Who Knows? Information Differences Between Trader Types

Albert J. Menkveld[*] Ion Lucas Saru[†]

August 25, 2023

Abstract

We study informativeness of agent and principal trades. Order informativeness depends on the horizon and frequency we analyze. In line with the literature on high-frequency trading, principals are more informed than agents at the highest frequency, as measured by the contribution of the respective order flow series to the variance of efficient price innovations. Once we move to lower frequencies, price discovery is dominated by agents, while the share of principals goes to zero. This is reflected in gross trading revenues of agents and principals at different frequencies. Our results hold across market conditions as measured by the VIX.

1 Introduction

The consensus results of the #fincap project show a simultaneous decline of market efficiency and the share of agency volume in overall trading volume over the past decades (Menkveld et al., forthcoming). We address the question to which extent these findings are connected by analyzing the trading behavior of different types of traders – agent and principal traders – in terms of their contribution to price discovery and price impact at different horizons. Thus we are asking the questions: What are the gross trading revenues that both principals and agents realize? How does this depend on the horizon? We link trading revenues to the informativeness of order flow. That is, we analyze which orders move prices permanently and which orders contribute to pricing errors and ask, whether a trader's orders being informed at a high frequency translates into order informativeness at a lower frequency.

Our analysis is based on nine years of trading data for Euro STOXX 50 futures, one of the most actively traded futures contracts. Today's futures markets are leading in terms of price discovery over the underlying (see, for

---

[*] Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands, albertjmenkveld@gmail.com, +31 20 598 6130, and Tinbergen Institute.

[†] Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands, saru.ionlucas@gmail.com, +31 20 598 6060, and Tinbergen Institute.

example, Hasbrouck (2003)). Thus, futures markets are not merely a market for hedging, but more than that, they are a fundamental part of today's market architecture. The prices that emerge as a result of trading in these markets are relevant for a wide range of market participants and capital allocation decisions.

Also, there is a wide interest in identifying informed traders. For academics this is relevant to understand price discovery and trading behavior of informed traders, and for practitioners for optimal liquidity provision and protection of their quotes from adverse selection. The academic literature focusing on informed trading usually relies on an endogenous classification of who is informed, or focuses on a subset of arguably informed traders (Collin-Dufresne and Fos, 2015; Kacperczyk and Pagnotta, 2019). However, who is informed is likely endogenous to the market conditions and the horizon that is analyzed. We propose to link the endogenous informativeness of an order with exogenous information on whether the trade pertains to an agent or a principal. As informedness of orders may depend on the horizon, we ask the questions: How does economy-wide information get incorporated into prices? Are traders that appear informed at a high frequency also informed at a lower frequency? How does information incorporation into prices depend on market conditions?

European equity derivatives markets were at the center of margin call events 2020 during the Covid-pandemic (ESRB, 2020). In light of this, it is not only relevant to analyze which traders are active in these markets, but more than that, how they contribute to the efficiency of these markets. We address this point by analyzing differences in price discovery between agent and principal traders at different horizons in Euro STOXX 50 futures.

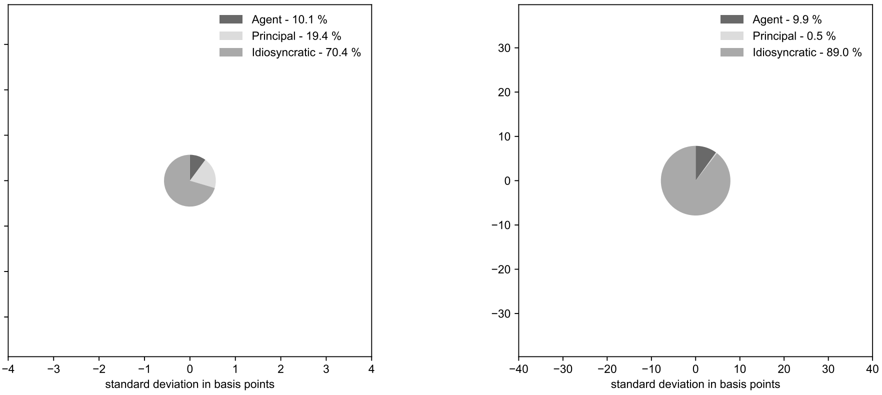
We propose a state space model to estimate the contribution of orders by different traders to price discovery. Our model controls for differences in trading volume. Thus, even though trading volume between agents and principals differs, we estimate the information content in their order flow scaled by volume. The state space framework allows us to disentangle transitory and permanent movements in security prices. Based on this, we identify whether price changes are due to information or price pressure. Based on this, order informativeness can be expressed as the contribution to efficient price variation. Even if agents and principals use the same order type, the order's contribution to price discovery differs.

We exhibit our main results in Figure 1. At a 1-second frequency, agent's aggressive order account for 10.1% of the variance of innovations in efficient prices, while principal's aggressive orders account for 19.4%. Once we move to a lower frequency, the share of agent's order increases, while the share of principal's order flow drops to virtually zero. Our findings indicate that agents contribute to price discovery at lower frequencies, while principal's contribution to price discovery is limited to the highest frequencies. This is consistent with the model of Foucault, Hombert, and Roşu, 2016 where high-frequency traders' orders are more correlated with short-term price changes.

We provide evidence that our results are indeed driven by an information channel by comparing the relative frequency of how often agents and principals trade on the right side of the market on macroeconomic news announcement

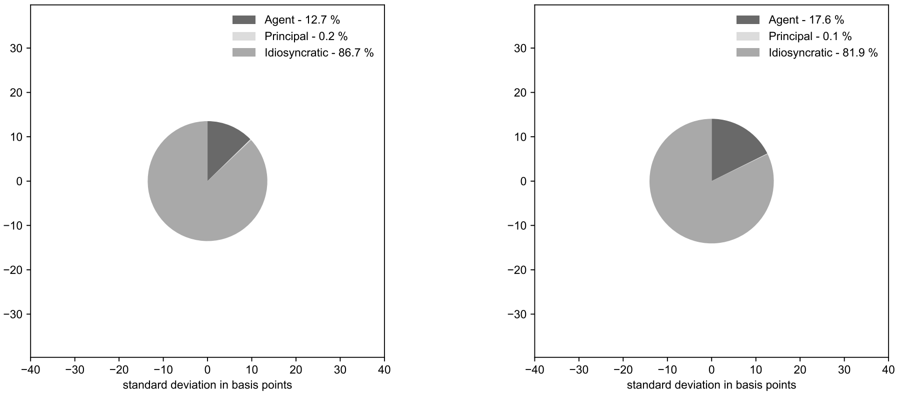
Figure 1: Variances of efficient price innovations at different frequencies

The figure plots variances of efficient price innovations as well as the share of agents' and principals' aggressive orders in these variances for different frequencies. The radius corresponds to the standard deviation of efficient price innovations in *bp*. Note the different scales for second frequency as well as daily and weekly frequencies.



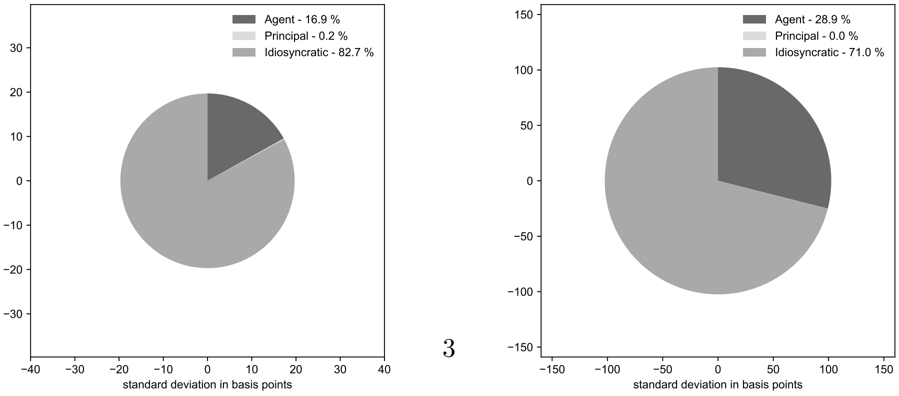
(a) 1 second

(b) 5 minutes



(c) 15 minutes

(d) 30 minutes



(e) 1 hour

(f) 1 day

days. We reject the null hypothesis of no differences between agents and principals for news announcement days but not for roll days (low information days).

Order informativeness at different frequencies translates into gross trading revenues. Using a frequency domain decomposition of trading profits we show that principals record positive trading profits for the highest frequencies – the same frequencies for which they appear informed based on our state space results. Agents, on the other hand, make positive trading profits at lower frequencies. Again, this corresponds to the frequencies at which they are more informed then principals.

Our analysis focuses on trading in Euro STOXX 50 index futures, capturing the 50 largest companies in the Eurozone from eight countries. We use data on futures trading as futures are leading other instruments in terms of price discovery (Kawaller, Koch, and Koch, 1987; Stoll and Whaley, 1990; Hasbrouck, 2003; Tse, Bandyopadhyay, and Shen, 2006). Since we are asking the question how different traders contribute to price discovery, we analyze the instrument that is leading in price discovery.[1]

High-frequency traders are a subset of the principal traders in our sample. We show that our results for the highest frequencies closely replicate the results of Brogaard, Hendershott, and Riordan (2014). We interpret this as reassuring evidence that our results do not only apply to derivatives markets exclusively, but are also aligned with the existing literature on equity markets.

Brogaard, Hendershott, and Riordan (2014) and Hasbrouck (2021) analyze price discovery at the highest frequencies. We complement this literature by analyzing how price discovery and order informativeness at high frequencies relates to informativeness at lower frequencies. We find that the classification of who is more informed depends on the frequency, with principals being more informed at high frequencies and agents being more informed at lower frequencies.

In contrast to existing studies on the use of order types by informed traders that focus on traders with stock-specific information (Collin-Dufresne and Fos, 2015; Kacperczyk and Pagnotta, 2019), we focus on information regarding a basket of stocks. It is not clear that existing results extend to instruments that reflect a basket of multiple stocks. Prices of futures, such as the Euro STOXX 50 futures, reflect next to information on the indexes constituents also information on the state of the "European" economy.[2] We address the question how traders with this macro-type of information trade upon their information and provide liquidity using different orders. We find that both information on the trader type as well as the order used matters for the expected informational content of the order, but depends on the horizon. Thus, with our work we contribute to the question of price efficiency and information incorporation into prices.

---

[1] When analyzing an instrument that is lagging in price discovery, traders might incorporate information into prices that has been revealed in other instruments before. For this, however, the pure speed of the trader matters and the analysis might confound different effects.

[2] Traders with stock-specific information can achieve a greater exposure to a company by trading its shares or derivatives on its shares rather that trading futures on the index that contains the share.

Our work relates to a vast literature studying information content of different order types as well as order submission by informed traders. In recent work, Li, Ye, and Zheng (2021) study different order types at the NYSE and evaluate their performance, also in terms of price discovery. However, they do not observe the traders submitting the orders. It is possible that the same trader uses different order types, depending on the market circumstances and the informational horizon. While we do not observe the exact order type, we observe whether a trade record stems from an aggressive or a passive order. However, we have information on the type of trader utilizing the order. Thus, observing how orders perform at different horizons depending on which type of trader uses them allows us to infer valuable information on the trader's information. We find a pecking order of order types depending on market volatility. Our findings reveal that aggressive orders are relatively more informative in high-volatility periods and relatively less informed in low-volatility periods.

We address the question which orders are informed. The early literature assumes that informed traders use market order (Harris, 1998). Recent evidence shows that informed investors use limit orders. For example, Collin-Dufresne and Fos (2015) find that 13D activist traders use limit orders. Parlour (1998) and Foucault (1999) study limit and market order submission by uninformed traders. Hollifield, Miller, and Sandås (2004) theoretically and and empirically analyze optimal limit order submissions. In early empirical work, Biais, Hillion, and Spatt (1995) study limit order book dynamics and interactions between market and limit orders.[3] Bloomfield, O'Hara, and Saar (2005), Baruch, Panayides, and Venkataraman (2017) and Kacperczyk and Pagnotta (2019) find that informed traders use limit orders.[4] Other work focuses on order choice by informed and uninformed traders, respectively. Bloomfield, O'Hara, and Saar (2015) find that informed trader's order choice depends on whether they can hide liquidity and that they use more limit orders if liquidity not visible. Kaniel and Liu (2006) present a theoretical model in which informed trades can submit limit orders. In equilibrium, limit orders may contain more information than market orders, dependent on the horizon of the information. Our research analyzes how aggressive and passive orders by agents and principals differ in their informativeness dependent on the frequency we are analyzing. Overall, we find more market orders to be more informed. At the same time, our evidence suggests that agents' limit orders contain information at longer horizons and in times of low market volatility.

Several studies that analyze the trading behavior of groups of traders focus only on subgroups. Kelley and Tetlock (2013) study the usage of limit and market orders by retail investors. In contrast to us, they analyze trading and return patterns at solely a lower frequency and relate daily imbalances

---

[3] Goettler, Parlour, and Rajan (2009) develop a model of the choice of acquiring information and choosing the order type and Hoffmann (2014) studies order choice in a market with HFTs.

[4] Brogaard, Hendershott, and Riordan (2019) document that HFTs' limit orders contribute to price discovery, and Fleming, Mizrach, and Nguyen (2018) provide evidence consistent with price discovery through limit orders.

to monthly returns. They find that retail investors using market orders trade on new information. Also limit orders by retail investors provide liquidity and some limit orders may be informed. Hendershott, Livdan, and Schürhoff (2015) study price discovery of institutional investors and find that they contribute to price discovery regarding news events. However, they do not distinguish between order types. In a recent study, Beason and Wahal (2020) study institutional trading algorithms and find that they mainly use limit orders. Hence, to the extent that institutional investors are informed, they are incorporating their information using limit orders. Biais, Declerck, and Moinas (2016) study how agents and principals provide liquidity through their orders and to which extent their orders are subject to adverse selection. Anand, Chakravarty, and Martell (2005) analyze the intraday pattern of orders usage by liquidity traders and institutional traders. They find that institutional traders use market orders early during the day and limit orders later during the day. Our analysis incorporates information on different trader types (agents and principals) as well as information on the orders used. This allows us to analyze differences in the information content of orders scaled by volume.

Barber et al. (2009) study the Taiwanese market during the period 1995 – 1999. They distinguish between what they call aggressive and non-aggressive orders as the market only allowed for limit orders. They find that individual traders make losses through "aggressive orders" while aggressive and non-aggressive orders of institutions are profitable. We show that the profits or orders depend on the frequency at which we analyze gross trading revenues. Differences in trading revenues at different frequencies translate into differences in order informativeness based on the state space model.

Our results further relate to Menkveld, Sarkar, and Van der Wel (2012), Czech et al., 2021 and Jurkatis et al., 2022 who study the performance of orders of agents' and principals' orders in government and corporate bond markets. Our results complement their results by showing that agent's orders appear informed at lower frequencies in equity derivatives markets.

The remainder of the paper is structured as follows. In Section 2 we provide institutional details on Euro STOXX 50 futures trading before we describe the data in Section 3. Section 4 discusses our state space methodology. We present our main results on order informativeness at different frequencies in Section 5. Finally, Section 6 concludes.

2 Institutional Background

Our analysis focuses on Euro STOXX 50 futures, one of the most actively traded futures contracts in the world. The Euro STOXX 50 is the index for the largest companies of the Eurozone with the constituents registered in Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands, and Spain.[5] The constituents are not all listed at the same exchange, however the trading hours are aligned. Trading takes place in an electronic limit order book.

---

[5] `https://www.stoxx.com/index-details?symbol=SX5E`

In our sample period, trading in Euro STOXX 50 futures takes place from 8 in the morning to 10 in the evening, Central European Time. Trading starts with an opening auction in the morning. After continuous trading terminates in the evening, it is followed by a call phase of at least three minutes, before a closing auction is held.

Account roles at Eurex distinguish between principal accounts (exchange members) and agency accounts (non-exchange members).[6] Exchange members include internationally operating (dealer) banks as well a nationally operating European banks, proprietary trading firms, high frequency trading firms, global asset management companies, and other institutional investors. Non-exchange members are by definition the remaining traders and include other institutional investors and pension funds. Exchange member accounts distinguish again between proprietary trading accounts and market maker accounts. From our conversations with Eurex we understand that this distinction is not economically meaningful and solely relevant for the calculation of fees and rebates. Inspection of the data shows, for example, a substantial use of market orders and marketable limit orders among market maker accounts. Thus, in the main analysis, we group both proprietary and market maker accounts together and label them as principals.

