PUEB PRESS
Economics and Business Review
https://www.ebr.edu.pl
Vol. 11 (4), 2025: 0–0
https://doi.org/10.18559/ebr.2025.4.2330

Algorithmic trading, liquidity and volatility: Evidence from Poland

Henryk Gurgul[1] Robert Syrek[2]

Abstract

The aim of this paper is to examine the causality between pairs of measures that describe the intensity of algorithmic trading, market liquidity and volatility for selected blue-chip companies from the Warsaw Stock Exchange, which were permanently included in the WIG20 index from January 1, 2020, to August 31, 2023. In the study, both daily and high-frequency intraday data are used. The research is based on fundamental concepts of information theory, namely entropy and transfer entropy. Additionally, Rényi entropy is used to examine the causal relationships between extreme values of the variables. Our results, based on Shannon's transfer entropy, suggest that algorithmic trading affects liquidity and volatility. The main finding is that if the frequency increases, the number of companies for which information transfer is significant also grows. However, this relationship is not observed for extreme values, for which Rényi entropy is applied.

Keywords

- algorithmic trading intensity
- liquidity
- volatility
- transfer entropy

JEL codes: G10, G19

Article received 7 June 2025, accepted 10 November 2025.

Suggested citation: Gurgul, H., & Syrek, R. (2025). Algorithmic trading, liquidity and volatility: Evidence from Poland. *Economics and Business Review*, *11*(4), 0–0. https://doi.org/10.18559/ebr.2025.4.2330

[1] AGH University of Krakow, Department of Applications of Mathematics in Economics, ul. Al. Mickiewicza 30, 30-059 Kraków, Poland, corresponding author, henryk.gurgul@gmail.com, https://orcid.org/0000-0002-6192-2995.

[2] Jagiellonian University, Faculty of Management and Social Communication, Institute of Economics, Finance and Management, ul. prof. St. Łojasiewicza 4, 30-348 Kraków, Poland, robert.syrek@uj.edu.pl, https://orcid.org/0000-0002-8212-8248.

Introduction

New directions in financial research stem from the rapid growth of algorithmic trading (AT). This phenomenon has been observed since the 1990s (Mestel et al., 2018). Trading in financial markets has experienced an extremely strong shift towards AT.

The first studies to be conducted in this field, e.g., by Hendershott et al. (2011), indicated that AT improves liquidity and enhances the informativeness of quotes. However, they did not use the exact AT data, but the rate of electronic message traffic as a proxy for the amount of AT taking place. Although Desagre et al. (2022) observed a general improvement in liquidity on the Euronext stock exchange between 2002 and 2006, they noted that those stocks that are most heavily traded algorithmically show the weakest growth in liquidity and lose the liquidity advantage observed before the advent of AT.

In stable conditions, AT increases liquidity by reducing the bid-ask spread and increasing the depth of the order book. However, it may happen that algorithms cancel orders before they are executed, thus creating apparent liquidity (which happens in conditions of high volatility). Regulators in some countries including the United States, the European Union, China, and Japan have been concerned about the possible negative effects of AT on market quality. Therefore, they have taken measures to limit its expansion.

The first exact high-frequency trading (HFT) data on trades and quotes from 26 high-frequency trading firms across 120 stocks was provided by NASDAQ and covered 2008 and 2009. Using the supplied dataset, Brogaard et al. (2014) examined the role of high-frequency traders (HFTs) in price discovery and price efficiency. NASDAQ's 120 stock sample (with explicit HFT flags) from 2008–2009 has become the most popular dataset for training purposes. Other authors have access to proprietary account-level data on trades produced by AT firms. Most of these first exact AT datasets covered only short sample periods, not more than several months (e.g., for Germany only 13 trading days, Hendershott & Riordan, 2013). We note that "algorithmic" and "high-frequency" are not synonyms, but they are closely related notions. All high-frequency data is used in an algorithmic way, but not all algorithmic processing involves high-frequency data. HFT is a specialised subset of algorithmic trading.

The research is based on unique intraday data for the WSE, which is the largest stock exchange in Central and Eastern Europe. To the best of the authors' knowledge, the paper is the first in the literature to empirically establish the relationships between algorithmic trading intensity (*ATI*), volatility, and liquidity for different data frequencies. Moreover, our data (trades and quotes) contains an algorithmic trade indicator, which allows us to estimate *ATI* directly (without using proxies for AT).

The main purpose of this study is to identify the direction of dependence between *ATI*, market liquidity and volatility for daily and intraday data at different time frequencies for selected blue-chips from the Warsaw Stock Exchange (WSE). Identifying causal relationships can help answer questions such as whether *ATI* has an impact on market quality, which is primarily determined by liquidity and risk (measured, e.g., by volatility). These questions provide a basis for constructing dynamic econometric models with regressors that identify a causal relationship.

Our empirical research utilises information-theoretic methods related to entropy and transfer entropy. The study was conducted for pairs of variables formed from *ATI*, volatility, and liquidity. The concept of Rényi entropy is also used to investigate causality between extreme values of variables. Our results, based on Shannon's transfer entropy, suggest that *ATI* affects both liquidity and volatility.

The rest of the paper is organised as follows: Section 1 is devoted to the literature review and hypotheses development. This is followed by Section 2, which includes an outline of research methodology. In Section 3, the dataset and variables are described. In Section 4, the empirical results are presented and discussed. The final section concludes the paper.

1. Literature review and hypotheses development

The literature indicates that AT typically improves liquidity and reduces the volatility of stock returns. However, there are studies, particularly those concerning emerging stock markets, whose results differ from those obtained for developed stock markets. In emerging or developing markets, some researchers have observed a decrease in liquidity and an increase in volatility as AT's share in stock trading grows. Due to the limited length of this study, however, only a few examples of studies representing these groups of results are provided.

