}^{K} \mathbb{P}_i^k \log\left(\frac{\mathbb{P}_i^k}{L_i^k}\right). \tag{16}$$

S_i is the degree of change in the prior belief because of the realization of the event at time i; thus, it is an explicit measure of how informative trading is at a particular moment. Beyond the estimate of the probability concerning the presence of an agent-type, $\mathbb{P}_i^k$, our modeling provides also an assessment of how much this estimate is affected by market conditions ($L_i^k | \boldsymbol{J}_i$) or by the realization of a trade ($H(\chi_i; \boldsymbol{\tau}_k)$). Higher values for S_i indicate that the realization (duration) of the last trade affected more $\mathbb{P}_i^k$ than the market conditions $\boldsymbol{J}_i$.

This dissection is a major implication of our modeling because it enables an assessment of how informative trading speed is; an inextricable feature of HFT (O'Hara, 2015). Higher values for S_i suggest that conventional metrics $\boldsymbol{J}_i$, such as trading volume, do not describe perfectly the composition of agent-types, but it is significantly complemented by their arrival rates; a feature that becomes more prevalent during monre intense endogenous trading. That makes S_i a measure of trading endogeneity; How much trading depends on trading signals, contained in $\boldsymbol{J}_i$, compared to arrival rates. When $\boldsymbol{J}_i$ is sufficient, the prior estimate of the probability is adequate, $\boldsymbol{J}_i \rightarrow \boldsymbol{W}_i \rightarrow L_i$, and, the observed durations match the expected. When, however, $\boldsymbol{J}_i$ is not sufficient, the prior is biased and the observed durations are different from expected. The estimate of the (posterior) probability is revised according to new durations (relative speed) and thus trading becomes more endogenous. Consequently, S_i can be thought of as an indirect local measure of "reflexivity" (Soros, 1994), which measures the degree at which trading is affected by trading itself. Naturally, higher reflexivity (higher S_i) should be correlated with microstructural issues related to information propagation, such as variance (Easley and O'hara, 1992), order flow toxicity (Easley et al., 2014) and/or Ultra-fast Extreme Events (Johnson et al., 2013).

3 Empirical Application: iRP and Information

Eq.(14) expresses the relative proportion of a specific agent-type as a function of the cumulative hazard function that describes her 'on average' trading characteristics. Consequently, it is essential to map distinctive trading characteristics to distinctively shaped hazard functions. Eqq.(3)-(10) propose an empirical framework that links the magnitude of observable variables with different values of shape/scale parameters of the distribution of durations that determine the shape of the hazard function. Therefore, Eqq.(3)-(14) can identify a plethora of different agent-types, as long as their trading behavior (i) has a material impact on some observable variables (ii) in a manner that is associated with a distinctively shaped hazard function. There is no restriction in the number of variables or in the shapes of hazard functions and, thus, the modeling here can theoretically identify the existence of any number of agent-types. Unlike previous approaches, iRP is estimated in the limit when $\Delta t \rightarrow 0$ and not over an interval, which is more appropriate for an HFT environment.

This is a significant generalization over previous approaches and it constitutes the major contribution of this study. To demonstrate the relative merits of iRP, we illustrate how it can be used in detecting the presence of information in this section (probability of informed trading), while in Section 4 and Section 5 we present its empirical and theoretical properties, while presenting briefly an alternative two ways that it can be used: i) to investigate the agent composition of the market and ii) to assess the degree of market reflexivity.

3.1 (i)ntensity-based (P)robability of (IN)formed trading (iPIN)

One of the most well-documented concepts in the literature (Easley and O'Hara, 1995; O'Hara, 2003), is the presence of private information. This information refers to shifts in fundamental valuations and it is revealed gradually through the actions of traders that have a timing advantage on it (O'Hara, 2015). These agents, called "informed", as opposed to "uninformed", exploit their information benefit and the direction of their trading would deviate systematically from a random arrival of order flow. Easley et al. (1996, 1997b) introduce a measure of the (P)robability of (IN)formed trading (PIN) derived from these deviations. The basic notion in PIN is that when there is no information only liquidity traders exist in the market and the direction of their trading should be random with a probability of 50%. In contrast, when there is private information, informed agents align their demand with the direction of the signal and this creates an order imbalance. The PIN interprets the magnitude of these deviations as increased presence of informed agents. Accordingly, the PIN is defined as, $\frac{\alpha\mu}{(\alpha\mu+e)}$, where a is the probability of the existence of private information and μ and e are the arrival rates of informed and uninformed agents, accordingly. This intuitive measure estimates the proportion of informed agents, $(\alpha\mu)$ relative to the number of all trades $(\alpha\mu + e)$, which is equivalent to a time invariant version of Eq. (14).

Albeit insightful, PIN has a notable limitation. It is derived from the trade direction, which is usually latent in raw data. To mitigate the issue of noise in trade direction classification algorithms (Ellis et al., 2000), PIN is estimated over a period of time, without a clear selection criteria for the optimal interval length (Andersen and Bondarenko, 2014).[11]. This makes it a rather slow interval estimate, inapt for using it a higher trading frequencies required in algorithmic trading (O'Hara, 2015). This is further exacerbated by the fact that PIN is a time invariant estimate. Several approaches propose time variant probabilities conditional on either daily trade imbalances (Lee, 1989) or trade durations (Easley et al., 2008). These approaches account for differential arrival rates of buys and sells on a daily scale, but they still require a trade classification algorithm and the selection of an optimal interval to mitigate the impact of classification bias; thus, it is relatively slow for HFT standards.

Instead, Eq.(14) shifts the focus from the aggregated properties of order flow to the aggregation rate itself to estimate the instantaneous probability of a trade to be informed $\mathbb{P}_i^{inf}$ and their arrival rates $H^{inf}(\chi_i;\boldsymbol{\tau}^{inf}|\mathcal{F}_{i-1})$. Then using Eq.(14), PIN can be transformed into an intensity-based equivalent, $\text{iPIN}_i = \text{iRP}_i^{\text{inf}}|\mathcal{F}_{i-1}$:

$$\text{iPIN}_t = \frac{(\alpha_t|\mathcal{F}_s)\mu}{(\alpha_t|\mathcal{F}_s)\mu + (e_t|\mathcal{F}_s)} = \frac{\mathbb{E}\left(\left(p_t^{inf}\middle|\mathcal{F}_s\right)\int_s^t \lambda^{inf}(u)\,du\right)}{\sum_{k=1}^{K}\mathbb{E}\left(\left(p_t^k|\mathcal{F}_s\right)\int_s^t \lambda^k(u)\,du\right)} = \frac{\left(p_t^{inf}\middle|\mathcal{F}_s\right)H^{inf}(t|\mathcal{F}_s)}{\sum_{k=1}^{K}\left(p_t^k|\mathcal{F}_s\right)H^k(t|\mathcal{F}_s)} \quad (17)$$

which becomes in terms of parameters $\boldsymbol{\tau}^k$:

$$\text{iPIN}_i = \frac{\mathbb{E}\left(L_i^{inf}(\boldsymbol{W}_i)H^{inf}(\chi_i;\boldsymbol{\tau}^{inf})\right)}{\sum_{k=1}^{K}\mathbb{E}\left(L_i^k(\boldsymbol{W}_i)H^k(\chi_i;\boldsymbol{\tau}^k)\right)} = \frac{\sum_{Q\bigotimes(m:k)}\left\{\prod_{Q\bigotimes m:k} W_{m,i}^q H^{Q\bigotimes(m:k)}(\chi_i;\boldsymbol{\tau}^{Q\bigotimes(m:k)}|\mathcal{F}_{i-1})\right\}}{\sum_{Q\bigotimes M}\left\{\prod_{Q\bigotimes M} W_{m,i}^q H^{Q\bigotimes M}(\chi_i;\boldsymbol{\tau}^{Q\bigotimes M}|\mathcal{F}_{i-1})\right\}}$$

The first line of Eq. (17) express PIN as a function of the prior probabilities $p_t^{inf}|\mathcal{F}_s$ and the hazard functions $H^{inf}(t|\mathcal{F}_s)$, as in Eq. (14). The numerator is the number of trades initiated by informed agents and the denominator is the total number of trades.

[11] Easley et al. (1996, 1997b) suggest that one month worth of data produces is sufficient for a relatively accurate estimate of PIN; a claim that has been debated in the literature (Pöppe et al., 2016)

This is a definition identical to conventional PIN, but the second line, of Eq. (17) provides an instantaneous estimate of PIN that can be estimated conditionally over any interval. More precisely, the parameters in τ^k determine the shape of the hazard function that is then matched to the trading characteristics of different agent-types. The parameters in W_i capture the probability of belonging to different regimes, m, and consequently, the prior probabilities that a trade is initiated by a particular agent-type. According to relevant literature (De Luca and Zuccolotto, 2006, Bowsher, 2007, Kalaitzoglou and Ibrahim, 2013), informed traders are more likely to exhibit a downward slopping hazard functions, due to their incentive to act prior to their informational advantage becoming public information.

A matching shape (decreasing) in Eq.(11) would be observed when $\tau^{q=1} \geq 1$ and $\tau^{q=2} \leq 1$. The regimes, m of $\tau^{q=1}$ $\tau^{q=2}$ that lead to a decreasing shape of the hazard function, identify informed trading $(m : k = inf)$. Then, according to Eq.(8), the aggregated number of informed agents over time is the sum of all the probabilities, i.e., $\prod_{Q \bigotimes M} W^q_{m,i}$, of intersections $Q \bigotimes (m : k) \equiv Q \bigotimes (m\,(\tau^{q=1}_m \geq 1, \tau^{q=2}_m \leq 1) : k)$ in the contingency table, where $\tau^{q=1} \geq 1$ and $\tau^{q=2} \leq 1$, times the respective cumulative hazard functions of these intersections, i.e., $H^{Q \bigotimes \left(m\left(\tau^{q=1}_m \geq 1, \tau^{q=2}_m \leq 1\right):k\right)}(\chi_i; \boldsymbol{\tau}^{Q \bigotimes \left(m\left(\tau^{q=1}_m \geq 1, \tau^{q=2}_m \leq 1\right):k\right)} | \mathcal{F}_{i-1})$. This is then compared to the expected number of the trades of all agent-types, defined explicitly in the denominator. A specific version of Eq.(17), as an example, is discussed in Section 4.

3.2 VPIN: A High(er) Frequency PIN and iVPIN

A fundamental difference between iPIN and PIN is that it derives the probability not from the aggregated sign of trades, but from trading frequency and volume. This is in line with previous studies that account for the time (Easley et al., 2008) and the volume (Easley et al., 2011, 2012) dimensions. PIN is a rather noisy measure due to its reliance on the noisy signal of trade initiation and Easley et al. (2011) suggest a more HFT-friendly estimate of PIN, based on fixed buckets of volume or time, as well as from price changes (rather than trade imbalances).[12] The new measure, VPIN, is an explicit recognition that the speed of volume accumulation might be better associated with information. The aggregated signed volume (Volume Imbalance, VI_i) approaches $\alpha\mu$, while the aggregated volume accounts for all trading $(\alpha\mu + e)$. Using a rolling window of length n, VPIN_t can be estimated as $\mathrm{VPIN}_{t_{bucket}} = \frac{\sum_{t_{bucket}-n}^{t_{bucket}} |V^B_{t_{bucket}} - V^S_{t_{bucket}}|}{\sum_{t_{bucket}-n}^{t_{bucket}} V_{t_{bucket}}}$, with $\frac{\mathbb{E}(\|V^B - V^S\|)}{\mathbb{E}(V^B + V^S)} \rightarrow \frac{\alpha\mu}{(\alpha\mu+e)}$, where $\sum_{t_{bucket}-n}^{t_{bucket}} \left|V^B_{t_{bucket}} - V^S_{t_{bucket}}\right|$ is the $\mathrm{VI}_{t_{bucket}}$ and $\sum_{t_{bucket}-n}^{t_{bucket}} V_{t_{bucket}}$ is the total volume over the last n buckets prior to t_{bucket} .

This bucket-ing mechanism makes it rather slow for HFT. Whereas VPIN approaches informed trading from the perspective of aggregated signed volume, Eq.(14) focuses on the aggregation process itself. In particular, $N(t)$ in iPIN (Eq.(17)) counts the total number of trades and a time transformation is required to account for the volume dimension. In line with price (Engle and Russell, 1998) or volume (Kalaitzoglou and Ibrahim, 2015) durations, which measure the arrival rate of a "differently" defined event (a unit of price change or volume), we transform $t \rightarrow t^*$, where the counting process $N(t^*) = \sum_{i \geq 1} 1(T_i < t^*)$ counts how many events (units of) occur over a time interval. Define $\chi_i^* = t_i^* - t_{i-1}^*$ as the waiting time for a given magnitude of volume. $N(t^*)$ counts how much volume is traded over a time interval and, following Eq.(2), it can be traced back to the agent-type who initiated it. $N^k(t^*)$ is defined as $N^k(t^*) = \sum_{i \geq 1} 1(T_i < t*) = \sum_{k=1}^{K} \sum_{i \geq 1} 1(T < t*)(Z_i = k)$ and counts the volume traded by each agent-type k. This way, the (volume) trading activity of each agent-type is modelled and its relative proportion to the total trading activity can be computed according to Eq.(14). The intensity-based VPIN (iVPIN) can be formulated as:

$$\mathrm{iVPIN}_i = \frac{\sum\limits_{Q\bigotimes\left(m\left(\tau_m^{q=1}\geq 1, \tau_m^{q=2}\leq 1\right):k\right)} \left\{ \prod\limits_{Q\bigotimes\left(m\left(\tau_m^{q=1}\geq 1, \tau_m^{q=2}\leq 1\right):k\right)} W^q_{m,i} H^{Q\bigotimes(m:k)}(\chi_i^*; \boldsymbol{\tau}^{Q\bigotimes(m:k)} | \mathcal{F}_{i-1}) \right\}}{\sum\limits_{Q\bigotimes M} \prod\limits_{Q\bigotimes M} W^q_{m,i} H^{Q\bigotimes M}(\chi_i^*; \boldsymbol{\tau}^{Q\bigotimes M} | \mathcal{F}_{i-1}))} \tag{18}$$

Eq. (18) differs from Eq.(17) in the way the conditional intensities are considered. In-

[12] Andersen and Bondarenko (2014) report that different time interval lengths change its distributional properties rather than mitigating the issue

stead of the $\lambda^k(t)$ that account for the arrival rate of the transactions instigated by agent-type k, Eq.(18) is based on $\lambda^k(t^*)$ that account for the arrival rate of volume traded by agent-type k. This way, the regimes, m of $\tau^{q=1}$ and of $\tau^{q=2}$ that lead to a decreasing shape of the hazard function, identify informed trading, by considering the arrival rate of their trading volume. Then the numerator defines the aggregated volume traded by informed agents over time as the sum of all the probabilities, i.e., $\prod_{Q \bigotimes M} W^q_{m,i}$, of intersections $Q \bigotimes (m\,(\tau^{q=1}_m \geq 1, \tau^{q=2}_m \leq 1) : k)$ in the contingency table, where $\tau^{q=1}_m \geq 1$ and $\tau^{q=2}_m \leq 1$, times the respective cumulative hazard functions of these intersections, i.e., $H^{Q \bigotimes (m:k)}(\chi^*_i; \boldsymbol{\tau}^{Q \bigotimes (m:k)}|\mathcal{F}_{i-1})$. This is then compared to the aggregated total volume that is defined in the denominator. This is an alternative measure to VPIN in continuous time that does not suffer from sampling bias.[13]

3.3 From iRP to iPIN and iVPIN. An Empirical Specification

Eqq (17)-(18) present the general definitions for iPIN and iVPIN, but how could this be applied in a real-world scenario? The following specification of iRP is employed as an indicative example to show how both can be linked to estimable parameters:

Table 2: Indicative Empirical Specification Example

$$\chi_i = \psi_i \epsilon_i,\ \psi_i = \mathbb{E}\left(\chi_i|\mathcal{F}_{i-1}, \omega, \beta, \phi, \delta\right) = \omega + \beta\psi_{i-1} + (\chi_i - \beta\chi_{i-1}) - (\tilde{\chi}_i - \phi\tilde{\chi}_{i-1})$$

$$f\left(\chi_i|\mathcal{F}_{i-1}; A_i = \left[\Gamma\left(1+\frac{1}{\tau_i^{q=2}}\right)^{-\tau_i^{q=2}} /\psi_i\right], \tau^{q=1}, \tau_i^{q=2}\right) = (2-\tau^{q=1})\frac{\tau_i^{q=2}}{\chi_i}\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}} e_q\left(-\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}}\right)$$

$$\tau_i^{q=2} = \underbrace{\left(\overbrace{G^{q=2}_{m=1,i}}^{1} - G^{q=2}_{m=2,i}\right)}_{W^{q=2}_{m=1,i}}\tau^{q=2}_{m=1} + \underbrace{\left(G^{q=2}_{m=2,i} - \overbrace{G^{q=2}_{m=3,i}}^{0}\right)}_{W^{q=2}_{m=2,i}}\tau^{q=2}_{m=2}$$

$$\text{for}\quad G^{q=2}_{m=2,i} = \left(1 + e^{-g^{q=2}_{m=2}(J_i - j^{q=2}_{m=2})}\right)^{-1},\ \text{with } J_i = ti_i = (duration_i * K(volume_i))^{-1}$$

[13] The empirical properties of VPIN depend on the selection (Andersen and Bondarenko, 2014) of the time (Easley et al., 2011, 2012) or volume (Easley et al., 2016) bucket size, as well as on trade classification. iVPIN, instead of selecting an "optimal" sampling frequency that would account for the aggregation properties of information, it models the aggregation process itself with the intensities of the different agent-types. This way, iVPIN, unlike VPIN, derives the presence of private information from the relative speed of volume accumulation without suffering from sampling frequency bias.

In Table 2, χ_i is defined as $\chi_i = \Delta t_i$, when a trade-"clock" is employed (relative to PIN), or as as $\chi_i = \Delta t_i^*$, when a volume-"clock" is employed (relative to VPIN). The conditional mean follows a FIACD$(1, \delta, 1)$ (Jasiak, 1999) specification, where $\tilde{\chi}_i = (1-L)^\delta \chi_i$ is a fractional difference (L is the lag operator) of χ with a fractional differentiating parameter δ.[14] $\tau^{q=1}$ is time invariant and $\tau_m^{q=2}$ has a dimension of $M = 2$ regimes, identified by different levels of the threshold variable J_i. $J_i = ti_i$ is the inverse of volume weighted duration, which is estimated as the product of diurnally adjusted durations and $K(volume_i) = e^{\left(-\frac{volume_i - \overline{volume}}{\sigma_{volume}}\right)}$, where $\overline{volume}$ and σ_{volume} are the sample mean and standard deviation of volume, estimated per stock. J_i is an increasing variable of trading intensity. This specification models the inter-trade or volume weighted inter-trade waiting times as an infinite mixture of two q-Weibull distributions that are identified by a sole time-varying shape parameter, $\tau_i^{q=2}$.[15]

To articulate this more firmly, trading intensity is employed as a classification variable, J_i, different levels of which are associated with differently shaped hazard functions and thus, with different agent-types. In line with prior literature (Easley and O'hara, 1992) higher levels of trading intensity, i.e., $J_i^{q=2} > j_{m=2}^{q=2}$, are expected to be associated with decreasing hazard functions, i.e., $\tau_{m=2}^{q=2} < 1$, and, therefore, with a higher probability of this trade to have been initiated by an informed agent, captured by a higher value of $W_{m=2,i}^{q=2}$. Accordingly, $W_{m=2,i}^{q=2}$ is an estimate of the probability of the existence of private information and $W_{m=2,i}^{q=2}\lambda_0^{k=inf}(t, \tau_{m=2}^{q=2})$ is the conditional instantaneous probability of the arrival of an informed trader. Then, assuming $\tau^{q=1} \to 1$, $\text{iVPIN}_i = \text{iVPIN}_{i:t_i \to t_i+\Delta t}$ can be written as:

$$\text{iVPIN}_{i:t_i \to t_i+\Delta t} = \frac{W_{m:\text{inf},i}^{q=2}(A_i((t_i+\Delta t)-t_i))^{\tau_{m:\text{inf}}^{q=2}}}{\sum_{m=1}^{3} W_{m,i}^{q=2}(A_i((t_i+\Delta t)-t_i))^{\tau_m^{q=2}}}$$

[14] This specification is preferred over more conventional specifications (Engle and Russell, 1998) due to the potentially long memory of durations. δ is the degree of decay that captures, in a sense, the life span of trading information. This allows distant past events to affect the conditional expectation of χ, $\psi_i = \mathbb{E}(\chi_i|\mathcal{F}_{i-1})$, and its conditional distribution, $f(\chi_i|\mathcal{F}_{i-1}; A_i)$. Variations in $f(\chi_i|\mathcal{F}_{i-1}; A_i)$, captured by $\boldsymbol{\tau}_m^q$, depend on ψ_i and, thus, on the long memory, i.e., δ, and its impact, i.e., ϕ, on ψ_i. This way, the identification of the different agents is done in a way that accounts for market reflexivity.

[15] When $\tau^{q=2} = 1$ the hazard function is flat matching the time-invariant arrival rate of uninformed agents, while when $\tau^{q=2} < 1$ ($\tau^{q=2} > 1$) the hazard function is decreasing (increasing) matching the characteristics of informed (technical) traders (Kalaitzoglou and Ibrahim, 2013).

4 Empirical Properties of the iRP

Section 4 presents a real-world application in order to illustrate the empirical performance of iRP. Section 4.1 presents the data sample employed, while Section 4.2 discusses the estimation procedure for all the metrics used. Section 4.3 discusses the empirical findings on detecting information, the probability of informed trading and market reflexivity.

