UNIVERSITY OF TARTU
School of Economics and Business Administration

Kristin Lass

# STOCK PRICE RELATION TO NEWS: THE CASE OF THE TALLINN STOCK EXCHANGE

Master's Thesis

Supervisors: Kristjan Pulk (PhD) and Mustafa Hakan Eratalay (PhD)

Tartu 2025

I prepared the work independently. I have cited all works of other authors, fundamental positions, and data from literary sources and elsewhere used in the preparation of the work.

Kristin Lass

## Abstract

This article examines the relationship between media sentiment and stock price movements using the example of 12 companies listed on the Tallinn Stock Exchange. The study uses sentiment analyzes of Estonian-language financial news using the ARMAX model to quantify the impact of news sentiment on daily stock returns. The dataset comprises 1,930 news articles and trading data from 2019 to 2023. The results show that media sentiment has a statistically significant impact on short-term stock price movements, especially in the case of strongly negative news. This supports the view that market movements cannot be explained solely by fundamental analysis, but also indicate behavioral factors, such as attentional bias and overreaction, that affect stock price formation. The results contribute to media-based market analysis in small markets and highlights the need for considering the influence of media sentiment in financial modeling.

**Keywords:** *behavioral finance, sentiment analysis, Tallinn Stock Exchange, ARMA model, media influence, stock price prediction*

**Academic discipline code S180:** Economics, econometrics, economic theory, economic systems, economic policy

## 1. Introduction

Debates about the media's role in the economy, especially in stock price formation, have persisted for decades. Understanding the behavior of financial markets and investors is one of the most significant and challenging finance tasks. Investor behavior has been studied over time through various theories and practical approaches. Traditional economic models assume that investors act deliberately and aim to optimize profit. However, the behavioral model is based on the idea that investor behavior is not entirely rational but is influenced by psychological factors such as overconfidence, herd instinct, or heuristics. At the same time, growing attention has been paid to the influence of the media on investor decision-making. Several research articles confirm that nowadays, media coverage has become the primary information channel in financial markets (Bushee et al., 2010; Strycharz et al., 2018; Dong et al., 2022). Despite this, the influence of media sentiment on stock price formation is quite varied. Given the increasing speed of media and scope of the surrounding media, examining the extent to which media coverage shapes investor market behavior, in addition to traditional fundamental stock analysis, is increasingly necessary.

Media coverage diverges from conventionally composed analytical reports, being faster and broader, featuring shorter and more accessible news stories (Bushee et al., 2010). On one hand, it has been found that the media disseminates new information about companies, reducing information asymmetry between companies and investors (Tetlock, 2010; 2011). On the other hand, companies may manipulate media news (Ahern & Sosyura, 2014), use sentimentally oriented language, and exaggerate or omit views that can influence investors' cognitive biases (Hermida et al., 2012). Based on these aspects and as argued by Guldiken et al. (2017), it is crucial to consider the impact of media sentiment in studies of stock prices and investor behavior.

Majority of previous studies have found a correlation between news and changes in stock prices. Although the results vary, negative news often affect changes in stock prices more than positive news (Hautsch & Klußmann, 2011; Uhl, 2014; Liu et al., 2023). There is a weak but statistically significant relationship between the news sentiment and movements in stock prices (Schumaker et al., 2012; Wan et al., 2020). It can be argued that previous studies were convinced of the connection between media coverage and stock price relationships. Despite this, the applications of different models, differences in dictionaries, and the particularities of languages often yield different results in the context of this work. The content and results of previous studies are further elaborated in the present thesis, the related works section.

Sentiment analysis is a widely used method for detecting the polarity of text, whether it conveys positive, negative, or neutral sentiment, under specific conditions (Mehta & Pandya, 2020). In financial research, it has increasingly been used to examine the link between media tone and market behavior. This thesis applies sentiment analysis in combination with ARMA time series modelling to investigate whether news sentiment can be used to predict stock price movements on the Tallinn Stock Exchange. While such methods have been widely applied to major global markets, the Estonian market context remains underexplored. It is commonly believed that the Tallinn Stock Exchange reacts slowly to news and events (Lill, 2025), yet little academic attention has been devoted to empirically testing this assumption.