It is important to note that all trades that pertain to an agent are recorded to an agency account, even if the trade was internalized by a principal. Also, trades that are executed using a principal's agency algorithm on behalf of an agent are labelled as agency trades. Thus, exchange members have no discretion whether they record a trade to an agency or a principal account. According to Eurex, there is little migration of agents to principal accounts and retail order flow in agency order flow is negligible.

3 Data

Our focus is on studying the information in order flow and liquidity provision by different trader types. We use trading data on Euro STOXX 50 futures from Eurex. The sample period spans from January 4, 2010 to December 7, 2018.[7] The trading data contains both legs of a trade. We do not observe quotes or trader IDs. We focus on futures trading as futures are leading in terms of price discovery (see Hasbrouck (2003), among others).[8] Central to our analysis is the

---

[6] A list of exchange members can be found on the Eurex website. We provide a list of all unique exchange members listed on Eurex' website in Appendix A.

[7] We restrict our overall sample that end on December 31, 2018, to end on December 7, 2018 in order to avoid any confounding effects with a trading hour extension in the morning on December 8, 2018.

[8] An alternative is analyzing patterns in price discovery based on ETF trading data. Hasbrouck (2003) finds that ETFs contribute significantly to price discovery as well, with the ETFs' contribution varying between instruments. Furthermore, Menkveld and Yueshen (2019) document a cross-market non-arbitrage relationship between S&P 500 futures and ETFs (E-mini and SPY). Moreover, the futures market is less fragmented than the ETF market and very liquid.

information on the account role of a trader as well as whether the trader utilizes an aggressive order or not.[9] The data contains the following information:

- Expiration of the futures contract.
- Indicator whether the trade is a buy or sell.
- Trade size.
- Execution price.
- Aggressor flag, whether the trade stems to a market or limit order. In the following, we label trades that pertain to a market or marketable limit order as *aggressive* trades and trades that stem to a limit order as *passive* trades.
- The account role of the trader, whether the trader is a principal or an agent.
- Indicator whether the order was fully or partially executed in the trade.

The analysis is based on trading in the contract with the highest trading volume. This is usually the nearest-to-maturity contract.[10] Also, we focus on continuous trading as only during continuous trading there is a limit order for every market order. Our focus is on informational differences between traders at different frequencies. We thus distinguish between their aggressive and passive orders. This distinction is only meaningful during continuous trading. Furthermore, this assures that there are no confound effects with dynamics during the opening, closing, and intraday auctions (Bogousslavsky and Muravyev, 2021; Comerton-Forde and Rindi, 2021).

In the next section we describe details on the data cleaning procedure before we provide summary statistics for the data.

3.1 Data Cleaning

We identify continuous trading by requiring that for every timestamp, the volume of market orders equals the volume of limit orders. This classification is performed based on the aggressor flag: the volume of aggressive orders has to equal the volume of passive orders. This is feasible for the period January 1, 2010 – May 7, 2013. On May 8, 2013, Eurex migrated its products to the T7 trading system. This causes some imprecision in the timestamps in the data of the following form. Orders that were executed against each other are not necessarily recorded at the same timestamp but at consecutive timestamps where the

---

[9] We do not observe the exact order used. Orders at Eurex include next to market and limit orders also stop orders, orders for the closing auction, as well as one-cancels-other and book-or-cancel restrictions for limit orders (Eurex, 2021).

[10] Similarly, Huang (2018) uses the contract with the highest volume. Our approach yields similar results to using the front contract and rolling over to the next contract a pre-specified number of days before expiration (Andersen et al., 2007).

difference between the records is usually within the range of a few tens of milliseconds.[11] Thus, requiring the volume of market orders to equal the volume of limit orders at every timestamp to identify continuous trading is not feasible.[12]

We address this problem using an event time approach.[13] Trades that are recorded at the same price within a defined time-interval are grouped together. Then, the total volume of limit orders and the total volume of market orders over the grouped trades are computed. If the total volume of limit orders equals the total volume of market orders, the trades are labeled as continuous trading and included in the main analysis. This approach allows to effectively filter for auctions in the trade data.

The algorithm starts at the beginning of each trading day. For each price-timestamp combination, a time window starting with that trade record is initialized.[14] All following trades that are executed within the time window at the same trading price are grouped together and assigned the timestamp of the first recorded trade in that group. Once a trade is executed at the same trading price but does not fall within the time window, or at a different trading price, a new time window starting from that trade record is initialized. Again, all trades that occur within the time window at the same execution price are grouped together. This procedure continues until the end of the trading day.

The only parameter that has to be chosen is the length of the time window. Choosing the window length trades off two factors. On the one hand, choosing a longer window length assures that all corresponding trades are grouped together even if there is substantial imprecision in the timestamps and high trading activity. On the other hand, if the window length is chosen too long and the volume of market orders and limit orders is *not* equal, substantial volume is excluded form the main analysis. We consider the possibilities of 100ms and 500ms (as well as 2s and 4s for robustness checks). In the main analysis, we focus on data that has been cleaned using a window length of 100ms. Our findings are robust to using a different window length.

3.2 Variable Construction

We run our analysis at different frequencies. To each time interval, the last observed trading price within that time interval is assigned. For order flow, we sign volume using the trade direction indicator from the data. Then, we sum signed volume for every time interval. Thus, order flow has the same frequency as the prices that we observe. We confirm that signed volume is a stationary series.

---

[11] Also, it appears that this "noise" in the timestamps increases in times of high trading activity.

[12] Also, market clearing does not hold for every timestamp as a results of the imprecision in timestamps.

[13] Here we outline the main considerations and move a detailed discussion in Appendix 3.1. We discuss the advantages of an event-time approach over a wall-clock time approach in which all trades within a unit of time are grouped together.

[14] This approach is comparable to the methodology developed in Aquilina, Budish, and O'Neill (2021), Ernst (2020), and Ernst, Sokobin, and Spatt (2021).

We are focusing on differences in price discovery and liquidity provision of the same order type between different trader types and frequencies. In terms of order characteristics, we distinguish between aggressive (market and marketable limit) orders and passive (limit) orders based on the aggressor flag. This is in line with the classification in Brogaard, Hendershott, and Riordan (2014) who distinguish between liquidity demanding and liquidity supplying orders. In terms of trader types, we distinguish between principal and agent traders based on the account role information in the data set. Trades must be identified as principal or agent trades with the distinction not being arbitrary as described in Section 2.[15]

We use the information on the aggressor flag and account role to create order flow variables for every aggressor flag-account role combination. This yields several account role-aggressor flag combinations:

1. Agent flow;
2. Principal flow;
3. Aggressive agent flow;
4. Passive agent flow;
5. Aggressive principal flow;
6. Passive principal flow;
7. Aggressive flow;
8. Passive flow.

1 and 2, 3 – 6, as well as 7 and 8 clear the market, thus, in the empirical analysis at least one of the respective account role-aggressor flag combinations has to be omitted from the model for estimation.

4 Methodology

The center of our analysis is to study differences in the information content in order flow and liquidity provision by different traders using a state space model. We also analyze trading profits at different frequencies in the frequency domain, which we discuss in more detail in Section 5.1 and Appendix E.

---

[15] For accounts classified by Eurex as agency accounts (account code "A"), we retain the label. We label accounts classified by Eurex as proprietary (account code "P") and market maker (account code "M") as principal accounts.

4.1 State Space Model

We build on Hasbrouck's (1993) approach as well as on the state space framework developed in Menkveld, Koopman, and Lucas (2007). The framework decomposes observed security prices into two latent components: an efficient price component as well as a deviation from the efficient price, the pricing error. Our approach allows us to decompose observed prices into both components and to get estimates of both the efficient price series as well as the size of the pricing error.

Following Hasbrouck (1993) and Campbell et al. (1998), efficient prices are modeled to follow a martingale and observed prices are the sum of the efficient price and the pricing error

$$p_t = m_t + s_t \tag{1}$$

$$m_t = m_{t-1} + w_t \tag{2}$$

with p_t denoting log prices, $s_t \sim \mathcal{N}(0, \sigma_s^2)$ being the pricing error, m_t the efficient price, and $w_t \sim \mathcal{N}(0, \sigma_w^2)$ innovations in the efficient price. Identification of pricing errors in this standard model relies on either the assumption of independent pricing errors and innovation in the efficient price series or fixing the correlation between pricing errors and innovations in efficient prices to a specific value (George and Hwang, 2001; Menkveld, Koopman, and Lucas, 2007). Imbalances in order flow help explaining pricing errors and innovations in order flow contain information (Brandt and Kavajecz, 2004; Pasquariello and Vega, 2007; Evans and Lyons, 2008; Hendershott and Menkveld, 2014). We thus use information on signed order flow to identify pricing errors and innovations in efficient prices. It follows the full state space model

$$p_t = m_t + s_t \tag{3}$$

$$m_t = m_{t-1} + \sum_{s \in \mathcal{S}} \gamma_s \tilde{x}_{s,t} + w_t \tag{4}$$

$$s_t = \phi s_{t-1} + \sum_{s \in \mathcal{S}} \delta_s x_{s,t} + \varepsilon_t \tag{5}$$

with $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. $x_{s,t}$ denotes order flow by account role-aggressor flag combination s and $\tilde{x}_{s,t}$ are surprises in order flow. $\mathcal{S}$ denotes the account role-aggressor flag combinations included in the estimation. The identifying assumption is that w_t and ε_t are uncorrelated (Durbin and Koopman, 2012; Hendershott and Menkveld, 2014; Brogaard, Hendershott, and Riordan, 2014). Similar models have been applied in Menkveld, Koopman, and Lucas (2007), Menkveld (2013), Hendershott and Menkveld (2014), Brogaard, Hendershott, and Riordan (2014), Chordia, Green, and Kottimukkalur (2018), and Yueshen and Zhang (2020), among others. Economically, the correlation between pricing errors and innovations in order flow is due to trading activity in the market that is captured by the order imbalance. Thus, once we control for order flow and innovations in order flow, the orthogonal part is arguably independent.

The pricing error captures all transitory deviations of the observed price series from the estimated efficient price series. This includes next to liquidity-driven price deviations also deviations that result from constraints on market makers as a result of inventory holding costs, limited risk bearing capacity as well as dealer market power.

4.2 Order Flow Series

Surprises in order flow are obtained as the residual from an VAR model for all order flow series included in the specification.

Incorporating information on account role and aggressor flag yields several account role-aggressor flag combinations, as discussed in in Section 3.2.[16] As total order flow clears the market, at least one of the respective account role-aggressor flag combinations has to be omitted from the model for estimation. We thus estimate differences between the included order flow series. Given that our focus is on differences between agents and principals, we use aggressive and passive order flow, respectively, and distinguish between the trader types, similar to Brogaard, Hendershott, and Riordan (2014).[17]

As we control for trading volume in euros, our estimates on order flow and surprises in order flow account for differences in order flow and measure *relative* differences in order flow. We thus estimate information scaled by volume. Additionally, we quantify the contribution of an account role-aggressor flag combination to price discovery by expressing the variation in efficient prices that can be explained by innovations in the respective order flow series relative to the total variation in efficient prices. This yields

$$\frac{\gamma_s^2 var(\tilde{x}_s)}{\gamma' \mathbf{\Sigma} \gamma + \sigma_w^2}$$

where s denotes the account role-aggressor flag combination, γ is the vector of estimated coefficient on innovations in order flow and $\mathbf{\Sigma}$ is the covariance matrix of innovations in order flow.

4.3 Model Estimation

Our state space model can be mapped into the standard state space representation (Durbin and Koopman, 2012) and standard estimation techniques apply. We describe the mapping in appendix C.

The model is estimated by maximum-likelihood estimation and the Kalman filter is used to evaluate the likelihood function. The Kalman filter requires initial conditions for the state variables, given by a prior mean and a prior variance. Since the efficient price series is assumed to follow a martingale, the

---

[16] These are agent flow, principal flow, aggressive agent flow, passive agent flow, aggressive principal flow, passive principal flow, aggressive flow, passive flow.

[17] Thus, for the specification with aggressive order flow omitted, we include agent passive and principal passive order flow. For the specification with passive order flow omitted, we include agent aggressive and agent passive order flow.

state for the efficient price series is initialized as diffuse. Therefore, the prior variance is set to κ with $\kappa \to \infty$. The prior for the pricing errors are initialized as stationary using the unconditional variance.

Based on the estimation results, the Kalman smoother is used to obtain estimates of the unobserved states, conditional on all observations. This allows us to obtain estimates of the efficient price series as well as of pricing errors at every point in time. Starting values for the maximum likelihood estimation are obtained in three steps. First, we obtain starting values for a reduced form model excluding order flow and innovations in order flow

$$\begin{aligned} p_t &= m_t + s_t \\ m_t &= m_{t-1} + w_t \\ s_t &= \phi s_{t-1} + \varepsilon_t \end{aligned}$$

with $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. We obtain the starting values based on return variances and autocovariances, details are provided in Appendix D. Second, the estimation results from this reduced form model are used as starting values for the full state space model, but with the coefficients on the order flow variables estimated as states rather than parameters in MLE.[18] Finally, estimation results from the second model are used to estimate the full model by maximum likelihood.[19] Inference is based on robust, quasi-maximum likelihood standard errors (Harvey, 1990) such that inference is still valid under misspecification. Estimation is implemented using the state space models package within statsmodels in python (Seabold and Perktold, 2010; Fulton, 2015).

The methodology has several advantages over alternative approaches, as discussed in Menkveld, Koopman, and Lucas (2007), Menkveld (2013), Hendershott and Menkveld (2014), and Brogaard, Hendershott, and Riordan (2014). We observe only trade prices during the trading hours, rather than trades and quotes. Thus, missing observations have to be dealt with. The Kalman filter deals in a tractable manner with missing observations by extrapolating the state vector from the last observation, while the Kalman smoother interpolates between observations (Durbin and Koopman, 2012). This allows obtaining estimates of the states even for periods without observations. Also, the model can incorporate level shifts and structural breaks in the time series. This is important given that we study almost a decade of trading data. Also, estimation using maximum likelihood is efficient and unbiased.

As our methodology builds on the approach developed in Hasbrouck (1993), it should be noted that similar results can be obtained based on a VAR approach. Representing Hasbrouck's (1993) classical approach in state space form has several advantages when dealing with trade data only as we do. We can deal with missing overnight overnight observations in the state space model which is not possible in a VAR model. This allows us estimate the model over consecutive

---

[18] Therefore, the coefficients are introduced as other latent state variables, for example $\delta_t = \delta_{t-1}$, and initialized as diffuse states.

[19] We perform simulation exercises revealing that this approach yields reliable convergence of the maximum likelihood estimation to known parameters.

trading days. Also, our model given by (3) – (5) is a non-linear function of past observations. A VAR model, however, assumes the model to be linear in past observations. Thus, we require a state space framework to implement the full model incorporating order flow and innovations in order flow.

5 Results

In this section we present results of a frequency domain decomposition of trading profits by agents and principals. Then, we turn to our state space model at a high frequency, before moving to lower frequencies. Finally, we presents results that provide evidence that our results are indeed driven by information and explore differences by market conditions.

5.1 Frequency Domain Decomposition of Trading Profits

We first present results of a frequency domain decomposition of the trading profits pertaining to agency and principal traders. Trading profits are the product of the positioning change and the price change. Intuitively, the frequency domain decomposition allows us to investigate if order flow changes and price changes are in-phase (i.e., a positive contribution to trading profits) or out-of-phase (i.e., a negative contribution to trading profits). Technical details on the implementation can be found in Appendix E. We use a NFFT length of 12 weeks motivated by the evidence presented in Hendershott et al. (2021) and the time difference between roll days. Thus, trading profits in the lowest frequency bracket cover profits between one day and twelve weeks. Our results are robust to using four weeks as well as one week.