4.1 Data

The specification defined in Eq.(17) (as well as some extensions in Appendix B.3) is indicatively evaluated on a one-year, 2/1/2019-6/12/2019, sample of all constituents of Dow Jones Industrial Average (aka DOW30). The primary objective is to evaluate the performance of iRP against interval-based metrics, which would be better facilitated by a sample with minimal market-specific stylized factors or other types of trading biases. DOW30 consists of liquid, large-cap stocks and, from this perspective, is less likely to suffer from market specific biases. In addition, the constituents do not change during the sample period and, therefore, the trading activity is unlikely to be affected by portfolio rebalances. Also, the sample stops prior to the COVID-19 news and, thus, it should not be affected by it.[16]

For each transaction we record the date, time-stamp (millisecond), price ($) and trading volume (# stocks). The trade direction is inferred by applying the ”EMO” (Ellis et al., 2000) rule. All observations outside the “normal trading hours” as well as the first transaction of each day are omitted. Furthermore, all trades with identical time stamp, price and trade initiation are considered as one segmented trade with aggregated volume. This accounts for passive splitting and mitigates the information loss of the trade classification algorithm. In addition, it reduces the proportion of trades with zero duration or zero price change, which, beyond the computational benefits, creates a sample that is focused on the time evolution of

[16] All our findings remain robust to sampling bias. We test our estimates to two one-month samples from 2024 (April) and 2025 (October) and the findings remain qualitatively the same. We report the full sample estimates from 2019, because it appears to be a sample with minimum macro-economic bias.

Table 3: Descriptive Statistics: Full Sample

	Full Sample #230,084,293				Simulation 1 asset 100 days x #100,000			
	Avg.	Min.	Max.	Std.	Avg.	Min.	Max.	Std.
Return	0	-9.76	9.77	0.04	0	-8.72	14.01	0.04
Volume	205.8	1	88,000	538.94	355.48	1	112,000	355.78
Duration	0.90	2.1E-5	4,500	2.34	0.85	0	3,856	1.99

Table 3 presents the descriptive statistics, i.e., the average (mean), the maximum (max), the minimum (min) and the standard deviation (std) of duration (in seconds), trading volume (in number of stocks) and price change (in $'s). The left panel presents the cross-sectional estimates of the statistics for the full sample that consist of all constituents of DJ30 (full table in Appendix). The left panel presents the cross-sectional estimates of the simulation that generates a calendar quarter worth of data (100 days, assuming 100,000 observations per day) the # sign is the count of observations per stock.

volume and price change without thinning the data. Moreover, duration has been computed as the time between two consecutive trades excluding the overnight period and has been diurnally adjusted (Engle and Russell, 1998). This results in a panel dataset of all filtered transactions of 30 firms with 230,084,293 unique observations.

The basic statistics of the final sample are presented in Table 3 (per-stock in Appendix A). The sample consists of liquid stocks with an average duration of less than 1 second; around 0.9 seconds. The average volume per trade is just over 200 stocks and it is over-dispersed, indicating a wide range of values. Price changes exhibit some moderate variation with a standard deviation of 0.04. Our sample is relatively homogeneous, with adequate variation, consistent with prior literature.

4.2 Empirical Estimation and Performance

In order to get comparable estimates, for both intensity-based and interval measures, the focus of the analysis is time, rather than volume. This is not an issue for iVPIN, but for VPIN the time and length of interval are parameters of choice. Although fixed-volume buckets might perform better (Pöppe et al., 2016), the empirical setup here employs fixed-time buckets in order to create contemporaneous estimates. The sample is split into fixed-time buckets of different time lengths, bucketsize $= (1'', 5'', 15'', 30'', 1', 5', 15', 30', 60')'$, and all measures are estimated at the end of each bucket, at a time noted as t_{bucket}.[17] Then VPIN and iVPIN are estimated using the same sampling frequency, over 10 lags, following that: $\mathbb{E}\left(\frac{\alpha_{t_{bucket}}\mu_{t_{bucket}}}{(\alpha_{t_{bucket}}\mu_{t_{bucket}}+e_{t_{bucket}})}\right) \rightarrow \frac{\alpha\mu}{(\alpha\mu+e)}$.

VPIN. The benchmark metric for comparison is the conventional VPIN, which is estimated based on the principle (Easley et al., 2008) that $\mathbb{E}\left(\left\|V^B - V^S\right\|\right) \rightarrow \alpha\mu$ and $\mathbb{E}\left(\left\|V^B + V^S\right\|\right) \rightarrow \alpha\mu + e$. The aggregated buy $V^B \rightarrow V^B_{t_{bucket}}$ and sell $V^S \rightarrow V^S_{t_{bucket}}$ volumes are computed over different time intervals, i.e., bucketsize, with t_{bucket} marking the time at the end of each bucket, acting as a time identification. Then, $\text{VPIN}_{t_{bucket}}$ is computed over a rolling window of $n = 10$ lags as: $\text{VPIN}_{t_{bucket}} = \frac{\sum_{t_{bucket}-n}^{t_{bucket}}\left|V^B_{t_{bucket}} - V^S_{t_{bucket}}\right|}{\sum_{t_{bucket}-n}^{t_{bucket}} V_{t_{bucket}}}$.[18] In this metric, the most important element is the identification of trade initiation; a variable that is absent in raw data. This, has been shown to be an important issue (Pöppe et al., 2016), affecting the performance of VPIN, especially in combination with different sampling frequencies. The trade classification rules that exhibit superior performance are the "EMO" (Ellis et al., 2000) and the Bulk Volume (hereafter BV: Easley et al., 2012) classification. In a misuse of the term, the estimate of VPIN using the EMO classification is called PIN, due to its resemblance to the lower frequency estimate of PIN. In parallel, the BV classification derives trade initiation solely from price changes. In particular, using the notation in this paper, $V^B_{t_{bucket}}$ and $V^S_{t_{bucket}}$ are defined as $V^B_{t_{bucket}} = V_{t_{bucket}} Z\left(\frac{Price_{t_{bucket}} - Price_{t_{bucket}-1}}{\sigma_{\Delta Price}}\right)$ and $V^B_{t_{bucket}} = V_{t_{bucket}}\left(1 - Z\left(\frac{Price_{t_{bucket}} - Price_{t_{bucket}-1}}{\sigma_{\Delta Price}}\right)\right)$.

iVPIN. As an example of iRP, in Eq.(14), the conventional VPIN is compared to

[17] Pöppe et al. (2016) suggests that volume buckets are more relevant in identifying information because they are based on the speed of accumulation of the same magnitude of information, captured by a unit of volume. This would create buckets of variant time intervals, which would undermine comparability with iRP. Instead, the comparison is based on fixed intervals that vary from from 1" to 1h.

[18] The length of the rolling window of 10 is selected in order to facilitate the investigation of the performance of intensity-based measures in HFT. when bucketsize $= 1''$ the rolling window of 10 corresponds to 10". This interval is sufficiently short for algorithmic trading standards in order to evaluate whether the VPIN metric is of relevance at this sampling frequency, while it is also sufficiently long enough to avoid missing observations due to lack of data. On the opposite side, a rolling window of bucketsize $= 60'$ corresponds to a full trading day. Previous literature (Pöppe et al., 2016) shows that this interval provides reasonable estimates for VPIN, while it is still relevant to algorithmic trading, whose trading horizons rarely exceed one trading day.

an intensity-based alternative, coined iVPIN. In order to create a comparable estimate, iVPIN is estimated on a per-trade basis, considering the inter-trade durations, i.e., $\chi_{i+1} = (t_{i+1} - t_i)$ for trades or $(t^*_{i+1} - t^*_i)$ for volume weighted durations, as the waiting time, i.e.: $\text{iVPIN}_{i:\chi_{i+1}} = \frac{W^{q=2}_{m:\inf,i}(A_i\chi_{i+1})^{\tau^{q=2}_{m:\inf}}}{\sum_{m=1}^{3} W^{q=2}_{m,i}(A_i\chi_{i+1})^{\tau^{q=2}_m}}$. This provides a rather granular estimate of the expected level of VPIN, that can be, then, estimated at any desirable interval. VPIN is estimated for each bucketsize $= (1'' \, to \, 60')'$ and then an average over a time interval $n = 10$. The same approach is applied in the case of $\text{iVPIN}_{t_{bucket}}$, which is estimated as: $\text{iVPIN}_{t_{bucket}} = \frac{\sum_{t_{bucket}-n}^{t_{bucket}} \frac{\sum_i^{\#trades_{t_{bucket}}} \text{iVPIN}_i}{\#trades_{t_{bucket}}}}{n}$. For reference, when $\chi_{i+1} = (t_{i+1} - t_i)$, the intensity-based estimate is named iPIN, while, when $\chi_{i+1} = (t^*_{i+1} - t^*_i)$, it is named iVPIN.

iVPIN vs VPIN.After computing VPIN and iVPIN their performance is evaluated based on their forecasting ability on subsequent variance and the existence of UEE's. Following Andersen and Bondarenko (2014) this is evaluated based on the following regression:

$$Q_{t_{bucket}} = c_0 + c_1 \text{Metric}_{t_{bucket}-n} + \boldsymbol{c}\boldsymbol{CV}_{t_{bucket}-n} + \text{f.e.} + \epsilon_{t_{bucket}} \tag{19}$$

where, $Q_{t_{bucket}} = (\text{RV}_{t_{bucket}}, \text{UEE}_{t_{bucket}})'$. The values are multiplied $x100$ in order to adjust the decimal places of the estimated coefficients. $\text{Metric} = (\text{PIN}, \text{VPIN}, \text{iPIN}, \text{iVPIN})'$. $\text{RV}_{t_{bucket}}$ is the average realized volatility of each bucket over the time interval n. $\text{UEE}_{t_{bucket}}$ is the average number of UEE's: (Johnson et al., 2013) of each bucket over the time interval n.[19] $\boldsymbol{CV} = (\text{RV}, \text{spread}, \text{orders}, \text{average duration})'$ is a collection of standard market microstructure variables that are introduced to control for varying market conditions and the sensitivity of VPIN to them (Chakrabarty et al., 2015). f.e. is company fixed effects.

Market Reflexivity. According to Eq. (16), S_i can capture the information gain due to durations $(H^k(\chi_i, \boldsymbol{\tau}^k|\mathcal{F}_{i-1}))$ on top of market conditions $(\boldsymbol{J}_i)$. The information gain should be expected to be more intense during periods that the realization of events carries more impact on subsequent trading. We consider that, in HFT, this captures the "learning" of

[19] UEE's are defined as periods that last less than 1,500ms, they follow a run (same trade direction) that exceeds 10 trades, during which prices change by more than 0.8% of the price at the beginning of the run.

AT from the actions of other AT and, thus, we link it explicitly to market reflexivity (Soros, 1994) in HFT. According to Johnson et al. (2013), when the intensity of AT towards a specific direction increases, it creates the conditions for toxicity (Easley et al., 2014), and, ultimately UEE's (Johnson et al., 2013). Should this be the case, S_i should be correlated with these metrics, because AT "learns" more from duration during these periods and, thus, S_i should exhibit higher values.

We compute the information gain $S_i = \sum_k \mathbb{P}_i^k \log\left(\frac{\mathbb{P}_i^k}{L_i^k}\right)$, as a proxy for market reflexivity, where, $k = (inf, unin)'$, $\mathbb{P}_i = (\text{iVPIN}_{i:\chi_{i+1}} = \frac{W_{m:\text{inf},i}^{q=2}(A_i\chi_{i+1})^{\tau_{m:\text{inf}}^{q=2}}}{\sum_{m=1}^{3} W_{m,i}^{q=2}(A_i\chi_{i+1})^{\tau_m^{q=2}}}, (1 - \text{iVPIN}_{i:\chi_{i+1}}))'$ and $L_i = (W_{m:\text{inf},i}^{q=2}, ((1 - W_{m:\text{inf},i}^{q=2}))'$. We distinguish between iPIN and iVPIN, by considering $\chi_{i+1} = (t_{i+1} - t_i)$ for the former and $\chi_{i+1} = (t_{i+1}^* - t_i^*)$ for the latter. Then in a similar fashion to $\text{iVPIN}_{t_{bucket}}$, we compute, the average information gain, as a metric of the intensity of market reflexivity over a period= *bucket*, as $S_{t_{bucket}} = \frac{\sum_{t_{bucket}-n}^{t_{bucket}} \frac{\sum_i^{\#trades_{t_{bucket}}} S_i}{\#trades_{t_{bucket}}}}{n}$.

Then, we investigate whether this explicit measure of information gain is related to market microstructure systematic risk, by assessing its impact on the occurrence of UEE's and on price change volatility. We add S_i in Eq. 19 in order to capture how/whether market reflexivity interacts with information at a micro level. S_i captures how much the duration of a trade (arrival of an agent) affects the belief about the probability of a specific agent-type entering the market. When the revision is high, this specific trade influences the market-wide perception about the probability and it is expected to affect subsequent trading. This, in a sense, measures how much the market learns from a trade and how much it increases reflexivity. Should this coincide with the presence of information, higher S_i would mean higher learning rate and, thus, faster price discovery (Madhavan, 2000). However, when this is the result of higher endogenous trading (higher market reflexivity), rather than of information, order flow might become toxic (Easley et al., 2014) and prices might deviate from fundamental valuations (Ibrahim and Kalaitzoglou, 2025). Metric captures information, while S_i captures reflexivity.

4.3 Empirical Findings

4.3.1 Trading and Information

Table 4: Estimation Results

	iPIN	iVPIN		iPIN		iVPIN	
				low	high	low	high
ω	0.6298	0.5305	q	1.3069		1.0144	
	(0.02)	(0.01)		(0.01)		(0.02)	
β	0.3073	0.5376	$(\tau\|ti)$	1.1480	0.6376	1.1853	0.5945
	(0.03)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)
ϕ	0.4898	0.3695	*g(ti)*	1.0077		0.9951	
	(0.01)	(0.02)		(0.00)		(0.02)	
δ	0.1565	0.2475	*j(ti)*	1.1388		1.0639	
	(0.02)	(0.02)		(0.03)		(0.03)	

The left panel of Table 4 presents the estimation results for the conditional mean specification parameters, assuming a FI-ACD specification $\omega + \beta\psi_{i-1} + (\chi_i - \beta\chi_{i-1}) - (\tilde{\chi}_i - \phi\tilde{\chi}_{i-1})$. The right panel presents the distribution parameter estimates, assuming a $q-$Weibull distribution for χ, i.e., $f\left(\chi_i|\mathcal{F}_{i-1}\right) = (2-\tau^{q=1})^{\tau_i^{q=2}}/\chi_i \left[\chi_i/A_i\right]^{\tau_i^{q=2}} e_q\left(-\left[\chi_i/A_i\right]^{\tau_i^{q=2}}\right)$, where $A_i = \left[\Gamma\left(1+1/\tau_i^{q=2}\right)^{-\tau_i^{q=2}}/\psi_i\right]$, $\tau_i^{q=2} = \left(G^{q=2}_{m=1,i} - G^{q=2}_{m=2,i}\right)\tau^{q=2}_{m=1} + G^{q=2}_{m=2,i}\tau^{q=2}_{m=2}$ and $G^{q=2}_{m=2,i} = \left(1 + e^{-g^{q=2}_{m=2}(J_i - j^{q=2}_{m=2})}\right)^{-1}$. The estimates are for the iPIN and iVPIN specifications in section 3.3 All estimates are cross-sectional averages, with standard deviations in (:).

Table 4 presents the estimates of the parameters of the specification in Table 2. The left panel refer to the conditional mean specification. δ takes the value of 0.1565 (0.2475 for VPIN), suggesting a long memory (Jasiak, 1999), but not a fully integrated process. The values for β and ϕ indicate strong persistence, but at the same time satisfy the positivity constraints (Caporin, 2003) $\beta - \delta \leq \phi \leq \frac{2-\delta}{3}$ and $\delta\left(\phi - \frac{1-\delta}{2}\right) \leq \beta(\delta - \beta + \phi)$.

The right panel reports the estimates of the parameters of the distribution, The entropy parameter q is converging to 1 and this shows a convergence to a Weibull distribution. The smoothness parameters, $g^{q=2}_{m=2}$ noted as $g(ti)$'s, are very close to one, while the threshold values, $j^{q=2}_{m=2}$ noted as $s(ti)$'s, are very close to the unconditional means. This shows, a rather smooth transition from one regime to the other. These, data identified, regimes exhibit two distinct trading groups, as they are captured by the shape parameters, $\tau^{q=2}_{m=1}$

and $\tau_{m=2}^{q=2}$, noted as $\tau_{ti<j(ti)}$ (low) and $\tau_{ti>j(ti)}$ (high), respectively. $\tau_{ti<j(ti)}$'s are consistently higher than 1 and $\tau_{ti>j(ti)}$'s are consistently less than 1. This is consistent with previous literature (Kalaitzoglou and Ibrahim, 2013, 2015, 2023) and indicates that higher trading intensity is associated with an increasing probability of informed trading. According to the estimates, when the threshold variable, J_i, takes values that are higher than 1.1388 (1.0639), the iPIN (iVPIN) metric indicates that there is an increasing probability that this trade was instigated by an informed trader. This is inferred by its sample-wide post trade impact that is reflected on a decreasing hazard function. The higher J_i is, the closer the shape parameter is to $\tau_{m=2}^{q=2}$ (0.6376 for PIN and 0.5945 for VPIN) and, thus, the sharper is the decreasing shape of the hazard function (indicating an accelerating market and, thus, a trade with high(er) post-trade impact), which is interpreted as more informative. Heuristically, this indicates that higher trading intensity is associated with higher presence of private information. This finding is consistent with Easley and O'hara (1992) but the novelty in iVPIN is that it provides an empirical framework to assess the exact probability. In addition, the smoothness parameter is an indirect measure of how easy is to infer information from trading signals and, thus, it is an estimate of market opacity.[20]

4.3.2 Intensity-Based versus Interval Estimates

Furthermore, in order to provide a direct comparison between the intensity-based versus the interval-based metrics of PIN, their relative performance is tested in forecasting realized volatility (Pöppe et al., 2016) and UEE's (Johnson et al., 2013). Table 5 presents the estimations results of Eq.(19) for variance and Table 6 for UEE's. The top panel presents the estimates of the parameter c_1, with t-statistics in (:) and it is followed by the R^2 and (M)ean (S)quared (E)rror (MSE). All estimations include asset fixed effects and a set of microstrure controlled variables, the estimates of which are not reported for brevity. Each section named,

[20] A sharper (smoother) transition, captured by a higher (lower) value of the smoothness parameter, $g_{m=2}^{q=2}$, would indicate a clearer (less clear) distinction between the two regimes. Considering that the presence of different agent-types is latent information, a clearer (less clear) distinction would indicate greater (lower) market opacity because it is easier (more difficult) to extract latent information from observable signals.

Table 5: Real Data: Performance of metrics-Variance

	1"	5"	15"	30"	1'	5'	15'	30'	60'
					Estimates				
PIN	-0.0054	-0.0084	-0.0103	-0.0073	-0.0059	-0.0058	-0.0025	-0.0026	-0.0024
(t)	(-1.44)	(-74.27)	(-64.22)	(-56.63)	(-37.03)	(-34.24)	(-20.36)	(-3.63)	(-2.29)
R^2	0.6688	0.5749	0.5249	0.4951	0.3238	0.3220	0.2726	0.2581	0.2436
MSE	0.0198	0.0163	0.0154	0.0132	0.0128	0.0120	0.0088	0.0068	0.0065
iPIN	0.6816	0.4112	0.3926	0.3239	0.2722	0.2326	0.1950	0.1306	0.0933
(t)	(124.44)	(76.17)	(66.62)	(61.91)	(40.64)	(37.30)	(22.01)	(10.60)	(8.51)
R^2	0.6817	0.5819	0.5285	0.4980	0.3256	0.3229	0.2731	0.2646	0.2601
MSE	0.0160	0.0133	0.0130	0.0113	0.0111	0.0106	0.0080	0.0061	0.0052
VPIN	-0.0017	-0.0077	-0.0093	-0.0053	-0.0051	-0.0498	-0.0036	-0.0032	-0.0032
(t)	(-0.32)	(-52.26)	(-62.15)	(-48.83)	(-36.52)	(-32.14)	(-20.01)	(-8.46)	(-6.00)
R^2	0.6683	0.5740	0.5231	0.4947	0.3233	0.3212	0.2721	0.2643	0.2587
MSE	0.0257	0.0187	0.0165	0.0147	0.0130	0.0121	0.0091	0.0068	0.0056
iVPIN	0.8395	0.7551	0.6949	0.4712	0.4216	0.3743	0.2758	0.2213	0.2105
(t)	(133.90)	(92.07)	(55.75)	(52.53)	(50.47)	(40.58)	(26.93)	(12.81)	(9.21)
R^2	0.6848	0.5849	0.5312	0.5007	0.3281	0.3252	0.2753	0.2668	0.2609
MSE	0.0158	0.0124	0.0123	0.0110	0.0102	0.0100	0.0074	0.0061	0.0051
					MSE				
PIN/iPIN	1.2377	1.2277	1.1863	1.1660	1.1466	1.1308	1.1017	1.1135	1.2512
VPIN/iVPIN	1.6330	1.5074	1.3401	1.3370	1.2710	1.2071	1.2295	1.1109	1.0951
iPIN/iVPIN	1.0138	1.0725	1.0549	1.0314	1.0907	1.0526	1.0782	1.0019	1.0239
PIN/VPIN	0.7684	0.8735	0.9339	0.8995	0.9839	0.9861	0.9662	1.0041	1.1699
PIN/iVPIN	1.2548	1.3167	1.2515	1.2026	1.2506	1.1903	1.1879	1.1155	1.2811
					R2				
PIN/iPIN	0.9810	0.9879	0.9933	0.9940	0.9945	0.9972	0.9981	0.9755	0.9368
VPIN/iVPIN	0.9760	0.9813	0.9847	0.9878	0.9854	0.9878	0.9883	0.9907	0.9918
iPIN/iVPIN	0.9955	0.9950	0.9947	0.9946	0.9924	0.9928	0.9919	0.9916	0.9969
PIN/VPIN	1.0007	1.0016	1.0035	1.0009	1.0016	1.0022	1.0017	0.9764	0.9417
PIN/iVPIN	0.9955	0.9950	0.9947	0.9946	0.9924	0.9928	0.9919	0.9916	0.9969

Table 5 presents the estimation results for Eq. (19), where the dependent variable is Realized volatility. All estimations include the same control variables and company fixed effects. The top panel presents the estimates of the coefficients with t-stats in (), as well as the adjusted R^2 and the (M)ean (S)quared (E)rror (MSE). The bottom two panels report the ratios of MSE and R^2 for the pairs indicated on the left.