Transfer entropy has been previously employed in the economic context by numerous scholars (Abdi & Ranaldo, 2017; Będowska-Sójka & Kliber, 2021; Brauneis & Mestel, 2018; Diaz & Escribano, 2020; Dionisio et al., 2004; Garma, Klass, 1980; He & Shang, 2017; Leone & Kwabi, 2019; Lesmond, 2005), and especially more recent contributions (Ao & Li, 2024; Banerjee & Nawn, 2024; Lacava et al., 2023; Mestel et al., 2024).

Our paper focuses on the relationships between *ATI*, volatility, and liquidity in the emerging stock market. Empirical research on these relationships requires using different measures of these three variables, and the key works on this topic devote considerable attention to these measures and their properties.

Popular measures of volatility are Garman–Klass (1980) volatility, realised volatility and bi-power variation (Aggarwal &Thomas, 2014), while widely used measures of liquidity include proportional effective spread or quoted spread (Mestel et al., 2018) and realised Amihud illiquidity (Lacava et al., 2023). The authors of the latter paper investigated the theoretical and empirical properties of a refinement of the classic daily Amihud measure. They suggested two measures of realised illiquidity: realised Amihud and high-low Amihud.

Liquidity is commonly measured based on different daily proxies versus benchmarks related to high-frequency data (Abdi & Ranaldo, 2017; Lesmond, 2005). In the literature, a rich body of approaches to approximate liquidity and volatility were proposed by Diaz and Escribano (2020). Moreover, Dionisio et al. (2004) used the concept of mutual information to measure both linear and nonlinear interdependence in financial time series.

The seminal paper by Aggarwal and Thomas (2014) provides evidence of the causal impact of AT on the stability of prices and liquidity. According to these authors, policy makers and regulators often exhibit concerns that the higher level of liquidity is transient because AT exits the market rapidly when unexpected news appears. Their main criticism is that AT causes a higher probability of extreme drops and reversals over a very short period of time during the trading day. The results showed that AT lowered the intraday liquidity risk. It is also demonstrated that higher AT leads to a lower incidence of extreme price movements during the trading day. This paper's contribution lies in its moving towards a causal analysis of the impact of AT upon market quality. The analysis used a high-quality dataset with a long time span and numerous securities. In addition to the well-studied measures of liquidity and volatility, the paper also provides evidence about intra-day ash crashes and intra-day liquidity risk.

In some papers, the correlation between several liquidity proxies and benchmarks is investigated. This stream of research was undertaken and significantly extended by Będowska-Sójka and Kliber (2021). The main aim of their paper was to compare the mutual information shared by various liquidity and volatility estimators within each group separately for a sample period from January 2006 to December 2016. The proxies were computed using either daily data from www.stooq.pl or transaction data from the WSE directly. In this way, daily measures of liquidity and volatility were obtained. The authors found that in terms of their information content, volatility measures are much more coherent, while liquidity ones are more dispersed. The Garman–Klass volatility estimator seemed to be the broadest measure of volatility, while Amihud illiquidity and volatility over volume shared the highest amount of mutual information among liquidity proxies.

Jain et al.'s (2021) findings reflect a new perspective on the impact of algorithmic traders on liquidity provision. These authors examined the importance of AT across a sample of stocks listed on the NYSE between 2001 and 2005, demonstrating that the role of algorithmic traders as liquidity providers declines during periods of high information asymmetry.

Ekinci and Ersan (2022) conducted a study on a sample of 30 blue-chip companies listed on Borsa Istanbul between December 2015 and March 2017. They found a negative impact of high-frequency trading on market quality (high liquidity, narrow bid-ask spread, low transaction costs, efficient price discovery i.e. the speed and accuracy with which the market incorporates new information into asset prices), despite its minor role on Borsa Istanbul. The authors emphasised that the provision of liquidity by entities outside FX markets is significantly reduced as the extent of high-frequency trading increases.

Ramos and Perlin (2020) provided the first evidence of AT reducing liquidity in the Brazilian equities market. Their results were contrary to most studies that found a positive relationship between AT and liquidity. However, this was based not on actual AT data, but on two types of AT proxies. In contrast, Dubey et al. (2021) found conclusively, using Indian data, that a rise in AT led to significant improvements in liquidity and reduced market volatility, especially for large-cap stocks.

In their study, Courdent & McClelland (2022) used a set of AT indicators for South Africa: average trade size, odd-lot volume ratio, and trade-to-order volume ratio. Panel regressions were used to determine the relationship between these indicators and two measures of market quality, namely market liquidity and short-term volatility. The study found a strong positive relationship between market liquidity and average trade size, but the opposite relationship for the other two AT indicators. The study points to a strong positive relationship with short-term volatility. In general, AT has a positive impact on market quality, despite the risk of volatility in some markets. Mestel et al. (2018) found that an increase in the market share of AT causes a reduction in quoted and effective spreads while quoted depth and price impacts are unaffected. These findings are consistent with algorithmic traders, on average acting as market makers.

According to Mestel et al. (2024), the relationship between algorithm intensity and volatility can be both positive and negative. Under normal conditions, high-frequency trading reduces volatility. However, during periods of market shocks, algorithms may react synchronously (e.g., by withdrawing from the market), causing price jumps. High frequency, in turn, can lead to more frequent “mini-flash crashes”. In their empirical paper, the authors studied 144 mini flash crashes on the Austrian stock market between 2011 and 2015. Using panel logit models, they tried to relate mini flash crashes to AT for the Vienna Stock Exchange. The authors addressed endogeneity by using a control function approach, but found no evidence that AT significantly affects mini flash crashes.