The results from the frequency domain decomposition are presented in Table 1. Principals' trades are profitable at the highest frequencies, that is at frequencies higher that 5 seconds. This is consistent with the idea that principals' profits are the result of them turning over positions at a high frequency (see, for example, Hasbrouck and Sofianos (1993)). Also, comparing the results from the frequency domain composition with the features of dealer banks and high-frequency traders documented in the literature shows widely consistent results. Given that that high frequency traders have been shown to contribute to price discovery at high frequencies, we would also expect the principals in our sample to be informed, at least at the highest frequency (Menkveld, 2013; Brogaard, Hendershott, and Riordan, 2014; Brogaard, Hendershott, and Riordan, 2019). We will address this question using our state space framework. At the same time, agents' orders are profitable at all lower frequencies. However, accumulated profits of principal flow are positive even at lower frequencies.

Considering aggressive versus passive orders, we find that passive orders are profitable at the highest frequency, while aggressive orders' profitability stems from lower frequencies. This may be explained by several factors. First, we only observe trade prices and no quotes. At a high frequency, bid-ask-bounces in the trade prices may be relevant, especially for aggressive orders. Second,

Table 1: Frequency domain decomposition of trading profits

This table presents results from a frequency domain decomposition of trading profits pertaining to different order flow series. Profits are in 1,000 EUR. Note that the contract value per index point is 10 EUR. The estimation uses NFFT length segments of 12 weeks motivated by the evidence presented in Hendershott et al. (2021). Thus, the last row of trading profits at a frequency lower than a day captures trading profits up to 12 weeks.

	aggressive	passive	agent	principal
$\leq$ 5 sec	−134 833.19	134 833.82	−38 232.95	38 233.59
(5 sec, 30 sec]	110 953.27	−110 953.35	6877.54	−6877.62
(30 sec, 1 min]	13 155.52	−13 155.54	462.80	−462.82
(1 min, 15 min]	14 025.77	−14 025.81	2234.95	−2234.98
(15 min, 30 min]	621.48	−621.47	267.33	−267.32
(30 min, 1 hours]	317.87	−317.87	160.03	−160.04
(1 hour, 1 day]	307.74	−307.73	180.73	−180.72
$>$ 1 day	26.69	−26.69	33.34	−33.34

the frequency domain decomposition assumes that orders are turned over at the respective frequency. In practice, however, this is not necessarily the case.

We next turn to our state space model to address the question to which extent higher gross trading revenues by agents and principals at a specific horizon translate into order informativeness.

5.2 High-frequency Results

We first estimate our state space model day-by-day at a high frequency. We choose the same frequencies as Brogaard, Hendershott, and Riordan (2014): a second frequency and event time (trade-by-trade) frequency. The event time analysis is based on the data that has been cleaned using a window length of 100ms (see Section 3.1 for more details). High-frequency traders are a subset of the principals in our sample. This raises the question whether Brogaard, Hendershott, and Riordan's (2014) results for high-frequency traders also hold for the principals in our sample. Also, we analyze a futures market rather than equity markets as Brogaard, Hendershott, and Riordan (2014). With this we address the question to which extent our results are aligned with the literature on equity markets.

The results are presented in Table 2. Statistical inference is based on standard errors that are robust to autocorrelation across 20 days in the parameter estimates. Our results are qualitatively similar to the results of Brogaard, Hendershott, and Riordan (2014) and lie quantitatively between their results for large and medium sized stocks. The major differences that we document are that we find agents to trade in the direction on pricing errors using their aggressive orders rather than against pricing errors and principals to trade in the opposite direction of pricing errors using their passive orders rather than in the direction of pricing errors. These differences are consistent between the event time and calendar time analyses and work in the direction of larger differences between agents and principals. Thus, these results for the highest frequency are arguably even stronger than the results of Brogaard, Hendershott, and Riordan (2014). The negative correlation between principals' passive trade flow and transitory price movements is consistent with principals correcting pricing errors rather than risk management.

The results for passive order flow indicate that innovations in order flow are more negatively correlated with innovations in efficient prices for agents than for principals. However, while we confirm the difference between the estimates for agents and principals to be statistically significant for the other coefficients, we do not find a statistically significant difference for the coefficients on innovations in passive order flow.

These results alleviate the concern that our results are solely driven by our focus on futures markets. The analysis of futures markets has several advantages for the application at hand, for example, it being leading in terms of price discovery (Hasbrouck, 2003) as well as a long, attrition-free data sample. Nevertheless, our results are aligned with the literature on equity markets as well.

Overall, our results for the highest frequency suggest that principals' aggressive orders are more informed than agents' aggressive orders. Based on the analysis at a second-frequency we find the agent share in efficient price innovations to be roughly 10.1%, while the principal share is approximately 19.4%. Comparing the results from the state space model to the results from the frequency domain analysis yields interesting insights. Principals not only contribute to price discovery using their aggressive orders at the highest frequency, but also their trading profits are mainly due to their trading at the highest frequencies (frequencies less than 5 seconds). Furthermore, principals trade against the direction of pricing errors using both their aggressive and passive orders. Thus, positive trading profits at high frequencies are likely due to both information and the correction of pricing errors.

5.3 State Space Results at Lower Frequencies

In this section we present results from estimating the full state space model at different frequencies. We estimate the model at several intraday frequencies before estimating the model at a daily frequency and a weekly frequency. With this we address the question whether traders being informed at a high frequency translates into the same group of traders being informed at a lower frequency as well.

We estimate the state space model for aggressive flow (Table 3) and passive flow (Table 4) at different frequencies. This specification allows us to compare relative differences between trader types – agents and principals – that are using the same order type at different frequencies.

Table 2: High-frequency results for the state space model with order flow

This table presents estimation results for the full state space model

$$\begin{aligned} p_t &= m_t + s_t \\ m_t &= m_{t-1} + \sum_{s \in \mathcal{S}} \gamma_s \tilde{x}_{s,t} + w_t \\ s_t &= \phi s_{t-1} + \sum_{s \in \mathcal{S}} \delta_s x_{s,t} + \varepsilon_t \end{aligned}$$

at a second frequency and in event time. The model is estimated day-by-day. x_t is order flow and $\tilde{x}_t$ are surprises in order flow obtained as the residual from a VAR model. The subscripts on γ and δ denote the account role with c (p) standing for agent (principal) and the aggressor flag with a (n) standing for aggressive (passive) order flow. Standard deviations are in bp and δ as well as γ in $bp/1,000,000$ EUR. Standard errors robust to autocorrelation in daily estimates of 20 days in parentheses. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.

	calendar time		event time	
	aggressive	passive	aggressive	passive
σ_w	0.4316***	0.4313***	0.5927***	0.5928***
	(0.0152)	(0.0151)	(0.0145)	(0.0146)
σ_ϵ	2.0860***	2.0861***	2.0331***	2.0342***
	(0.0231)	(0.0232)	(0.0224)	(0.0224)
ϕ	0.0998***	0.1001***	0.0026**	0.0033***
	(0.0026)	(0.0025)	(0.0012)	(0.0012)
efficient price				
$\gamma_{c,a}$	1.2850***		1.2772***	
	(0.0549)		(0.0608)	
$\gamma_{p,a}$	1.4841***		1.5395***	
	(0.0656)		(0.0709)	
$\gamma_{c,n}$		−1.5159***		−1.5278***
		(0.0475)		(0.0503)
$\gamma_{p,n}$		−1.3974***		−1.4416***
		(0.0689)		(0.0752)
pricing error				
$\delta_{c,a}$	0.2765***		0.6325***	
	(0.0145)		(0.0250)	
$\delta_{p,a}$	−0.1206***		−0.2207***	
	(0.0085)		(0.0147)	
$\delta_{c,n}$		0.3511***		0.6074***
		(0.0176)		(0.0251)
$\delta_{p,n}$		−0.1652***		−0.3458***
		(0.0118)		(0.0170)

Table 3: State space model for aggressive order flow

This table presents estimation results for the full state space model

$$\begin{aligned} p_t &= m_t + s_t \\ m_t &= m_{t-1} + \sum_{s \in \mathcal{S}} \gamma_s \tilde{x}_{s,t} + w_t \\ s_t &= \phi s_{t-1} + \sum_{s \in \mathcal{S}} \delta_s x_{s,t} + \varepsilon_t \end{aligned}$$

at different frequencies. x_t is order flow and $\tilde{x}_t$ are surprises in order flow obtained as the residual from a VAR model. The subscripts on γ and δ denote the account role with c (p) standing for agent (principal) and the aggressor flag with a standing for aggressive order flow. Standard deviations are in *bp* and δ as well as γ in $bp/1,000,000$ EUR. Robust standard errors in parentheses. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.

	5min	15min	30min	60min	1day	1week
σ_w	7.4155***	12.5893***	12.6485***	17.7380***	86.3586***	193.4271***
	(0.0114)	(0.0275)	(0.1325)	(0.2311)	(3.3313)	(18.3423)
σ_ϵ	3.8318***	4.6784***	13.7246***	18.9428***	41.6842***	84.7849***
	(0.2215)	(0.4673)	(0.2926)	(0.4612)	(6.8496)	(30.1210)
ϕ	0.2073***	0.0885	0.9403***	0.8853***	–0.0069	0.0905
	(0.0529)	(0.0819)	(0.0073)	(0.0144)	(0.2141)	(0.2926)
efficient price						
$\gamma_{c,a}$	0.9257***	0.9399***	0.8282***	0.7775***	0.7516***	0.5724***
	(0.0932)	(0.0850)	(0.0550)	(0.0463)	(0.0612)	(0.0945)
$\gamma_{p,a}$	0.1437***	0.0984***	0.0389	0.0390	–0.0069	–0.1338*
	(0.0287)	(0.0256)	(0.0329)	(0.0325)	(0.0509)	(0.0754)
pricing error						
$\delta_{c,a}$	0.0717***	0.0642***	0.1740***	0.1696***	0.1139***	0.1520**
	(0.0256)	(0.0162)	(0.0443)	(0.0415)	(0.0404)	(0.0725)
$\delta_{p,a}$	0.0210**	0.0011	0.0396	0.0169	–0.0522	–0.0304
	(0.0089)	(0.0069)	(0.0294)	(0.0272)	(0.0333)	(0.0516)
# Obs	382,849	127,909	64,154	32,277	2,278	462

Recall that aggressive orders are either market orders or marketable limit orders. In line with the previous results, changes in efficient prices load positively on innovations in aggressive agent and principal flow for higher frequencies. The coefficients on aggressive agent flow are statistically highly significant and economically relevant for all frequencies. The same does not hold true for aggressive principal flow. While the coefficients are positive and significant for a 5-minute and a 15-minute frequency, we cannot reject the null hypothesis that innovations in efficient prices do not correlate with innovations in aggressive principal flow once we move to a frequency lower than 30 minutes.

Kaniel and Liu (2006) argue that if information is more short-lived, traders rather use market orders. Our findings suggest that this hold especially true for principal traders as their aggressive orders are informed at the highest frequencies while the informativeness of their orders decreases for lower frequencies. The finding that agents' aggressive orders are informative also at lower frequencies suggests that agents also trade on longer-term information using their aggressive orders.

Overall, the coefficients in the efficient price equation in Table 3 decrease in size once we move to a lower frequency. However, at the same time, the variation of innovations in principal and agent flow may also differ by frequency. Thus, we express the contribution to price discovery as the fraction of variation in efficient prices that can be explained by the variance in aggressive agent and principal flow, that is

$$\frac{\gamma_{c,a}^2 var(\tilde{x}_{c,i})}{\gamma'\mathbf{\Sigma}\gamma + \sigma_w^2}$$

where γ is the vector of estimated coefficient on innovations in order flow and $\mathbf{\Sigma}$ is the covariance matrix of innovations in order flow. At a 5-minute frequency, this share equals roughly 9.9%, remarkably close to the share of 10.1% that we found for a second frequency. Rather than declining as we move to lower frequencies, the agent share in efficient price innovations increases.

We exhibit the share of agents' and principals' aggressive orders in the variation in efficient price innovations in Figure 2. For agents, we find that this share is positive and economically relevant for all intraday frequencies as well as for a daily and weekly frequency. The same does not hold true for principals. Despite being positive, the principal share is not economically sizeable for intraday frequencies except at a 1-second frequency.

These differences in contribution to price discovery are economically meaningful. Our results suggest that agents contribute to price discovery at lower frequencies, while price discovery at high frequency is driven by principal flow.

In general, our results are consistent with informed agents using market orders (Harris, 1998). We find that agents are relatively more informed at all except the highest frequency. Principals may infer information from order flow or agency accounts (Evans and Lyons, 2002; Menkveld, Sarkar, and Van der Wel, 2012). For example, principal dealers observe the order flow of their customers and may also deduce information from the state of the order book. Hortaçsu and Kastl (2012) argue that dealers may extract information from the orders of their customers to either compete with their customers or deduce fundamental information from order flow. We cannot explicitly test this channel. However, at lower frequencies, principals do not add incremental information once we control for aggressive agent flow.

The finding that agent flow is relatively more informative at lower frequencies is in line with the findings of Menkveld, Sarkar, and Van der Wel (2012) for the US treasury market. Furthermore, it complements the findings of Czech et al. (2021) and Jurkatis et al. (2022) for government and corporate bond markets, respectively. While their findings are related to relationship discounts, our results show that agents trade on their information profitably in the European futures market in a modern electronic limit order book setting.

For the transient price component – the pricing error equation – we find that pricing errors load positively on agent flow for all frequencies. This is consistent with prices overreacting to information in agent flow. Also, as Brogaard, Hendershott, and Riordan (2014) point out, a positive association between order flow and transitory price movements is associated with risk management. While we find that principals trade in the opposite direction of transitory price movements for the highest frequency, thus correcting mispricing, principal trading is largely uncorrelated with changes in pricing errors for lower frequencies. This suggests not only that principals' contribution to price discovery is concentrated at the highest frequencies, but more than that, also their ability to correct mispricing is limited to the highest frequency.

The estimation results for agents and principals using passive – i.e., limit – orders (Table 4) reveal that efficient price changes load negatively on innovations in passive agent and principal flow for most frequencies. This is consistent with traders using limit orders being adversely selected (see, for example, Gârleanu and Pedersen (2004) and Linnainmaa (2010)). While agents are stronger subject to adverse selection for higher frequencies (5 minutes and 15 minutes), principals are stronger subject to adverse selection for frequencies of 30 minutes and lower. Together with the results on aggressive agent and principal flow, this suggests that on aggregate both informed agents and principals use market orders, while those traders that use limit orders and are adversely selected.

Principals' passive order flow is uncorrelated with transient changes in prices for most intraday frequencies. However, at a daily and a weekly frequency, we find principals to trade in the opposite direction of pricing errors. Thus, they correct transitory mispricing using their passive orders at lower frequencies. The pattern for agents is reversed. Their passive order flow is negatively correlated with pricing errors for higher intraday frequencies (5 to 30 minutes), but uncorrelated with changes in pricing errors for lower frequencies.

A potential mechanism causing informed traders to use limit orders instead of market orders are rebates for supplying liquidity as well as limit orders earning the spread. If rebates are sufficiently high compared to the execution risk of limit orders relative to market orders, informed traders prefer to submit limit orders. Our results indicate that this is, in general, not the case as limit orders are on average subject to adverse selection while efficient price innovations load positively on market orders.

The results for principal traders are consistent with principals making the market and offering quotes to other traders. Agents' passive orders perform worse than agents' aggressive orders, but better than principal's passive orders. Together with the results on aggressive order flow, a possible explanation is that agents with a longer information horizon not only trade using market orders, but also use limit orders, consistent with Kaniel and Liu (2006). An-

Table 4: State space model for passive order flow

This table presents estimation results for the full state space model

$$\begin{aligned} p_t &= m_t + s_t \\ m_t &= m_{t-1} + \sum_{s \in \mathcal{S}} \gamma_s \tilde{x}_{s,t} + w_t \\ s_t &= \phi s_{t-1} + \sum_{s \in \mathcal{S}} \delta_s x_{s,t} + \varepsilon_t \end{aligned}$$

at different frequencies. x_t is order flow and $\tilde{x}_t$ are surprises in order flow obtained as the residual from a VAR model. The subscripts on γ and δ denote the account role with c (p) standing for agent (principal) and the aggressor flag with n standing for passive order flow. Standard deviations are in *bp* and δ as well as γ in $bp/1,000,000$ EUR. Robust standard errors in parentheses. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.