PIN, iPIN, VPIN and iVPIN are separate estimations with each metric being considered independently of the others. The comparison is based on R^2 and the MSE and the bottom panel presents the ratios for direct commparison. Finally, both tables are organized into columns according to the interval frequency, bucketsize $= (1'', 5'', 15'', 30'', 1', 5', 15', 30', 60')'$.

The empirical findings using both metrics are rather consistent and show that the intensity-based metrics perform better, especially in higher sampling frequencies. iVPIN and iPIN exhibit a higher R^2 and a lower MSE than VPIN and PIN, in this order. This is more

Table 6: Real Data: Performance of metrics-UEE's

	1"	5"	15"	30"	1'	5'	15'	30'	60'
					Estimates				
PIN	0.0352	0.0542	0.0666	0.0473	0.0381	0.0372	0.0159	0.0166	0.0158
(t)	(4.84)	(36.27)	(31.36)	(27.65)	(18.08)	(16.72)	(9.94)	(7.77)	(4.90)
R^2	0.0679	0.0576	0.0524	0.0496	0.0324	0.0324	0.0273	0.0258	0.0243
MSE	0.5025	0.4976	0.0492	0.4312	0.0321	0.0319	0.2834	0.0255	0.0268
iPIN	0.0573	0.0346	0.0330	0.0272	0.0229	0.0196	0.0164	0.0110	0.0079
(t)	(49.45)	(39.23)	(34.31)	(31.89)	(20.93)	(19.21)	(14.34)	(9.91)	(7.96)
R^2	0.0686	0.0586	0.0532	0.0501	0.0328	0.0325	0.0275	0.0266	0.0262
MSE	0.4944	0.4618	0.4516	0.3943	0.3865	0.3669	0.2764	0.2120	0.1792
VPIN	0.0367	0.1613	0.1955	0.1111	0.1075	1.0465	0.0761	0.0671	0.0664
(t)	(3.39)	(35.55)	(42.28)	(33.22)	(24.84)	(21.87)	(13.61)	(8.76)	(6.21)
R^2	0.0648	0.0573	0.0522	0.0493	0.0321	0.0322	0.0271	0.0263	0.0257
MSE	0.5763	0.0583	0.0527	0.0501	0.0326	0.0326	0.0276	0.0254	0.0229
iVPIN	0.0598	0.0433	0.0398	0.0270	0.0242	0.0215	0.0158	0.0127	0.0121
(t)	(51.07)	(41.29)	(32.20)	(27.17)	(19.07)	(16.55)	(15.17)	(13.84)	(10.57)
R^2	0.0696	0.0594	0.0540	0.0509	0.0333	0.0330	0.0280	0.0271	0.0265
MSE	0.4726	0.4314	0.4291	0.3828	0.3552	0.3498	0.2577	0.2125	0.1767
					MSE				
PIN/iPIN	1.0164	1.0766	0.1104	1.0955	0.0833	0.0871	1.0301	0.1208	0.1496
VPIN/iVPIN	1.2196	0.1347	0.1241	0.1303	0.0926	0.0927	0.1058	0.1202	0.1314
iPIN/iVPIN	1.0462	1.0714	1.0459	1.0281	1.0869	1.0479	1.0676	0.9954	1.0195
PIN/VPIN	0.8719	8.5642	0.9303	8.6444	0.9787	0.9843	10.3904	1.0003	1.1606
PIN/iVPIN	1.0634	1.1535	0.1155	1.1262	0.0906	0.0913	1.0997	0.1202	0.1526
					R2				
PIN/iPIN	0.9893	0.9813	0.9843	0.9843	0.9898	0.9897	0.9979	0.9660	0.9276
VPIN/iVPIN	0.9311	0.9591	0.9657	0.9629	0.9669	0.9683	0.9718	0.9693	0.9695
iPIN/iVPIN	0.9861	0.9856	0.9853	0.9852	0.9831	0.9834	0.9825	0.9822	0.9875
PIN/VPIN	1.0478	1.0084	1.0043	1.0071	1.0063	1.0051	1.0088	0.9789	0.9447
PIN/iVPIN	0.9861	0.9856	0.9853	0.9852	0.9831	0.9834	0.9825	0.9822	0.9875

Table 6 presents the estimation results for Eq. (19), where the dependent variable is UEE. All estimations include the same control variables and company fixed effects. The top panel presents the estimates of the coefficients with t-stats in (), as well as the adjusted R^2 and the (M)ean (S)quared (E)rror (MSE). The bottom two panels report the ratios of MSE and R^2 for the pairs indicated on the left.

pronounced in higher sampling frequencies, especially when the interval is closer to 1" (more relevant for HFT). This highlights that interval-based measures might not be adequately adopted for HFT, even when the volume-clock is considered in VPIN. The ratios reported at the bottom of the two tables show that in lower frequencies, e.g., 30' or 60', the performance differences persist, but all the metrics seem to converge. The biggest differences are observed in higher sampling frequencies, e.g., 1" to 15", highlighting that, in HFT, the aggregation process itself (hazard functions), constitute a stronger signal, compared to the aggregated values (trade direction or price change). This finding is consistent cross-sectionally and provides strong evidence in favour of the underlying concept in this study, that in HFT, where AT acts in sub-human attention speeds, interval estimates are inapt in capturing the properties of the data. We show that this is even more pronounced as sampling frequencies increase. Instead, the modeling of the arrival rates performs notably and consistently better, because it focuses on the granular properties of the data.

In addition, trading volume enhances the performance of all metrics. In intensity-based metrics the volume enhanced iVPIN outperforms consistently the iPIN that is based solely on the arrival rate of trades, rather than of volume. In interval-based metrics, VPIN exhibits some significant noise in high sampling frequencies, 1" to 1', but it outperforms PIN in intervals that exceed 15'. This highlights that the initial intuition behind VPIN that the noise present in the trade direction might be mitigated by focusing on the volume-clock, is toward the right direction. However, classification according to trade direction is still noisy, independently on whether is based on trade direction (PIN) or price change (VPIN). Instead, focusing on the aggregation properties of volume leads to a consistently higher performance, that persist even at lower trading frequencies.

4.3.3 Market Reflexivity and Systematic Risk

S_i, in Section 2.6.1, is an indirect assessment of how endogenous trading is at a particular moment, which in the spirit of Filimonov and Sornette (2012), Hardiman et al. (2013) and Filimonov and Sornette (2015), is a measure of market reflexivity. In their work, they provide an interval estimate for the degree of reflexivity and they observe that it is directly linked to higher variance and the realization of UEE's (Johnson et al., 2013). S_i follows the same rationale, but it also provides a point estimate of reflexivity. In this section, we investigate its empirical properties and its correlation with price change variance and UEE's.

Table 7 presents the sample averages for the top and bottom 50% of S_i focusing on mar-

Table 7: Market statistics at different degrees of reflexivity

	bottom 50%	Top 50%	t-statistic	p
L_i	0.10	0.22	43.99	(0.00)
$\mathbb{P}_i$	0.12	0.69	59.98	(0.00)
VPIN	0.10	0.47	64.72	(0.00)
Bid-Ask Spread ($'c)	0.02	0.03	10.31	(0.00)
Algorithmic Trading (%)	0.57	0.97	76.34	(0.00)
Trading Intensity	1.34	0.63	-42.17	(0.00)
Duration (seconds)	1.14	0.09	-33.76	(0.00)
Volume ($\#shares$)	213.07	234.13	2.52	(0.01)

Table 7 presents the average prior (L_i) and posterior ($\mathbb{P}_i$) probabilities, as well as VPIN, Bid-Ask spread, the proportion of Algorithmic Trading and Trading Intensity, as a the volume weighted durations, for two sub-samples, defined by the distributional characteristics of reflexivity. The last column present the t-statistics and the p-values in (:) of the top-bottom differences.

ket conditions. The first notable observation is that higher reflexivity is due to a higher, not lower, posterior probability. This confirms our intuition that a sizable revision in the posterior probability is associated with "more" rather than "less" trading. This is also reflected on a significantly higher trading intensity (0.63 seconds for a share to be traded versus 1.34), that is primarily related to more (shorter durations), rather than bigger (volume is only marginally higher) trading. In line with Section 4.3.1, we observe that higher trading intensity accelerates subsequent trading and thus it is linked with higher presence of private information. Naturally, AT would intensify its presence and this is observed in the top 50%. According to Johnson et al. (2013), this has the potential to put significant directional pressure on the order flow (Easley et al., 2011) and crowd the market (Ibrahim and Kalaitzoglou, 2025), eventually rendering it toxic (Easley et al., 2014) to an extend that might lead to UEE's (Johnson et al., 2013). We do observe higher spreads, more AT and higher toxicity (VPIN) during periods of higher S_i. This is evidence that S_i captures periods of more endogenous trading, potentially due to AT, and thus it can capture reflexivity.

In order to assess whether this exhibits a significant impact on subsequent trading, we investigate whether S_i is correlated with price change variance (Table 8) and the realization

Table 8: Reflexivity and subsequent trading: Price Change Variance

		1"	5"	15"	30"	1'	5'	15'	30'	60'
Variance	S	0.0733	0.0666	0.0413	0.0351	0.0315	0.0271	0.0164	0.0128	0.0087
	(t)	(46.41)	(35.23)	(29.81)	(20.18)	(16.67)	(5.36)	(1.42)	(0.95)	(0.84)
	R^2	0.0660	0.0559	0.0509	0.0480	0.0315	0.0281	0.0237	0.0230	0.0225
	PIN	-0.0995	-0.0954	-0.0115	-0.0075	-0.0071	-0.0068	-0.0032	-0.0029	-0.0028
	(t)	(-18.14)	(-63.33)	(-58.82)	(-48.19)	(-37.01)	(-36.63)	(-21.85)	(-3.65)	(-2.31)
	S	0.0772	0.0704	0.0434	0.0370	0.0330	0.0285	0.0173	0.0135	0.0092
	(t)	(48.90)	(37.26)	(31.36)	(21.26)	(17.50)	(5.64)	(1.50)	(1.00)	(0.89)
	R^2	0.6705	0.5773	0.5302	0.4954	0.3256	0.3250	0.2726	0.2582	0.2435
	iPIN	0.5819	0.3575	0.3363	0.2866	0.2335	0.2069	0.1688	0.1120	0.0805
	(t)	(86.77)	(67.40)	(58.94)	(54.27)	(35.92)	(33.10)	(19.55)	(9.03)	(7.50)
	S	0.0652	0.0614	0.0530	0.0315	0.0305	0.0133	0.0072	0.0059	0.0015
	(t)	(18.89)	(14.30)	(11.33)	(11.07)	(10.39)	(5.99)	(1.21)	(0.91)	(0.46)
	R^2	0.6857	0.5824	0.5330	0.4996	0.3270	0.3237	0.2736	0.2666	0.2618
	VPIN	-0.0832	-0.2839	-0.0336	-0.0176	-0.0200	-0.1905	-0.0154	-0.0117	-0.0118
	(t)	(-44.03)	(-62.08)	(-79.31)	(-57.89)	(-50.86)	(-47.91)	(-29.92)	(-4.12)	(-2.93)
	S	0.0742	0.0682	0.0422	0.0357	0.0321	0.0276	0.0166	0.0129	0.0089
	(t)	(36.35)	(27.82)	(23.23)	(16.01)	(13.26)	(4.05)	(1.10)	(0.76)	(0.67)
	R^2	0.6709	0.5751	0.5264	0.4967	0.3239	0.3221	0.2722	0.2644	0.2589
	iVPIN	0.7577	0.7051	0.6616	0.4254	0.3848	0.3378	0.2617	0.2037	0.1965
	(t)	(94.85)	(85.83)	(52.59)	(49.44)	(45.87)	(36.78)	(24.82)	(11.33)	(8.56)
	S	0.0651	0.0608	0.0529	0.0312	0.0304	0.0132	0.0072	0.0058	0.0015
	(t)	(18.71)	(14.26)	(11.23)	(11.03)	(10.32)	(5.95)	(1.20)	(0.91)	(0.46)
	R^2	0.6912	0.5872	0.5323	0.5043	0.3307	0.3278	0.2772	0.2670	0.2613

Table 8 presents the estimates of the model in Eq. 19, $Variance_{t_{bucket}} = c_0 + c_1 \text{Metric}_{t_{bucket}-n} + \boldsymbol{cCV}_{t_{bucket}-n} + \text{f.e.} + \epsilon_{t_{bucket}}$, where $Matric = (S_i, \text{PIN}, \text{VPIN}, \text{iPIN}, \text{iVPIN})'$ and $\boldsymbol{CV} = (\text{spread}, \#\text{orders}, \text{average duration})'$. Variance is the sum of squared price changes. Each panel presents, the estimates, the t-statistics in (:) and the R^2.

of UEE's (Table 9). We observe that S_i is strongly and positively correlated with both, but its impact decays fast with time. Interval frequencies longer than 5' seem to have a minimal impact, suggesting that reflexivity might become a short term systematic risk. Looking only at the first panel of these tables, higher reflexivity leads to significantly higher variance and UEE's, within maximum 15'.[21]. This effect does not disappear when we consider private information (e.g., PIN, VPIN, iPIN, iVPIN). Heuristically, our findings suggest that when the market depends more on the arrival of events (durations) rather than on other observable market conditions ($\boldsymbol{J}_i$), it is more likely to enter into a higher reflexivity state,

[21] Following Johnson et al. (2013) and Aquilina et al. (2022), this time frame is rather long for AT standards, which might even reach micro-second durations. However, it is rather consistent with the time length of significant "flash"-events, like the May 2010 (US) or the May 2022 (Europe) flash-crashes, which lasted less than 20', but longer than the trading frequencies of AT.

Table 9: Reflexivity and subsequent trading: UEE's

UEE's	S	1.3551	0.4076	0.5690	0.4728	0.2833	0.2485	0.1852	0.1586	0.0523
	(t)	(37.62)	(32.90)	(26.60)	(14.78)	(4.88)	(4.03)	(0.92)	(0.33)	(0.22)
	R^2	0.0630	0.0535	0.0489	0.0460	0.0303	0.0299	0.0253	0.0246	0.0240
	PIN	0.0281	0.0432	0.0529	0.0379	0.0304	0.0297	0.0142	0.0153	0.0135
	(t)	(3.86)	(28.89)	(24.87)	(22.02)	(14.34)	(13.36)	(8.91)	(7.18)	(4.89)
	S	1.2988	0.3893	0.5466	0.4566	0.2720	0.2402	0.1762	0.1533	0.0503
	(t)	(36.10)	(31.26)	(25.50)	(14.05)	(4.65)	(3.85)	(0.88)	(0.32)	(0.21)
	R^2	0.0681	0.0576	0.0527	0.0498	0.0327	0.0326	0.0275	0.0259	0.0243
	$iPIN$	0.0513	0.0309	0.0296	0.0245	0.0205	0.0176	0.0146	0.0099	0.0070
	(t)	(44.34)	(35.13)	(30.85)	(28.49)	(18.70)	(17.19)	(12.78)	(8.85)	(7.14)
	S	1.1538	0.3527	0.4860	0.4091	0.2422	0.2128	0.1598	0.1364	0.0452
	(t)	(32.15)	(28.52)	(22.85)	(12.70)	(4.21)	(3.48)	(0.79)	(0.29)	(0.19)
	R^2	0.0706	0.0603	0.0546	0.0514	0.0335	0.0334	0.0282	0.0272	0.0269
	$VPIN$	0.0310	0.1359	0.1647	0.0936	0.0910	0.8869	0.0707	0.0670	0.0656
	(t)	(2.85)	(30.17)	(35.88)	(28.04)	(20.96)	(18.50)	(12.56)	(7.99)	(5.97)
	S	1.2255	0.3674	0.5215	0.4307	0.2583	0.2240	0.1700	0.1438	0.0474
	(t)	(34.11)	(29.75)	(23.99)	(13.51)	(4.45)	(3.68)	(0.84)	(0.30)	(0.20)
	R^2	0.0657	0.0583	0.0531	0.0500	0.0324	0.0327	0.0274	0.0266	0.0260
	$iVPIN$	0.0566	0.0408	0.0375	0.0255	0.0229	0.0203	0.0149	0.0119	0.0114
	(t)	(48.42)	(39.20)	(30.51)	(25.72)	(18.08)	(15.58)	(14.33)	(13.14)	(9.99)
	S	1.0196	0.3099	0.4345	0.3588	0.2130	0.1886	0.1411	0.1192	0.0397
	(t)	(28.27)	(25.15)	(20.29)	(11.14)	(3.67)	(3.04)	(0.70)	(0.25)	(0.17)
	R^2	0.0733	0.0627	0.0569	0.0539	0.0352	0.0349	0.0295	0.0285	0.0281

Table 9 presents the estimates, the t-statistics in (:) and the R^2 of the model in Eq. 19, $UEE_{t_{bucket}} = c_0 + c_1\text{Metric}_{t_{bucket}-n} + \boldsymbol{cCV}_{t_{bucket}-n} + \text{f.e.} + \epsilon_{t_{bucket}}$, where $Matric = (S_i, \text{PIN}, \text{VPIN}, \text{iPIN}, \text{iVPIN})'$ and $\boldsymbol{CV} = (\text{spread}, \#\text{orders}, \text{average duration})'$. UEE's are defined according to Johnson et al. (2013).

which eventually can increase the variance and lead to UEE's. Consequently, S_i can be used as and HFT-adopted metric for reflexivity, which can act as a sign for when the market might enter into a state where market conditions are not fully reflective of information and trading (speed) might become endogenous. To an extend that might lead to intraday crashes.

5 Simulations and Alternative Uses

Section 4 provides some empirical evidence in favor of the intensity-based estimates. In order to assert that this inference is not circumstantial, Section 5 investigates their robustness in simulated data, where the "true" probabilities are known. We focus on the impact of mis-specification in the main text and develop the theoretical properties in Appendix C.[22]

The regressions in Section 4.1, are easily interpretable when both the regression model and the specification in Table 2 are well-specified. Then the estimators for iPIN/iVPIN differ from the latent PIN due to the sample variation of the estimated coefficients, but converge faster than their interval-based counterparts (Appendix C.1). When, however, the model is misspecified, this comparison becomes a bit more complicated, as it also depends on the (potentially non-asymptotically vanishing) dependence between the regression error, and the misspecification error provided by the difference between the latent true iPIN (resp. iVPIN) and the model iPIN (resp. iVPIN) evaluated at the pseudo-true value of the parameter.

We investigate the impact of various types of misspecificiation on the estimate of the probabilities. conducting the following experiments. In Section 5.1 we assess the perforance of the intensity- versus interval-based estimates in the case when the conditional mean specification for the DGP of duration is misspecified. In Section 5.2 we test the robustness of the findings when the assumed number of agents is misspecified and we discuss alternative uses of iRP. In Section 5.3 we investigate mis-specification in the type of mixture of agents, allowing a linear (MOD) versus non-linear (MoE) specification.

5.1 Misspecification and Trade Arrival Data Contamination

In order to disentangle the evaluation of the quality of the approximation of the latent PIN by the intensity-based procedures compared to the interval-based ones, from the correct specification of a predictive regression, we perform a Monte Carlo experiment that enables control of the latent PIN. The duration process ("volume clock" for VPIN) (χ_i) is assumed to conform to the specification of Table 2. The latent PIN is thus explicitly known. We construct a subordinated (Ghysels et al., 1995) to (χ_i) stochastic volatility process for the

[22] Appendix C investigates the statistical theory of the plug-in estimator for the iRP, through the limit theory of the maximum likelihood estimator. Section C.1 provides (pseudo-) consistency considerations and shows that the application of iRP captures adequately well the true PIN under correct specification. Section C.2 discusses rates and asymptotic distributions that enable the validity of asymptotic inference. This is highly relevant in HFT because of the scale difference between the intensity- and interval-based estimators.

logarithmic returns of the underlying asset in the spirit of Feng et al. (2015) as:

$$\ln P_i - \ln P_{i-1} = \alpha_0 + a_1\chi_i + \exp(\omega_0 + \omega_1\chi_i + V_i)z_i, \tag{20}$$

$$V_i = bV_{i-1} + \eta_{i-1},\ (z_i, \eta_i)^T \sim N(\mathbf{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}),\ |\rho| < 1, |b| < 1. \tag{21}$$

The Gaussian random vector $(z_i, \eta_i)^T$, is also considered independent of ϵ_{i-j}, for all i, j. Given initial values P_0, V_0, and for realistic sample sizes T, and choice of parameter values, artificial sample paths of (p_i, χ_i) can be created. Given such a path, PIN, VPIN, iPIN and iVPIN are estimated according to Section 4.1 and are evaluated according to their MSE's. Specifically, T is allowed to assume the values of 100,000 events (trades), which is the equivalent of 1 day trading in the most liquid asset in the sample, i.e., AAPL. 100 paths are considered, which is the equivalent of 100 trading days. In total, this resembles the trading activity of a very liquid asset over a calendar quarter. This, according to previous literature, is a reasonable time frame that does not introduce bias in the PIN or VPIN estimates. Furthermore, the true underlying parameter values are chosen according to the estimates in Table 4. For the price change and price change variance parameters, the following values are assumed, $\alpha_0 = -0.001, \alpha_1 = 0.9, \omega_0 = -3.5, \omega_1 = 0.05$ and $\beta = 0.98$.[23]

Table 10 presents the results of the simulation, following the main scenario. The first line of top panel of the table presents the on-average "true" values of the (P)robability of (IN)formed trading and it is organized in columns according to the sampling frequency that varies from 1" to 60'. The following 4 lines present the average estimates of all the metrics considered, following to empirical procedure described in Section 4.2, namely PIN, VPIN, iPIN and iVPIN. The second (middle) panel presents the MSE for each on the

[23]The robustness of the findings are tested against different setups. For threshold values between 1 and 1.9. The findings remain qualitatively the same, but are stronger when a higher threshold is employed. In addition, different values for T are considered, as well as a different number of sample paths. The number of sample paths does not change qualitatively the results, but a higher number of observations exhibits slightly less (but still) significant differences. The main scenario considers a strong negative correlation $\rho = -0.8$. The results do not change with different values, but are relatively weaker with lower correlation.