The topic is relevant, as the public interest in investment is steadily increasing, and media coverage in the financial world plays an increasingly significant role in influencing investors' decision-making (Brunswick Group, 2023). This background is also related to the advantage of online news sources, i.e., news sources are more interactive and immediately available (Karlsson, 2011). Interactivity enables the masses' tendency to consume more news related to their interests, and instant availability supports the need to be up to date with the news immediately (Kohut et al., 2010). This has also given rise to the study of the influence of media sentiment on the movement of stock prices.

The aim of this thesis is to evaluate the relationship between media sentiment and stock market movements on the Tallinn Stock Exchange. In this paper, financial news published in one of Estonia's famous economic newspapers, Äripäev, is analyzed to elucidate the relationship between the sentiment of news articles about companies listed on the Tallinn Stock Exchange and the movement of their stock prices. A total of 1,930 news articles related to 12 companies listed on the Tallinn Stock Exchange were collected for the analysis, forming the empirical basis of this study. As is typical for non-English source texts, the author translated the news into English for analysis. To the authors' best knowledge, financial news in Estonian has not been studied for this purpose before in this way and for this purpose.

The empirical results show that general news sentiment statistically affects stock performance on the same day. Negative sentiment has a more substantial effect than positive sentiment, which indicates possible asymmetric market behavior. The effect was weaker in the case of weekly data, but some lagged effects were observed. The results support the view that sentiment is an essential source of additional information in understanding short-term market dynamics.

This work complements the existing literature by applying the ARMAX model with outlier detection and asymmetric sentiment analysis to a small, sensitive market. The results offer practical value to investors and analysts, emphasizing the need to monitor news content in real time even in smaller economic environments.

# 2. Literature review

## 2.1 Rational models of investor behavior

In traditional economic theory, rationality is a fundamental assumption underpinning many financial models. Statman (2014) outlines four main pillars of classical finance: investors are rational, markets are efficient, portfolios are constructed according to mean-variance optimization, and asset pricing models explain expected returns. This framework assumes that investors make informed, logical, and consistent decisions. Thus, Fromlet (2001) has found that rational decision-making means investors' ability to assess probabilities to make reasoned and optimal choices.

This view aligns with the Efficient Market Hypothesis (EMH), which posits that all available information is entirely and immediately reflected in asset prices, making it impossible to consistently achieve above-average returns through analysis or forecasting (Fama, 1970; Malkiel, 2003). The assumption of investor rationality has long been central to classical finance. However, it has been challenged by behavioral economists who emphasize cognitive limitations and the role of biases in decision-making (Simon, 1955; Barberis & Thaler, 2003).

For much of the 20th century, the dominant paradigm in finance held that investors act rationally and that markets are efficient. This perspective is encapsulated in Fama's (1970) formulation of the EMH, which asserts that asset prices fully incorporate all available information. According to the EMH, investors process information objectively, act in their economic self-interest, and prices adjust instantly to new information, rendering both technical and fundamental analysis ineffective in producing consistent excess returns.

Fama's theory, developed in the 1970s, claims that all relevant market-related information, historical and publicly available, is immediately and wholly reflected in security prices (Bernstein, 2012). Under this hypothesis, prices adjust to new information when it becomes available (Fama, 1965). As a result, prices always reflect fair value, and predicting future prices based on market data is impossible. Fama proposed three forms of market efficiency – weak, semi-strong, and strong (Fama, 1997). The weak form suggests that market prices incorporate all past trading information, including prices and volumes, rendering technical analysis ineffective. The most widely accepted semi-strong form incorporates all publicly available information, such as financial statements and corporate disclosures. The strong form asserts that prices reflect public and historical data and all insider information. Under this form, individuals with access to non-public, material information are legally prohibited from trading on it (Read, 2013). However, it is notable that insider trading varies across countries. Market fairness and investor confidence are also affected by market regulatory frameworks and their differences in the definition, enforcement, and penalties of insider trading (Thompson, 2013).

This means that, according to the EMH, investor behavior is unaffected by emotions or psychological factors. The hypothesis cannot adequately explain market bubbles, crashes, arbitrage opportunities, or the growing influence of media coverage

(Malkiel, 2003). Despite these limitations, it is argued that over the long term, markets correct inefficiencies and stock prices tend to gravitate toward their fair value (Gupta et al., 2014).