	5min	15min	30min	60min	1day	1week
σ_w	8.1219***	13.9051***	14.1660***	19.6001***	75.7713***	188.5614***
	(0.0085)	(0.0232)	(0.1399)	(0.2655)	(7.4915)	(20.2580)
σ_ϵ	3.4521***	4.3400***	14.5700***	20.4950***	76.3027***	107.9518***
	(0.1631)	(0.4020)	(0.2080)	(0.3829)	(7.7649)	(25.9318)
ϕ	0.2271***	0.1553*	0.9438***	0.8940***	0.6337***	0.1173
	(0.0446)	(0.0926)	(0.0058)	(0.0124)	(0.0826)	(0.2159)
efficient price						
$\gamma_{c,n}$	−0.5370***	−0.5132***	−0.2839***	−0.2782***	−0.0481***	0.2612
	(0.1031)	(0.0881)	(0.0741)	(0.0697)	(0.1064)	(0.1257)
$\gamma_{p,n}$	−0.3835***	−0.3976***	−0.4220***	−0.4107***	−0.3920***	−0.3571***
	(0.0514)	(0.0510)	(0.0406)	(0.0375)	(0.0553)	(0.0681)
pricing error						
$\delta_{c,n}$	−0.1027***	−0.0430**	−0.1871***	−0.0908	0.1139	0.1366
	(0.0228)	(0.0204)	(0.0608)	(0.0570)	(0.0998)	(0.1010)
$\delta_{p,n}$	−0.0108	−0.0135	−0.0248	−0.0500	−0.1784***	−0.1177**
	(0.0108)	(0.0094)	(0.0344)	(0.0338)	(0.0490)	(0.0507)
#Observations	382,849	127,909	64,150	32,277	2,278	462

other interpretation that is consistent with our results is that the group of agent traders is diverse. On the one hand, the group consists of uninformed traders whose limit orders are adversely selected and who do not contribute to price discovery. On the other hand, the group consists of informed traders with long-lived information who are using limit orders that contribute to price discovery, as documented by Bloomfield, O'Hara, and Saar (2005), Collin-Dufresne and Fos, 2015, Baruch, Panayides, and Venkataraman (2017), and Kacperczyk and Pagnotta (2019). Furthermore, our results suggest that principals do not learn enough from the orders of their customers to prevent their limit orders from being adversely selected (Hortaçsu and Kastl, 2012).

Linking the results from the state space models for aggressive and passive order flow to the frequency domain decomposition of trading profits (Table 1) yields the following insights. In the frequency decomposition of the trading profits we found that principals' profits are mainly due to the positions they take at the highest frequency (5 seconds and higher), while agents' profits mainly pertain to lower frequencies (5 seconds and lower). This is reflected in the results from our state space model. At the highest frequency, principals trade in the direction of efficient price changes and against pricing errors using aggressive orders. At the same time, their passive orders are subject to adverse selection but trade against pricing errors. These features do not translate to lower frequencies. At lower frequencies, the correlation between principals' aggressive orders and efficient price changes decreases. Also, they are not negatively correlated with pricing errors anymore. At the same time, their passive orders are still subject to adverse selection. Thus, the features of principal order flow that likely explain their their positive trading revenues at high frequencies do not translate to lower frequencies.

Agents' aggressive orders, however, are still positively associated with changes in efficient prices. This likely accounts for the positive trading revenues of agents at lower frequencies, despite their passive orders being subject to adverse selection.

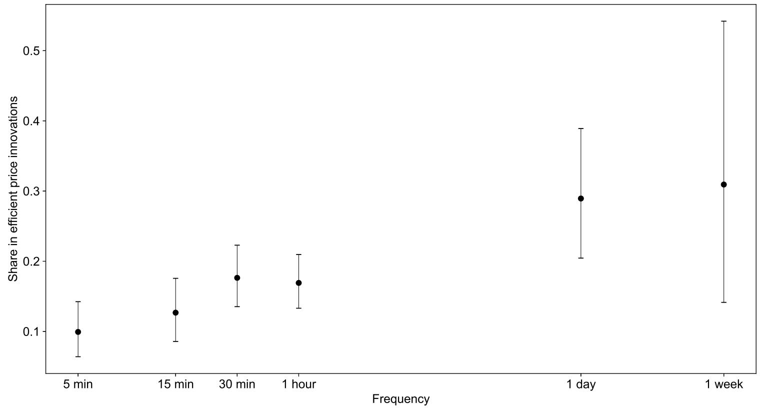
Overall, our findings suggest that a traders' orders being informed at a high frequency does not translate into order informativeness at a lower frequency. Differences between frequencies are likely due to differences in trading strategies and the type of information both agents and principals trade on. Our results are consistent with the theoretical model of Foucault, Hombert, and Roşu (2016) where the fast traders' order flow is more correlated with short-term price movements. Our findings in terms of order informativeness and trading against pricing errors from the state space model translate into differences in trading profits at different frequencies, as we show based on a frequency domain decomposition.

5.4 Evidence on Information Channel

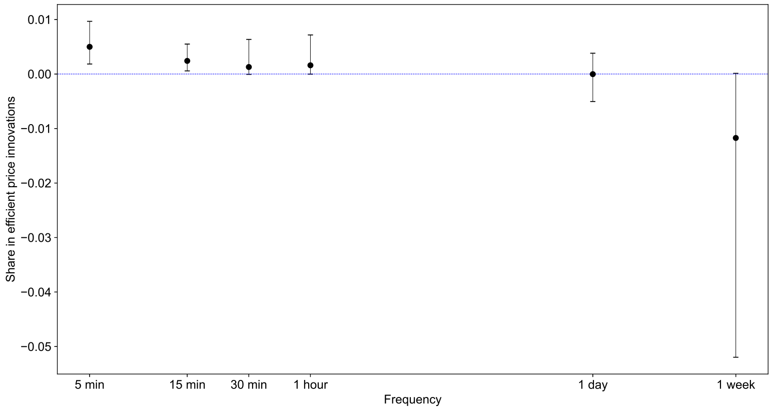
The evidence presented in the previous section is consistent with agents being more informed at longer horizons. This is especially true for those using aggressive orders. In this section we present evidence that is consistent with these

Figure 2: Agent and principal shares in efficient price innovations at different frequencies

The figure plots the fraction of variation in efficient prices that can be explained by the variance in aggressive agent and principal flow for different frequencies. The whiskers depict 95% confidence intervals based on the estimated coefficients on innovations in order flow from the state space model.



(a) Agents



(b) Principals

results being driven by an information channel rather than by non-informational reasons.

The type of information that is likely most relevant for trading in Euro STOXX 50 futures is macroeconomic information.[20] Thus, we identify high-information days as macroeconomic news announcement days in our sample. In line with the literature on public news announcements, we focus on the announcement of employment data as well as producer price index (PPI) and consumer price index (CPI) reports (see, for example, Ederington and Lee (1993) and Fleming and Remolona (1999)). We compute the relative frequencies of agents and principals trading on the right side of the market. That is, we compute the frequency of days in which agents and principals buy on days on which the price increases and sell on days on which the price decreases. Then, we compute the relative frequency (relative risk) of agents being on the right side relative to principals being on the right side of the market

$$RF = \frac{Pr(correct|agent)}{Pr(correct|principal)},$$

where "correct" refers to agents and principals, respectively, trading on the right side of the market and $Pr(\cdot)$ denotes the empirical probability. Odds ratios and relative frequencies are commonly used in medical and health sciences to study associations, with relative frequencies generally yielding more conservative estimates (McNutt et al., 2003; Schmidt and Kohlmann, 2008). Under the null hypothesis of no informational differences, we expect $H_0 : RF = 1$. A relative frequency that is significantly higher than 1 is evidence for agents being more informed while a relative frequency that is significantly lower that 1 is evidence for principals being more informed.

Figure 3 depicts relative frequencies for EU news days as well as for EU news days and constituent country news days.[21] The results show that we can soundly reject the null hypothesis of agents and principals trading equally often on the right side of the market. On average, agents trade twice as often on the right side of the market as principals. This holds both for EU news days and EU and constituent country news days. We interpret these results as consistent with an information channel. We repeat the analysis using odds ratios instead of relative frequencies and find consistent results.

Next, we focus on low information days by ex-post identifying roll days. Anecdotal evidence suggests that futures contracts are rolled over on the trading day eight calendar days before expiration.[22] Hence, we focus on both the trading day eight calendar days before expiration as well as the period eight to one calendar days before expiration. We again compute relative frequencies of agents trading on the right side of the market relative to principals trading on the right

---

[20] Also stock-specific information may drive trading in Euro STOXX 50 futures as long as it affects (a subset of stocks) with a sufficiently large weight.

[21] The Euro STOXX 50 contains constituents from eight different countries; Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands, and Spain.

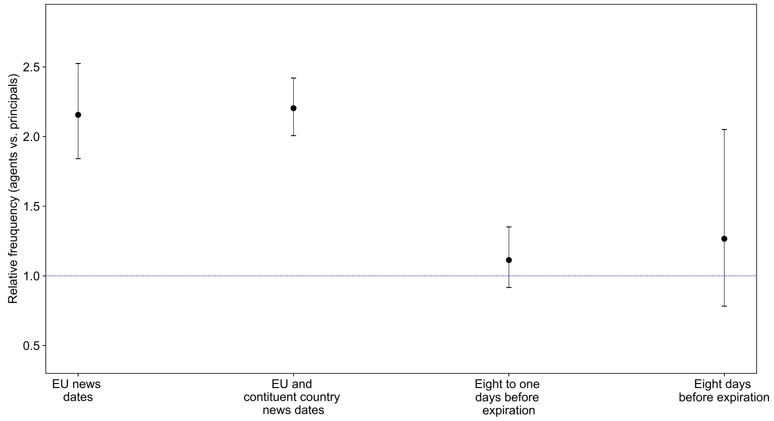
[22] See, for example, CME Group: https://www.cmegroup.com/trading/equity-index/rolldates.html.

Figure 3: Relative frequencies on news announcement days and roll days

The figure plots the relative frequency of agents trading on the right side of the market relative to principals trading on the right side of the market,

$$RF = \frac{Pr(correct|agent)}{Pr(correct|principal)}.$$

The six plots correspond to different event days in the sample. The whiskers depict 95% confidence intervals.



side of the market. Given that most of the trading volume on roll days is driven by non-informational reasons, i.e., rolling over to the next futures contract, we expect $H_0 : RF = 1$.

The results are exhibited in Figure 3. The results show that for both cases we cannot reject the null hypothesis of agents trading equally often on the right side of the market compared with principals on low information days. This provides evidence that our results are not driven by a non-information channel. We find consistent results when using odds ratios instead of relative frequencies.

5.5 Estimation by market volatility

A natural question that arises is to whether our previous results are driven by specific market conditions. We next turn to how trading patterns change depending on the market conditions. Therefore, we include information of the CBOE's volatility index (VIX). The VIX serves as a background variable that is calculated based on options on the S&P 500. At the same time, it captures general market conditions that influence trading in Euro STOXX 50 futures. Thus, we prefer this specification over alternative specifications using measures such as returns in certain time intervals or realized volatilities as these are endogenous to the trading process.

We augment our state space model as follows. We include an indicator variable that equals one if the VIX on the respective trading day falls in the lowest and highest decile, respectively, of the distribution over our sample period and interact it with the order flow variables.[23] As trading in Euro STOXX 50 futures might not instantaneously react to changes in the VIX and our focus is on price discovery at both high and low frequencies, we assign the indicator based on daily VIX levels. Then it follows for the state space model:

$$p_t = m_t + s_t \tag{6}$$

$$m_t = m_{t-1} + \sum_{s \in \mathcal{S}} \gamma_s \tilde{x}_{s,t} + \sum_{s \in \mathcal{S}} \gamma_{s,VIX} \mathbb{1}(VIX_t \in \mathcal{D}) \tilde{x}_{s,t} + w_t \tag{7}$$

$$s_t = \phi s_{t-1} + \sum_{s \in \mathcal{S}} \delta_s x_{s,t} + \sum_{s \in \mathcal{S}} \delta_{s,VIX} \mathbb{1}(VIX_t \in \mathcal{D}) x_{s,t} + \varepsilon_t. \tag{8}$$

In the specification, $\mathcal{D}$ denotes the respective decile. This specification captures differences in price discovery and liquidity provision in normal periods versus high-VIX and low-VIX periods. Again, we ask the question which traders are informed and how they are trading on their information.

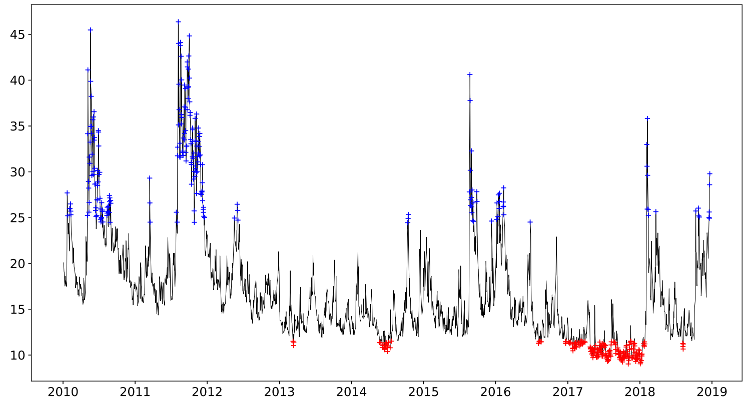
Observations in the bottom and top decile of the VIX distribution are distributed unevenly over our sample period (Figure 4). Most of the the observations in the bottom decile of the sample distribution cluster in 2017. Most

---

[23] We compute the distribution of the VIX over our sample period based on daily closing prices. Then, using each daily closing price, we assign the indicator variable to all observations on the respective trading day, depending on in which percentile of the sample distribution the respective closing price falls.

Figure 4: Daily VIX closing prices over the sample period

The figure plots daily VIX closing prices over the sample period. Blue crosses indicate observations in the top decile of the sample distribution and red crosses indicate observations in the bottom decile of the sample distribution.



observations that fall in the top decile of of the sample distribution cluster early in our sample period in the years 2010 and 2011.

As in the previous analysis, we run the state space model for aggressive and passive order flow separately and include order flow from agents and principals. Here we present results for the state space model estimated at an hourly frequency. Overall, our results indicate differences in trading patterns in times of high and low market volatility. Also, the differences are larger for agents than for principals.

The results for aggressive flow are presented in Table 5. In comparison to the results presented in the previous section, our results are remarkably stable, indicating the robustness of our results. In low-volatility regimes, the relative contribution of aggressive agent flow to price discovery is lower than over the whole sample period. Still, agents' contribution to price discovery using market and marketable limit orders dominates the contribution of principals using the same order type. The coefficient on agent flow in the pricing error equation, δ, is lower in low-volatility periods.

While we find differences in the informativeness of agents' aggressive orders between low-volatility periods and the overall sample, our main conclusions remain unchanged. At lower frequencies, agents contribute relatively more to price discovery using their aggressive orders than principals do. For principals using aggressive orders, most of the coefficients for the low-volatility dummy are insignificant. Overall, there is no evidence for a higher contribution of principals' aggressive orders to price discovery in low-volatility periods.

For high-volatility periods, the patterns are reversed in comparison to low-volatility periods – except that we find small changes for principals. Efficient price innovations load stronger on innovations in aggressive agent flow in high-volatility periods relative to the overall sample. That is, aggressive agent flow contains more information scaled by volume and contributes more to price discovery in high-VIX periods than in the overall sample. A plausible explanation for this pattern is that informed agents rather use market or marketable limit orders in high-volatility periods while they rely on limit orders in low volatility periods. This is consistent with a pecking order of order types dependent on the market conditions, in a spirit of Menkveld, Yueshen, and Zhu (2017).

The impact of agent flow on pricing errors increases – if anything – in high volatility periods, reinforcing our results for the overall sample. We interpret this result as overreactions to information that are especially pronounced in highly volatile periods.

For principal trades, the results are mostly unchanged. Efficient price innovations do not load significantly on innovations in aggressive principal flow in high-VIX periods. This is consistent with the previous finding that most information is incorporated into prices through agent flow. In high volatility periods, the variance in agent flow accounts for roughly 45% of the variance in efficient price innovations, while the variance in principal flow accounts for less than 1%. We do not find evidence that in highly volatile times, principals are better able to extract information from the state of the order book and trade on this information (Parlour, 1998), as principals do not add incremental information once we include agent flow.