Table 10: Simulation: Performance of metrics-Probability of Informed Trading

	1"	5"	15"	30"	1'	5'	15'	30'	60'
				Average Estimates					
TRUE	0.3065	0.3384	0.3453	0.3471	0.3480	0.3487	0.3489	0.3489	0.3488
PIN	0.2659	0.2108	0.2053	0.2056	0.2059	0.2062	0.2062	0.2062	0.2062
iPIN	0.2957	0.3262	0.3326	0.3343	0.3351	0.3357	0.3359	0.3358	0.3358
VPIN	0.2372	0.2285	0.2254	0.2249	0.2248	0.2282	0.2417	0.2640	0.2933
iVPIN	0.2633	0.2904	0.2962	0.2977	0.2985	0.2991	0.2992	0.2992	0.2991
				MSE					
PIN	0.0719	0.0658	0.0585	0.0576	0.0571	0.0567	0.0566	0.0566	0.0565
iPIN	0.0307	0.0288	0.0264	0.0261	0.0260	0.0259	0.0259	0.0258	0.0258
VPIN	0.1129	0.0983	0.0792	0.0779	0.0768	0.0742	0.0642	0.0518	0.0357
iVPIN	0.0207	0.0187	0.0161	0.0161	0.0157	0.0156	0.0156	0.0156	0.0156
				MSE Ratios					
PIN/iPIN	2.3413	2.2857	2.2142	2.2065	2.1988	2.1909	2.1898	2.1882	2.1855
VPIN/iVPIN	5.4629	5.2443	4.9083	4.8283	4.8986	4.7669	4.1276	3.3302	2.2906
iPIN/iVPIN	1.4869	1.5347	1.6363	1.6174	1.6554	1.6617	1.6608	1.6602	1.6588
PIN/VPIN	0.6373	0.6689	0.7381	0.7391	0.7431	0.7637	0.8811	1.0909	1.5826
PIN/iVPIN	3.4813	3.5078	3.6230	3.5687	3.6401	3.6405	3.6369	3.6327	3.6253

Table 10 presents the performance of PIN, VPIN, iPIN and iVPIN in capturing the real PIN. The first line is the true average proportion of informed agents in the simulated data, estimated based on different sampling frequencies (in columns). The following four lines are the average estimates according to PIN, iPIN, VPIN and iVPIN. The second panel of Table 10 presents the (M)ean (S)quared (E)rror (MSE) of each estimate and the bottom panel presents the ratio of MSE's of the different metrics.

metrics in each sampling frequency. The bottom panel reports the ratios of MSE's for visual reference. The findings reported here are fully in line and confirm the basic findings of the empirical analysis presented in Section 4.1. In summary, the intensity-based metrics outperform consistently the interval-based metrics in accurately capturing the "true" PIN, by a factor of at least 2. This is consistent in all sampling frequencies, and, although it shows a decreasing trend, the intensity-based metrics are by far a better measure for private information. The second finding that is confirmed, is that volume does indeed contribute to capturing private information, but only when its aggregation properties are considered (iVPIN is consistently better than iPIN). Otherwise, it might introduce noise (VPIN in sampling frequencies>15').

5.2 Misspecification in the Number of agent-types

The empirical setup proposed in Eqq.(3)-(10) can be used to either estimate the probability of a specific agent-type to enter the market, or, alternatively to assess the number and the type of agent present in the market. Section 4 estimates the proportion of informed agents as an example of the former and Section 5.1 investigates its performance under mis-specification in the DGP of duration. In this section we illustrate the latter use of iRP by investigating whether it can i)detect the right number of agents in the case that there are more (assumed) regimes than real agents (i.e., $K > K_{real}$) and ii) assess properly the probability of an agent-type when the (assumed) regimes are less than the real number of agents ($K < K_{real}$).

Table 11 presents the estimates of the specification in Table 2, with $q = 2$ and $q = 3$ regimes, assuming an equivalent number of agents. There are two scenarios, where the true DGP consists of 2 or 3 agents. The model might under- $((q = 2) < (K_{True} = 3))$ or over- $((q = 3) < (K_{True} = 2))$ estimate the true number of agents. The former investigates the performance of our modeling in assessing the true probability of an agent-type, when K is less than actual. The latter, investigates an overfit scenario, where the true number of agents is lower than the one imposed by the model. iRP performs very well in all scenarios.

More specifically, when the model under-estimates the number of agents (second column) one of the regimes becomes redundant. In this particular scenario, the middle regime becomes largely indistinguishable from the third. The Wald tests at the bottom show that neither the threshold value ($j_1 \neq j_2$) nor the shape parameter ($\tau_{m=2} \neq \tau_{m=3}$) are significantly different. The value of the shape parameters for the (true) two regimes are somewhat diluted, but they remain within the bounds that define similarly shaped hazard functions (identify two agent-types). The first (second) regime gest an estimate of $\tau \approx 1$ ($\tau < 1$), which is associated with a flat (decreasing) hazard function and thus, with uninformed (informed) trading. The derivation of PIN in the right panel shows that the intensity-based estimates, especially when the volume-clock is employed, clearly outperform their interval-based counterparts, even under mis-specification in the number of agents $((q = 3) > (K_{True} = 2))$.

Table 11: Mis-specification in the number of agents

DGP	$q=2$		$q=3$	
Model	$q=2$ $=K_{True}$	$q=3$ $>K_{True}$	$q=2$ $<K_{True}$	$q=3$ $=K_{True}$
$\tau_{m=1}$	1.3828	1.4103	1.3816	1.0113
	(0.03)	(0.04)	(0.03)	(0.04)
$\tau_{m=2}$		0.6390		5.6133
		(0.05)		(0.51)
$\tau_{m=3}$	0.5296	0.5079	0.5619	0.5066
	(0.11)	(0.04)	(0.10)	(0.04)
g_1	1.0852	1.0832	0.8528	1.0822
	(0.08)	(0.03)	(0.06)	(0.03)
g_2		1.0294		1.0337
		(0.03)		(0.03)
j_1		1.1968		0.9636
		(0.24)		(0.01)
j_2	1.3883	1.3944	1.3057	1.3803
	(0.03)	(0.03)	(0.04)	(0.03)
$\tau_{m=2} \neq \tau_{m=3}$		1.67		9.86
		(0.05)		(0.00)
$j_1 \neq j_2$		0.98		6.49
		(0.16)		(0.00)

	DGP	$q=2$		$q=3$	
	Model	$q=2$ $=K_{True}$	$q=3$ $>K_{True}$	$q=2$ $<K_{True}$	$q=3$ $=K_{True}$
Uninformed	TRUE	0.4876		0.4211	
	PIN	0.5322		0.6031	
	MSE	0.0582		0.0962	
	VPIN	0.5025		0.5641	
	MSE	0.0550		0.0899	
	iPIN	0.4680	0.5481	0.5591	0.5012
	MSE	0.0217	0.0474	0.0648	0.0227
	iVPIN	0.4859	0.4468	0.4554	0.4196
	MSE	0.0167	0.0386	0.0587	0.0107
Informed	TRUE	0.5124		0.48242	
	PIN	0.4678		0.3969	
	MSE	0.0512		0.0633	
	VPIN	0.4975		0.4359	
	MSE	0.0467		0.0521	
	iPIN	0.5320	0.4519	0.4409	0.4988
	MSE	0.0218	0.0406	0.0601	0.0213
	iVPIN	0.5141	0.4610	0.4538	0.4837
	MSE	0.0177	0.0398	0.0584	0.0123

Table 11 presents the estimation results for the regimes and the performance of VPIN and iVPIN. Each panel has 2 sections with 2 columns each. The DGP stands for the true data generation process of duration. The conditional mean specification follows a FIACD according to Table 2. The true distribution consists of either 2 ($q=2$) or 3 ($q=3$) agent-types (i.e., 3 regimes). The model might be mis-specified, either assuming a higher ($q > K_{real}$) or a lower ($q < K_{real}$) number of agents. The left panel presents the estimates of the distributional parameters with three regimes. For each market setup ($q=1$ or $q=3$) there are two estimations; one assuming 2 agent-types ($\tau_{m=1} = \tau_{J_i<j_1}$ and $\tau_{m=3} = \tau_{J_i>j_2}$) and one assuming 3 ($\tau_{m=1} = \tau_{J_i<j_1}$, $\tau_{m=2} = \tau_{j_i<J_i<j_2}$ and $\tau_{m=3} = \tau_{J_i>j_2}$). The estimates are cross-sectional averages and the values in () are cross-sectional standard deviations. The bottom part presents wald tests with p-values in (). The right panel presents the true values (probabilities) for uninformed (top) and informed (bottom) agents, as well as the estimates with PIN, VPIN, iPIN and iVPIN. MSE is the cross-sectional mean squared error.

The third and fourth columns, present the estimates for the scenario when the true number of agents is higher than the number of regimes. In this case, the threshold value is somewhat lower, but still distinguishes a top regime, where the shape parameter identifies informed trading ($\tau < 1$). The estimates of both regimes are biased estimates, but not far from the estimates of the model, when the correct ($q=3$) number of regimes is employed. The Wald tests show a significant difference between the middle and top regime, suggesting that the model identifies correctly the number of agent-types. This is also reflected in the performance of the intensity-based estimates of PIN, which, even under wrong specification outperform the inteval-based estimates.

The implications of these findings are twofold. First, they show that the interval based estimates are more suitable for estimating the probability of an agent-type in HFT, since they exhibit a superior performance, even under mis-specification. Second, our estimation framework is flexible enough to capture the number of agents in the market, as well as their probability. This highlights an alternative use for our model, which can be used in a completely agnostic manner, where there is no prior in the number or type of agents present in the market. Using a higher number of regimes ($q > K_{True}$) can identify the number of agent-types and then the differentials in the hazard functions can capture the type of the agent. iRP can then estimate the probability.

5.3 Mis-specification in the Mixture of Agents Types

Eq.(2) considers that all agent-types interact in a highly non-linear manner, either with their presence, or their absence, and their combined actions shape the overall distribution of the market, as this is reflected on the shape parameter of the distribution of market durations. In order to assess this empirically, a certain distribution must be assumed for the DGP of market durations, in order to link its shape/scale parameters, $\boldsymbol{\tau}_i$, to certain agent-types as in Eq.(6). This approach has the notable benefit of being completely data driven, but at the expense of requiring a specific distribution function, like the one suggested in Eq.(11). In order to investigate the impact of mis-specification in the distributional assumptions, we consider a scenario where, the point process of the market is the superposed process of K agent-types of the form: $\mathbb{E}(N(t)-N(s)|\mathcal{F}_s) = \mathbb{E}\left(\int_s^t \lambda(u)\,du \middle| \mathcal{F}_s\right) = \sum_{k=1}^{K} \mathbb{E}\left(\int_s^t \lambda^k(u)\,du \middle| \mathcal{F}_s\right) = \sum_{k=1}^{K} \mathbb{E}\left(\int_s^t (p_t^k|\mathcal{F}_s)\lambda_0^k(u)\,du\right)$. We refer to this specification as a (M)ixture (o)f (D)istributions (MoD), as opposed to MoE, where the counting of the market events ($\mathbb{E}(N(t)-N(s)|\mathcal{F}_s)$) is the superposed process ($\mathbb{E}\left(\int_s^t \lambda(u)\,du \middle| \mathcal{F}_s\right)$), defined as a linear combination (sum: $\sum_{k=1}^{K} \mathbb{E}\left(\int_s^t \lambda^k(u)\,du \middle| \mathcal{F}_s\right)$) of the arrival rates of k agent-types ($\lambda^k(u)$). To arrive at $\lambda^k(u)$ in a data driven manner, $\lambda^k(u)$ is defined as $(p_t^k|\mathcal{F}_s)\lambda_0^k(u)$. $(p_t^k|\mathcal{F}_s)$ is the weighting, which acts as the prior probability (L_i in our setup) and $\lambda_0^k(u)$ is the baseline intensity of agent-type k. In both MoD and MoE we assume a FIACD specification. For MoE we follow Table 2 and in MoD, we assume an infinite mixture of two Weibul distributions with shape parameters 1 (flat hazard function) for the uninformed and 0.6 (decreasing hazard function) for the informed. Then we estimate iVPIN following Eq.(18).

Table 12: Mixture of Experts versus Mixture of Distributions

			1"	5"	15"	30"	1'	5'	15'	30'	60'
True: MoE	MEAN	TRUE	0.3065	0.3384	0.3453	0.3471	0.3480	0.3487	0.3489	0.3489	0.3488
		iVPIN_{MoE}	0.2633	0.2904	0.2962	0.2977	0.2985	0.2991	0.2992	0.2992	0.2991
		iVPIN_{MoD}	0.5603	0.5920	0.5989	0.6070	0.6016	0.6024	0.6026	0.6025	0.6025
	MSE	iVPIN_{MoE}	0.0207	0.0187	0.0161	0.0161	0.0157	0.0156	0.0156	0.0156	0.0156
		iVPIN_{MoD}	0.0528	0.0330	0.0298	0.0291	0.0297	0.0284	0.0283	0.0288	0.0284
		$^{\text{iVPIN}_{\text{MoD}}}/_{\text{iVPIN}_{\text{MoE}}}$	2.5555	1.7576	1.8467	1.8008	1.8943	1.8247	1.8207	1.8518	1.8195
True: MoD	MEAN	TRUE	0.4369	0.4788	0.4882	0.4908	0.4921	0.4929	0.4913	0.4933	0.4934
		iVPIN_{MoE}	0.4032	0.4378	0.4454	0.4475	0.4485	0.4491	0.4493	0.4495	0.4496
		iVPIN_{MoD}	0.4395	0.4616	0.4663	0.4776	0.4688	0.4685	0.4689	0.4690	0.4691
	MSE	iVPIN_{MoE}	0.0216	0.0192	0.0176	0.0159	0.0158	0.0157	0.0145	0.0137	0.0138
		iVPIN_{MoD}	0.0165	0.0142	0.0137	0.0137	0.0136	0.0136	0.0136	0.0136	0.0136
		$^{\text{iVPIN}_{\text{MoD}}}/_{\text{iVPIN}_{\text{MoE}}}$	0.7617	0.7376	0.7773	0.8575	0.8636	0.8671	0.9350	0.9932	0.9862

Table 12 presents the true probability of informed trading (PIN), as well as the estimates with the MoD ($\text{iVPIN}_{\text{MoD}}$) and MoE ($\text{iVPIN}_{\text{MoE}}$) approaches. Mean is the average across seeds, ane MSE is the mean squared error. $^{\text{iVPIN}_{\text{MoD}}}/_{\text{iVPIN}_{\text{MoE}}}$ is the ratio of the MSE's of each approach. The sampling frequency varies from 1" to 1h, according to Section 4.2. The true DGP is a MoE in the top and a MoD in the bottom.

Table 12 presents the performance of iRP when the mixture of agents is mis-specified. The MEAN and MSE estimates show that iVPIN performs better when the distributional assumptions match the true DGP. More precisely, iVPIN_{MoE} (iVPIN_{MoD}) performs better when the market distribution follows a MoE (MoD) specification, since it exhibits average estimates closer to the TRUE values and lower MSE's. However, when the MoE framework is mis-specified, the ratio $^{\text{iVPIN}_{\text{MoD}}}/_{\text{iVPIN}_{\text{MoE}}}$ takes values that are proportionally smaller compared to when the MOD framework is mis-specified. Especially on lower sampling frequencies. When the sampling frequency is beyond 1', iVPIN_{MoE} outperforms iVPIN_{MoD} by a factor of around 1.8, while in the opposite case it under-performs by a factor of around 0.9. This shows that the non-linearity imposed by the MoE framework is flexible enough to capture the true PIN, even with wrong distributional assumptions.

6 Conclusion

Who is trading in the market? This latent piece of information becomes an increasingly relevant question as trading frequencies increase. At HFT algorithms dominate trading and their reaction times are far shorter than those of humans (Johnson et al., 2013). This acceleration challenges the human-centric notion of efficient markets, which implies that fundamental information is incorporated into prices (Fama, 1965) instantaneously and unanimously. At these trading speeds, either due to market frictions (Madhavan, 2000), or simply because there is not enough time for the conventional information acquisition/processing/acting-upon cycle, AT relies on "trading" information as a noisy signal of "fundamental" information.

But what costitutes a trading signal? Early literature (Madhavan, 2000) points towards trade direction for market makers to infer the informational content of trades and condition their liquidity and adverse selection (Madhavan et al., 1997) costs. However, with the continuing rise of AT, market reflexivity (Soros, 1994, 2013) intensifies too, which makes trade direction less relevant due to signal propagation. So, ''*If one cannot learn from buys and sells, what should be looked at to infer underlying information? The high frequency world gives clues in that HFT algorithms draw inferences from trade sequences and time patterns (...)*". In these words, O'Hara (2015) argues that "time" itself becomes a source of information concerning the trading motives behind market events.

Motivated by this view, we introduce a new empirical framework to estimate the number, the type and the probability of the different agent-types at any point in time, as well as at any interval. We introduce the idea that the intrinsic trading motives (e.g., learning patterns, technology used, access to information, etc.) of each agent-type are reflected on her tangible actions and, in particular on her arrival rate, which we capture with the conditional intensity (hazard function) of a point process Daley and Vere-Jones (2003). We assume that this is time invariant and thus it becomes detectable. Furthermore, we consider that each agent-type contributes to the overall market activity as "experts", either with their presence (intensity function) or their absence (survival function), in a manner that, collectively, they resolve incoming information through liquidity patterns. We model their interactions as a function of market conditions (they enter the market upon suitable market conditions (time variant) at an time invariant rate). This creates a highly non-linear infinite ”Mixture of Experts” that accounts for their intrinsic motivation and for market reflexivity, simultaneously, which exhibits superior theoretical properties and empirical performance.

This approach exhibits several advantages. First, unlike previous literature (Easley et al., 2014), which focuses on the aggregated characteristics of observable variables to detect the presence of various agent-types, our approach focuses on the aggregation process itself and, thus, it provides point, rather than interval estimates, without suffering from sampling bias. In addition, it links the presence of agent-types to a statistical measure (intensities), rather than to a variable (Easley and O'hara, 1992) and therefore, it can be used, not only to estimate the probability, but also the number and the type of different agents, as long as, their motivation is reflected on tangible actions. Finally, since AT intensifies market reflexivity (Soros, 2013), it might propagate noise (Johnson et al., 2013) too, with detrimental effects to market stability (Zhang and Zhang, 2025), because they might render the order flow toxic (Easley et al., 2014), ultimately leading to higher variance and/or extreme events (Johnson et al., 2013), such as flash crashes. In the spirit of Filimonov and Sornette (2012, 2015), Hardiman et al. (2013), our approach enables a direct assessment of how much trading is affected by arrival rates (endogeneity) versus other market conditions, based on which we develop a local measure for HFT reflexivity that is found to be highly correlated with both variance and extreme events.

Collectively, we extract information about the, otherwise hidden, presence of different agent-types, based on time, and from this, we derive a metric for the propensity of the market to enter into a ”stress” state due to reflexivity. To the question: ”*Who trades in the market?*”, we respond:

Time will tell!...

References

Admati, A. R. and P. Pfleiderer (1986): "A monopolistic market for information," *J. Econ. Theory*, 39, 400–438.

Amari, S.-i. (2016): *Information geometry and its applications*, vol. 194, Springer.

Andersen, T. G. and O. Bondarenko (2014): "VPIN and the flash crash," *J. Fin. Mark.*, 17, 1–46.

Andrews, D. W. (1999): "Estimation when a parameter is on a boundary," *Econometrica*, 67, 1341–1383.

Aquilina, M., E. Budish, and P. O'neill (2022): "Quantifying the high-frequency trading "arms race"," *The Quarterly Journal of Economics*, 137, 493–564.

Arthur, W. B. (2013): "Complexity economics," *Complexity and the Economy*.

Austern, M. and P. Orbanz (2022): "Limit theorems for distributions invariant under groups of transformations," *The Annals of Statistics*, 50, 1960–1991.

Bacry, E., I. Mastromatteo, and J.-F. Muzy (2015): "Hawkes processes in finance," *Market Microstructure and Liquidity*, 1, 1550005.

Barndorff-Nielsen, O. E. and A. N. Shiryaev (2015): *Change of time and change of measure*, vol. 21, World Scientific Publishing Company.

Bauwens, L. and P. Giot (2003): "Asymmetric ACD models: Introducing price information in ACD models," *Empir. Econ.*, 28, 709–731.

Blasques, F., P. Gorgi, S. J. Koopman, and O. Wintenberger (2018): "Feasible invertibility conditions and maximum likelihood estimation for observation-driven models," .

Bowsher, C. G. (2007): "Modelling security market events in continuous time: Intensity based, multivariate point process models," *J. Econom.*, 141, 876–912.

Box, G. E. P. and D. R. Cox (1964): "An analysis of transformations," *J. R. Stat. Soc.*, 26, 211–243.

Brogaard, J., T. Hendershott, and R. Riordan (2014): "High-frequency trading and price discovery," *The Review of Financial Studies*, 27, 2267–2306.

——— (2019): "Price discovery without trading: Evidence from limit orders," *The Journal of Finance*, 74, 1621–1658.

Brogaard, J., M. C. Ringgenberg, and D. Roesch (2025): "Does floor trading matter?" *The Journal of Finance*, 80, 375–414.

Caporin, M. (2003): "Identification of long memory in GARCH models," *Statistical Methods and Applications*, 12, 133–151.

Chaboud, A. P., B. Chiquoine, E. Hjalmarsson, and C. Vega (2014): "Rise of the machines: Algorithmic trading in the foreign exchange market," *The Journal of Finance*, 69, 2045–2084.