Banerjee and Nawn (2024) provide the first direct evidence of the behaviour of proprietary algorithmic traders during the COVID-19 pandemic. In turn, research by Ao and Li (2024) shows that the performance of trading algorithms that refer to directional changes in stock markets may not be as expected.

Arumugam et al. (2023) analysed the impact of trading by algorithmic (AT) and non-algorithmic (NAT) traders on volatility. They also investigated the possible inverse relationship, i.e. the impact of volatility shocks on AT and NAT. ATs are classified as high-frequency traders (HFTs) and buy-side algorithmic traders (BATs). Using spike-resistant volatility estimates, the authors concluded that abnormal directional and non-directional trading by BATs and HFTs increases volatility, while trading by NATs slightly decreases it. One hour after a volatility shock, all traders increase their non-directional trading. The authors reported that BATs engage in more directional trading during volatility shock, while HFTs retreat from such activity.

The ambiguous results presented in the literature demonstrate the difficulty in analysing the relationships between measures of algorithmic / high-frequency trading and liquidity or volatility, indicating that these results may depend not only on the stock markets or data periods but also the explanatory variables used in the analysis.

Based on the literature review, especially approaches and results presented in the papers by Mestel et al. (2018), Aggarwal and Thomas (2014), Będowska-Sójka and Kliber (2021) and Dionisio et al. (2004), we formulate:

Hypothesis 1: There exists a significant pairwise causal dependence between *ATI*, market liquidity and volatility.

The second hypothesis refers to some extent to the Epps effect (Gurgul & Machno, 2017). The Epps effect describes how the correlation between the returns of two different stocks decreases with the length of the interval for which the price changes are measured. By analogy, we suppose that with higher data frequency, the number of companies observed with causal relationships between the economic variables under consideration may increase, i.e. that holds true the conjecture

Hypothesis 2: With the increase in data frequency, the causality patterns are depicted in more companies.

In the case of stock markets, researchers are particularly interested in the relationships, or lack thereof, for extreme values of financial variables, i.e. in the tails of probability distributions (crisis phase or boom phase on the stock markets). These relationships are often more pronounced in this area than in the rest of the distribution domain (Gurgul & Syrek, 2023). This strengthened dependence in the tails of the distributions explains why the occurrence of a crisis in one country implies its rapid spread to other countries (contagion effect). Rényi entropy is used to examine causality in the case of extreme values. We presume that for extreme values, similar relationships exist as those discussed in hypotheses 1 and 2. Therefore, we formulate the following hypothesis:

Hypothesis 3: Extreme values of the variables under consideration exhibit causal relationships more often than for the entire distribution.

2. Research methodology

The linear Granger causality test does not appear to be suitable for testing the proposed research hypotheses, the reason being that the dependencies of economic processes are usually non-linear. Different versions of the linear Granger causality test expand the application possibilities. The Hiemstra–Jones test with Panchenko correction can be used in the case of non-linear causality. The Toda–Yamamoto test allows the examination of causality between variables with different degrees of integration. The description of these tests, conditions of their applicability with reference to source papers can be found in Gurgul et al. (2012) and Gurgul and Lach (2012), for example. However, time series data, especially those with high-frequency data, still do not simultaneously meet all the conditions for the applicability of individual tests. Given the limitations of the article's length, the authors decided to use only entropy-based methods.

The connections between causality and transfer entropy are discussed in papers, e.g., by Hlaváčková-Schindler et al. (2007) and Syczewska and Struzik (2015).

To test the research hypotheses, we use the transfer entropy methodology based on the properties of stationary Markov processes and information theory, which measures the directional transfer of information between two variables. It shows the extent to which past values of one variable influence the future values of the other variable. Transfer entropy expresses how much information one variable transfers to another. This statistic is used in various fields, primarily in biology, engineering, and finance (to analyse the influence and predict market behaviour).

Most economists use the Diebold–Yilmaz (dynamic connectedness, based on VAR model) methodology in their analysis (Diebold & Yilmaz, 2012, 2014). A prerequisite is the correct estimation of the VAR model (selection of lags, stationarity of variables), which can prove difficult in practice. With a larger number of variables, the VAR becomes very complex, and parameter estimation can be unstable ("curse of dimensionality"). The Diebold–Yilmaz procedure helps us to understand how different markets are connected and how changes in one market can affect others, which is useful in risk management and investment strategies. This procedure evaluates connections, not directional interactions. However, since we do not focus on simultaneous relationships between different markets but on causal relationships with specific directions within the same market, the more appropriate methodology in our study is transfer entropy.

Suppose that X and Y are stationary Markov processes of order k and l, respectively:

$$P_X(X_{t+1} \mid X_t, \ldots, X_{t-k+1}) = P_X(X_{t+1} \mid X_t, \ldots, X_{t-k}) \tag{1}$$

and

$$P_Y(Y_{t+1} \mid Y_t, \ldots, Y_{t-l+1}) = P_Y(Y_{t+1} \mid Y_t, \ldots, Y_{t-l}) \tag{2}$$

The average number of bits needed to encode the observation of X in time $t+1$, once the previous k values are known (Behrendt et al., 2019; Będowska--Sójka & Kliber, 2021) and the average number of bits needed to encode the observation of Y once the previous l values are known, are given by:

$$h_X(k) = -\sum_{x_{t+1},\, x_t^{(k)}} P\left(X_{t+1}, X_t^{(k)}\right) \log_2 P\left(X_{t+1} \mid X_t^{(k)}\right) \tag{3}$$

$$h_Y(l) = -\sum_{y_{t+1},\, y_t^{(l)}} P\left(Y_{t+1}, Y_t^{(l)}\right) \log_2 P\left(Y_{t+1} \mid Y_t^{(l)}\right) \tag{4}$$

where $X_t^{(k)} = (X_t, \ldots, X_{t-k+1})$ and $Y_t^{(l)} = (Y_t, \ldots, Y_{t-l+1})$.