For example, analysis of leading global financial news sources revealed a 67% correlation between positive sentiment trends and stock price movements, supporting the semi-strong to strong forms of EMH (Chowdhury et al., 2014). One advantage of the traditional framework is its 'elegant simplicity'; however, by the turn of the century, it became increasingly evident that individual trading strategies could not be fully understood within this model (Barberis & Thaler, 2003). There is no clear consensus on whether stock prices follow a truly random walk or if they can be predicted. Nearly every study supporting the EMH is countered by another study that disputes it. Testing the efficiency of markets has become increasingly complex due to constantly changing market and economic conditions (Ţiţan, 2015).

This rational approach was the foundation for many traditional models and return theories. It also assumed that markets are self-correcting, that deviations are tied to insider information, and that arbitrage quickly resolves them. Thus, the EMH shaped academic research and real-world investing practices for decades. At the same time, all of the aforementioned assumptions have been subject to fundamental critique by behavioral scientists, making them a logical starting point for understanding behavioral finance.

## 2.2 Behavioral models of investor behavior

Behavioral finance introduced a new approach to financial analysis and investor behavior, one that is grounded in psychology and its effects on financial markets. Since the late 1980s, when the first studies began to suggest that financial markets may not be entirely rational, the field of behavioral finance has developed tools to demonstrate this impact (Uhl, 2014). The core assertion of behavioral finance is that investors are ordinary people who may behave inefficiently. In order to better understand market dynamics, behavioral finance integrates concepts from traditional economics and psychology to explain why individuals often make poor financial decisions (Byrne & Brooks, 2008; Madaan & Singh, 2019). Accordingly, behavioral finance posits that psychological factors influence investors and may not be capable of making entirely rational investment decisions.

Psychologists Amos Tversky and Daniel Kahneman argued in their landmark 1979 study that cognitive biases, or heuristics, play a critical role in decision-making under uncertainty. Their insights have been the basis for subsequent research into behavioral finance. Examples of such heuristics include imitation, where investors follow others instead of analyzing available information; loss aversion, the tendency to be more sensitive to losses than to equivalent gains; and regret aversion, where investors hesitate to make decisions for fear of future regret (Gupta et al., 2014). These heuristics systematically influence investor decisions in uncertain conditions. For instance, herd behavior – investors imitating others without independent analysis – can lead to impulsive buying or panic selling, and is well-documented in financial markets (Singh & Sharma, 2024; Wang & Liu, 2023).

Another extensively studied bias is loss aversion, which describes the more substantial emotional impact of losses compared to similar gains. This can cause investors to hold on to declining stocks for too long or sell appreciating stocks prematurely (Chapman & Roberts, 2022; Vandolder, 2023). Investment decisions are also influenced by the fear of regretting an investment decision later. Such fear favors decision-making avoidance, leading the investor not to decide at all or to trade too cautiously (Kumar & Gupta, 2023; Zhu & Li, 2024). Interpreting this type of anomaly is also an important part of explaining what happens in financial markets.

As early as 1981, Kahneman and Tversky demonstrated that individuals tend to overreact to new information – a behavior inconsistent with the notion of perfectly rational market participants. In the same year, Shiller (1981) showed that stock prices exhibit excess volatility beyond what would be expected based on fundamental changes, directly challenging the Efficient Market Hypothesis (EMH).

Subsequent empirical studies in the mid-1980s further questioned the validity of the EMH. De Bondt and Thaler (1985) conducted one of the earliest studies attempting to link exogenous variables to market behavior. Their research in experimental psychology indicated that overreaction primarily occurs in response to unexpected and dramatic news, resulting in predictable reversals – clear violations of market efficiency. For example, Barakat and Barakat (2024) tested the overreaction hypothesis in developed markets using data from 2000 to 2020. They found that stocks considered "losers" tend to outperform "winners" in the long run, even in global crises. This suggests that overreaction as a behavioral pattern has been a persistent phenomenon over time, and behavioral finance theory is necessary to explain it.