Chakrabarty, B. and R. Pascual (2023): "Stock liquidity and algorithmic market making during the COVID-19 crisis," *Journal of Banking & Finance*, 147, 106415.

Chakrabarty, B., R. Pascual, and A. Shkilko (2015): "Evaluating trade classification algorithms: Bulk volume classification versus the tick rule and the Lee-Ready algorithm," *Journal of Financial Markets*, 25, 52–79.

Chordia, T., B. Miao, and J. Noh (2024): "Do Low Latency Traders Destabilize Prices? Evidence from News Releases," *Evidence from News Releases (June 06, 2024)*.

Daley, D. J. and D. Vere-Jones (2003): *An introduction to the theory of point processes: volume I: elementary theory and methods*, Springer.

De Luca, G. and P. Zuccolotto (2006): "Regime-switching Pareto distributions for ACD models," *Comput. Stat. Data Anal.*, 51, 2179–2191.

Easley, D., M. L. de Prado, and M. O'Hara (2016): "Discerning information from trade data," *J. Financ. Econ.*, 120, 269–285.

Easley, D., M. L. De Prado, and M. O'Hara (2011): "The microstructure of the Flash Crash," *Journal of Portfolio Management*, 37, 118–128.

Easley, D., M. M. L. de Prado, and M. O'Hara (2014): "VPIN and the flash crash: A rejoinder," *Journal of Financial Markets*, 17, 47–52.

Easley, D., R. F. Engle, M. O'Hara, and L. Wu (2008): "Time-varying arrival rates of informed and uninformed trades," *Journal of Financial Econometrics*, 6, 171–207.

Easley, D., N. M. Kiefer, and M. O'Hara (1997a): "The information content of the trading process," *J. Empir. Finance*, 4, 159–186.

——— (1997b): "One day in the life of a very common stock," *Rev. Financ. Stud.*, 10, 805–835.

Easley, D., N. M. Kiefer, M. O'hara, and J. B. Paperman (1996): "Liquidity, information, and infrequently traded stocks," *J. Finance*, 51, 1405–1436.

Easley, D., M. M. López de Prado, and M. O'Hara (2012): "Flow toxicity and liquidity in a high-frequency world," *Rev. Financ. Stud.*, 25, 1457–1493.

Easley, D. and M. O'hara (1992): "Time and the process of security price adjustment," *J. Finance*, 47, 577–605.

Easley, D. and M. O'Hara (1995): "Market microstructure," *Handbooks in operations research and management science*, 9, 357–383.

Ellis, K., R. Michaely, and M. O'Hara (2000): "The accuracy of trade classification rules: Evidence from NASDAQ," *J. Fin. Quant. Anal.*, 35, 529.

Engle, R. F. (2000): "The econometrics of ultra-high-frequency data," *Econometrica*, 68, 1–22.

Engle, R. F. and J. R. Russell (1998): "Autoregressive conditional duration: A new model for irregularly spaced transaction data," *Econometrica*, 66, 1127.

Fama, E. F. (1965): "The behavior of stock-market prices," *J. Bus.*, 38, 34–105.

Feng, D., P. X.-K. Song, and T. S. Wirjanto (2015): "Time-deformation modeling of stock returns directed by duration processes," *Econometric Reviews*, 34, 480–511.

Filimonov, V. and D. Sornette (2012): "Quantifying reflexivity in financial markets: Toward a prediction of flash crashes," *Physical Review E*, 85, 056108.

——— (2015): "Apparent criticality and calibration issues in the Hawkes self-excited point process model: application to high-frequency financial data," *Quantitative Finance*, 15, 1293–1314.

Foucault, T., O. Kadan, and E. Kandel (2005): "Limit order book as a market for liquidity," *Rev. Financ. Stud.*, 18, 1171–1217.

Ghysels, E., C. Gouriéroux, and J. Jasiak (1995): "Market time and asset price movements: Theory and estimation," *Cahier de recherche*.

Haldane, A. G. (2012): "The race to zero," in *The Global Macro Economy and Finance*, London: Palgrave Macmillan UK, 245–270.

Hardiman, S. J., N. Bercot, and J.-P. Bouchaud (2013): "Critical reflexivity in financial markets: a Hawkes process analysis," *The European Physical Journal B*, 86, 442.

Hasbrouck, J. and G. Saar (2013): "Low-latency trading," *Journal of Financial Markets*, 16, 646–679.

Hasbrouck, J. and R. A. Schwartz (1988): "Liquidity and execution costs in equity markets," *Journal of Portfolio Management*, 14, 10.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011): "Does algorithmic trading improve liquidity?" *The Journal of finance*, 66, 1–33.

Hujer, R. and S. Vuletić (2007): "Econometric analysis of financial trade processes by discrete mixture duration models," *J. Econ. Dyn. Control*, 31, 635–667.

Ibragimov, I. A. and Y. V. Linnik (1971): *Independent and stationary sequences of random variables*, Wolters-Noordhoff.

Ibrahim, B. M. and I. A. Kalaitzoglou (2025): "Crowdedness, mispricing, crashes, and spikes," *Journal of Banking & Finance*, 107485.

Jacobs, R. A., M. I. Jordan, S. J. Nowlan, and G. E. Hinton (1991): "Adaptive Mixtures of Local Experts," *Neural Computation*, 3, 79–87.

Jakubowski, A. (2012): "Principle of Conditioning revisited," *Demonstratio Mathematica*, 45, 325–36.

Jasiak, J. (1999): "Persistence in intertrade durations," *Finance*, 16, 166–195.

Jenkinson, O. (2019): "Ergodic optimization in dynamical systems," *Ergodic Theory and Dynamical Systems*, 39, 2593–2618.

Johnson, N., G. Zhao, E. Hunsader, H. Qi, N. Johnson, J. Meng, and B. Tivnan (2013): "Abrupt rise of new machine ecology beyond human response time," *Sci. Rep.*, 3, 2627.

Johnson, N. F., P. Jefferies, and P. M. Hui (2003): *Financial market complexity*, Oxford Finance Series, London, England: Oxford University Press.

Jovanovic, B. and R. W. Rosenthal (1988): "Anonymous sequential games," *J. Math. Econ.*, 17, 77–87.

Kalaitzoglou, I. and B. M. Ibrahim (2013): "Does order flow in the European Carbon Futures Market reveal information?" *J. Fin. Mark.*, 16, 604–635.

Kalaitzoglou, I. A. and B. M. Ibrahim (2015): "Liquidity and resolution of uncertainty in the European carbon futures market," *Int. Rev. Fin. Anal.*, 37, 89–102.

——— (2023): "Market conditions and order-type preference," *Int. Rev. Fin. Anal.*, 87, 102559.

Keim, D. B. and A. Madhavan (1995): "Anatomy of the trading process empirical evidence on the behavior of institutional traders," *J. Financ. Econ.*, 37, 371–398.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun (2017): "The flash crash: High-frequency trading in an electronic market," *The Journal of Finance*, 72, 967–998.

Kirilenko, A. A. and A. W. Lo (2013): "Moore's law versus murphy's law: Algorithmic trading and its discontents," *Journal of Economic Perspectives*, 27, 51–72.

Kullback, S. and R. A. Leibler (1951): "On information and sufficiency," *The annals of mathematical statistics*, 22, 79–86.

Kyle, A. S. (1985): "Continuous auctions and insider trading," *Econometrica*, 53, 1315.

Lee, P. M. (1989): *Bayesian statistics*, Oxford University Press London:.

Li, S., X. Wang, and M. Ye (2021): "Who provides liquidity, and when?" *Journal of financial economics*, 141, 968–980.

Lof, M. (2012): "Heterogeneity in stock prices: A STAR model with multivariate transition function," *J. Econ. Dyn. Control*, 36, 1845–1854.

MacKay, D. J. (2003): *Information theory, inference and learning algorithms*, Cambridge university press.

Madhavan, A. (2000): "Market microstructure: A survey," *J. Fin. Mark.*, 3, 205–258.

Madhavan, A., M. Richardson, and M. Roomans (1997): "Why do security prices change? A transaction-level analysis of NYSE stocks," *Rev. Financ. Stud.*, 10, 1035–1064.

Nadarajah, S. and S. Kotz (2007): "On the -type distributions," *Physica A*, 377, 465–468.

O'Hara, M. (2003): "Presidential address: Liquidity and price discovery," *J. Finance*, 58, 1335–1354.

——— (2015): "High frequency market microstructure," *J. Financ. Econ.*, 116, 257–270.

Patterson, D. M. and V. Sharma (2010): "The incidence of informational cascades and the behavior of trade interarrival times during the stock market bubble," *Macroecon. Dyn.*, 14, 111–136.

Pöppe, T., S. Moos, and D. Schiereck (2016): "The sensitivity of VPIN to the choice of trade classification algorithm," *J. Bank. Financ.*, 73, 165–181.

Ruijsenaars, S. N. (2000): "On Barnes' multiple zeta and gamma functions," *Advances in Mathematics*, 156, 107–132.

Sarkar, A. and R. A. Schwartz (2009): "Market sidedness: Insights into motives for trade initiation," *J. Finance*, 64, 375–423.

Shannon, C. E. (1948): "A mathematical theory of communication," *Bell Syst. Tech. J.*, 27, 379–423.

Shannon, C. E. and W. Weaver (1949): *The mathematical theory of communication, by CE Shannon (and recent contributions to the mathematical theory of communication), W. Weaver*, University of illinois Press.

Soros, G. (1994): *The theory of reflexivity*, Soros Fund Management New York.

——— (2013): "Fallibility, reflexivity, and the human uncertainty principle," *Journal of Economic Methodology*, 20, 309–329.

Straumann, D. (2006): *Estimation in conditionally heteroscedastic time series models*, vol. 181, Springer Science & Business Media.

Teräsvirta, T. (1994): "Specification, estimation, and evaluation of smooth transition autoregressive models," *J. Am. Stat. Assoc.*, 89, 208–218.

Tsallis, C. (1988): "Possible generalization of Boltzmann-Gibbs statistics," *J. Stat. Phys.*, 52, 479–487.

Vaart, A. v. d. and J. A. Wellner (2023): "Statistical Applications," in *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer, 385–591.

Yadav, Y. (2015): "How algorithmic trading undermines efficiency in capital markets," *Vand. L. Rev.*, 68, 1607.

Zhang, L. and X. Zhang (2025): "Mispricing and algorithm trading," *Information Systems Research*, 36, 21–40.

Appendix A Supplementary Tables

This appendix presents more information on the descriptive statistics of the dataset employed in the empirical application of the iRP metric, as well as the estimation results for further applications of the iRP, introduced in Appendix B.3.

- Table A.1 presents the descriptive statistics for the full sample period.
- Table A.2 presents the estimation results for two further applications of the iRP introduced in sections B.3.1 and B.3.2.

The sample consists of rather liquid stocks like AAPL with average duration 0.21 sec (1.21 minus 1 sec added for computational reasons) to relatively less liquid stocks, like AXP with average duration of 1.41 sec (2.41 minus 1 sec). The average volume per trade also varies from relatively low values, like in TRV with 115.7 stocks per trade, to almost triple volume, like in PFE with 393.2 stocks per trade. In addition, price change variance exhibits a wide range of values from 0.01 in KO or VZ to 0.12 in BA covering stocks with different intensities of price discovery. Besides the relative variation that is to be expected due to the presence of cross-sectional fixed effects, no major outliers are observed, while min and max values are comparable across stocks. This implies that the sample is relatively homogeneous, but with adequate variation, in order to provide a sample with minimal trading biases or extreme events.

Table A.1: Descriptive Statistics

	D	V	R	D	V	R	D	V	R	D	V	R	D	V	R
	AAPL #25,598,759			AXP #3,836,091			BA #8,058,521			CAT #4,801,058			CSCO #10,808,593		
max	324.55	22499	4.57	2577.12	29200	1.02	840.91	12500	9.77	3211.30	28500	2.29	1439.39	53910	0.92
mean	1.21	212.10	0.00	2.41	160.58	0.00	1.66	134.46	0.00	2.14	167.88	0.00	1.48	346.35	0.00
min	1.00	1.00	-4.57	1.00	1.00	-1.02	1.00	1.00	-9.76	1.00	1.00	-2.29	1.00	1.00	-0.92
std	0.48	510.42	0.05	3.64	394.25	0.02	1.60	296.68	0.12	3.09	389.42	0.03	1.35	975.34	0.01
	CVX #6,661,469			DIS #10,296,242			DOW #3,080,034			DWDP #2,052,370			GS #3,779,536		
max	2548.29	32595	2.12	888.23	31708	1.83	4482.68	71100	1.27	154.91	78320	1.54	2543.17	19441	2.01
mean	1.81	167.86	0.00	1.52	185.19	0.00	2.31	225.44	0.00	1.70	289.63	0.00	2.44	137.43	0.00
min	1.00	1.00	-2.12	1.00	1.00	-1.83	1.00	1.00	-1.28	1.00	1.00	-1.54	1.00	1.00	-2.01
std	2.27	430.45	0.03	1.16	493.48	0.03	4.25	819.56	0.02	1.38	858.93	0.02	3.46	294.93	0.04
	HD #5,840,233			IBM #4,735,123			INTC #11,300,740			JNJ #7,406,086			JPM #10,205,497		
max	3795.37	20785	3.44	3472.45	19811	2.25	1245.15	61898	1.01	1948.77	27900	2.24	1796.65	25000	1.92
mean	1.90	128.54	0.00	2.16	145.18	0.00	1.47	363.21	0.00	1.73	166.24	0.00	1.52	218.66	0.00
min	1.00	1.00	-3.44	1.00	1.00	-2.25	1.00	1.00	-1.01	1.00	1.00	-2.24	1.00	1.00	-1.92
std	2.54	302.61	0.06	2.84	311.96	0.04	1.32	1032.44	0.01	1.99	411.58	0.04	1.42	486.70	0.03
	KO #7,288,916			MCD #4,721,261			MRK #7,413,220			MSFT #18,867,325			NKE #5,536,203		
max	1393.04	66002	0.50	2925.53	25680	2.31	2112.19	49900	1.11	762.11	27514	2.03	2478.18	49900	0.89
mean	1.74	299.68	0.00	2.14	130.48	0.00	1.71	230.84	0.00	1.29	237.32	0.00	1.99	209.79	0.00
min	1.00	1.00	-0.50	1.00	1.00	-2.31	1.00	1.00	-1.11	1.00	1.00	-2.03	1.00	1.00	-0.89
std	1.74	909.80	0.01	2.72	327.95	0.04	2.05	586.03	0.02	0.68	569.69	0.03	2.46	582.56	0.01
	PFE #10,421,158			PG #6,640,793			TRV #2,064,457			UNH #5,833,071			UTX #4,066,815		
max	1520.34	88100	0.56	2886.51	34495	1.59	2428.43	20400	1.18	4220.96	24900	5.18	4410.39	36000	3.87
mean	1.51	393.16	0.00	1.82	194.57	0.00	3.60	115.71	0.00	1.94	146.00	0.00	2.32	153.62	0.00
min	1.00	1.00	-0.56	1.00	1.00	-1.59	1.00	1.00	-1.18	1.00	1.00	-5.18	1.00	1.00	-3.87
std	1.47	1186.96	0.01	2.21	486.70	0.02	5.58	268.54	0.03	2.88	362.48	0.09	3.83	435.15	0.04
	V #7,998,274			VZ #8,121,704			WBA #4,625,492			WMT #6,213,580			XOM #7,911,163		
max	2632.40	19859	2.19	1504.48	56250	0.84	4147.27	49900	0.81	789.80	24957	1.10	1265.12	35000	1.15
mean	1.69	150.36	0.00	1.64	280.82	0.00	2.16	234.04	0.00	1.86	179.76	0.00	1.67	251.62	0.00
min	1.00	1.00	-2.20	1.00	1.00	-0.84	1.00	1.00	-0.81	1.00	1.00	-1.10	1.00	1.00	-1.16
std	1.76	323.16	0.03	1.71	786.02	0.01	3.62	601.91	0.01	1.78	411.65	0.02	1.63	588.10	0.02

Table A.1 presents the descriptive statistics, i.e., the maximum (max), the average (mean), the minimum (min) and the standard deviation (std) Duration (D in seconds; 1 second is added to all observations for computational reasons), Trading Volume (V in number of stocks) and of price change (R in $'s). The statistics are presented for the full sample period for all stocks. The numbers following the # sign is the count of observations per stock.

Table A.2: Estimation Results

	iPTT			iPTT (wating)	
	low	high		low	high
ω	0.6746		ω	0.6962	
	(0.01)			(0.01)	
β	0.4295		β	0.4183	
	(0.01)			(0.01)	
ϕ	0.4324		ϕ	0.4419	
	(0.01)			(0.01)	
δ	0.0656		δ	0.0811	
	(0.01)			(0.01)	
$(q\|E(rt))$	0.8503	1.0807	$(q\|E(rt))$	0.7242	1.4298
	(0.03)	(0.02)		(0.02)	(0.03)
$(\tau\|ti)$	1.1275	0.9522	$(\tau\|ti)_{duration<j_{duration}}$	1.0942	0.6065
	(0.01)	(0.05)		(0.01)	(0.06)
			$(\tau\|ti)_{duration<j_{duration}}$	1.6323	0.9048
				(0.02)	(0.06)
g_{ti}	0.9999		g_{ti}	0.9958	
	(0.02)			(0.02)	
			$g_{duration}$	1.0001	
				(0.02)	
$g_{E(rt)}$	0.9665		$g_{E(rt)}$	0.9950	
	(0.02)			(0.03)	
j_{ti}	1.0046		j_{ti}	1.0023	
	(0.03)			(0.05)	
			$j_{duration}$	0.9966	
				(0.04)	
$j_{E(rt)}$	1.0090		$j_{E(rt)}$	1.1052	
	(0.04)			(0.02)	

The top panel of Table A.2 presents the estimation results for the conditional mean specification parameters, assuming a FI-ACD specification $\omega + \beta\psi_{i-1} + (\chi_i - \beta\chi_{i-1}) - (\tilde{\chi}_i - \phi\tilde{\chi}_{i-1})$. The bottom panel presents the distribution parameter estimates, assuming a $q-$Weibull distribution for χ, i.e., $f\left(\chi_i|\mathcal{F}_{i-1}\right) = (2 - \tau^{q=1})\frac{\tau_i^{q=2}}{\chi_i}\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}} e_q\left(-\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}}\right)$, where $A_i = \left[\Gamma\left(1+\frac{1}{\tau_i^{q=2}}\right)^{-\tau_i^{q=2}} / \psi_i\right]$, $\tau_i^{q=2} = \left(G^{q=2}_{m=1,i} - G^{q=2}_{m=2,i}\right)\tau^{q=2}_{m=1} + G^{q=2}_{m=2,i}\tau^{q=2}_{m=2}$ and $G^{q=2}_{m=2,i} = \left(1 + e^{-g^{q=2}_{m=2}(J_i - j^{q=2}_{m=2})}\right)^{-1}$. The collumns of Table A.2 present the estimates for the specifications of iPTT introduced in Section B.3. All estimates are cross-sectional averages, with standard deviations in (:).

Appendix B iRP and Further Applications

This appendix provides supplementary material that highlights the qualitative characteristics of iRP, as well as it discusses further potential applications.

- Section B.1 discusses how the new metric is aligned with the attributes of the marked viewed as a complex system.
- Section B.2 presents the theoretical properties of our measure for informational gain.
- Section B.3 presents further ”indicative” applications of iRP. Section B.3.1 presents an empirical specification that can identify the presence of technical trading alongside information, while Section B.3.2 presents another customization that can identify a finer (sub-)classification of technical trading that considers waiting costs.

B.1 The market as a complex system

The formulation in Eq.(2) is also inspired and consistent with designing the market as a complex system. Johnson et al. (2003) argue that financial markets are more than "complicated" systems. They are "complex" systems, in the sense that the observable outputs of the interactions of market participants (agent-types) are more than the sum of their actions; a concept known as emergence properties. This becomes increasingly relevant at higher trading frequencies, where trading is dominated by AI-agents. These agents observe the market continuously and infer "fundamental" information from observable "trading" information, depending on their processing capacity. In parallel, their actions also generate "trading" information (information propagation) and the market is a dynamic equilibrium between information storage and propagation (MacKay, 2003). Shannon and Weaver (1949) suggests that there is an optimal balance between information propagation and information storage capacity that maximizes reward per unit of effort, a concept known as maximum informational entropy. Arthur (2013) argues that this optimal level is an emergence property, which should be modelled alongside the actions of individual agents.

This specific point is a major contribution over previous approaches of Eq.(2), which indirectly addresses all the properties of a complex system (e.g., Johnson et al., 2003):

Multiple interacting agents: Eq.(2) proposes a market design where multiple agents with intensities $\lambda_0^k(t)$ interact and their actions collectively create the market wide intensity $\lambda^k(t|\mathcal{F}_s)$. Consequently, the overall market activity is the output of the actions of all agents present in the market, $\lambda_0^k(t)$, as well as their interactions, $(p_t^k|\mathcal{F}_s)$, which are conditional on the market as a whole. This is because the information set that determines the overall market activity does so through the conditional probabilities of each agent entering the market.

Adaptation: Adaptation refers to the ability of individual agents to adjust their behavior in order to improve their performance. This is directly modelled in Eq.(2) with the weighting probabilities, $(p_t^k|\mathcal{F}_s)$. $(p_t^k|\mathcal{F}_s)$'s are conditional on market conditions, $(:|\mathcal{F}_s)$, and define the probability of a specific agent-type with intensity $\lambda_0^k(t)$ to enter the market. This is done in a manner conditional on market activity and not independently of it and, therefore, different agent-types can adapt their behavior, i.e., adjust the weight of $\lambda_0^k(t)$, according to market conditions, $(:|\mathcal{F}_s)$, which include the actions of all agents.