In the bivariate case, information flows from process to process is measured by quantifying the deviation from the generalised Markov property:

$$P\left(X_{t+1} \mid X_t^{(k)}\right) = P\left(X_{t+1} \mid X_t^{(k)}, Y_t^{(l)}\right) \tag{5}$$

relying on the Kullback–Leibler distance. The formula for Shannon transfer entropy is given by:

$$T_{Y \to X} = \sum_{x_{t+1},\, x_t^{(k)},\, y_t^{(l)}} P\left(X_{t+1}, X_t^{(k)}, Y_t^{(l)}\right) \log_2 \frac{P\left(X_{t+1} \mid X_t^{(k)}, Y_t^{(l)}\right)}{P\left(X_{t+1} \mid X_t^{(k)}\right)} \tag{6}$$

and measures the information flow from Y to X (whereas similarly defined $T_{X \to Y}$ measures the information flow from Y to X). The difference between these measures gives information about the dominant direction (net information flow). Transfer entropy can also be based on Rényi entropy to model the dependencies between events with low probability, i.e. in the tails of distributions. For the discrete random variable $X = \{x_i\}_{i=1}^{n}$ Rényi entropy is defined as:

$$H_X^q(X) = \frac{1}{1-q}\log_2 \sum_{i=1}^{n} [p_i]^q \quad (7)$$

where q is a positive weighting parameter and $p_i = P(X = x_i)$.

Rényi entropy converges to Shannon entropy as $q \to 1$. If $0 < q < 1$, then events that have a low probability of occurring receive more weight, while for $q > 1$, the weights induce a preference for outcomes X with a higher initial probability.

3. Data and variable description

The data, which was made available by the Warsaw Stock Exchange, was subject to a non-disclose agreement prohibiting its being sharing with third parties. It only indicates whether a given transaction was the result of AT. It does not include information on the known types of AT strategies, such as arbitrage, mean reversion, market timing, trend following, momentum, high frequency trading, index found rebalancing and breakout trading.

Moreover, in the case of (a few) other stock exchanges that register exact AT data, e.g., the Vienna Stock Exchange, AT types are not provided. Due to data availability, the dataset contains the trades and quotes of 12 companies (PZU, KGH, CDR, SPL, JSW, OPL, PKO, PKN, PEO, CPS, PGE, DNP) which were permanently included in WIG20 index in the period from 1 January 2020 to 31 August 2023. We use data that are time-stamped to the microsecond and contain an algorithmic trade indicator which specifies whether or not transactions were executed as a result of AT (defined in Article 4(1)(39) of Directive 2014/65/EU)). We consider trades and quotes only for a continuous trading phase and compute commonly used measures of liquidity and volatility along with the *ATI*. Most existing papers on Polish and foreign stock markets are based on these proxies.

Following Aggarwal and Thomas (2014), we used *ATI* over a fixed interval (day, 1-, 5- and 10-minutes frequencies) for each company, which is expressed as:

$$ATI_t = \frac{TTV_t^{AT}}{TTV_t} \quad (8)$$

where TTV_t^{AT} is the sum of the values of algorithmic trades and TTV_t is the sum of the values of all trades.

For all frequencies under consideration, we calculated Garman–Klass volatility (Garman & Klass, 1980) as:

$$GK_t = \sqrt{0.5\left[\log\left(\frac{H_t}{L_t}\right)\right]^2 - \left(2\log(2)-1\right)\cdot\left[\log\left(\frac{C_t}{O_t}\right)\right]^2} \tag{9}$$

where O_t, H_t, C_t, L_t are the open, highest, close and lowest prices in a given period (this measure is computed for either a one-day period or intradaily).

We also incorporated two popular measures of daily volatility: square root of realised variance and bi-power variation given by the following formulas:

$$RV_t = \sqrt{\sum_{i=1}^{m} r_{t,i}^2} \tag{10}$$

$$BPV_t = \sqrt{\sum_{i=2}^{m} \left|r_{t,i-1}\right| \cdot \left|r_{t,i}\right|} \tag{11}$$

where $r_{t,i}$ are five-minute returns within the day. Parameter is equal to 94, since continuous trading on the WSE is between 9:00 a.m. and 4:50 p.m.

For each frequency we computed two popular measures of liquidity, namely, proportional effective spread and proportional quoted spread of the form:

$$EFF_t = \frac{2 \cdot D_t \cdot (P_t - MID_t)}{MID_t} \tag{12}$$

$$QUO_t = \frac{ASK_t - BID_t}{MID_t} \tag{13}$$

where D_t is equal to 1 (–1) if trade at time t was buy (sell) $MID_t = \frac{BID_t + ASK_t}{2}$ is midpoint at time t and P_t is trade price. For each frequency, we computed the size-weighted effective spread and time-weighted quoted spread (Mestel et al., 2018).

In addition, for daily frequency, we computed the recently introduced measure of illiquidity (Lacava et al., 2023), that is, realised Amihud illiquidity:

$$RAI_t = \frac{RPV_t}{V_t} \tag{14}$$

where $RPV_t = \sum_{i=1}^{m} \left|r_i\right|$ is realised absolute variation based on log-returns and $V_t = \sum_{i=1}^{m} v_i$ is trading volume generated in the same period. We used five-minute returns, which gave $m = 94$.

4. Empirical results

4.1. Descriptive statistics

For all intraday time series of liquidity, volatility and *ATI* we reject the null of lack of autocorrelation (Ljung–Box test) and observe a departure from normality (Jarque–Bera test). For two companies' (CDR and PEO) daily series of *ATI*, we do not reject the null in the Ljung–Box test at the 10% significance level.