Next, we turn to the results for passive order flow (Table 6). In low-volatility periods, innovations in efficient prices load stronger on passive agent flow than over the entire sample period. Recall that over the whole sample, changes in efficient prices load negatively on passive agent flow, indicating that agents' limit orders are adversely selected. In low volatility periods, agents are thus less exposed to adverse selection. We do not find evidence that principals' limit order are less exposed to adverse selection in low volatility periods in comparison with the overall sample period.

These findings may be explained by agents being better able to manage their orders in times of low volatility. Also, some informed agents may use passive orders rather than aggressive orders in less volatile times. As a result, on average, agents are less exposed to adverse selection. At the same time, principal traders may mainly provide liquidity and are subject to adverse selection.

The results for the high-VIX periods are in line with this intuition. In comparison to the overall sample period, efficient price innovations load more negative on both agents' and principals' passive orders. That indicates that limit orders of both trader types are stronger subject to adverse selection in volatile times than in the overall sample period. This result mirrors the results for aggressive orders in high-VIX periods: efficient price innovations load stronger on aggressive agent flow. This may be because informed traders use predominantly aggressive orders in periods of high volatility. Traders supplying liquidity, in contrast, trade for non-informational reasons (Kaniel and Liu, 2006).

Table 5: State space model for aggressive order flow including VIX

This table presents estimation results for the full state space model including VIX

$$
\begin{aligned}
p_t &= m_t + s_t \\
m_t &= m_{t-1} + \sum_{s \in \mathcal{S}} \gamma_s \tilde{x}_{s,t} + \sum_{s \in \mathcal{S}} \gamma_{s,VIX} \mathbb{1}(VIX_t \in \mathcal{D}) \tilde{x}_{s,t} + w_t \\
s_t &= \phi s_{t-1} + \sum_{s \in \mathcal{S}} \delta_s x_{s,t} + \sum_{s \in \mathcal{S}} \delta_{s,VIX} \mathbb{1}(VIX_t \in \mathcal{D}) x_{s,t} + \varepsilon_t
\end{aligned}
$$

at an hourly frequency. x_t is order flow and $\tilde{x}_t$ are surprises in order flow obtained as the residual from a VAR model. $\mathbb{1}(VIX_t \in \mathcal{D})$ is an indicator that equals one if the closing VIX on the respective trading day is in the lowest decile or highest decile, respectively, of the distribution over the sample period. Passive order flow is omitted from the specification. The subscripts on γ and δ denote the account role with c (p) standing for agent (principal) and the aggressor flag with a standing for aggressive order flow. Standard deviations are in *bp* and δ as well as γ in *bp*$/1,000,000$ EUR. Robust standard errors in parentheses. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.

	low	high
σ_w	17.5000*** (0.2370)	17.1452*** (0.2382)
σ_ϵ	18.8085*** (0.4708)	18.5078*** (0.3984)
ϕ	0.8859*** (0.0150)	0.8901*** (0.0133)
efficient price		
$\gamma_{c,a}$	0.8285*** (0.0508)	0.7126*** (0.0453)
$\gamma_{p,a}$	0.0376 (0.0345)	-0.0076 (0.0294)
$\gamma_{c,a,VIX}$	-0.2932^{***} (0.0865)	0.6286*** (0.1391)
$\gamma_{p,a,VIX}$	-0.1533^{***} (0.0574)	0.1736 (0.1516)
pricing error		
$\delta_{c,a}$	0.1395*** (0.0408)	0.1028*** (0.0330)
$\delta_{p,a}$	0.0360 (0.0292)	0.0540** (0.0260)
$\delta_{c,a,VIX}$	-0.2400^{***} (0.0705)	0.2673** (0.1236)
$\delta_{p,a,VIX}$	0.0813 (0.0522)	0.1298 (0.1412)
#Obs	32,223	32,223

Table 6: State space model for passive order flow including VIX

This table presents estimation results for the full state space model including VIX

$$
\begin{aligned}
p_t &= m_t + s_t \\
m_t &= m_{t-1} + \sum_{s\in\mathcal{S}} \gamma_s \tilde{x}_{s,t} + \sum_{s\in\mathcal{S}} \gamma_{s,VIX} \mathbb{1}(VIX_t \in \mathcal{D}) \tilde{x}_{s,t} + w_t \\
s_t &= \phi s_{t-1} + \sum_{s\in\mathcal{S}} \delta_s x_{s,t} + \sum_{s\in\mathcal{S}} \delta_{s,VIX} \mathbb{1}(VIX_t \in \mathcal{D}) x_{s,t} + \varepsilon_t
\end{aligned}
$$

at an hourly frequency. x_t is order flow and $\tilde{x}_t$ are surprises in order flow obtained as the residual from a VAR model. $\mathbb{1}(VIX_t \in \mathcal{D})$ is an indicator that equals one if the closing VIX on the respective trading day is in the lowest decile or highest decile, respectively, of the distribution over the sample period. Aggressive order flow is omitted from the specification. The subscripts on γ and δ denote the account role with c (p) standing for agent (principal) and the aggressor flag with n standing for passive order flow. Standard deviations are in *bp* and δ as well as γ in $bp/1,000,000$ EUR. Robust standard errors in parentheses. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.

	low	high
σ_w	19.7916*** (0.2371)	19.3578*** (0.2363)
σ_ϵ	20.0336*** (0.3998)	19.7299*** (0.3703)
ϕ	0.8846*** (0.0136)	0.8875*** (0.0127)
efficient price		
$\gamma_{c,n}$	−0.3269*** (0.0810)	−0.0912 (0.0632)
$\gamma_{p,n}$	−0.4340*** (0.0442)	−0.3949*** (0.0385)
$\gamma_{c,n,VIX}$	0.2691** (0.1104)	−0.7087*** (0.2409)
$\gamma_{p,n,VIX}$	0.0508 (0.0745)	−0.4234*** (0.1434)
pricing error		
$\delta_{c,n}$	−0.0849 (0.0616)	−0.1415*** (0.0482)
$\delta_{p,n}$	−0.0372 (0.0336)	−0.0159 (0.0275)
$\delta_{c,n,VIX}$	0.1962** (0.0982)	−0.2140 (0.2324)
$\delta_{p,n,VIX}$	0.1249* (0.0648)	−0.1927 (0.1332)
#Obs	32,221	32,221

In high-volatility periods, the change in exposure to adverse selection is higher for agents than for principals. In such periods, principals might have to post limit orders to provide liquidity within the exchange's requirements. At the same time, they might infer information from the order book and limit their exposure to adverse selection (in line with Hortaçsu and Kastl (2012)). Our results provide suggestive evidence in line with the latter channel. They also suggest that the finding that agent's passive orders are adversely selected is mainly due to high-volatility periods.

These results suggest a pecking order of order types dependent on the market conditions, akin to Menkveld, Yueshen, and Zhu (2017). Also, these results are consistent with the intuition of Kaniel and Liu (2006). Higher market volatility can be interpreted as decreasing the horizon on which investors can trade on their information. As markets are volatile, movements in the disadvantage of a trader may occur more frequent and are less predictable. A reduction in their horizon causes informed traders to use market orders rather than limit orders.

The results of Collin-Dufresne and Fos (2015) and Kacperczyk and Pagnotta (2019) show that informed traders are using limit orders. Their results speak to insiders that possess firm-specific information. We analyze trading in futures on a Pan-European index, the Euro STOXX 50. Thus, even though traders active in these futures contracts may be motivated by firm-specific information on the constituents, prices of Euro STOXX 50 futures also reflect information on the state of the "European" economy. Our results suggest that also traders possessing information of this nature use limit orders, dependent on the market conditions. Thus, our evidence is consistent with the results of Collin-Dufresne and Fos (2015) and Kacperczyk and Pagnotta (2019) extending to a wider set of information and asset classes.

6 Conclusions

How does information in futures markets get incorporated into prices? What are the gross trading revenues that agents and principals realize at different horizons? Does a trader's orders being informed at a high frequency translate into informedness at a lower frequency? How does this depend on market conditions? We address these questions based on almost a decade of trading records in Euro STOXX 50 futures and a state space framework.

Our results indicate that principals' aggressive orders are more informed than agents' aggressive orders at the highest frequencies, consistent with the literature on high-frequency trading (Brogaard, Hendershott, and Riordan, 2014). However, this does not translate into order informativeness at lower frequencies. Once we move to lower intraday frequencies as well as a daily and weekly frequency, we find that the share of agents' order in efficient price innovations increases to up to 30% while the share of principals' orders turns economically insignificant. This is reflected in gross trading revenues for agents and principals. Principals' gross trading revenues are positive for the highest frequencies before they turn negative for lower frequencies. The opposite holds true for agents' gross trading revenues.

Our results are consistent with the theoretical model of Foucault, Hombert, and Roşu (2016) where fast traders' orders are more correlated with short-term price changes. At lower frequencies, principals do not add incremental information beyond what is contained in agent flow.

Comparing different market conditions as measured by the VIX indicates a pecking order of (informed) traders' order choice. This complements the pecking order of trading venues documented in Menkveld, Yueshen, and Zhu (2017). In low-volatility regimes, agents' passive orders are less subject to adverse selection than in normal times and in high-volatility regimes, aggressive orders are relatively more informative. The results suggest that some informed traders use limit orders not only to trade on stock-specific information, but also trade on economy-wide information in futures markets.

References

Anand, Amber, Sugato Chakravarty, and Terrence Martell (2005). "Empirical Evidence on the Evolution of Liquidity: Choice of Market versus Limit Orders by Informed and Uninformed Traders". In: *Journal of Financial Markets* 8.3, pp. 288–308.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Clara Vega (2007). "Real-time Price Discovery in Global Stock, Bond and Foreign Exchange Markets". In: *Journal of International Economics* 73.2, pp. 251–277.

Aquilina, Matteo, Eric Budish, and Peter O'Neill (July 2021). *Quantifying the High-Frequency Trading "Arms Race"*. NBER Working Papers 29011. National Bureau of Economic Research, Inc.

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean (2009). "Just How Much Do Individual Investors Lose by Trading?" In: *Review of Financial Studies* 22.2, pp. 609–632.

Baruch, Shmuel, Marios Panayides, and Kumar Venkataraman (2017). "Informed Trading and Price Discovery Before Corporate Events". In: *Journal of Financial Economics* 125.3, pp. 561–588.

Beason, Tyler and Sunil Wahal (2020). *The Anatomy of Trading Algorithms.* Working Paper.

Biais, Bruno, Fany Declerck, and Sophie Moinas (2016). *Who Supplies Liquidity, How and When?* Working Paper.

Biais, Bruno, Pierre Hillion, and Chester Spatt (1995). "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse". In: *Journal of Finance* 50.5, pp. 1655–1689.

Bloomfield, Peter (2004). *Fourier Analysis of Time Series: An Introduction.* John Wiley & Sons.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar (2005). "The "Make or Take" Decision in an Electronic Market: Evidence on the Evolution of Liquidity". In: *Journal of Financial Economics* 75.1, pp. 165–199.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar (2015). "Hidden Liquidity: Some New Light on Dark Trading". In: *Journal of Finance* 70.5, pp. 2227–2274.

Bogousslavsky, Vincent and Dmitriy Muravyev (2021). *Who Trades at the Close? Implications for Price Discovery and Liquidity.* Working Paper.

Brandt, Michael W. and Kenneth A. Kavajecz (2004). "Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve". In: *Journal of Finance* 59.6, pp. 2623–2654.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2014). "High-Frequency Trading and Price Discovery". In: *Review of Financial Studies* 27.8, pp. 2267–2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2019). "Price Discovery without Trading: Evidence from Limit Orders". In: *Journal of Finance* 74.4, pp. 1621–1658.

Campbell, John Y., Andrew W. Lo, A. Craig MacKinlay, and Robert F. Whitelaw (1998). “The Econometrics of Financial Markets”. In: *Macroeconomic Dynamics* 2.4, pp. 559–562.

Chordia, Tarun, T Clifton Green, and Badrinath Kottimukkalur (2018). “Rent Seeking by Low-Latency Traders: Evidence from Trading on Macroeconomic Announcements”. In: *Review of Financial Studies* 31.12, pp. 4650–4687.

Collin-Dufresne, Pierre and Vyacheslav Fos (2015). “Do Prices Reveal the Presence of Informed Trading?” In: *Journal of Finance* 70.4, pp. 1555–1582.

Comerton-Forde, Carole and Barbara Rindi (2021). *Trading @ the Close*. Working Paper.

Czech, Robert, Shiyang Huang, Dong Lou, and Tianyu Wang (2021). “Informed trading in government bond markets”. In: *Journal of Financial Economics* 142.3, pp. 1253–1274.

Durbin, James and Siem Jan Koopman (2012). *Time Series Analysis by State Space Methods*. OUP Catalogue 9780199641178. Oxford University Press.

Ederington, Louis H and Jae Ha Lee (1993). “How Markets Process Information: News Releases and Volatility”. In: *Journal of Finance* 48.4, pp. 1161–1191.

Ernst, Thomas (2020). *Stock-Specific Price Discovery from ETFs*. Working Paper.

Ernst, Thomas, Jonathan Sokobin, and Chester Spatt (2021). *The Value of Off-Exchange Data*. Working Paper.

Eurex (Mar. 2021). *Conditions for Trading at Eurex Deutschland*. Tech. rep.

European Systemic Risk Board (June 2020). *Liquidity Risks Arising from Margin Calls*. Report.

Evans, Martin D. D. and Richard K. Lyons (2002). “Order Flow and Exchange Rate Dynamics”. In: *Journal of Political Economy* 110.1, pp. 170–180.

Evans, Martin D.D. and Richard K. Lyons (2008). “How is Macro News Transmitted to Exchange Rates?” In: *Journal of Financial Economics* 88.1, pp. 26–50.

Fleming, Michael J., Bruce Mizrach, and Giang Nguyen (2018). “The Microstructure of a U.S. Treasury ECN: The BrokerTec Platform”. In: *Journal of Financial Markets* 40.C, pp. 2–22.

Fleming, Michael J. and Eli M. Remolona (1999). “Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information”. In: *The Journal of Finance* 54.5, pp. 1901–1915.

Foucault, Thierry (1999). “Order Flow Composition and Trading Costs in a Dynamic Limit Order Market”. In: *Journal of Financial Markets* 2.2, pp. 99–134.

Foucault, Thierry, Johan Hombert, and Ioanid Roşu (2016). “News Trading and Speed”. In: *Journal of Finance* 71.1, pp. 335–382.

Fulton, Chad (2015). “Estimating Time Series Models by State Space Methods in Python: Statsmodels”. In:

George, Thomas J and Chuan-Yang Hwang (2001). “Information Flow and Pricing Errors: A Unified Approach to Estimation and Testing”. In: *Review of Financial Studies* 14.4, pp. 979–1020.

Goettler, Ronald L., Christine A. Parlour, and Uday Rajan (2009). “Informed Traders and Limit Order Markets”. In: *Journal of Financial Economics* 93.1, pp. 67–87.

Gârleanu, Nicolae and Lasse Heje Pedersen (2004). “Adverse Selection and the Required Return”. In: *Review of Financial Studies* 17.3, pp. 643–665.

Hamilton, James D. (1986). “State-Space Models”. In: *Handbook of Econometrics*. Ed. by R. F. Engle and D. McFadden. Vol. 4. Handbook of Econometrics. Elsevier. Chap. 50, pp. 3039–3080.

Harris, Lawrence (1998). “Optimal Dynamic Order Submission Strategies in Some Stylized Trading Problems”. In: *Financial Markets, Institutions & Instruments* 7.2, pp. 1–76.

Harvey, Andrew C. (1990). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Hasbrouck, Joel (1993). “Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement”. In: *Review of Financial Studies* 6.1, pp. 191–212.

Hasbrouck, Joel (2003). “Intraday Price Formation in U.S. Equity Index Markets”. In: *Journal of Finance* 58.6, pp. 2375–2400.

Hasbrouck, Joel (2021). “Price Discovery in High Resolution”. In: *Journal of Financial Econometrics* 19.3, pp. 395–430.

Hasbrouck, Joel and George Sofianos (1993). “The Trades of Market Makers: An Empirical Analysis of NYSE Specialists”. In: *Journal of Finance* 48.5, pp. 1565–1593.