Feedback: The adaptive learning described above is designed in way that also considers information propagation. $(p_t^k|\mathcal{F}_s)$'s are conditional on observable information, which according to Eq.(2) is necessarily the collective output of all interacting agents. This way, Eq.(2) establishes a feedback mechanism that considers learning and information processing through the conditional set, $(:\ F_s)$, as well as through the constraint that overall market activity, $\lambda^k(t|\mathcal{F}_s)$, is necessarily the output of individual agent actions, $\lambda_0^k(t)$, i.e., $\sum_{k=1}^{K}(p_t^k|\mathcal{F}_s) = 1$.

Evolution: The information set $(:|\mathcal{F}_s)$ does not necessarily contain only endogenous information, i.e., $\lambda^k(t|\mathcal{F}_s)$, but it can also contain other exogenous variables. This way, exogenous, as well as endogenous, shocks can affect the probabilities of different agent-types to enter the market, $(p_t^k|\mathcal{F}_s)$, without necessarily imposing a mean reversion property.[24] Consequently, the market activity, $\lambda(t|\mathcal{F}_s)$, can evolve, continuously searching for an equilibrium, and can even exhibit 'extreme behavior', such as crashes or bubbles. This would happen when $\lambda^k(t|\mathcal{F}_s)$ takes extreme values, which would then be also fed back to the system through $(:|\mathcal{F}_s)$'s, potentially leading to market failures.

Non-stationarity: This evolution does not have to be stationary, i.e., system properties observed in the past do not necessarily remain unchanged in the future. Eq.(2) operates at a trade-off with respect to stationarity. The agent-specific characteristics $\lambda_0^k(t)$, i.e., the way agent-types act given a specific attribute, such as learning, speed, processing capacity etc., are assumed to be stationary. This is done for traceability reasons and not because it represents better complex system properties. This is an explicit assumption, which implies

[24] Bacry et al. (2015) provides a martingale representation of Hawke's processes, while Engle and Russell (1998) suggest a way to model the the innovations of inter-event waiting times, i.e., durations. These two are the two most popular approaches employed in finance to model time and both have an innovation component embedded. This way endogenous or exogenous shocks can affect the arrival rate of either agent-specific or market-wide events. In this paper, due to its empirical focus, the approach of Engle and Russell (1998) is preferred due to the explicit modeling of the innovations, which will be the main tool to distinguish among different agent-types.

that an agent-type reacts to the same stimuli in a way that does not change over time. According to the formulation in Eq.(2), this is what makes her actions distinguishable and thus, traceable. This is definitely a constrain in the modeling, but Eq.(2) can accommodate time variant agent-type characteristics, albeit in a rigid way. Assuming that the information processing capacity of an agent-type changes due to a structural break, such as changes in technology. This would necessarily imply a change of the baseline intensity, $\lambda_0^{kA}(t) \rightarrow \lambda_0^{kB}(t)$. Eq.(2) would be able to identify $\lambda_0^{kB}(t)$ (i.e., as a new intensity) and also $(p_t^{kB}|\mathcal{F}_s)$, making at the same time $(p_t^{kA}|\mathcal{F}_s) = 0$.

B.2 Informational Gain

Figure 1: Information Gain

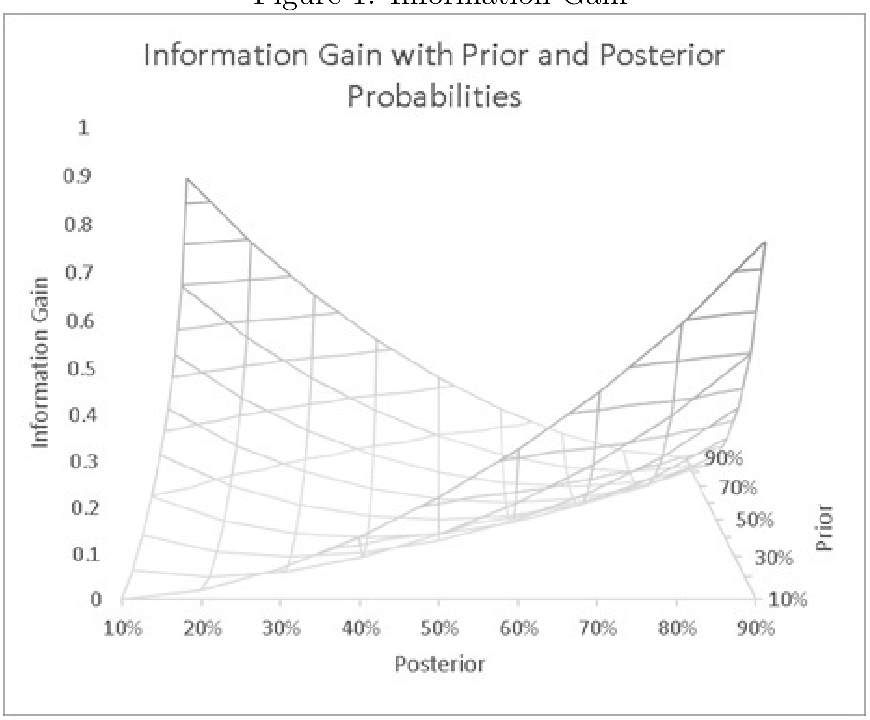

Figure 1 illustrates the value of Information Gain (S_i in Eq. 16) for different estimates of the prior (L_i^k) and posterior ($\mathbb{P}_i^k$) probabilities.

Figure 1 illustrates how S_i changes across different estimates of prior and posterior probabilities. In the cases where the prior over- or under-estimates the probability significantly,

S_i gets values closer to 1, indicating that trading does revise strongly the probability. We interpret this as a strong impact of arrival rates associated with market reflexivity.

B.3 Further Applications: Beyond Information

This section presents an application of the iRP through various extensions of the framework presented in Eq.(3)-(10). More precisely, Section 3.2 employs a different modelled variable, i.e., the speed of volume accumulation, in order to capture order flow toxicity, similar to VPIN. Section B.3.1 expands the interpretation of iPIN by dissecting the shape and the scale parameter simultaneously in order to investigate technical trading. Finally, Section B.3.2 dissects the shape parameter further, i.e., in two dimensions, in order to investigate the behavioral bias of ”patient traders” (e.g., Foucault et al., 2005). In brief, the following sections employ different specifications of Eq.(3)-(10) and link them indicatively to different agent-types, as an illustration of the flexibility of the framework.

B.3.1 Information and Market-Wide Learning: Technical Trading

Access to private information might not be synchronous to all traders and, therefore, the classification of informed versus uninformed (e.g., Kyle, 1985) might be unrealistic. Previous literature recognizes the existence of other groups under the general notion of discretionary liquidity traders (e.g., Admati and Pfleiderer, 1986), usually referred to as “technical” traders. Although there are many sub-classifications according to their way/speed of learning or access to information (e.g., O’Hara, 2015), they tend to exhibit some common characteristics. They are generally understood as a challenge to weak-form market efficiency. They observe the market and “learn”, in the sense of inferring exploitable patterns from trading history. They enter the market when they extract price-relevant information, in a discretionary way by selecting the timing and the volume of their trading. Unlike the uninformed agents, whose arrival rate is time invariant, these partially uninformed agents try to become better informed (learn). Consequently, their arrival rate depends on market conditions.

Kalaitzoglou and Ibrahim (2013) suggest that technical traders, due to their learning time-related cost, react to information with a delay compared to better informed agents. Consequently, their probability of entering the market, following an information signal increases with time, resulting in a hazard function with an increasing shape. Other studies, however, focus on market signals that might instigate technical trading, with expected return being probably the strongest indicator. For example, Bauwens and Giot (2003) imply that the trading activity of technical traders increases after large price changes are expected and therefore, their probability to enter the market is higher closer to the events that result in higher expected price changes and, then, it decreases over time. This implies that when $J_i := \mathbb{E}(R_i|\mathcal{F}_{i-1}) > j_{\text{technical}}$, then the hazard function should exhibit a decreasing shape. Both approaches imply a monotonic hazard function and are constrained by the use of the distributions employed, as well as the observable threshold variables.

Eq.(14) provides a versatile tool for identifying technical trading, using multiple observable factors. Combining both approaches, technical traders are understood to observe past price changes and formulate expectations about the presence of information signals that motivate their discretionary interaction with the market. Naturally, they should be expected to act as soon as they can, before the informational advantage expires by becoming public information. This should result in a hazard function with a decreasing shape. However, they cannot enter the market immediately after the arrival of new information because they do not possess it from the beginning. They have to extract it first and, therefore, their probability of entering the market after the arrival of information should increase over time, resulting in a hazard function with increasing shape. Consequently, the trading behavior of technical traders should be a combination of the two different shapes of the hazard function – first increasing (enter the market with a delay) and then decreasing (act timely on newly acquired information) until it reaches zero when information becomes public – implying a unimodal shape.[25]

[25] In the absence of information, the probability of technical traders should be low or zero. In the presence of "extracted" information they should act within its "life span", implying a decreasing hazard function.

Eq.(11) is flexible enough to capture non-monotonic hazard functions and the one that matches the unimodal shape is when $\tau^{q=1} \geq 1$ and $\tau^{q=2} \geq 1$. Following, the formulation of iPIN and iVPIN, the actions of technical traders are assumed to be expressed primarily in trading intensity, prior to any significant price change that would incorporate the new information. They act on information with a delay, because they have to "learn" first, and therefore, their hazard function should exhibit mostly an increasing shape, at least during the "life span" of the information. In line with the formulation of of iPIN and iVPIN, trading intensity is assumed to be a primary characteristic that expresses the actions of different agent-types and therefore, it is also used here as a major determinant of the shape of the hazard function, i.e., flat/increasing/decreasing. Following the notation in Eq. (11) and the example in Table 2 the shape parameter ,i.e., $\left(\tau_i^{q=2} \middle| J_{v=1,i}^{q=2} = ti_i\right)$, is conditional on trading intensity. Considering that a decreasing hazard function is associated with informed trading (e.g., iPIN), a milder or increasing (e.g., Kalaitzoglou and Ibrahim, 2013) shape of the hazard function should be associated with less informed trading that is more consistent with the presence of technical traders. This on its own is not sufficient and a secondary signal is also considered. The magnitude of the reaction of technical traders is proportional to the magnitude of the inferred information and this is in lined with the impact (extensivity) of the entropy parameter on the shape of the hazard function. The "informativeness" of the magnitude of an observation and its subsequent impact on the shape of the hazard function is captured by the entropy parameter, $\tau_i^{q=1}$, which is assumed to be conditional on the magnitude of the (information) signal, here assumed to be only (but this is only indicative) expected returns, i.e., $\left(\tau_i^{q=1} \middle| J_{v=1,i}^{q=1} = \mathbb{E}(R_i|\mathcal{F}_{i-1})\right)$. Higher values should be associated with higher presence of technical trading. It follows that the intensity-based probability of technical traders, i.e., iTT, can be defined following Eq.(14)

However, because they have to "learn" first, they cannot act as fast as the informed agents and their actions are expressed with a delay. This implies an increasing hazard function upon the arrival of information. When inforamtion is price-resolved there is no monetary benefit and, therefore, their probability of entering the market returns to zero. Collectively, the probability of technical traders follows a unimodal shape that cannot be captured by single-shape parameter distributions.

as: iTT $= \left(\text{iRP}_t^{\text{technical}} \middle| \mathcal{F}_s\right) = \frac{\mathbb{E}\left(\int_s^t \lambda^{Z=\text{technical}}(u)\, du \middle| \mathcal{F}_s\right)}{\sum_{k=1}^{K} \mathbb{E}\left(\int_s^t \lambda^{k}(u)\, du \middle| \mathcal{F}_s\right)} = \frac{\left(p_t^{Z=\text{technical}} \middle| \mathcal{F}_s\right) H^{Z=\text{technical}}(t|\mathcal{F}_s}{\sum_{k=1}^{K}(\left(p_t^{k} \middle| \mathcal{F}_s\right) H^{k}(t|\mathcal{F}_s)}$, which can then be expressed in terms of the regimes of shape parameters, τ_i^q, in Eq.(11) as:

$$\text{iTT}_i = \frac{\sum_{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2}>1\right):k\right)} \left\{\prod_{Q \bigotimes M} W_{m,i}^q H^{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2}>1\right):k\right)}(t)\right\}}{\sum_{Q \bigotimes M} \prod_{Q \bigotimes M} W_{m,i}^q H^{Q \bigotimes M}(t)} \tag{B.1}$$

In this formulation, the regimes, m of $\tau^{q=1}$ and $\tau^{q=2}$ that lead to a unimodal shape of the hazard function, identify technical trading (k = technical). Then, the aggregated number of technical traders is the sum of all the probabilities, i.e., $\prod_{Q \bigotimes M} W_{m,i}^q$, of intersections $Q \bigotimes (m\,(1 < \tau_m^{q=1} \leq 2, \tau_m^{q=2} > 1) : k)$ in the contingency table, where $1 < \tau_m^{q=1} \leq 2$ and $\tau_m^{q=2} > 1$, times the respective cumulative hazard functions of these intersections, i.e., $H^{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2}>1\right):k\right)}(t)$. This is then compared to the expected number of all trades. Technical trading in Eq.(B.1) can be estimated in parallel with iPIN. Different regimes of $\tau_m^{q=1}$ and $\tau_m^{q=1}$ that lead to a decreasing hazard function would identify informed trading. The regimes that exhibit a unimodal shape would indicate the presence of technical trading.

Table A.2 reports the cross-sectional estimates of the specification in this table that intends to identify technical trading. In particular, the interest lies in the middle panel under iPTT, which reports the estimates for the distribution of χ with both the scale and the shape parameter being an infinite mixture of two regimes. Focusing on the shape parameter $\tau_i^{q=2}$, noted as τ_{ti}, it takes values higher than one when $ti_i < j_{ti}$ (noted as $ti < j(ti)$, while it takes values less than 1 when $ti_i > j_{ti}$ (noted as $ti > j(ti)$. In consistence with the iPIN and iVPIN estimations higher trading intensity is associated with higher presence of private information (decreasing hazard function). However, lower trading intensity that leads to $\tau_{ti} > 1$ and higher $E(rt)$ that leads to $\tau_i^{q=1} > 1$ (noted as $q_{E(rt)>j(E(rt))}$) make the hazard function take a unimodal shape. This matches the characteristics of technical trading. At the bottom panel, their cross-sectional average over the sample period is estimated around 30%.

$$\psi_i = \omega + \beta\psi_{i-1} + (\chi_i - \beta\chi_{i-1}) - (\tilde{\chi}_i - \phi\tilde{\chi}_{i-1}), \text{ where } \chi_i = \Delta t_i^*$$

$$f(\chi_i|\mathcal{F}_{i-1}) = (2 - \tau_i^{q=1})\frac{\tau_i^{q=2}}{\chi_i}\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}} e_q\left(-\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}}\right), \text{ where } A_i = \left[\Gamma\left(1+\frac{1}{\tau_i^{q=2}}\right)^{-\tau_i^{q=2}} / \psi_i\right]$$

$$\tau_i^{q=1} = \left(1 - G_{m=2,i}^{q=1}\right)\tau_{m=1}^{q=1} + G_{m=2,i}^{q=1}\tau_{m=2}^{q=1} \text{ and } \tau_i^{q=2} = \left(1 - G_{m=2,i}^{q=2}\right)\tau_{m=1}^{q=2} + G_{m=2,i}^{q=2}\tau_{m=2}^{q=2}$$

$$\text{for} \quad G_{m=2,i}^{q=1} = \left(1 + e^{-g_{m=2}^{q=2}(E(rt)_i - j_{m=2}^{q=2})}\right)^{-1}, \text{ and } G_{m=2,i}^{q=2} = \left(1 + e^{-g_{m=2}^{q=2}(ti_i - j_{m=2}^{q=2})}\right)^{-1}$$

The remaining regimes that exhibit differently shaped hazard functions could be associated with different agent-types. As an indications, previous literature (e.g., Kalaitzoglou and Ibrahim, 2013) suggest that a flat hazard function can be associated with uninformed agents. Of course, there are different shapes and/or different combinations. The next section (Section B.3.2) discusses indicatively a variation of the potential shapes the hazard function can take.

B.3.2 Information Diffusion and Learning: Waiting Costs

The discussion above is based on market-wide learning and uniform access to information. However, technical traders are a diverse collection of agents (e.g., O'Hara, 2015) motivated by a variety of factors, such as investment style and behavioral biases. These differences would instigate different trading styles that thus, different accumulation rates (hazard functions). A potential (indicative) factor is waiting costs (e.g., Foucault et al., 2005). Lower-frequency traders ("patient") trade according to their portfolio-re-balancing needs (e.g., Keim and Madhavan, 1995). They are more likely to wait until a sufficiently strong signal or submit a limit order trying to secure a better price (e.g., Foucault et al., 2005); actions that lead to delayed execution. In contrast, faster, e.g., algorithmic, traders ("impatient") profit from accessing and acting on information faster than the remaining uninformed agents. They are more likely to submit a market order as soon as it is profitable. Both types can be considered uninformed and their actions are associated with a unimodal hazard function. However, (im-)patient traders face a (higher) lower waiting cost and therefore, the degree of the curvature of the hazard function should be (sharper) milder (e.g., Foucault et al., 2005), even monotonically (decreasing) increasing in the limit.

An extension of Eq.(B.1) can capture this dissection of technical traders. Consider the setup, in Eq.(B.1), i.e., $\left(\tau_i^{q=2} \middle| J_{v=2,i}^{q=2} = \frac{\text{volume}_i}{\text{duration}_i}\right)$ and $\left(\tau_i^{q=1} \middle| J_{v=1,i}^{q=1} = \mathbb{E}(R_i|\mathcal{F}_{i-1})\right)$. Impatient traders act faster than patient traders and this can be captured by distinguishing the trades that result into $\tau^{q=2} > 1$, i.e., technical trading, into relatively high or low duration. To capture this, $\tau^{q=2}$ is dissected into levels of trading intensity, i.e., $J_{v=2,i}^{q=2}$, as well as of duration, i.e., $J_{v=3,i}^{q=2} = \text{duration}_i$. This conditions the shape parameter on two variables, $\left(\tau_i^{q=2} \middle| J_{v=2,i}^{q=2}, J_{v=3,i}^{q=2}\right)$, and dissects it into four cases; high (low) trading intensity and long (short) durations.

$$\begin{aligned}
\tau_i^{q=2} = \Bigg[& \underbrace{\left(1 - G_{v=2,m=2,i}^{q=2}\left(J_{v=2,i}^{q=2}\right)\right)\left(1 - G_{v=3,m=2,i}^{q=2}\left(J_{v=3,i}^{q=2}\right)\right)}_{W_{\text{low } J_{v=2,i}^{q=2}, \text{ long } J_{v=3,i}^{q=2}, i}} \tau_{\text{low } J_{v=2,i}^{q=2}, \text{ long } J_{v=3,i}^{q=2}}^{q=2} \\
& + \underbrace{G_{v=2,m=2,i}^{q=2}\left(J_{v=2,i}^{q=2}\right)\left(1 - G_{v=3,m=2,i}^{q=2}\left(J_{v=3,i}^{q=2}\right)\right)}_{W_{\text{high } J_{v=2,i}^{q=2}, \text{ long } J_{v=3,i}^{q=2}, i}} \tau_{\text{high } J_{v=2,i}^{q=2}, \text{ long } J_{v=3,i}^{q=2}}^{q=2} \Bigg] \\
\Bigg[& \underbrace{\left(1 - G_{v=2,m=2,i}^{q=2}\left(J_{v=2,i}^{q=2}\right)\right) G_{v=3,m=2,i}^{q=2}\left(J_{v=3,i}^{q=2}\right)}_{W_{\text{low } J_{v=2,i}^{q=2}, \text{ short } J_{v=3,i}^{q=2}, i}} \tau_{\text{low } J_{v=2,i}^{q=2}, \text{ short } J_{v=3,i}^{q=2}}^{q=2} \\
& + \underbrace{G_{v=2,m=2,i}^{q=2}\left(J_{v=2,i}^{q=2}\right) G_{v=3,m=2,i}^{q=2}\left(J_{v=3,i}^{q=2}\right)}_{W_{\text{high } J_{v=2,i}^{q=2}, \text{ short } J_{v=3,i}^{q=2}, i}} \tau_{\text{high } J_{v=2,i}^{q=2}, \text{ short } J_{v=3,i}^{q=2}}^{q=2} \Bigg]
\end{aligned} \tag{B.2}$$

The formulation above dissects the shape parameter, $\tau_i^{q=2}$, into four regimes. Different combinations of trading intensity and duration will result in different levels of $\tau_i^{q=2}$ and, consequently, in different shapes of the hazard function. The shapes of interest are the variations of a unimodal shape, i.e., $\tau_i^{q=2} > 1$, that might vary from a marginally decreasing shape to a marginally increasing. In particular, when expected return is high, $\tau_i^{q=1}$ is expected to be $1 < \tau_i^{q=1} \leq 2$. In combination with a $\tau_i^{q=2} > 1$, this would lead to a unimodal shape. The identification that is pursued here goes one step further and distinguishes different levels of $\tau_i^{q=2}$ according to different levels of duration. In particular, (shorter) longer duration should be associated with (im-)patient trading, resulting in a hazard function that exhibits a (sharper, i.e., $\tau_i^{q=2} \xrightarrow{+} 1$) milder, i.e., $\tau_i^{q=2} > 1$, decrease, which in the limit could even reach a (decreasing) increasing shape. This identifies impatient technical traders (iITT) and patient technical traders (iPTT) (see e.g., Foucault et al., 2005), with a relative proportion that can be defined as:

$$
\begin{aligned}
\text{iITT}_i &= \frac{\sum_{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2} \xrightarrow{+} 1\right):k\right)} \left\{\prod_{Q \bigotimes M} W_{m,i}^q H^{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2} \xrightarrow{+} 1\right):k\right)}(t)\right\}}{\sum_{Q \bigotimes M} \prod_{Q \bigotimes M} W_{m,i}^q H^{Q \bigotimes M}(t)} \\
\text{iPTT}_i &= \frac{\sum_{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2} > 1\right):k\right)} \left\{\prod_{Q \bigotimes M} W_{m,i}^q H^{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2} > 1\right):k\right)}(t)\right\}}{\sum_{Q \bigotimes M} \prod_{Q \bigotimes M} W_{m,i}^q H^{Q \bigotimes M}(t)}
\end{aligned}
\tag{B.3}
$$

The regimes, m of $\tau^{q=1}$ and $\tau^{q=2}$ that lead to a unimodal shape of the hazard function identify the technical trading, while the different degrees of curvature are associated with the magnitude of waiting costs. The aggregated number of impatient (or patient in (:)) technical traders, is the sum of all the probabilities, i.e., $\prod_{Q \bigotimes M} W_{m,i}^q$, of intersections $Q \bigotimes \left(m\left(1 < \tau_m^{q=1} \leq 2, \tau_m^{q=2} \xrightarrow{+} 1\right) : k\right)$ (or $Q \bigotimes \left(m\left(1 < \tau_m^{q=1} \leq 2, \tau_m^{q=2} > 1\right) : k\right)$ for patient traders) in the contingency table, where $1 < \tau_m^{q=1} \leq 2$ and $\tau_m^{q=2} \xrightarrow{+} 1$ (or $1 < \tau_m^{q=1} \leq 2$ and $\tau_m^{q=2} > 1$, times the respective cumulative hazard functions, $H^{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2} \xrightarrow{+} 1\right):k\right)}(t)$ (or $H^{Q \bigotimes \left(m\left(1<\tau_m^{q=1}\leq 2, \tau_m^{q=2} > 1\right):k\right)}(t)$). This is then compared to the expected number of all trades.