We analysed all series and observed that the autocorrelation function decays more slowly than exponential decay. Moreover, for intraday data (especially for liquidity and volatility measures) we observe intraday seasonality. In Figure 1, we present as an example the autocorrelation function of proportional quoted spread, proportional effective spread, volatility and *ATI* for the company PGE for five-minute intervals. Intraday seasonality can be easily noticed (investors are characterised by different activity at different moments of the session, especially greater at the beginning and end of the trading day).

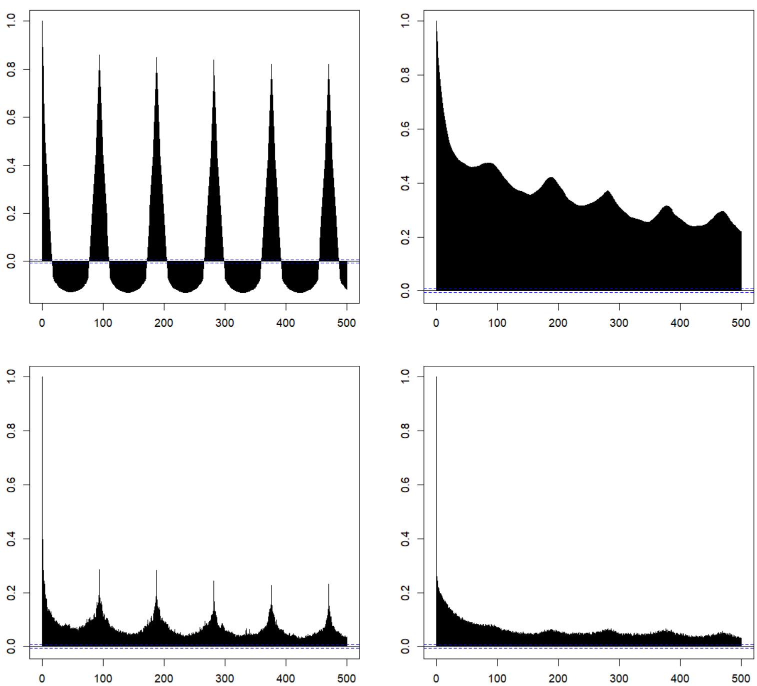

Figure 1. Autocorrelation function of proportional quoted spread (top left), proportional effective spread (top right), volatility (bottom left) and ATI (bottom right) for the company PGE (5-minute intervals)

Source: own calculations.

Table 1. Summary statistics for daily data

Variable	Mean	Standard deviation	Min	Max
ATI	0.49793	0.11687	0.15444	0.84805
EFF	0.00402	0.02227	0.00028	0.27606
QUO	0.00187	0.00082	0.00074	0.00737
RAI	0.00053	0.00060	0.00007	0.00876
GK	0.00051	0.00081	0.00003	0.01395
RV	0.00051	0.00081	0.00003	0.01395
BPV	0.00053	0.00060	0.00007	0.00876

Note: The table contains descriptive statistics calculated on the basis of trades and quotes data of 12 companies that in the period from 1 January 2020 to 31 August 2023 were permanently included in the WIG20 index.

Source: own calculations.

Table 2. Summary statistics for intraday data

Variable	Mean	Standard deviation	Min	Max
1-minute frequency				
ATI	0.452358000	0.367843900	0.000033229	1.000000000
EFF	0.003359015	0.021707740	0.000000001	0.369432100
QUO	0.001522686	0.001440292	0.000200780	0.031584196
GK	0.000000653	0.000004921	0.000000005	0.001370589
5-minute frequency				
ATI	0.539948072	0.228760851	0.000165937	1.000000000
EFF	0.003893397	0.022234950	0.000000001	0.342044900
QUO	0.001866829	0.001156481	0.000209468	0.018047837
GK	0.000004710	0.000017253	0.000000005	0.002113177
10-minute frequency				
ATI	0.531513442	0.194909261	0.000813558	1.000000000
EFF	0.003899087	0.022264020	0.000006655	0.335197400
QUO	0.001871816	0.001067395	0.000229696	0.011459379
GK	0.000010400	0.000035071	0.000000006	0.003143300

Note: The table contains descriptive statistics calculated on the basis of trades and quotes data of 12 companies that in the period from 1 January 2020 to 31 August 2023 were permanently included in the WIG20 index.

Source: own calculations.

The dataset covers the period from 1 January 2020 to 31 August 2023 (910 trading days). The sample contains trades and quotes data of 12 companies that in the mentioned period were permanently included in WIG20 index. In Table 1, we present descriptive statistics of variables for daily data (median values of means, standard deviations, minimums and maximums), whereas in Table 2 we provide the statistics for high frequency data. Daily RV is calculated by summing the squared intraday 5-minute returns over a single trading day. Also, BPV and RAI are daily data constructed from intraday data in the same way.

We remove intraday seasonality in a way described by Bińkowski and Lehalle (2018). Denoting the $x(d,\tau)$ value of variable x on day d at bin τ, we compute:

$$y(d,\tau)=\log x(d,\tau)-\log\overline{x}(d,\tau) \tag{15}$$

where $\log\overline{x}(d,\tau)$ denotes the mean value of $\log x(d,\tau)$ over all days.

Next, we set $m=n^{0.7}$ (where is the length of the series) and apply an exact local Whittle estimator of long memory (Shimotsu & Phillips, 2005). The results are mixed, but in general most of the series are stationary, with a long memory parameter less than 0.5.[3] We differentiate all series using the estimated parameters of the Whittle estimator.