Hau, Harald (2001). “Location Matters: An Examination of Trading Profits”. In: *Journal of Finance* 56.5, pp. 1959–1983.

Hendershott, Terrence, Dmitry Livdan, and Norman Schürhoff (2015). “Are Institutions Informed About News?” In: *Journal of Financial Economics* 117.2, pp. 249–287.

Hendershott, Terrence and Albert J. Menkveld (2014). “Price Pressures”. In: *Journal of Financial Economics* 114.3, pp. 405–423.

Hendershott, Terrence, Albert J Menkveld, Rémy Praz, and Mark Seasholes (Apr. 2021). “Asset Price Dynamics with Limited Attention”. In: *The Review of Financial Studies*.

Hoffmann, Peter (2014). “A Dynamic Limit Order Market With Fast and Slow Traders”. In: *Journal of Financial Economics* 113.1, pp. 156–169.

Hollifield, Burton, Robert A. Miller, and Patrik Sandås (2004). “Empirical Analysis of Limit Order Markets”. In: *Review of Economic Studies* 71.4, pp. 1027–1063.

Hortaçsu, Ali and Jakub Kastl (2012). “Valuing Dealers’ Informational Advantage: A Study of Canadian Treasury Auctions”. In: *Econometrica* 80.6, pp. 2511–2542.

Huang, Xin (2018). “Macroeconomic News Announcements, Systemic Risk, Financial Market Volatility, and Jumps”. In: *Journal of Futures Markets* 38.5, pp. 513–534.

Jurkatis, Simon, Andreas Schrimpf, Karamfil Todorov, and Nicholas Vause (2022). *Relationship Discounts in Corporate Bond Trading*. Working Paper.

Kacperczyk, Marcin and Emiliano S Pagnotta (2019). "Chasing Private Information". In: *Review of Financial Studies* 32.12, pp. 4997–5047.

Kaniel, Ron and Hong Liu (2006). "So What Orders Do Informed Traders Use?" In: *Journal of Business* 79.4, pp. 1867–1914.

Kawaller, Ira G, Paul D Koch, and Timothy W Koch (1987). "The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index". In: *Journal of Finance* 42.5, pp. 1309–1329.

Kelley, Eric K. and Paul C. Tetlock (2013). "How Wise Are Crowds? Insights from Retail Orders and Stock Returns". In: *Journal of Finance* 68.3, pp. 1229–1265.

Li, Sida, Mao Ye, and Miles Zheng (Feb. 2021). *Financial Regulation, Clientele Segmentation, and Stock Exchange Order Types.* NBER Working Papers 28515. National Bureau of Economic Research, Inc.

Linnainmaa, Juhani T. (2010). "Do Limit Orders Alter Inferences about Investor Performance and Behavior?" In: *Journal of Finance* 65.4, pp. 1473–1506.

McNutt, Louise-Anne, Chuntao Wu, Xiaonan Xue, and Jean Paul Hafner (May 2003). "Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes". In: *American Journal of Epidemiology* 157.10, pp. 940–943.

Menkveld, Albert J. (2013). "High Frequency Trading and the New-Market Makers". In: *Journal of Financial Markets* 16.4, pp. 712–740.

Menkveld, Albert J., Anna Dreber, Felix Holzmeister, Juergen Huber, Magnus Johanneson, Michael Kirchler, Michael Razen, Utz Weitzel, David Abad, Menachem (Meni) Abudy, Tobias Adrian, Yacine Ait-Sahalia, Olivier Akmansoy, Jamie Alcock, Vitali Alexeev, Arash Aloosh, Livia Amato, Diego Amaya, James J. Angel, Amadeus Bach, Edwin Baidoo, Gaetan Bakalli, Andrea Barbon, Oksana Bashchenko, Parampreet Christopher Bindra, Geir Hoidal Bjonnes, Jeffrey R. Black, Bernard S. Black, Santiago Bohorquez, Oleg Bondarenko, Charles S. Bos, Ciril Bosch-Rosa, Elie Bouri, Christian T. Brownlees, Anna Calamia, Viet Nga Cao, Gunther Capelle-Blancard, Laura Capera, Massimiliano Caporin, Allen Carrion, Tolga Caskurlu, Bidisha Chakrabarty, Mikhail Chernov, William Ming Yan Cheung, Ludwig B. Chincarini, Tarun Chordia, Sheung Chi Chow, Benjamin Clapham, Jean-Edouard Colliard, Carole Comerton-Forde, Edward Curran, Thong Dao, Wale Dare, Ryan J. Davies, Riccardo De Blasis, Gianluca De Nard, Fany Declerck, Oleg Deev, Hans Degryse, Solomon Deku, Christophe Desagre, Mathijs A. Van Dijk, Chukwuma Dim, Thomas Dimpfl, Yun Jiang Dong, Philip Drummond, Tom Dudda, Ariadna Dumitrescu, Teodor Dyakov, Anne Haubo Dyhrberg, Michał Dzieliński, Asli Eksi, Izidin El Kalak, Saskia ter Ellen, Nicolas Eugster, Martin D.D. Evans, Michael Farrell, Ester Félez-Viñas, Gerardo Ferrara, El Mehdi FERROUHI, Andrea Flori, Jonathan Fluharty-Jaidee, Sean Foley, Kingsley Y. L. Fong, Thierry Foucault, Tatiana Franus, Francesco A. Franzoni, Bart Frijns, Michael Frömmel, Servanna Mianjun Fu, Sascha Füllbrunn, Baoqing Gan, Thomas Gehrig, Dirk Gerritsen, Javier Gil-Bazo, Lawrence R. Glosten, Thomas Gomez, Arseny Gorbenko, Ufuk Güçbilmez, Joachim Grammig, Vincent Gregoire, Björn Hagströmer, Julien Hambuckers, Erik

Hapnes, Jeffrey H. Harris, Lawrence Harris, Simon Hartmann, Jean-Baptiste Hasse, Nikolaus Hautsch, Xue-Zhong 'Tony' He, Davidson Heath, Simon Hediger, Terrence J. Hendershott, Ann Marie Hibbert, Erik Hjalmarsson, Seth Hoelscher, Peter Hoffmann, Craig W. Holden, Alex R. Horenstein, Wenqian Huang, Da Huang, Christophe Hurlin, Alexey Ivashchenko, Subramanian R. Iyer, Hossein Jahanshahloo, Naji Jalkh, Charles M. Jones, Simon Jurkatis, Petri Jylha, Andreas Kaeck, Gabriel Kaiser, Arzé Karam, Egle Karmaziene, Bernhard Kassner, Markku Kaustia, Ekaterina Kazak, Fearghal Kearney, Vincent van Kervel, Saad Khan, Marta Khomyn, Tony Klein, Olga Klein, Alexander Klos, Michael Koetter, Jan Pieter Krahnen, Aleksey Kolokolov, Robert A. Korajczyk, Roman Kozhan, Amy Kwan, Quentin Lajaunie, Full Yet Eric Campbell Lam, Marie Lambert, Hugues Langlois, Jens Lausen, Tobias Lauter, Markus Leippold, Vladimir Levin, Yijie Li, (Michael) Hui Li, Chee Yoong Liew, Thomas Lindner, Oliver B. Linton, Jiacheng Liu, Anqi Liu, Jesus-Guillermo Llorente-Alvarez, Matthijs Lof, Ariel Lohr, Francis A. Longstaff, Alejandro Lopez-Lira, Shawn Mankad, Nicola Mano, Alexis Marchal, Charles Martineau, Francesco Mazzola, Debrah C Meloso, Roxana Mihet, Vijay Mohan, Sophie Moinas, David Moore, Liangyi Mu, Dmitriy Muravyev, Dermot Murphy, Gabor Neszveda, Christian Neumeier, Ulf Nielsson, Mahendrarajah Nimalendran, Sven Nolte, Lars L. Nordén, Peter O'Neill, Khaled Obaid, Bernt Arne Ødegaard, Per Östberg, Marcus Painter, Stefan Palan, Imon Palit, Andreas Park, Roberto Pascual Gascó, Paolo Pasquariello, Lubos Pastor, Vinay Patel, Andrew J. Patton, Neil D. Pearson, Loriana Pelizzon, Loriana Pelizzon, Matthias Pelster, Christophe Pérignon, Cameron Pfiffer, Richard Philip, Tomáš Plíhal, Puneet Prakash, Oliver-Alexander Press, Tina Prodromou, Talis J. Putnins, Gaurav Raizada, David A. Rakowski, Angelo Ranaldo, Luca Regis, Stefan Reitz, Thomas Renault, Renjie Wang, Roberto Renò, Steven Riddiough, Steven Riddiough, Kalle Rinne, Paul Rintamäki, Ryan Riordan, Thomas Rittmannsberger, Iñaki Rodríguez Longarela, Dominik Rösch, Lavinia Rognone, Brian Roseman, Ioanid Rosu, Saurabh Roy, Nicolas Rudolf, Stephen Rush, Khaladdin Rzayev, Aleksandra Rzeźnik, Anthony Sanford, Harikumar Sankaran, Asani Sarkar, Lucio Sarno, Olivier Scaillet, Stefan Scharnowski, Klaus Reiner Schenk-Hoppé, Andrea Schertler, Michael Schneider, Florian Schroeder, Norman Schürhoff, Philipp Schuster, Marco A. Schwarz, Mark S. Seasholes, Norman Seeger, Or Shachar, Andriy Shkilko, Jessica Shui, Mario Sikic, Giorgia Simion, Lee A. Smales, Paul Söderlind, Elvira Sojli, Konstantin Sokolov, Laima Spokeviciute, Denitsa Stefanova, Marti G. Subrahmanyam, Sebastian Neusüss, Barnabas Szaszi, Oleksandr Talavera, Yuehua Tang, Nicholas Taylor, Wing Wah Tham, Erik Theissen, Julian Thimme, Ian Tonks, Hai Tran, Luca Trapin, Anders B. Trolle, Maria Vaduva, Giorgio Valente, Robert A. Van Ness, Aurelio Vasquez, Thanos Verousis, Patrick Verwijmeren, Anders Vilhelmsson, Grigory Vilkov, Vladimir Vladimirov, Sebastian Vogel, Stefan Voigt, Wolf Wagner, Thomas Walther, Patrick Weiss, Patrick Weiss, Michel van der Wel, Ingrid M. Werner, P. Joakim Westerholm, Christian Westheide, Evert Wipplinger, Michael Wolf, Christian C. P. Wolff, Leonard

Wolk, Wing Keung Wong, Jan Wrampelmeyer, Zhen-Xing Wu, Shuo Xia, Dacheng Xiu, Ke Xu, Caihong Xu, Pradeep K. Yadav, José Yagüe, Cheng Yan, Antti Yang, Woongsun Yoo, Wenjia Yu, Shihao Yu, Shihao Yu, Bart Zhou Yueshen, Darya Yuferova, Marcin Zamojski, Abalfazl Zareei, Stefan Zeisberger, Sarah Zhang, Xiaoyu Zhang, Zhuo Zhong, Z. Ivy Zhou, Chen Zhou, Xingyu Zhu, Xingyu Zhu, Xingyu Zhu, Marius Zoican, Remco C.J. Zwinkels, Jian Chen, Teodor Duevski, Ge Gao, Roland Gemayel, Dudley Gilder, Paul Kuhle, Emiliano Pagnotta, Emiliano Pagnotta, Michele Pelli, Jantje Sönksen, Lu Zhang, Konrad Ilczuk, Dimitar Bogoev, Ya Qian, Hans C. Wika, Yihe Yu, Lu Zhao, Michael Mi, and Li Bao (forthcoming). "Non-Standard Errors". In: *Journal of Finance*.

Menkveld, Albert J., Siem Jan Koopman, and Andre Lucas (2007). "Modeling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Methods". In: *Journal of Business & Economic Statistics* 25, pp. 213–225.

Menkveld, Albert J., Asani Sarkar, and Michel Van der Wel (2012). "Customer Order Flow, Intermediaries, and Discovery of the Equilibrium Risk-Free Rate". In: *Journal of Financial and Quantitative Analysis* 47.4, 821–849.

Menkveld, Albert J. and Bart Zhou Yueshen (2019). "The Flash Crash: A Cautionary Tale About Highly Fragmented Markets". In: *Management Science* 65.10, pp. 4470–4488.

Menkveld, Albert J., Bart Zhou Yueshen, and Haoxiang Zhu (2017). "Shades of Darkness: A Pecking Order of Trading Venues". In: *Journal of Financial Economics* 124.3, pp. 503–534.

Parlour, Christine A (1998). "Price Dynamics in Limit Order Markets". In: *Review of Financial Studies* 11.4, pp. 789–816.

Pasquariello, Paolo and Clara Vega (2007). "Informed and Strategic Order Flow in the Bond Markets". In: *Review of Financial Studies* 20.6, pp. 1975–2019.

Schmidt, Carsten Oliver and Thomas Kohlmann (2008). "When to Use the Odds Ratio or the Relative Risk?" In: *International Journal of Public Health* 53.3, 165–167.

Seabold, Skipper and Josef Perktold (2010). "Statsmodels: Econometric and Statistical Modeling with Python". In: *Proceedings of the 9th Python in Science Conference*. Vol. 57. Austin, TX, p. 61.

Stoll, Hans R. and Robert E. Whaley (1990). "The Dynamics of Stock Index and Stock Index Futures Returns". In: *Journal of Financial and Quantitative Analysis* 25.4, pp. 441–468.

Tse, Yiuman, Paramita Bandyopadhyay, and Yang-Pin Shen (2006). "Intraday Price Discovery in the DJIA Index Markets". In: *Journal of Business Finance & Accounting* 33.9-10, pp. 1572–1585.

Yueshen, Bart Z and Jinyuan Zhang (2020). *Dynamic Trade Informativeness*. Working Paper.

A Exchange Members

The following list contains all unique exchange members listed on the Eurex website.

3Red Partners LLC
Aardvark Trading, L.L.C.
ABC arbitrage Asset Management
ABN AMRO Bank N.V.
ABN AMRO Clearing Bank N.V.
ADG Europe Ltd
ADG Market Making LLP
ADG Markets Ltd.
ADM Investor Services Inc.
ADM Investor Services International Ltd.
Advantage Futures LLC
AFS Equity & Derivatives B.V.
All Options International B.V.
Allston Capital LLC
Allston Trading LLC
Allston Trading UK Limited
AlphaGrep Pte Ltd
Altura Markets Sociedad de Valores SA
AMP Global Clearing LLC
AP Capital Investment Limited
ARB Trading Group North, LP
Atlantic Trading London Limited
Aurel BGC
Auriga Capital Limited
B. Metzler seel. Sohn & Co. Aktiengesellschaft
Baader Bank Aktiengesellschaft
Banca Akros Spa
Banca Profilo SPA
Banca Sella Holding S.p.A.
Banca Simetica S.p.A.
Banco Bilbao Vizcaya Argentaria S.A.
Banco Comercial Portugues S.A.
Banco Santander SA
Bank J. Safra Sarasin AG
Bank Julius Bär & Co. AG
Bank Vontobel AG
Bankhaus Lampe KG
Bankinter
Banque de Luxembourg
Banque Lombard Odier & Cie SA
Banque Pictet & Cie SA

Barak Capital G.T. LTD.
Barclays Bank Ireland Plc
Barclays Bank PLC
Barclays Capital Securities Ltd.
Basler Kantonalbank
Bayerische Landesbank
BCS Prime Brokerage Limited
Belfius Banque SA
Berner Kantonalbank AG
Bethmann Bank AG
BGC Brokers L.P.
Blue Fire Capital LLC
Bluefin Capital Management, LLC
BNP Paribas
BNP Paribas (Suisse) SA
BNP PARIBAS Arbitrage SNC
BNP Paribas Fortis SA/NV
BNP Paribas S.A. Niederlassung Deutschland
BNP Paribas Securities Services S.C.A. Zweigniederlassung Frankfurt
Boerboel Trading L.P.
BofA Securities Europe SA
BRED Banque Populaire
BSMA Limited
CACEIS Bank SA
Caixabank S.A.
Cantor Fitzgerald Europe
Capital Fund Management
Capital Futures Corp.
Capital MarketsTrading UK LLP
Capital Ventures International
Capitalead Pte. Ltd.
Cast Trading L.P.
Centercross B.V.
China Construction Bank Corporation Niederlassung Frankfurt
China Xin Yongan Futures Company Limited
Citadel Securities (Europe) Ltd.
Citadel Securities GCS (Ireland) Limited
Citigroup Global Markets Europe AG
Citigroup Global Markets Limited
CM Capital Markets Bolsa S.A. A.V.
CN FIRST INTERNATIONAL FUTURES LIMITED
Commerzbank AG
Concord Futures Corp.
Contech LP
Coöperatieve Rabobank U.A.
Corner Banca SA