As a final example, the specification below follows the one in Section B.3.1 but with a higher refinement of the shape parameter, which is now split across an additional dimension. More specifically, consider the following model:

$$\psi_i = \omega + \beta\psi_{i-1} + (\chi_i - \beta\chi_{i-1}) - (\tilde{\chi}_i - \phi\tilde{\chi}_{i-1}), \text{ where } \chi_i = \Delta t_i^*$$

$$f\left(\chi_i|\mathcal{F}_{i-1}\right) = (2-\tau_i^{q=1})\frac{\tau_i^{q=2}}{\chi_i}\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}} e_q\left(-\left[\frac{\chi_i}{A_i}\right]^{\tau_i^{q=2}}\right), \text{ where } A_i = \left[\Gamma\left(1+\frac{1}{\tau_i^{q=2}}\right)^{-\tau_i^{q=2}} / \psi_i\right]$$

$$\tau_i^{q=1} = \left(1-G_{m=2,i}^{q=1}\right)\tau_{m=1}^{q=1} + G_{m=2,i}^{q=1}\tau_{m=2}^{q=1}$$

$$\tau_i^{q=2} = \left(1-G_{v=2,i}^{q=2,a}\right)\left(1-G_{v=3,i}^{q=2,b}\right)\tau^{q=2}_{J^{q=2}_{v=2,m=1,i}J^{q=2}_{v=3,m=1,i}} + G_{v=2,i}^{q=2,a}G_{v=3,i}^{q=2,b}\tau^{q=2}_{J^{q=2}_{v=2,m=2,i},J^{q=2}_{v=3,m=2,i}}$$

$$+G_{v=2,i}^{q=2,a}\left(1-G_{v=3,i}^{q=2,b}\right)\tau^{q=2}_{J^{q=2,a}_{v=2,i},J^{q=2}_{v=3,m=1,i}} + \left(1-G_{v=2,i}^{q=2,b}\right)G_{v=3,i}^{q=2,a}\tau^{q=2}_{J^{q=2}_{v=2,m=1,i},J^{q=2}_{v=3,m=2,i}}$$

$$\text{for } \quad G_{m=2,i}^{q=1} = \left(1+e^{-g_{m=2}^{q=2}(E(rt)_i - j_{m=2}^{q=2})}\right)^{-1},$$

$$G_{m=2,i}^{q=2,a} = \left(1+e^{-g_{v=2,m=2}^{q=2}(ti_{i-1} - j_{v=2,m=2}^{q=2})}\right)^{-1} \text{ and } G_{m=2,i}^{q=2,b} = \left(1+e^{-g_{v=3,m=2}^{q=2}(\text{duration}_i - j_{v=3,m=2}^{q=2})}\right)^{-1}$$

In this specification, the entropy parameter, $q = \tau_i^{q=1}$, is split into two regimes, $m = 1, 2$, defined by the threshold variable, $E(rt)$, as it is compared to the threshold, $j_{m=2}^{q=1}$. The shape parameter though is split into a $2x2$ dimension matrix, defined by two variables, $J_{v=2,i}^{q=2} = \text{duration}_{i-1}$ and $J_{v=3,i}^{q=2} = ti_i$. The first one captures the market conditions that might be associated with technical trading (trading intensity associated with an increasing hazard function), while the second distinguishes this level of trading intensity into faster-bigger (shorter duration-higher volume) and slower-smaller (longer duration-lower volume) trades, which is intended to capture the execution strategy (fast or slow). Each threshold variable has its own threshold, $j_{v=2,m=2}^{q=2}$ for $J_{v=2,i}^{q=2}$ and . The magnitude of the threshold variables, $J_{v=2,i}^{q=2} = ti_{i-1}$ and $J_{v=3,i}^{q=2} = ti_i$, relative to their thresholds, $j_{v=2,m=2}^{q=2}$ and $j_{v=3,m=2}^{q=2}$, define four different combinations, with the shape of the hazard functions being defined by the related shape parameters:

	$J_{v=3,i}^{q=2} \leq j_{v=3,m=2}^{q=2}$	$J_{v=3,i}^{q=2} > j_{v=3,m=2}^{q=2}$
$J_{v=2,i}^{q=2} \leq j_{v=2,m=2}^{q=2}$	$\tau^{q=2}_{J^{q=2}_{v=2,m=1,i}J^{q=2}_{v=3,m=1,i}}$	$\tau^{q=2}_{J^{q=2}_{v=2,m=1,i},J^{q=2}_{v=3,m=2,i}}$
$J_{v=2,i}^{q=2} > j_{v=2,m=2}^{q=2}$	$\tau^{q=2}_{J^{q=2,a}_{v=2,i},J^{q=2}_{v=3,m=1,i}}$	$\tau^{q=2}_{J^{q=2}_{v=2,m=2,i},J^{q=2}_{v=3,m=2,i}}$

These parameters have been indicatively been estimated in the left side of Table A.2, in the right column under iPTT. The middle part of this column reports the estimates of the parameters of the distribution, which are inline with expectations and previous estimations. More precisely, in accordance with Section B.3.1, the values of $\tau^{q=2}$ are below 1 when $J^{q=2}_{v=2,m=2,i} > j_{v=2,m2}$ (noted as $ti > j(ti)$), which indicates informed trading. However, the shape parameter, $\tau^{q=2}$ takes values that are consistently higher than 1 when $J^{q=2}_{v=2,m=2,i} \leq j_{v=2,m2}$ (noted as $ti < j(ti)$), which in combination with a scale parameter, $\tau^{q=1} = q$, that takes values higher than 1 when $E(rt) > j_{v=3,m2}$ (noted as $E(rt) > j(E(rt))$), imply a unimodal shape for the hazard function. This is consistent with technical trading. On a deeper dissection, when duration is low (shaded areas) the shape parameter takes a value closer to 1, while longer durations are associated with higher values. This is consistent with impatient and patient technical trading. These quantities can be then estimated following Eq. (B.3) and they are reported in the bottom panel and take values aroung 15%.

Appendix C Technical Appendix

The definition and the limit theory of the maximum likelihood estimator (MLE) for the parameters of interest are presented and derived in this section. The statistical model is allowed to be misspecified. The point processes involved assume their values in an interval of the form $[0, T]$. The asymptotics operate as $T \to +\infty$. $\rightsquigarrow$ denotes convergence in distribution. "w.h.p." abbreviates the phrase "with probability tending to one".

The latent conditional duration process $(\psi_i)_{i\in\mathbb{Z}}$ is assumed to be a solution of a stochastic recurrence equation (SRE-see Ch.2 of Straumann, 2006) of the form $\psi_i = \Psi(\theta_0, \psi_{i-l}, \chi_{i-j}, l = 1, \ldots, q,\ j = 1, \ldots, p)$, where Ψ is a real function, θ_0 is an unknown value of a Euclidean parameter θ that belongs to some known $\Theta \subseteq \mathbb{R}^l$, while p is allowed to assume the value $+\infty$ to incorporate ARCH(∞)-type of elements associated for example with the FIACD model used in the empirical section (e.g., Section 4). We also have that for each $i \in \mathbb{Z}$, χ_{i-l}, $\varepsilon_{i-l}, \boldsymbol{J}_{i-l+1}$, $l \geq 1$, are measurable w.r.t. $\mathcal{F}_{i-1}$. The parameters of interest are collected in the Euclidean vector $\phi := (\theta^T, \tau^T)^T$, that assumes its values in a known subset Φ of the Euclidean space $\mathbb{R}^{l+3MQ}$, where $\tau := \mathrm{vec}(\tau_m^q, g_m^q, j_m^q)_{m=1,\ldots,M}^{q=1,\ldots,Q}$. The sample realizations of the observable processes $y := (\chi_i, \boldsymbol{J}_i)_{i=1,\ldots T}$ are known, with $T \geq \max(p^\star, q)$, for $p^\star$ a finite truncation of any non trivial ARCH(∞) component, and a-potentially random-initialization $\hat{\psi}_0(\theta)$, $\theta \in \Theta$ constructs a recurrent filter of the latent conditional duration defined by:

$$\hat{\psi}_i(\theta) := \begin{cases} \hat{\psi}_0(\theta),\ i = 0, \\ \Psi^\star((\theta; \hat{\psi}_{i-l}, x_{i-j}, l = 1, \ldots, \min(i, \max(p^\star, q)),\ j = 1, \ldots, \min(i, p^\star)),\ 0 < i \leq T, \end{cases}$$

where for the modified SRE that appears in the previous display we have that $\Psi^\star((\theta; \hat{\psi}_{i-l}, \chi_{i-j}, l = 1, \ldots, \max(p^\star, q),\ j = 1, \ldots, \min(i, p^\star)) := \Psi((\theta; \hat{\psi}_{i-l}, {}^{\chi_{i-j}}/_{\hat{\psi}_{i-j}}, l = 1, \ldots, \min(i, q),\ j = 1, \ldots, \min(i, p^\star))$. Then, the log-likelihood function is defined by $\ell_T(y; \phi) := {}^1/_T \sum_{i=1}^T \ell_i(\hat{\psi}_i, \tau_i; \phi)$, where the likelihood contributions are $\ell_i(\hat{\psi}_i, \tau_i; \phi) := \ln f(\chi_i; \hat{\psi}_i, \tau_i; \phi))$. The MLE, say ϕ_T, is defined via the variational problem

$$\ell_T(y; \phi_T) \leq \inf_{\Phi} \ell_T(y; \phi) + \epsilon_T,$$

with ϵ_T almost surely non-negative that admits the role of optimization error. The specification appearing in Table 2 is readily conformable to the above.

Maximization of the likelihood function produces the MLE ϕ_T as well as the filtered estimate $\hat{\psi}_{s+1}(\phi_T)$ for the latent duration. Plugging those into Eq.(**??**) or Eq.(18) produces estimators for the iPIN and iVPIN respectively (empirical iRP). For example, the resulting estimator of $\text{iPIN}_{t,s}$ at time t, given $\mathcal{F}_s$, is denoted $\text{iPIN}_{t,s}(\phi_T, \hat{\psi}_{s+1})$. The quality of the approximation of $\text{iPIN}_{t,s}$ by its empirical counterpart $\text{iPIN}_{t,s}(\phi_T, \hat{\psi}_{s+1})$ is derived in the following paragraphs via the asymptotic properties of the MLE and of the empirical filter. Similar derivations can be made for the plug-in estimator of the iVPIN, which are discussed here briefly for the sake of brevity, as well as for other estimators in Appendix C.4.

C.1 (Pseudo-)Consistency of the new estimator

This section examines (pseudo-)consistency of the plug-in estimator of iRP. The framework employed ensures the existence of an approximating likelihood function with stationary and ergodic contributions, so that locally uniform versions of the ergodic Law of Large Numbers (LLN) are applicable. The assumptions also posit that the limiting likelihood is uniquely minimized at a parameter value $\phi^\star$ that implies minimization of the Kullback-Leibler (KL) Kullback and Leibler (1951) divergence between the true conditional density of χ_i and $f(\chi_i; \psi_i, \tau_i; \phi^\star)$. These conditions enable the investigation of the limiting behavior of the MLE estimate of iRP, even in cases of model misspecification. Consistency is derived when the intensity-based statistical model at hand is well-specified. A general framework is thus provided, that ensures the approximation of the true latent PIN by the empirical iRP. The framework is exemplified for the specification that appears in Table 2. In this context, consider the following assumptions:

A1: *The joint process* $(\chi_i, \boldsymbol{J}_i)_{i\in\mathbb{Z}}$ *is stationary and ergodic. The density of* χ_i *conditionally on* $\mathcal{F}_{i-1}$ *exists and has an integrable logarithm.*

Remark 1. Stationarity and ergodicity for $(\chi_i)_{i\in\mathbb{Z}}$ would follow from stationarity and ergodicity of $(\varepsilon_i)_{i\in\mathbb{Z}}$ and conditions that ensure existence and uniqueness (up to modification) of a solution to the Ψ-SRE defined via an almost sure limit of backward substitutions (Ch.2 of Straumann, 2006). If the specification appearing in Table 2 is correct, this would follow as long as $\max(|\beta_0|, |\phi_0|) < 1$ and $\delta_0 \in [0, 1/2)$, where β_0, ϕ_0, and δ_0 denote the unique true values for the auto-regressive and the fractional differencing parameters respectively-see the assumption that follows (Jasiak, 1999). If the remaining parts of the process satisfy SREs, then analogous considerations would suffice. If the elements of the remaining components are measurable transformations, fixed with respect to i, of the underlying stationary and ergodic processes, then these components will also exhibit stationarity and ergodicity.

A2: *There exists a* $\phi^\star \in \Phi$ *such that* $\mathbb{E}(\ell_0(\phi^\star)) > \mathbb{E}(\ell_0(\phi))$ *for all* $\Phi \ni \phi \neq \phi^\star$.

Remark 2. This is an identification condition for the (pseudo-) true value of the parameter involved. In the case of the specification in Table 2, and if the model is well specified, this follows from the fact that the conditional distribution of χ_i has a density (Par. 5.4.2 of Straumann, 2006). In the case of misspecification, the previous would also suffice due to Proposition 2.3 of Jenkinson (2019).

In what follows, $\Theta^\star$ denotes an arbitrary compact subset of Θ.

A3: *Suppose that: (i).* $\mathbb{E}(\sup_{\theta\in\Theta^\star} |\Psi^\star(\cdot)|) < +\infty$. *(ii).* $\Psi^\star(\theta; \psi_{i-l}, x_{i-j}, l = 1, \ldots, \max(p, q), j = 1, \ldots, p)$ *is almost surely Lipschitz continuous in* $(\psi_{i-l},\ l = 1, \ldots, \max(p, q))$*, with Lipschitz coefficient* $\Lambda_i(\theta)$*, and such that, (a). the map* $\Theta^\star \ni \theta \to \Lambda_0(\theta)$ *is almost surely continuous and, (b).* $\mathbb{E}(\sup_{\theta\in\Theta^\star} \log^+ \Lambda_0(\theta)) < 0$*, where* $\log^+$ *is the positive part of the logarithmic function.*

Remark 3. The assumption implies the continuous invertibility (Par. 3 and Prop. 3.1 of

Blasques et al., 2018) of the filter $(\hat{\psi}_i(\theta))$; this is equivalent to the existence of a stationary and ergodic process (say $(f_i(\theta))_{i\in\mathbb{Z}}$), that approximates appropriately fast, almost surely, and (locally) uniformly over (δ,θ), the original duration process as $i\to\infty$. In the specification of Table 2, it is ensured whenever $\max(|\beta_0|,|\phi_0|)$ is bounded below 1.[26]

A4: *$\Phi^\star$ denotes any compact subset of Φ such that if $\Phi^\star\ni\phi=(\theta,\eta)$, then $\theta\in\Theta^\star$. Suppose that there exists a stationary non-negative process process $(m_i)_{i\in\mathbb{Z}}$, with $\mathbb{E}(\log^+ m_0)<+\infty$, such that almost surely $\sup_{\phi\in\Phi}|\ell_i(\hat{\psi}_i,\tau_i;\phi)-\ell_i(f_i,\tau_i;\phi)|\leq m_i\sup_{\theta\in\Theta^\star}|\hat{\psi}_i-f_i|$ for any $i\in\mathbb{N}$, where $f_i(\theta)$ as in the previous remark.*

Remark 4. Given the continuous invertibility, the assumption allows the approximation of the likelihood function by a stationary and ergodic version that is constructed via the limiting filtering process $(f_i)_{i\in\mathbb{Z}}$. In the specification of Table 2 it holds when the elements of the (τ_i) process are almost surely bounded from above and away from zero, x_0 has a logarithmic moment (see also 1), and f_i is uniformly over θ bounded away from zero (Jasiak, 1999).

A5: *The elements of τ_0 are almost surely continuous and bounded on Ψ, and $\mathbb{E}(\sup_{\Phi^\star}\max(\ell_0(\chi_0;f_0,\tau_0;\phi),0))<+\infty$.*

Remark 5. Given the stationary and ergodic version of the likelihood function, the assumption implies the applicability of the locally uniform version of Birkhoff's LLN so that the function converges almost surely to its expectation which is well defined. Similarly to the previous remark, in the specification of Table 2, or more generally in models where the (q)-Weibull density is used, it holds whenever the elements of the (τ_i) process are almost surely bounded from above and away from zero, and χ_0 has enough moments; moment orders that approximate from above the essential supremum of the τ_0 suffice.

A1-**A5** imply then pseudo consistency; the estimator converges almost surely to the unique pseudo-true value described in **A2**:

[26]For more complicated filters this may not be the case (Par. 6 of Blasques et al., 2018).

Theorem 1. *Suppose that **A1**-**A5** hold, there exists a* $\gamma > 1$ *such that* $\gamma^T |p^\star - p| \to 0$*, and* $\epsilon_T \to 0$, $\mathbb{P}$ *a.s. Then, (i). the expectation of the Kullback-Liebler divergence between the density in **A1** and* $f(\chi_i; \psi_i, \tau_i; \phi)$ *is well defined, and uniquely minimized at* $\phi^\star$*. (ii).* $\phi_T \to \phi^\star$, $\mathbb{P}$ *a.s. for any* $\hat{\psi}_0$*.*

The existence of a $\gamma > 1$ that validates the condition $\gamma^T |p^\star - p| \to 0$ in the case of the specification of Table 2 is ensured by the meromorphic continuation of the Barnes' Zeta function (Ruijsenaars, 2000) and the fact that that δ is not allowed to lie outside $[0, 1/2)$.

The result along with the identification Assumption **A2** imply that the pseudo-true value $\phi_\star$ has a variational characterization; it is the unique minimum of the expected Kullback-Liebler divergence (Amari, 2016) between the conditional distributions that appear in the statistical moment at hand, and the DGP distribution of χ_i.

When the model $\{f(\chi_i; \psi_i, \tau_i; \phi); \phi \in \Phi\}$, is well-specified, Theorem 1 implies that there exists some $\phi_0 \in \Phi$ such that the density in **A1** is $f(\chi_i; \psi_i, \tau_i; \phi_0)$. Then necessarily $\phi^\star = \phi_0$:

Corollary 1. *If the model is well-specified then* $\phi_T \to \phi_0$, $\mathbb{P}$ *a.s. for any* $\hat{\psi}_0$*.*

Consequently, the empirical duration process assumed (e.g., in Table 2) converges to the true DGP of duration. This is a fundamental block in discussing the convergence properties of iRP because, unlike interval estimators, iRP is derived from the aggregation properties of variables used to construct it, and the conditional intensity is the means by which this becomes feasible. The following result regarding iPIN is easily established via the continuous mapping theorem (CMT) and Corollary 1:

Proposition 1. *Suppose that (i). assumptions **A1**-**A5** hold and* $\epsilon_T \to 0, \mathbb{P}$ *a.s., (ii). each element of* τ_i *is almost surely continuous in* ϕ*, (iii). the cumulative hazard functions employed are continuous functions of the shape and scale parameters, (iv)* $\mathbb{E}(\log^+ \sup_\theta f_0(\theta)) < +\infty$ *and (v). the statistical model is well specified. Then as* $s + 1 \leq T \to \infty$, $|\mathrm{iPIN}_{t,s}(\phi_T, \hat{\psi}_{s+1}) - \mathrm{iPIN}_{t,s}| \to 0$, $\mathbb{P}$ *a.s. conditionally on* $\mathcal{F}_s$*.*

The continuity properties for the conditional shape parameters as well as condition (iv) of the proposition, hold trivially for the specification in Table 2. The compactness of Φ implies a uniform integrability argument that in turn implies that $\mathrm{iPIN}_{t,s}(\phi_T, \hat{\psi}_{s+1})$ converges to $\mathrm{iPIN}_{t,s}$ in the L^2 mode, conditionally on $\mathcal{F}_s$. An analogous result holds for iVPIN due time deformation and the volume clock (e.g., Barndorff-Nielsen and Shiryaev, 2015, Kalaitzoglou and Ibrahim, 2015) and shows that the empirical iRP can be a consistent estimator of the latent probability of informed trading. It follows that parametric statistics, such as t-tests, can be employed for goodness-of-fit and comparability reasons (see also Section C.2 for a discussion of the derivation of confidence sets).[27]

C.2 Rate of Convergence and Weak Gaussian Approximation

The present section derives the rate of convergence and a subsequent Gaussian approximation in distribution for the MLE and subsequently for the empirical iRP. The derivations validate the inference in Section 4.2. The analysis is based on the asymptotics of the first order conditions for the optimization of the log-likelihood function. In a similar manner to the case of consistency, invertibility issues for the SREs that might emerge via differentiation of the filters that appear in the likelihood are addressed by introducing the following assumptions:

B1: *$\phi^\star$ lies in the interior of Φ.*

Remark 6. This enables the w.h.p. use of f.o.c.s. for the analysis. It can be easily dis-

[27] The same is true for the smoothed versions of the conventional estimates of PIN and VPIN for fixed n. In order to illustrate that the intensity-based estimates are better than their interval-based counterparts it would suffice to show that they exhibit a higher convergence rate (the rate of convergence for iRP is discussed in Section C.2). This would be, in principle, the case due to the different scale in the number of observations. Under correct specification and negligible data contamination, and if the interval based estimators have standard rates, the intensity measures would converge more rapidly to the latent probability of informed trading when the competing interval measures utilize sample sizes of smaller order due to the use of large intervals. The intensity based measures could also be better approximations of the latent probability even in cases of misspecification (see section 5.1) for the duration processes, whenever the data used in the interval measures are adequately contaminated. The analysis of these situations and the development of inferential methodologies for comparing the effectiveness of empirical intensity-based versus empirical interval-based measures in approximating the probability of informed trading are deferred to future research. However, although a reasonable comparison between the two would require either a correct specification or a precise measurement of the impact of noise in the estimates, still a Monte Carlo simulation (Section 5.1) can provide meaningful insides on their performance/convergence, even when the models are misspecified.

carded (Andrews, 1999) if-among others-the parameter space can be locally approximated by a convex cone. Even though the assumption essentially precludes the weak dependence case ($\delta_0 = 0$) in the model appearing in Table 2-if well specified, we do not pursue this generalization to avoid clutter.