4.2. Transfer entropy

In this section, we establish whether there is any information transfer between *ATI*, liquidity and volatility. We use both Shannon transfer entropy and the Rényi approach. The number of lags used in both approaches is equal to 1 for both variables in a pair. We choose such values to compare our results with Będowska-Sójka & Kliber (2021). The RTransferEntropy package was used for the calculations (Behrendt et al., 2019).

It is recommended to assign q a low value in order to give more weight to extreme events, as in the case of financial time series, the most important information comes in the tails. Following (Będowska-Sójka & Kliber, 2021), we set $q=0.1$ (default value in the function from the RTransferEntropy package) and this highlights the information in the tails. In Table 3, we present the number of companies of significant information transfer from (to) liquidity and volatility and *ATI* at a significance level of 0.05. The results show that for daily data there is no information transfer from *ATI* to either liquidity or volatility in the case of Shannon entropy. The same holds true for the opposite direction for all variables. When considering the tails of the Rényi approach, the results are quite similar. Table 3 shows the number of companies with significant information transfer from *ATI* to liquidity and in the opposite direction, and from *ATI* to volatility and in the opposite direction.

[3] More detailed results of estimation are available upon request.

Table 3. Number of companies of significant information transfer between *ATI* and liquidity, and between *ATI* and volatility for daily data

***ATI* → liquidity**			
Variable	*QUO*	*EFF*	*RAI*
Shannon	3	1	1
Rényi	2	0	1
liquidity → *ATI*			
Variable	*QUO*	*EFF*	*RAI*
Shannon	1	0	0
Rényi	3	1	0
***ATI* → volatility**			
Variable	*QUO*	*EFF*	*RAI*
Shannon	0	0	1
Rényi	0	1	1
volatility → *ATI*			
Variable	*QUO*	*EFF*	*RAI*
Shannon	0	2	0
Rényi	0	2	1

Note: Significance level is equal to 0.05 using either the Shannon or Rényi approach. The number of lags used in both approaches is equal to 1 for both variables in a pair. In the case of Rényi transfer entropy, we set $q = 0.1$. *RV*, *BPV* and *RAI* are based on 5-minute log returns.

Source: own calculations.

We apply the same procedure to investigate the information transfer between volatility and liquidity. The results are presented in Table 4. In most companies, information transfer exists between volatility and liquidity, especially when liquidity is proxied by *EFF*. When considering Rényi transfer entropy, we do not find information transfer in both directions (with an exception for the direction from volatility to liquidity for the measure of illiquidity). Note that *RV* offers a good approximation of unobservable volatility, but it is sensitive to price spikes, while *BPV* is robust to them. The differences in the results for both estimators are small (maximum two companies). This may indicate a lack of essential jumps.

We compute estimators of and for intraday data (1-, 5-, 10-minute frequencies) to check whether the choice of frequency has an impact on the results obtained. Table 5 contains the number of companies of significant information transfer for high frequency data.

From Shannon transfer entropy, we observe that as the frequency increases, the number of companies of significant information transfer also increases

Table 4. Number of companies with significant information transfer from volatility to liquidity, and from liquidity to volatility for daily data

volatility → liquidity						
Entropy	Shannon			Rényi		
Variable	*QUO*	*EFF*	*RAI*	*QUO*	*EFF*	*RAI*
GK	4	6	2	0	0	6
RV	6	9	1	0	0	2
BPV	8	9	0	0	0	4
liquidity → volatility						
Entropy	Shannon			Rényi		
Variable	*QUO*	*EFF*	*RAI*	*QUO*	*EFF*	*RAI*
GK	4	11	1	0	0	0
RV	4	11	0	0	0	1
BPV	6	10	1	0	0	2

Note: Significance level is equal to 0.05 using either the Shannon or Rényi approach. The number of lags used in both approaches is equal to 1 for both variables in a pair. In the case of Rényi transfer entropy, we set $q = 0.1$. *RV*, *BPV* and *RAI* are based on 5-minute log returns.

Source: own calculations.

Table 5. Number of companies of significant information transfer for intraday data

Pair	*ATI- GK*		*ATI- QUO*		*ATI- EFF*		*GK- QUO*		*GK-EFF*	
Direction	→	←	→	←	→	←	→	←	→	←
1-minute frequency										
Shannon	12	12	12	12	12	12	12	12	12	12
Rényi	11	3	4	0	7	2	3	12	12	12
5-minute frequency										
Shannon	12	12	12	12	12	12	12	12	12	12
Rényi	1	3	0	1	0	1	2	2	3	1
10-minute frequency										
Shannon	11	11	8	8	12	12	12	12	9	7
Rényi	0	0	0	0	0	0	2	1	1	2

Note: Significance level is equal to 0.05 using either the Shannon or Rényi approach. The number of lags used in both approaches is equal to 1 for both variables in a pair. In the case of Rényi transfer entropy, we set $q = 0.1$.

Source: own calculations.

for all variables under consideration. For 1-minute and 5-minute frequency, we observe directional dependence for all companies. When considering the tails (Rényi approach), there is a huge difference for 1- and 5-minute (10-minute) frequencies, and we do not observe as many companies of significant information transfer, although their number increases as frequency increases (Rényi approach compared to Shannon approach). In Table 6, we present net information flow, that is, the difference $T_{X \to Y} - T_{Y \to X}$.

Table 6. Net information flow for 1- and 5-minute data (Shannon entropy)

Variable X	Variable Y	1-minute frequency	5-minute frequency
ATI	*GK*	0.0006	0.0014
ATI	*QUO*	0.0003	0.0005
ATI	*EFF*	–0.0062	–0.0007
GK	*QUO*	–0.0008	–0.0008
GK	*EFF*	–0.0011	–0.0008

Note: This table shows the net information flow for 1- and 5-minute frequencies for which significant information transfer was detected for all companies (Shannon entropy).