Research into behavioral biases in financial markets has produced strong evidence that investor behavior reflects bounded rationality, challenging the EMH assumptions. Examples include overreaction to news, herd behavior, and loss aversion (Thaler, 2005; Barberis, Greenwood, Jin, & Shleifer, 2018; Chen et al., 2021; Lin et al., 2022). As Simon (1955) proposed, bounded rationality suggests that while investors aim to make rational decisions, their cognitive limitations, emotional influences, and information processing constraints lead to systematic deviations from optimal behavior. Studies have identified six primary behavioral biases influencing investors: overconfidence, the disposition effect, herd behavior, representativeness heuristic, anchoring, and availability heuristic (Elhuseini & Abdelgadir, 2020). Behavioral finance also challenges the strong form of the EMH, which asserts that no relevant information can be used to predict stock returns. In contrast, research has shown that financial market movements are not entirely random and contain predictable components. Specific valuation metrics and historical return indicators – such as the price-to-earnings ratio (P/E), book-to-market ratio, dividend yield, and momentum – have demonstrated some degree of predictive power for future returns (Malkiel, 2005; Campbell & Shiller, 1988; Jegadeesh & Titman, 1993). These findings suggest that markets may systematically underreact or overreact to information, allowing patterns to persist and be exploited by investors, which contradicts the strong-form EMH.

Some studies also show that investors place excessive weight on recent experiences, a phenomenon known as memory bias. Investors tend to underestimate likely outcomes compared to certain ones and overestimate the accuracy of their own beliefs, often described as overconfidence (Kumar, 2016). This means that investors are more likely to take risks after incurring losses, while those who have made gains tend to become more risk averse. Although the common principle for earning returns in the stock market is to 'buy low and sell high,' individual investors often struggle to assess market conditions accurately. A key reason is the influence of psychological biases, which are particularly detrimental during periods of heightened emotion and volatility (Nofsinger, 2017). These emotional reactions are frequently triggered by recent news. When investors are exposed to new positive or negative information, it can disproportionately shape their market perception – a phenomenon explained by recency bias and the availability heuristic (Tversky & Kahneman, 1973). In other words, more emotionally charged news stories are more likely to influence investment decisions, regardless of the objective relevance of the news content. This is an important aspect to emphasize in shaping investor behavior and leads to the next important topic, which provides an overview of the relationship between news sentiment and market dynamics.

## 2.3 Theoretical approaches and models of media influence on investors

Since the beginning of the 21st century, there has been growing evidence of the increasingly close connection between the financial sector and the media. Media coverage reflects financial developments and influences investor attention and market dynamics (Engelberg & Parsons, 2011; Dougal et al., 2012). The value of financial products is increasingly shaped by media narratives, which can drive sentiment, amplify market trends, and influence public perception (Shiller, 2017).

This means that the media coverage of a listed company affects the fundamental value of the stock as well as the perceived value. More and more researchers have begun to pay attention to the impact of media sentiment on asset pricing (Tetlock et al., 2008; Loughran & McDonald, 2011; Yang et al., 2022). In recent years, several financial market anomalies have been studied, such as the underestimation or overestimation of financial news, the effect of days of the week on investor reactions, and the effect of advertising on investor decisions. Therefore, there are many studies that investigate which instruments receive attention, how corporate advertisements affect the level of attention of investors, and how investors follow news published by companies or the media in general. (Siering, 2012) However, investor attention is not solely determined by the objective availability of information – it is also shaped by cognitive limitations such as memory bias. Investors may recall recent or emotionally intense news more vividly than older or more neutral information, leading to disproportionate reactions or neglect of relevant context (Tversky & Kahneman, 1973; Barberis et al., 1998). This selective recall can contribute to anomalies like overreaction to short-term news and persistent mispricing.

Media coverage is important for attracting investor attention and increasing stock returns for listed companies (Bushee & Miller, 2012). Thus, financial news, market announcements, company news, or analysts' forecasts shape investors' expectations and opinions and are reflected in the volatile reactions of stock markets

(Tetlock, 2014). The media thus plays an important role in disseminating information to interest groups (investors) and influencing their decision-making processes (Donaldson & Preston, 1995).

As mentioned earlier, emotions, the herd effect, and irrational behavior shape and influence investors' trading decisions and, in turn, price movements in financial markets (Nofsinger, 2005; Oberlechner & Hocking, 2004; Shiller, 2005; Kräussl & Mirgorodskaya, 2014). It has been established that the media plays an important role in inducing "herd effect" behavior (Thompson, 2013). The wide availability of financial information has induced market participants to respond similarly to new information (Shiller, 2005; Davis, 2006). Thus, investors tend to watch how the market reacts to certain news, thereby trying to predict how the herd will react by taking advantage of the expected direction of market movement.

Media coverage of a company affects the company's share price to a greater extent than the financial information published by the company itself (Lehavy & Solan, 2008). Similarly, firms that are less visible to investors have been found to be less likely to attract investor attention, even if they disclose information about the firm (Bushee & Miller, 2012). This is the concept of attention-grabbing stocks (Barber & Odean, 2008). More media coverage means more attention, making the stock more attractive to investors.