Corretaje e Información Monetaria y de Divisas, Sociedad de Valores SA
Credit Agricole Corporate and Investment Bank
Crédit Industriel et Commercial
Credit Suisse (Schweiz) AG
Credit Suisse AG
Credit Suisse Bank (Europe) SA.
Credit Suisse International
Criterion Arbitrage & Trading BV
CSC Futures (HK) Limited
CTC London Limited
Cunningham Commodities LLC
D. E. Shaw Asymptote Portfolios LLC
Da Vinci Derivatives B.V.
Daiwa Capital Markets Europe Limited
Danske Bank A/S
De Riva Asia Limited
DekaBank Deutsche Girozentrale
Deutsche Bank AG
Directa SIM
Dom Group AG
Donner & Reuschel Aktiengesellschaft
Dorman Trading L.L.C.
DRW Europe B.V.
DRW Europe Derivatives B.V.
DRW Global Markets Ltd
DRW Investments (UK) Limited
DRW Investments LLC
DRW Singapore Pte Ltd
DV Trading LLC
DZ BANK AG Deutsche Zentral-Genossenschaftsbank
DZ Privatbank S.A.
E D & F Man Capital Markets MENA Limited
Eagle Labs (HK) Limited
Eagle Seven LLC
ED & F Man Capital Markets Ltd
EFG Bank AG
Epoch Capital Pty Ltd
Equita Societa Di Intermediazione Mobiliare SPA
Erste Group Bank AG
Exane Derivatives
Exane S.A.
FCT Europe Limited
Fenics GO Holdings Limited
Fermion Investments Limited
Financial Market Engineering Limited
FinecoBank Banca Fineco S.p.A

Finovesta GmbH
flatexDEGIRO Bank AG
Flow Traders Asia Pte Ltd
Flow Traders B.V.
Flow Traders U.S. LLC
Freeman Commodities Limited
Fubon Futures Co., Ltd
G. H. Financials Ltd.
Gallardo Securities Limited
Gelber Coöperatief U.A.
Gelber Group LLC
Geneva Ireland Financial Trading Ltd.
GFI Securities Ltd.
Global Execution Limited
Goldman Sachs Bank Europe SE
Goldman Sachs International
Grammont Finance SA
GTS Securities Europe Ltd
Hamburg Commercial Bank AG
Hamburger Sparkasse AG
Hard Eight Futures LLC
Hardcastle Trading AG
Hauck & Aufhäuser Privatbankiers AG
HC Technologies LLC
Headlands Technologies Europe B.V.
Headlands Technologies LLC
HGNH INTERNATIONAL FUTURES CO. LIMITED
HNK ALPHA PTE. LTD.
HPC S.A.
HRTEU Limited
HSBC Bank plc
HSBC Continental Europe
HSBC Trinkaus & Burkhardt AG
Hudson River Trading Europe Ltd.
IBKR Financial Services AG
IBROKER GLOBAL MARKETS, S.V., S.A.
IBVV Trading DMCC
ICAP CORPORATES LLC
IMC Trading B.V.
ING Bank N.V.
Ingensoma Arbitrage PTE LTD
Interkapital vrijednosni papiri d.o.o.
Intermonte SIM S.p.A.
Intesa Sanpaolo S.p.A.
Invest Banca SPA
J.P. Morgan AG

Jane Street Capital, LLC
JB DRAX HONORE (UK) LIMITED
Jefferies GmbH
Jefferies International Ltd.
Joh. Berenberg Gossler & Co. KG
Jump Trading Europe B.V.
Jump Trading Futures LLC
Jump Trading Pacific Pte Ltd
KBC Bank N.V.
Kemp Trading B.V. ta Nino Options
Kepler Chevreux (Suisse) SA
Kerdos Investment-AG TGV
KGI Futures Co. Ltd.
Korea Investment & Securities Co. Ltd.
Kreissparkasse Köln
Kutxabank S.A.
Kyte Broking Limited
Landesbank Baden-Württemberg
Landesbank Berlin AG
Landesbank Hessen-Thüringen Girozentrale
Lang & Schwarz AG
Lang & Schwarz TradeCenter AG & Co. KG
Leonteq Securities AG
Liquid Capital Australia Pty. Ltd.
Liquid Capital Markets Ltd.
LR Financial LLC
M.M. Warburg & CO (AG & Co.) Kommanditgesellschaft auf Aktien
Macquarie Bank Europe Designated Activity Company
Mako Derivatives Amsterdam B.V.
Mako Financial Markets Partnership LLP
Mako Global Derivatives Partnership LLP
Marex Financial
Marex North America LLC
Marex Spectron Europe Limited
Mariana UFP LLP
Market Securities (FRANCE) SA
Market Wizards BV
Maven Derivatives Amsterdam B.V
Maven Europe Limited
Mediobanca Banca di Credito Finanziario S.p.A
Melanion Volatility Fund
Mercury Derivatives Trading Limited
Merrill Lynch International
Method Investments & Advisory LTD
Mint Tower Capital Management B.V.
Mizuho Securities USA LLC

MMX Trading B.V.
Morgan Stanley & Co. International PLC
Morgan Stanley Europe SE
Mosaic Finance SAS
MUFG Securities (Europe) N.V.
MUFG Securities EMEA plc
National Bank of Greece SA
Natixis
Natwest Markets NV
Natwest Markets Plc
NH FUTURES CO. LTD.
Nomura Financial Products Europe GmbH
Nomura International plc.
Norddeutsche Landesbank - Girozentrale
Nordea Bank Abp
NRW.BANK
Nyenburgh Holding B.V.
ODDO BHF Aktiengesellschaft
ODDO BHF SCA
Old Mission Capital, LLC
Optiver Australia Pty Limited
Optiver V.O.F.
Panthera Investment GmbH
Phillip Capital Inc.
PNT Financial LLC
Prime Trading, LLC
Q1E LP
Quant.Capital Verwaltungs GmbH
QuantRes Fund SPC
Qube Research & Technologies Limited
Quintet Private Bank (Europe) S.A.
R.J. O'Brien Limited
R.J.O Brien France S.A.S.
Radix Trading Europe B.V.
Radix Trading LLC
Raiffeisen Bank International AG
Raiffeisen Centrobank AG
Raiffeisenlandesbank Oberösterreich Aktiengesellschaft
RBC Capital Markets (Europe) GmbH
RBC Europe Limited
RCUBE ASSET MANAGEMENT
RSJ Securities a.s.
Saccade Capital Limited
Scotiabank Europe Plc
Sea Cliff Investments Limited
Sequoia Capital LLP

SIB (Cyprus) Limited
Sigma Broking Limited
Skandinaviska Enskilda Banken AB
Société Générale
Sparkasse Pforzheim Calw
Squarepoint Master Fund Limited
SSW-Trading GmbH
St. Galler Kantonalbank AG
Star Beta Pty Ltd
StoneX Financial Europe S.A.
StoneX Financial Inc.
StoneX Financial Ltd
Sucden Financial Limited
Sunrise Futures LLC
Susquehanna International Securities Ltd.
Swedbank AB
Swissquote Bank S.A.
Tanius Technology LLC
Tensor Technologies AG
Teza Capital Management LLC
TFS Derivatives HK Ltd
TFS Derivatives Ltd.
Tibra Trading Europe Limited
TMG Trading FZE
Tower Research Capital Europe B.V.
Tower Research Capital Europe Limited
TP ICAP (Europe) SA
TP ICAP Markets Limited
Tradegate AG Wertpapierhandelsbank
TradeLink LLC
TradeLink Worldwide Limited
TradeWeb Europe Ltd
Tradition Securities and Derivatives Inc
Tradition Securities and Futures S.A.
Transtrend B.V.
TTG Capital Limited
Tullett Prebon (Securities) Limited
Tullett Prebon Financial Services LLC
Tyler Capital Ltd.
UBS AG
UBS Europe SE
UniCredit Bank AG
UniCredit S.p.A.
Vallum Trading LLC
Vantage Capital Markets HK Limited
Vantage Capital Markets LLP

Vatic Fund I LLC
Vectalis
Vector Trading LLC
Vegasoul Opus Fund SPC High Street Segregated Portfolio
Virtu Financial Ireland Limited
Virtu Financial Singapore Pte. Ltd.
Volatility Performance Fund SA
VOLKSBANK WIEN AG
Vortex Street Fund Limited
VTB Capital plc
WEBB Traders B.V.
Wedbush Securities Inc.
Wells Fargo Securities International Limited
Wells Fargo Securities, LLC
WH Trading LLC
Whitney Capital Series Fund LLC
Wolfgang Steubing AG Wertpapierdienstleister
Xconnect Market Maker LLP
XConnect Trading Limited
XR Trading EU B.V.
XR Trading LLC
XTX Markets Limited
XTX Markets SAS
Yuanta Futures Co. Ltd.
Zürcher Kantonalbank

B Data Cleaning

In this section we describe details on the data cleaning procedure and discuss our approach in comparison to alternative approaches.

Between May 6, 2013 and May 13, 2013, Eurex migrated its products to its T7 trading architecture. Euro STOXX 50 futures were migrated on May 8, 2013. From this day onward, there is imprecision (or "noise") in the timestamps. In particular, orders that appear to be executed against each other are not necessarily recorded at the same timestamp. Rather, orders are recorded at consecutive timestamps, with the difference usually being within a few tens of milliseconds. As a result, for each timestamp, buying and selling volume as well as aggressive and non-aggressive volume do not necessarily satisfy market clearing. As our analysis focuses on continuous trading only (Section 3) and we are inferring continuous trading periods from the data, we have to deal with the imprecision in the timestamps in an efficient manner.

We clean the data using an event-time approach that is akin to the methodology of Aquilina, Budish, and O'Neill (2021), Ernst (2020), and Ernst, Sokobin, and Spatt (2021). A natural way to solve the problems arising from the imprecision in the timestamps is grouping trade records with the same execution price that are recorded closely together. This can be addressed in both clock time and event time. A clock time approach, however, comes with the caveat that trades that are executing against each other and recorded in different intervals, for example seconds, are not grouped together and thus neither of the trades would enter the main analysis.[24] The problem can be addressed by allowing volumes of market and limit order to deviate up to a threshold within each time interval. This threshold, however, is arbitrary and hard to infer from the data and the classification remains noisy.

An event time approach offers a tractable and more precise solution. Therefore, trades that are recorded at the same price within a short period of time are grouped together. Then, the total volume of limit orders and the total volume of market orders over these trades are computed. If the total volume of limit orders equals the total volume of market orders, the trades are labeled "continuous trading" and included in the main analysis. The algorithm for classifying trades runs from the start of each trading day. For each trade s_0 that is executed at price p_i, a time window starting with that trade record is initialized. All following trades $s_1, s_2, \dots$ that are executed within the time window at the same trading price p_i are grouped together and assigned the same time. Once a trade s_{break} is executed at the same trading price p_i but does not fall within the time window, a new time window starting from that trade record is defined. Again, all trades that occur within the time window at the same execution price are grouped together with this trade. This procedure continues until the end of the trading day.

The only parameter that has to be chosen is the length of the time window. In general, choosing the window length trades off two factors. On the one hand, choosing a longer window length assures that all corresponding trades are grouped together even if there is substantial noise in the timestamps and high trading activity, that might further delay recording of some of the trades. On the other hand, by choosing a shorter window only trades that were actually executing against each other are captured. If the window length is chosen too long and the volume of market orders and limit orders does not equal, a substantial volume does not enter the main analysis. We consider the possibilities of 100ms and 500ms (as well as 2s and 4s for robustness checks). In our particular dataset, inspection reveals that the imprecision is usually within the magnitude of a few tens of microseconds. Thus, in comparison to the clock-time approach discussed before, our approach allows to determine the appropriate choice to the parameter based on the data.

Another alternative to clean the data is by removing opening and closing auctions. This does however not account for potential intraday auctions that are held and thus have to be identified differently. Also, the identification of auctions is not clear cut due to the "noise" in the timestamps.[25]

---

[24] Suppose, for example trades are grouped by seconds. The first trade is recorded at $t.900000$ with t denoting seconds, and the next is recorded at $t + 1.100000$. In an clock time approach, market clearing would not be satisfied for either of the seconds.

[25] There are several ways to do so. First, opening auctions can be defined as the first trade record on each trading day and closing auctions as the last trade record on each trading day.

C State Space Representation

In this Section we show how the state space model presented in Section 4 can be mapped into the standard linear state space model form. Our state space model is given by

$$p_t = m_t + s_t \tag{9}$$

$$m_t = m_{t-1} + \sum_{s\in\mathcal{S}} \gamma_s \tilde{x}_{s,t} + w_t \tag{10}$$

$$s_t = \phi s_{t-1} + \sum_{s\in\mathcal{S}} \delta_s x_{s,t} + \varepsilon_t \tag{11}$$

with $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. The standard linear state space model is given by

$$\mathbf{y}_s = \mathbf{Z}_s \boldsymbol{\alpha}_s + \boldsymbol{\epsilon}_s, \tag{12}$$

$$\boldsymbol{\alpha}_{s+1} = \mathbf{T}_s \boldsymbol{\alpha}_s + \mathbf{R}_s \boldsymbol{\eta}_s, \tag{13}$$

with index $s = 1, \ldots, S$, and the disturbances $\boldsymbol{\epsilon}_s \sim \mathcal{N}(0, \mathbf{H}_s)$ and $\boldsymbol{\eta}_s \sim \mathcal{N}(0, \mathbf{Q}_s)$, following the notation of Durbin and Koopman (2012).

We follow Hamilton (1986) and include exogenous variables in the state vector. We collect the variables γ_s for $s \in \mathcal{S}$ in the $S \times 1$ vector $\boldsymbol{\gamma}$ and the variables δ_s for $s \in \mathcal{S}$ in the $S \times 1$ vector $\boldsymbol{\delta}$. Similarly, we collect order flow in the $S \times 1$ vector $\mathbf{x}_t$ and innovations in order flow in the $S \times 1$ vector $\tilde{\mathbf{x}}_t$. Note that the dimension S of the vectors depends on which order flow variables are included, as discussed in Secton 4.2. Then we obtain

$$\mathbf{y}_s = p_{t-1}, \tag{14}$$

$$\boldsymbol{\alpha}_s = (m_{t-1}, s_{t-1}, \boldsymbol{\gamma}'_{t-1}, \boldsymbol{\delta}'_{t-1})', \tag{15}$$

$$\boldsymbol{\eta}_s = (w_t, \varepsilon_t)', \tag{16}$$

It is set $\mathbf{H}_s \to \mathbf{0}$ and thus $\boldsymbol{\epsilon}_s = 0$. Then the design matrix[26] is given by

$$\mathbf{Z}_s = \begin{array}{cc} \begin{matrix} 1 & 1 & S & S \end{matrix} & \\ \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} & 1 \end{array} \tag{17}$$

This is, however, only feasible for the period before May 8, 2013, as after the first recorded trade is not necessarily the opening auction and opening auctions were occasionally recorded at several timestamps.

In principle, the aggressor flag for auction trades indicates that these trades pertain to a passive order. This suggests identifying auctions as timestamps for with only passive trades are recorded. In practice, there are timestamps around the (potential) opening or closing auctions for which trades with an aggressor flag indicating an aggressive order as well as trades with an aggressor flag indicating a passive order are recorded. Even if market clearing holds taking all orders into account, the volume of aggressive orders does not equal the volume of passive orders for these timestamps (this is, for example, the case on March 11, 2011). From the trade data, it is not possible to identify against which limit orders this market order was executed. Since the account flag is of first-order importance for the main analysis, such an identification would however be necessary if the market order were to be included.