Source: own calculations.

From Table 6, we conclude that more information flows from *ATI* to volatility than vice versa. The situation is similar when we consider quoted spread instead of volatility. However, for the effective spread (when we consider variables X = *ATI* and Y = *EFF*), the sign of net information flow is negative—the impact of *EFF* on *ATI* is greater than *ATI* on *EFF*. The situation is similar when we consider volatility as X and liquidity as Y. The conclusions are the same regardless of whether we consider 1- or 5-minute data.

Conclusions

We analysed causal dependencies between financial time series of *ATI*, liquidity and volatility. The paper employs a dataset for 12 companies listed on WSE that identifies transactions made through algorithmic trading in the period from 1 January 2020 to 31 August 2023. One limitation of the study is the sample size, limited to the most liquid companies listed on the Warsaw Stock Exchange. For daily data, we observe a lack of information transfer from *ATI* to liquidity and volatility regardless of whether we use Shannon or Rényi entropy. Regarding pair volatility-liquidity, we find information transfer from liquidity to volatility for almost all the companies considered, but only for proportional effective spread. This does not hold for tail dependencies when applying Rényi transfer entropy. Our result is in line with Będowska-Sójka and Kliber (2021), who used similar volatility estimators (realised variance and bi-power variation) and method (transfer entropy), but different liquidity measures. The results are similar regardless of whether we use a robust measure for price jumps (bi-power variation) or not (realised variance).

For intraday data, the results are quite different. Considering the entire distributions, we observe that as the data frequency increases, the number of companies with significant information transfer also increases (for all pairs of variables). For 1- and 5-minute frequencies, significant information transfer is observed for all companies under consideration.

Causal relationships of *ATI*, liquidity and volatility for high frequency data are identified. HFT has an impact on volatility and liquidity, and there is also a feedback loop. Given that the one lag is used in the transfer information approach, it shows the fast reaction of one variable to another. Trading bots that enter into transactions react to market signals in a short time, thus causing changes in market volatility and liquidity. The empirical results offer support for Hypotheses 1 and 2, according to which a significant pairwise causal dependence exists between *ATI*, market liquidity and volatility, and an increase in data frequency leads to more numerous causality patterns. However, these results are not comparable with previous studies because we use high-frequency data and investigate causality using changes in entropy.

Based on the results of net information flow, we conclude that more information flows from *ATI* to volatility rather than vice versa. The situation is similar when we consider quoted spread instead of volatility. However, the impact of effective spread on *ATI* is greater than vice versa. When considering volatility and liquidity, we observe a greater flow of information from liquidity to volatility. The conclusions are the same regardless of whether we consider 1- or 5-minute data.

When considering the tails (describing events that have low probability), there is a clear difference in the number of companies of significant information transfer for 1-minute and 5-minute (10-minute) frequencies, and we observe causal dependence only between volatility and the effective spread in both directions for highest frequency. The results based on Rényi entropy do not support Hypothesis 3, which posited more frequent causal relationships among variables in the tails of the distribution.

These interesting findings require further research into their causes and economic interpretation. By varying the number of lags in the computation of transfer entropy, we could determine the duration of information's influence on the market. Other types of entropy could also be used, e.g., Tsallis entropy, which is a versatile tool for measuring the degree of uncertainty in complex or non-linear systems.

From both the theoretical and practical point of view, it would be interesting to take into account the intra-day seasonal pattern when examining the relationships between variables. We intend to expand our sample to include mid-sized companies. The goal would be to examine how company size affects performance. The plan also includes investigating the impact of changes in the value of the Rényi entropy weighting parameter on the results. An interesting topic for future research concerns the shapes of the probability distributions of intraday values of the studied variables, their similarity and changes over time.

References

Abdi, F., & Ranaldo, A. (2017). A simple estimation of bid-ask spreads from daily close, high and low prices. *The Review of Financial Studies*, *30*(12), 4437–4480. https://doi.org/10.1093/rfs/hhx084

Aggarwal, N., & Thomas, S. (2014). *The causal impact of algorithmic trading on market quality*. Indira Gandhi Institute of Development Research, Mumbai Working Papers, 2014-023. http://www.igidr.ac.in/pdf/publication/WP-2014-023.pdf

Ao, H., & Li, M. (2024). Exploiting the potential of a directional changes-based trading algorithm in the stock market. *Finance Research Letters*, *60*, 104936. https://doi.org/10.1016/j.frl.2023.104936

Arumugam, D., Prasanna, P. K., & Marathe, R. R. (2023). Do algorithmic traders exploit volatility? *Journal of Behavioral and Experimental Finance*, *37*, 100778. https://doi.org/10.1016/j.jbef.2022.100778

Banerjee, A., & Nawn, S. (2024). Proprietary algorithmic traders and liquidity supply during the pandemic. *Finance Research Letters*, *61*, 105052. https://doi.org/10.1016/j.frl.2024.105052

Behrendt, S., Dimpfl, T., Peter, F. J., & Zimmermann, D. J. (2019). RTransferentropy—quantifying information flow between different time series using effective transfer entropy. *SoftwareX*, *10*, 100265. https://doi.org/10.1016/j.softx.2019.100265

Będowska-Sójka, B., & Kliber, A. (2021). Information content of liquidity and volatility measures. *Physica A: Statistical Mechanics and its Applications*, *563*, 125436. https://doi.org/10.1016/j.physa.2020.125436

Bińkowski, M., & Lehalle, C.A. (2022). Endogenous dynamics of intraday liquidity. *Journal of Portfolio Management Market Microstructure*, *48*(6), 145–169. https://doi.org/10.48550/arXiv.1811.03766