B2: *There exists an open neighborhood, say $B_{\phi^\star}$, of $\phi^\star$ such that: (i) $\Psi^\star$ is twice continuously differentiable w.r.t. (θ, ψ) on $B_{\phi^\star} \times \mathbb{R}^p$, for almost every value of its remaining arguments. τ_0 is twice continuously differentiable w.r.t. ϕ on $B_{\phi^\star}$ for almost every value of its remaining arguments. (ii). $\Psi^\star_{\partial\theta}$ denotes the SRE obtained by recursive differentiation of $\Psi^\star$ w.r.t. θ, partially via the chain rule through the derivatives of its ψ arguments w.r.t. θ. Then $\mathbb{E}(\sup_{\phi \in B_{\phi^\star}} ||\Psi^\star_{\partial\theta}(\cdot)||) < +\infty$, where $||\cdot||$ denotes the Euclidean norm. (A). $\Psi^\star_{\partial\theta}$ is almost surely Lipschitz continuous in $(\partial\theta\psi_{i-l},\ l = 1, \ldots, \max(p, q))$, with Lipschitz coefficient $\Lambda_i^{(\partial\theta)}(\theta)$, and such that, (B). the map $B_{\phi^\star} \ni \theta \to \Lambda_0^{(\partial\theta)}(\theta)$ is almost surely continuous and, (C). $\mathbb{E}(\sup_{\theta \in B_{\phi^\star}} \log^+ \Lambda_0^{(\partial\theta)}(\theta)) < 0$. (iii). Analogously, $\Psi^\star_{\partial\theta\partial\theta^T}$ denotes the SRE obtained by recursive differentiation of $\Psi^\star_{\partial\theta}$ w.r.t. θ, partially via the chain rule through the derivatives of its ψ and $\partial\theta\psi$ arguments w.r.t. θ. Then $\mathbb{E}(\sup_{\phi \in B_{\phi^\star}} ||\Psi^\star_{\partial\theta\partial\theta^T}(\cdot)||) < +\infty$, where $||\cdot||$ denotes the Frobenius norm. (A). $\Psi^\star_{\partial\theta\partial\theta^T}$ is almost surely Lipschitz continuous in $(\partial\theta\partial\theta^T\psi_{i-l},\ l = 1, \ldots, \max(p, q))$, with Lipschitz coefficient $\Lambda_i^{(\partial\theta\partial\theta^T)}(\theta)$, and such that, (B). the map $B_{\phi^\star} \ni \theta \to \Lambda_0^{(\partial\theta\partial\theta^T)}(\theta)$ is almost surely continuous and, (C). $\mathbb{E}(\sup_{\theta \in B_{\phi^\star}} \log^+ \Lambda_0^{(\partial\theta\partial\theta^T)}(\theta)) < 0$.*

Remark 7. The assumption implies among others the continuous invertibility of the filter $(\hat{(\psi)}_i(\theta))$ derivatives. In the specification of Table 2, it follows whenever $\max(|\beta_0|, |\phi_0|)$ is bounded below 1.

B4: *There exists a stationary non-negative process process $(m_{\partial_i})_{i \in \mathbb{Z}}$, with $\mathbb{E}(\log^+ m_{\partial_0}) < +\infty$, such that almost surely $\sup_{\phi \in B_{\Phi_\star}} ||\partial_\theta \ell_i(\hat{\psi}_i, \partial_\theta \hat{\psi}_i, \tau_i; \phi) - \partial_\theta \ell_i(\psi_{\theta i}, \partial_\theta \psi_{\theta i}, \tau_i; \phi)|| \leq m_{\partial_i} \sup_{\theta \in B(\phi^\star)}(|\hat{\psi}_i - \psi_i| + ||\partial_\theta \hat{\psi}_i - \partial_\theta \psi_i||)$ for any $i \in \mathbb{N}$, and, $\sup_{\phi \in B_{\Phi_\star}} ||\partial_\theta \partial_\theta^T \ell_i(\hat{\psi}_i, \partial_\theta \hat{\psi}_i, \partial_\theta \partial_\theta^T \hat{\psi}_i, \tau_i, \partial_\eta \tau_i; \phi) - \partial_\theta \partial_\theta^T \ell_i(\psi_{\theta i}, \partial_\theta \psi_{\theta i}, \partial_\theta \partial_\theta^T \hat{\psi}_i, \tau_i, \partial_\eta \tau_i; \phi)|| \leq m_{\partial_i} \sup_{\theta \in B(\phi^\star)}(|\hat{\psi}_i - \psi_i| + ||\partial_\theta \hat{\psi}_i - \partial_\theta \psi_i|| + ||\partial_\theta \partial_{\theta^T} \hat{\psi}_i - \partial_\theta \partial_{\theta^T} \psi_i||)$ for any $i \in \mathbb{N}$.*

Remark 8. Again, given the continuous invertibility, the assumption allows the approximation of the score and the Hessian of the likelihood function by a stationary and ergodic version that is constructed via the limiting filtering process and its derivatives. For the specification appearing in Table 2, it holds whenever the elements of the (τ_i) process are almost surely bounded from above and away from zero, x_0 has a logarithmic moment (see also 1), and f_i is uniformly over θ bounded away from zero-the latter holds trivially in the particular example.

B5: $\mathbb{E}(\sup_{\phi\in B_{\phi^\star}} \|\partial_\phi \ell_0)\| + \mathbb{E}(\sup_{\phi\in B_{\phi^\star}} \|\partial_\phi\partial_{\phi^T}\ell_0)\|) < +\infty$, *and for some* $\delta > 0$, $\mathbb{E}(\|\partial_\phi \ell_0(\phi^\star)\|^{1+\delta}) < +\infty$, *for the stationary and ergodic versions of the derivatives. Furthermore, the elements of the stationary and ergodic version of the Hessian are linearly algebraically independent.*

Remark 9. The assumption implies the applicability of the locally uniform version of Birkhoff's LLN on the stationary and ergodic version of the Hessian, the identification of the limiting focs via dominated convergence, and-in conjunction with **B1** and **B3** the applicability of the aforementioned CLT. Similarly to the previous remark, in the specification of Table 2, it holds whenever the elements of the (τ_i) process are almost surely bounded from above and away from zero, and x_0 has enough moments; orders that approximate from above the essential supremum of the τ_0 on the power of $1+\delta$ suffice.

Utilizing the aforementioned assumptions while regulating the speed at which the optimization error reduces to zero, we arrive at the subsequent result.

Theorem 2. *Under the premises of Theorem 1, and if moreover **B1-B5** hold and* $\sqrt{T}\epsilon_T \rightsquigarrow 0$, *then*

$$\sqrt{T}(\phi_T - \phi^\star) \rightsquigarrow N(\mathbf{0}, (\partial_\phi\partial_{\phi^T}\ell_0(\phi^\star))^{-1}(\partial_\phi\ell_0(\phi^\star)\partial_\phi\ell_0^T(\phi^\star))(\partial_\phi\partial_{\phi^T}\ell_0(\phi^\star))^{-1}),$$

for the stationary and ergodic versions of the associated derivatives.

The existence of γ, that ensures the exponentially fast approximation of the part of the truncated filter (that depends on the derivatives w.r.t. δ of the ARCH(∞) component) is ensured by the meromorphic continuation of the Barnes' Zeta function (Ruijsenaars, 2000), the fact that δ is not allowed to lie outside a compact subset of $(0, 1/2)$, and the asymptotic representation of the series' coefficients as $Cj^{1-\delta}$ for some $C > 0$ independent of j, δ.

The limit theory produces a standard rate and asymptotic normality with the usual sandwich form for the asymptotic variance. Consistent estimators of the terms that appear in there can be easily obtained via the non-stationary versions of the derivatives evaluated at the MLE, due to consistency and Assumptions **B1**, **B3-B5**. This can be useful for the construction of Wald-type tests for parameters of interest. When the statistical model is well-specified, then due to **B1-B5** and dominated convergence, the information matrix equality yields-as expected:

Corollary 2. *Under the premises of Theorem 2 and if the statistical model is well specified, then:*

$$\sqrt{T}(\phi_T - \phi_0) \rightsquigarrow N(\mathbf{0}, (\partial_\phi \ell_0(\phi_0) \partial_\phi \ell_0^T(\phi_0))^{-1}).$$

The above is then used in order to derive, using the Delta method, the limit theory of iPIN. In what follows d_W denotes any metric that metrizes weak convergence-see for example Par. 1.12 of Vaart and Wellner (2023):[28]

Proposition 2. *Under the premises of Theorem 2 and if the cumulative hazard functions employed are continuous functions of the shape and scale parameters, and the statistical model is well specified, then as* $s + 1 \leq T \to \infty$, $d_W(\sqrt{T}(\mathrm{iPIN}_{t,s}(\phi_T, \hat{\psi}_{s+1}(\phi_T)) - \mathrm{iPIN}_{t,s}), y) \to 0$, $\mathbb{P}$ *a.s. conditionally on* $\mathcal{F}_s$, *where* $y \sim N(0, \partial_\phi \mathrm{iPIN}(\phi_0)^T (\partial_\phi \ell_0(\phi_0) \partial_\phi \ell_0^T(\phi_0))^{-1} \partial_\phi \mathrm{iPIN}_{t,s}(\phi_0))$.

The latter can be used as the basis for the construction of confidence sets. An analogous derivation obviously holds for the estimated iVPIN. The construction of confidence sets for the plug-in estimators of iPIN and iVPIN can be also performed via Monte Carlo methods, due to the parametric nature of the statistical model at hand-when this is well specified.

[28]The notation used here is abusive; d_W is actually evaluated at the relevant distributions and not on the random elements that follow them.

Specifically, the limiting (unconditional) variance of the estimators, can be consistently estimated via the MC variance of the resulting iPIN (or iVPIN) estimator, when the DGP is evaluated at the originally estimated parameters. This is due to the CMT, the locally (w.r.t. ϕ) uniform convergencies mentioned above, and the consistency of MLE. A similar (yet possibly less accurate) approximation can be available from the cross sectional averaging of the iPIN (or iVPIN) estimators, when the cross sectional DGPs are similar, and satisfy some form of exchangeability property, or more generally invariance of the underlying joint distributions under groups of transformations (Austern and Orbanz, 2022).

C.3 Proofs

Proof of Theorem 1. The condition $\gamma^T|p^\star - p| \to 0$ ensures the exponentially fast almost sure approximation of the truncated at $p^\star$ filter $\Psi^\star(\theta; \psi_{i-l}, x_{i-j}, l = 1, \ldots, \max(p,q),\ j = 1, \ldots, p)$ by $\Psi^\star(\theta; \psi_{i-l}, x_{i-j}, l = 1, \ldots, \max(p,q),\ j = 1, \ldots, p)$ uniformly over Θ. The proof is then identical to the proof of Theorem 4.3 of Blasques et al. (2018). □

Proof of Proposition 1. Follows from the CMT, Corollary 1, and Proposition 3.2 of Blasques et al. (2018). □

Proof of Theorem 2. As in Theorem 1, the condition $\gamma^T|p^\star - p| \to 0$ ensures the exponentially fast almost sure approximation of the first and second derivatives of the truncated at $p^\star$ filter $\Psi^\star(\theta; \psi_{i-l}, x_{i-j}, l = 1, \ldots, \max(p,q),\ j = 1, \ldots, p)$ by the analogous derivatives of $\Psi^\star(\theta; \psi_{i-l}, x_{i-j}, l = 1, \ldots, \max(p,q),\ j = 1, \ldots, p)$ locally uniformly over Θ. Notice that irrespective of the pseudo-true value of q, the first derivative of the likelihood contributions evaluated at $\phi^\star$ conditionally on the filtration, lies in the normal domain of attraction of a zero mean Gaussian distribution (see Theorem 2.6.5 in Ibragimov and Linnik (1971)). Then stationarity and ergodicity and second order integrability of the limiting filter of the first derivatives as well as the almost sure boundedness of the derivatives of the remaining processes along with the principle of conditioning (see Jakubowski, 2012), implies $O_p(\sqrt{T})$

asymptotic tightness and limiting zero mean Gaussianity for the score. The locally uniform version of the ergodic theorem takes care the relevant a.s. convergence of the Hessian, and the result follows from a Mean Value expansion of the f.o.c.s. of the optimization problem that defines the estimator, which hold w.h.p. due to **B1**. □

C.4 Convergence of Alternative Agent-Types

The CMT and the Delta method, along with the limit theory derivations for the estimated parameters of interest, directly imply analogous results to Propositions 1, and 2 for the iTT and iPTT. Those are reported below for completeness of exposition:

Proposition 3 (C.2.1)**.** *Suppose that (i). assumptions* ***A1-A5*** *hold and* $\epsilon_T \to 0, \mathbb{P}$ *a.s., (ii). each element of* τ_i *is almost surely continuous in* ϕ*, (iii). the cumulative hazard functions employed are continuous functions of the shape and scale parameters, (iv)* $\mathbb{E}(\log^+ \sup_\theta f_0(\theta)) < +\infty$ *and (v). the statistical model is well specified. Then as* $s+1 \leq T \to \infty$*,* $|\mathrm{iTT}_{t,s}(\phi_T, \hat{\psi}_{s+1}) - \mathrm{iTT}_{t,s}(\phi_0)| + |\mathrm{iPTT}_{t,s}(\phi_T, \hat{\psi}_{s+1}) - \mathrm{iPTT}_{t,s}(\phi_0)| \to 0$*,* $\mathbb{P}$ *a.s. conditionally on* $\mathcal{F}_s$*.*

Proposition 4 (C.2.2)**.** *Under the premises of Theorem 2 and if the cumulative hazard functions employed are continuous functions of the shape and scale parameters, and the statistical model is well specified, then as* $s+1 \leq T \to \infty$*, we have for the* iTT *case that* $d_W(\sqrt{T}(\mathrm{iTT}_{t,s}(\phi_T, \hat{\psi}_{s+1}(\phi_T)) - \mathrm{iTT}_{t,s}(\phi_0)), N(0, \partial_\phi \mathrm{iTT}(\phi_0)^T (\partial_\phi \ell_0(\phi_0) \partial_\phi \ell_0^T(\phi_0))^{-1} \partial_\phi \mathrm{iTT}_{t,s}(\phi_0))) \to 0$*,* $\mathbb{P}$ *a.s. conditionally on* $\mathcal{F}_s$*, and furthermore likewise for the* iPTT *case we have that,* $d_W(\sqrt{T}(\mathrm{iPTT}_{t,s}(\phi_T, \hat{\psi}_{s+1}(\phi_T)) - \mathrm{iPTT}_{t,s}(\phi_0)), y^\star) \to 0$*,* $\mathbb{P}$ *a.s. conditionally on* $\mathcal{F}_s$*, where* $y^\star \sim N(0, \partial_\phi \mathrm{iPTT}(\phi_0)^T (\partial_\phi \ell_0(\phi_0) \partial_\phi \ell_0^T(\phi_0))^{-1} \partial_\phi \mathrm{iPTT}_{t,s}(\phi_0))$*.*

The arguments that lead to the proofs are analogous to the ones of Propositions 1, and 2 respectively. The continuity of the hazard functions assumed holds for the specification used above. The first proposition implies consistency for both the iTT and the iPTT under correct specification. The second implies standard rates and asymptotic normality that could be useful for statistical inference.

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

<u>2024</u>

- **01-24** Market Timing & Predictive Complexity, Stelios Arvanitis, Foteini Kyriazi, Dimitrios Thomakos
- **02-24** Multi-Objective Frequentistic Model Averaging with an Application to Economic Growth, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, Nikolas Topaloglou
- **03-24** State dependent fiscal multipliers in a Small Open Economy, Xiaoshan Chen, Jilei Huang, Petros Varthalitis
- **04-24** Public debt consolidation: Aggregate and distributional implications in a small open economy of the Euro Area, Eleftherios-Theodoros Roumpanis
- **05-24** Intangible investment during the Global Financial Crisis in the EU, Vassiliki Dimakopoulou, Stelios Sakkas and Petros Varthalitis
- **06-24** Time will tell! Towards the construction of instantaneous indicators of different agent-types, Iordanis Kalaitzoglou, Stelios Arvanitis
- **07-24** Norm Constrained Empirical Portfolio Optimization with Stochastic Dominance: Robust Optimization Non-Asymptotics, Stelios Arvanitis
- **08-24** Asymptotics of a QLR-type test for optimal predictive ability, Stelios Arvanitis
- **09-24** Strongly Equitable General Equilibrium Allocations, Angelos Angelopoulos
- **10-24** The Greek macroeconomy: A note on the current situation and future outlook, Apostolis Philippopoulos
- **11-24** Evolution of Greek Tax System, A Survey of Legislated Tax Changes from 1974 to 2018, Panagiotis Asimakopoulos
- **12-24** Macroeconomic Impact of Tax Changes, The case of Greece from 1974 to 2018, Panagiotis Asimakopoulos
- **13-24** \`Pareto, Edgeworth, Walras, Shapley' Equivalence in a Small Economy, Angelos Angelopoulos
- **14-24** Stimulating long-term growth and welfare in the U.S, James Malley and Apostolis Philippopoulos
- **15-24** Distributionally Conservative Stochastic Dominance via Subsampling, Stelios Arvanitis
- **16-24** Before and After the Political Transition of 1974. Institutions, Politics, and the Economy of Post-War Greece, George Alogoskoufis
- **17-24** Gaussian Stochastic Volatility, Misspecified Volatility Filters and Indirect Inference Estimation, Stelios Arvanitis, Antonis Demos
- **18-24** Endogenous Realtor Intermediation and Housing Market Liquidity, Miroslav Gabrovski, Ioannis Kospentaris, Victor Ortego-Marti
- **19-24** Universal Choice Spaces and Expected Utility: A Banach-type Functorial Fixed Point, Stelios Arvanitis, Pantelis Argyropoulos, Spyros Vassilakis

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

2025

01-25 Democracy, redistribution, and economic growth: Some evidence from post-1974 Greece, George C. Bitros
02-25 Assessing Downside Public Debt Risks in an Environment of Negative Interest Rates Growth Differentials, Yiannis Dendramis, Georgios Dimitrakopoulos and Elias Tzavalis
03-25 Wicksellian General Equilibrium, Angelos Angelopoulos
04-25 Institutions and Ethics, Angelos Angelopoulos
05-25 General Equilibrium in a Capitalist Economy, Angelos Angelopoulos
06-25 Public investment multipliers revisited: The role of production complementarities, Vasiliki Dimakopoulou, George Economides, and Apostolis Philippopoulos
07-25 The heterogeneous causal effects of the EU's Cohesion Fund, Angelos Alexopoulos, Ilias Kostarakos, Christos Mylonakis, and Petros Varthalitis
08-25 Gaussian Invariant Markov Chain Monte Carlo, Michalis K. Titsias, Angelos Alexopoulos , Siran Liu, and Petros Dellaportas
09-25 Financial Intermediation, Investment, and Monetary Policy, Ioannis Kospentaris and Florian Madison
10-25 Temperature as a premium factor. Insurance incorporating climate change. A comparative case study between Europe and North America, Nomikou-Lazarou Irene
11-25 How High will our Buildings Become? Angelos Angelopoulos
12-25 Emergency Room General Equilibrium, Angelos Angelopoulos
13-25 Updating Density Estimates using Conditional Information Projection: Stock Index Returns and Stochastic Dominance Relations, Stelios Arvanitis, Richard McGee, and Thierry Post
14-25 Measuring Textual Cohesion via Graph Curvature, Weighted Dominance, and Embedding-Based Coherence, Stelios Arvanitis
15-25 Universal Preference Spaces for Expected Utility via a Banach-type Functorial Fixed Point Stelios Arvanitis, Pantelis Argyropoulos, and Spyros Vassilakis
16-25 Planet Earth's General Equilibrium, Angelos Angelopoulos
17-25 Electable and Stable Insiders' Coalition Governments, Tryphon Kollintzas and Lambros Pechlivanos
18-25 Long Memory from Cheeger Bottlenecks and Long Cycles in Network Dynamics, Stelios Arvanitis
19-25 Robust Multiobjective Model Averaging in Predictive Regressions, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, and Nikolas Topaloglou
20-25 On robust Bayesian causal inference, Angelos Alexopoulos and Nikolaos Demiris
21-25 Leverage Effect, Volatility Feedback and the Influence of Jumps, Orestis Agapitos, Ioannis Papantonis, Leonidas Rompolis, and Elias Tzavalis
22-25 Leadership: The other deficit of post-1974 Greece, George C. Bitros

2026

01-26 The Tell-Tale Clock! Speed and Agent Composition in High Frequency Trading, Iordanis Kalaitzoglou and Stelios Arvanitis

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5