[26] For all matrices we denote the dimensions in the first row and last column

Furthermore, the transition matrix

$$\mathbf{T}_s = \begin{array}{cc} \begin{array}{cccc} 1 & 1 & S & S \end{array} & \\ \begin{bmatrix} 1 & 0 & \tilde{\mathbf{x}}_t' & 0 \\ 0 & \phi & 0 & \mathbf{x}_t' \\ 0 & 0 & \mathbf{I}_S & 0 \\ 0 & 0 & 0 & \mathbf{I}_S \end{bmatrix} & \begin{array}{c} 1 \\ 1 \\ S \\ S \end{array} \end{array} \tag{18}$$

is time-varying, with $\mathbf{I}_S$ being and $S \times S$ identity matrix. The selection matrix is given by

$$\mathbf{R}_s = \begin{array}{cc} \begin{array}{cc} 1 & 1 \end{array} & \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{array}{c} 1 \\ 1 \\ S \\ S \end{array} \end{array} , \tag{19}$$

and the state covariance matrix by

$$\mathbf{Q}_s = \begin{array}{cc} \begin{array}{cc} 1 & 1 \end{array} & \\ \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} & \begin{array}{c} 1 \\ 1 \end{array} \end{array} . \tag{20}$$

In the state space model given by (14) – (20), the parameters on order flow and innovations in order flow are estimated by state estimation. This model is used as second step when determining the starting values, as described in Section 4. For implementation, state estimation is replaced by parameter estimation such that the coefficients on the order flow variables are estimated by maximum likelihood estimation. Therefore, a constant 1 is assigned to exogenous variables. Both the exogenous variables as well as the parameters are included in the system matrices. This yields

$$\mathbf{y}_s = p_{t-1}, \tag{21}$$

$$\boldsymbol{\alpha}_s = (m_{t-1}, s_{t-1}, \boldsymbol{\iota}_S')', \tag{22}$$

$$\boldsymbol{\eta}_s = (w_t, \varepsilon_t)', \tag{23}$$

with $\boldsymbol{\iota}_S$ being an $S \times 1$ vector of ones. Again, it is set $\mathbf{H}_s \to \mathbf{0}$ and thus $\boldsymbol{\epsilon}_s = 0$. The design matrix is unchanged and given by (17). Furthermore, the transition matrix changes and is now given by

$$\mathbf{T}_s = \begin{array}{cc} \begin{array}{ccc} 1 & 1 & S \end{array} & \\ \begin{bmatrix} 1 & 0 & \boldsymbol{\gamma}' diag(\tilde{\mathbf{x}}_t) \\ 0 & \phi & \boldsymbol{\delta}' diag(\mathbf{x}_t) \\ 0 & 0 & \mathbf{I}_S \end{bmatrix} & \begin{array}{c} 1 \\ 1 \\ S \end{array} \end{array} . \tag{24}$$

is time-varying. The selection matrix is given by

$$\mathbf{R}_s = \begin{array}{cc} \begin{array}{cc} 1 & 1 \end{array} & \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} & \begin{array}{c} 1 \\ 1 \\ S \end{array} \end{array} , \tag{25}$$

and the state covariance matrix is unchanged and given by (20).

D Implementation of Estimation

This section describes details on the implementation of the model estimation described in Section 4. The model is estimated by maximum likelihood estimation and the Kalman filter recursion is used to evaluate the likelihood function. For the maximum likelihood estimation, starting values are required. Here, we describe how these starting values are obtained. Also, we discuss restrictions that we impose on the parameters for estimation.

Starting values for the maximum likelihood estimation are obtained in three steps. First, a simple state space model excluding order flow and innovations in order flow is estimated. Thus, the model is given by

$$p_t = m_t + s_t \tag{26}$$

$$m_t = m_{t-1} + w_t \tag{27}$$

$$s_t = \phi s_{t-1} + \varepsilon_t \tag{28}$$

with $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Since the model is estimated with log prices, return variances and autocovariances can be expressed as a function of the model's parameters. It can be shown that the autocovariances of the log returns are

$$\gamma(0) = \sigma_w^2 + \frac{2}{1+\phi}\sigma_\varepsilon^2 \tag{29}$$

$$\gamma(1) = \frac{\phi - 1}{1+\phi}\sigma_\varepsilon^2 \tag{30}$$

$$\gamma(2) = \frac{\phi(1-\phi)}{1+\phi}\sigma_\varepsilon^2. \tag{31}$$

Using this, starting values for the maximum likelihood estimation are given by

$$\phi = \frac{\gamma(2)}{\gamma(1)} \tag{32}$$

and using the starting value for ϕ it follows for σ_ε^2

$$\sigma_\varepsilon^2 = \gamma(1)\frac{1+\phi}{\phi - 1} \tag{33}$$

and finally for σ_w^2

$$\sigma_w^2 = \gamma(0) - \frac{2}{1+\phi}\sigma_\varepsilon^2. \tag{34}$$

Using these starting values, the reduced form state space model given by (26) – (28) is estimated. We estimate parameters for σ_ε and σ_w rather than for σ_ε^2

and σ_w^2. Thus, our starting values are given by the square root of (33) and (34), respectively. The estimates from the reduced form state space model are stored and used as starting values for estimating the full state space model, with parameter estimation replaced by state estimation. The model is discussed in Appendix C, (14) – (20). This model introduces the parameters on order flow and innovations in order flow as latent state variables given by

$$\delta_t = \delta_{t-1} \tag{35}$$

and

$$\gamma_t = \gamma_{t-1}. \tag{36}$$

For estimation, the states for the parameters are initialized as diffuse by setting a prior variance of κ with $\kappa \to \infty$.

After estimation, the estimated parameters for σ_ε, σ_w, and ϕ as well as the state estimates for δ and γ are stored and used as starting values for estimating the full state space model (3) – (5) by maximum likelihood.

The model contains parameters that are required to be positive (the standard deviations of the error terms in the state and observation equation) or to be in the interval $[-1, 1]$ (the autocorrelation in pricing errors). To ensure that these restrictions are satisfied in estimation, we transform the restricted parameters before optimization by applying the function $f(x)$ to parameter x. After optimization, we untransform the parameters by applying the function $g(y)$ to the transformed parameter y. For the variance parameters, we use

$$f(x) = x^2 \tag{37}$$

and

$$g(y) = y^{\frac{1}{2}}. \tag{38}$$

For the autocorrelations in pricing errors that are required to be in the interval $[-1, 1]$, the functions are given by

$$f(x) = tanh(x) \tag{39}$$

and

$$g(y) = arctanh(y). \tag{40}$$

E Frequency Domain Analysis

In this section we provide a discussion of the frequency domain decomposition of the trading profits presented in Section 5.1. The exposition is based on Hasbrouck and Sofianos, 1993, Bloomfield, 2004, Hau, 2001, and Menkveld, 2013.

Our goal is to decompose gross trading revenues by principals and agents into profits at different frequencies. Mark-to-market gross trading revenues for time t are given by

$$\Pi_t = x_{t-1}\Delta p_t \tag{41}$$

where x_{t-1} denotes order flow or the net trading balance at time $t-1$ and Δp_t is the price change from $t-1$ to t (Hasbrouck and Sofianos, 1993). Intuitively, the frequency domain decomposition allows to decompose trading revenues into revenues at different frequencies. If order flow and price changes are in-phase at a certain frequency, this results in positive trading profits. In contrast, if order flow and price changes are out-of-phase, this results in negative trading profits.

The decomposition is performed using Fourier transforms. For an equally spaced time series of length T, Fourier frequencies are given by $\omega_k = 2\pi k/T$ with $k = 0, 1, \ldots, T-1$. Then, the Fourier transform of X_t is given by

$$J_X(\omega_k) = \frac{1}{T}\sum_{t=1}^{T} X_t \exp(-i\omega_k t)$$

with $i = \sqrt{-1}$ and $J_X(\omega_k)$ being the Fourier component of X_t at frequency ω_k. The series X_t can be recovered using the inverse transform

$$X_t = \sum_{k=0}^{T-1} J_x(\omega_k) exp(i\omega_k t). \tag{42}$$

From this it can be seen that we can express the original time series X_t as the sum of $T-1$ frequency components. The cross product of the series X_t and Y_t is given by

$$\Pi = \frac{1}{T}\sum_{t=1}^{T} X_t Y_t \tag{43}$$

$$= \sum_{k=0}^{T-1} J_X(\omega_k)\overline{J_Y(\omega_k)} \tag{44}$$

where $\overline{J_Y(\omega_k)}$ is the complex conjugate of $J_Y(\omega_k)$. Given that the cospectrum at frequency ω_k is given by

$$Co_{XY}(\omega_k) = J_X(\omega_k)\overline{J_Y(\omega_k)}, \tag{45}$$

we can decompose the contribution of subsets of the Fourier frequencies $\omega_k = 2\pi k/T$ with $k = 0, 1, \ldots, T-1$ to the overall trading revenues.

We implement the analysis in with the matplotlib library in python.

F Reduced Form Results

In this section we present a reduced form version of our state space model omitting order flow and innovations in order flow, based on hourly data. We

Table 8: Estimation results of the reduced form state space model

We present estimation results for the reduced form state space model with autocorrelation in transitory pricing errors and omitting the order flow series given by

$$
\begin{aligned}
p_t &= m_t + s_t \\
m_t &= m_{t-1} + w_t \\
s_t &= \phi s_{t-1} + \varepsilon_t
\end{aligned}
$$

at an hourly frequency. Standard deviations are in *bp* per hour. Robust standard errors are computed. * denotes significance at the 10% level, ** denotes significance at the 5% level, and *** denotes significance at the 1% level.

Variable	Estimate
σ_w	22.1366***
	(0.4242)
σ_ε	21.4529***
	(0.4564)
ϕ	0.8650***
	(0.0088)
#Observations	32277

obtain estimates of the efficient price series and pricing errors as smoothed states from the model and relate the order flow variables to changes in efficient prices and pricing errors. Therefore, we estimate the model

$$
\begin{aligned}
p_t &= m_t + s_t \\
m_t &= m_{t-1} + w_t \\
s_t &= \phi s_{t-1} + \varepsilon_t
\end{aligned}
$$

with $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. As described in Section 4, the identifying assumption for estimation is that w_t and ε_t are uncorrelated. Estimation results are presented in Table 8.

Then, we relate changes in the efficient price series as well as pricing errors to order flow and innovations in order flow. Both order flow variables are expressed in EUR. We compute correlations between the order flow and price variables over time and plot them with the corresponding confidence intervals in Figures 5 and 6. For each month, quarter, and year in the sample, we compute correlations based on hourly data.[27]

---

[27] To compute the confidence intervals, we first apply a Fisher transformation to the correlation coefficients

$$
\begin{aligned}
z &= \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \\
&= \tanh^{-1}(r)
\end{aligned}
$$

The correlations presented in Figures 5 and 6 are quarterly. The patterns suggest that principals, even if using market orders, trade less in the direction of price pressures than agents. At the same time, agents' limit orders are overall positively correlated with pricing errors in the second half of the sample period. Similar patterns are true for innovations in order flow and changes in the efficient price series. Innovations in agent aggressive flow are positively correlated with changes in efficient prices. The correlations between principal aggressive flow and changes in efficient prices decreases over the sample period to zero. Passive order flow is negatively correlated with changes in the efficient price series over all account roles as well as for agents and principals.

Graphical inspection of the correlations suggests that aggressive agent flow contains information. The positive correlation between pricing errors and aggressive agent flow is both consistent with agents trading in the direction of price pressures – thus demanding liquidity – as well as with prices overreacting. This evidence both gives further motivation for the state space model presented in Section 4, but more than that, it shows the robustness of the results presented in the main section to a reduced form exhibition. Importantly, innovations in order flow and order flow itself are empirically highly correlated. Hence, pricing errors and innovations in efficient prices tend to be correlated. As a result, assuming that the innovations in a simple state space model without order flow variables are uncorrelated is not sufficient. Instead, order flow has to be accounted for as we do in the full results presented in the main section.

---

where r is the correlation. Then, two sided confidence limits are computed as

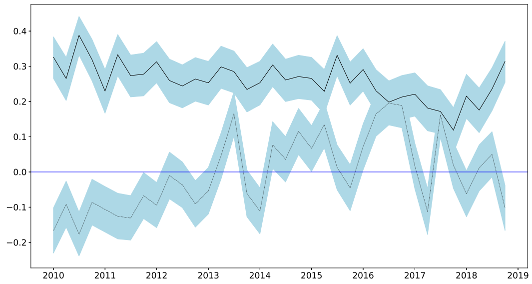
$$z_U = z + z_{1-\alpha/2}\sqrt{\frac{1}{N-3}}$$
$$z_L = z - z_{1-\alpha/2}\sqrt{\frac{1}{N-3}}$$

where N denotes the number of observations used to compute the correlation and $z_{1-\alpha/2}$ is the critical value of the normal distribution at an α significance level. Finally, critical values for the correlation are obtained by transforming the confidence limits

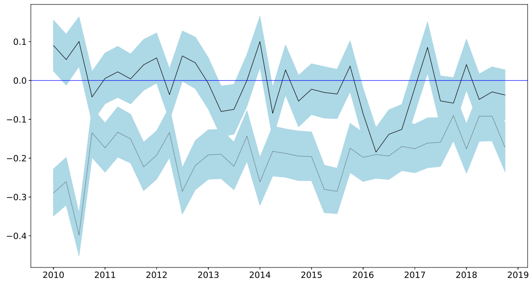
$$\begin{aligned} r_U &= \frac{\exp(2z_U - 1)}{\exp(2z_U + 1)} \\ &= \tanh(z_U) \\ r_L &= \frac{\exp(2z_L - 1)}{\exp(2z_L + 1)} \\ &= \tanh(z_L). \end{aligned}$$

Figure 5: Quarterly correlations between order flow and pricing errors

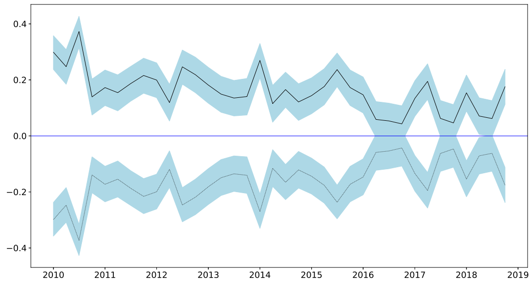
The figure plots quarterly correlations between order flow and pricing errors by account role. The solid line depicts aggressive order flow and the dotted line passive order flow. Blue areas are 95% confidence intervals. Note the different scales of the y-axes.



(a) Agent flow



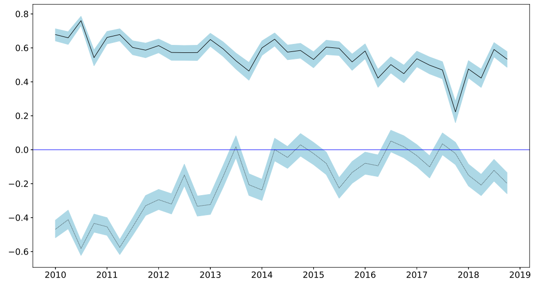
(b) Principal flow



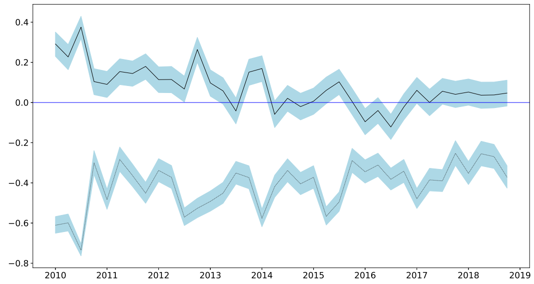
(c) All account roles

Figure 6: Quarterly correlations between innovations in order flow and efficient price changes

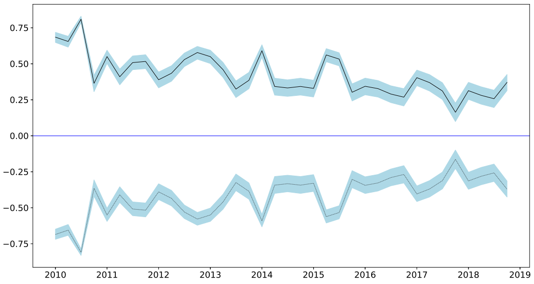
The figure plots quarterly correlations between innovations in order flow and changes in the efficient price series by account role. The solid line depicts aggressive order flow and the dotted line passive order flow. Blue areas are 95% confidence intervals. Note the different scales of the y-axes.



(a) Agent flow



(b) Principal flow



(c) All account roles