Brauneis, A., & Mestel, R. (2018). Price discovery of cryptocurrencies: bitcoin and beyond. *Economics Letters*, *165*, 58–61. https://doi.org/10.1016/j.econlet.2018.02.001

Brogaard, J., Hendershott,T., & Riordan, R. (2014). High frequency trading and price discovery. *The Review of Financial Studies*, *27*(8), 2267–2306. https://doi.org/10.1093/rfs/hhu032

Courdent, A., & McClelland, D. (2022). The impact of algorithmic trading on market quality: Evidence from the Johannesburg Stock Exchange. *Investment Analysts Journal*, *51*(3), 157–171. https://doi.org/10.1080/10293523.2022.2090056

Desagre, C., D'Hondt, C. D., Petitjean, M. (2022). The rise of fast trading: Curse or blessing for liquidity? *Finance*, *43*(3), 119–158. https://doi.org/10.2139/ssrn.3192597

Díaz, A., & Escribano, A. (2020). Measuring the multi-faceted dimension of liquidity in financial markets: A literature review. *Finance Research in International Business and Finance*, *51*, 101079. https://doi.org/10.1016/j.ribaf.2019.101079

Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*(1), 57–66. https://doi.org/10.1016/j.ijforecast.2011.02.006

Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, *182*(1), 119–134. https://doi.org/10.1016/j.jeconom.2014.04.012

Dionisio, A., Menezes, R., & Mendes, D. A. (2004). Mutual information: A measure of dependency for nonlinear time series. *Physica A: Statistical Mechanics and its Applications*, *344*(1), 326–329. https://doi.org/10.1016/j.physa.2004.06.144

Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU (recast). OJ L 173, 12.6.2014, p. 349. https://data.europa.eu/eli/dir/2014/65/oj

Dubey, R. K., Babu, A. S., Jha, R. R., & Varma, U. (2021). Algorithmic trading efficiency and its impact on market-quality. *Asia-Pacific Financial Markets*, *29*, 381–409. https://doi.org/10.1007/s10690-021-09353-5

Ekinci, C., & Ersan, O. (2022). High-frequency trading and market quality: The case of a "slightly exposed" market. *International Review of Financial Analysis*, *79*, 102004. https://doi.org/10.1016/j.irfa.2021.102004

Garman, M., & Klass, M. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, *53*(1), 67–78. https://doi.org/10.1086/296072

Gurgul, H., & Lach, Ł. (2012). The electricity consumption versus economic growth of the Polish economy. *Energy Economics*, *34*(2), 500–510. https://doi.org/10.1016/j.eneco.2011.10.017

Gurgul, H., Lach, Ł., & Mestel, R. (2012). The relationship between budgetary expenditure and economic growth in Poland. *Central European Journal of Operations Research*, *20*(1), 161–182. https://doi.org/10.1007/s10100-010-0186-z

Gurgul, H., & Machno, A. (2017). The impact of asynchronous trading on Epps effect on Warsaw Stock Exchange. *Central European Journal of Operations Research*, *25*, 287–301. https://doi.org/10.1007/s10100-016-0442-y

Gurgul, H., & Syrek, R. (2023). Contagion between selected European indexes during the COVID-19 pandemic*, Operations Research and Decisions*, *33*(1), 47–59. https://doi.org/10.37190/ord230104

He, J., & Shang, P. (2017). Comparison of transfer entropy methods for financial time series. *Physica A: Statistical Mechanics and its Applications*, *482*, 772–785. https://doi.org/10.1016/j.physa.2017.04.089

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, *66*(1), 1–33. https://doi.org/10.1111/j.1540-6261.2010.01624.x

Hendershott, T., & Riordan, R. (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, *48*(4), 1001–1024. https://doi.org/10.1017/S0022109013000471

Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., & Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Physics Reports*, *441*(1), 1–46. https://doi.org/10.1016/j.physrep.2006.12.004

Jain, A., Jain, C., & Khanapure, R.B. (2021). Do algorithmic traders improve liquidity when information asymmetry is high? *Quarterly Journal of Finance*, *11*(1), 2050015. https://doi.org/10.1142/S2010139220500159

Lacava, D., Ranaldo, A., & Santucci de Magistris, P. (2023). *Realized illiquidity*. Swiss Finance Institute Research Paper, 22–90. https://doi.org/10.2139/ssrn.4282541

Leone, V., & Kwabi, F. (2019). High frequency trading, price discovery and market efficiency in the FTSE100. *Economics Letters*, *181*, 174–177. https://doi.org/10.1016/j.econlet.2019.05.022

Lesmond, D.A. (2005). Liquidity of emerging markets. *Journal of Financial Economics*, *77*(2), 411–452. https://doi.org/10.1016/j.jfineco.2004.01.005

Mestel, R., Murg, M., & Theissen, E. (2018). Algorithmic trading and liquidity: Long term evidence from Austria. *Finance Research Letters*, *26*, 198–203. https://doi.org/10.1016/j.frl.2018.01.004

Mestel R., Steffen V., & Theissen E. (2024). Algorithmic trading and mini flash crashes: Evidence from Austria. *Economics Letters*, *244*, 111982. https://doi.org/10.1016/j.econlet.2024.111982

Ramos, H. P., & Perlin, M. S. (2020). Does algorithmic trading harm liquidity? Evidence from Brazil. *North American Journal of Economics and Finance*, *54*, 101243. https://doi.org/10.1016/j.najef.2020.101243

Shimotsu, K., & Phillips, P. C. B. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics*, *33*(4), 1890–1933. https://doi.org/10.1214/009053605000000309

Syczewska, E., & Struzik, Z. (2015). Granger causality and transfer entropy for financial returns. *Acta Physica Polonica A*, *127*(3A), 129–135. https://doi.org/10.12693/APhysPolA.127.A-129