Naturally, the sentiment of the news is also important. For example, the proportion of negative words in publicly traded company-specific news has been found to lead to lower future earnings forecasts, indicating that linguistic media content captures otherwise difficult-to-quantify aspects of company fundamentals that investors incorporate into stock prices (Tetlock et al., 2008). As we know from behavioral finance concepts, investors' moods can also change. Evidence shows that people's mood changes can be triggered by positive or negative content disseminated by mass media (Barber & Odean, 2008; Yang et al., 2017). However, in the financial market, not everything is so straightforward. Market participants who make decisions not based on fundamental analysis but rather on emotions, rumors, and speculations also have an impact. Some common speculation-based concepts are noise traders and buying rumors to sell the news.

## 2.4 Alternative trading theories

In this context, the terms "noise trader" and "buy the rumor, sell the news" are relevant to describe market investors. These two concepts are important for understanding what is happening in the financial market, as their effects are often harder to identify and predict than expected. Their importance is recognized, and more and more attempts are made to analyze them.

Noise traders make buy-sell decisions in the securities markets without professional advice or thorough fundamental or technical analysis. Several theoretical studies provide a framework in which the attention of noise traders can permanently affect asset pricing (De Long et al., 1990; Da et al., 2011; Herve et al., 2019). Thus, a Noise Trader is a market participant who makes investment decisions without using financial principles, has poor market timing, follows trends, and often tends to over- or under-react to good and bad news (Ramiah et al., 2015). Shleifer and Summers (1990) say noise traders chase trends and rumors. The development of financial technology has significantly influenced the growth of the activity of noise traders and increased the volatility of stock markets. Moreover, this increase in volatility is greater for smaller stocks than for larger stocks, which is consistent with the argument that stocks with low turnover (smaller stocks) are more affected by noise (Collins, 2020). In addition, the same study found that intraday returns have higher volatility than weekly or monthly returns, supporting the idea that noise traders move in and out of stock markets more quickly.

One change, however, is the emergence of so-called smart investors whose activities are opposed to noise traders. The attention of noise traders increases volatility, creating additional risk that is priced into the market. In contrast, the attention of smart investors reduces volatility because their presence stabilizes stock prices by reducing uncertainty. (Herve et al., 2019) Ultimately, the influence of noise traders disappears as informed investors bring prices into equilibrium. The main motivation for this study is the growing body of evidence in the financial literature on the impact of news on investors and, thus, on stock returns. DeLong et al. (1990) are among the first to discover that investors are the subject of the news. In their model, there are two groups of traders in financial markets: professional arbitrage traders, less experienced traders, or noise traders. They found that the unpredictability of noise traders causes market risk. Professional arbitrage traders react to the behavior of noise traders rather than following their fundamentals. In doing so, professional arbitrage traders consider pseudo-signals such as volume, price patterns, and news.

Another manipulation affecting stock prices is the spread of rumors (Tan et al., 2023). Several studies in different media spaces have shown that the market reacts relatively sensitively to rumors (Clarkson et al., 2006; Zhao et al., 2010; Ahern & Sosyura, 2014). If investors believe the rumor, they will likely adopt a "buy on the rumor, sell on the news" strategy (Peterson, 2002; Schmidt, 2020); otherwise, they will trade in the opposite direction.

"Buy on Rumor and Sell on News" (BRSN) describes a recurring price pattern in financial markets that results from the collective behavior of investors driven by the affective (emotion-based) experience. BRSN characterizes an increase in the price of securities before positive expected events and a decrease in price after the events have passed or a decrease after the event's outcome is equal to or better than expected. The BRSN pattern results from the collective behavior of investors driven by affective (emotion-based) experience (Richard & Peterson, 2006). Rumors are characterized by the uncontrollability of their content and unknown origin (Schmidt, 2020). One of the most important roles of the press is to check the accuracy of the content published in the media. In contrast, the likelihood of stock price manipulation through rumors increases with the level of company-related information, and the use of rumors to manipulate the market increases insider trading, which often leads to stock price reversals (Tan et al., 2023).

## 2.5 Previous literature on the effect of media coverage on stock price

The earliest studies on the impact of media coverage on asset pricing mainly focus on the amount of information conveyed and consider the media as a neutral information broker. This starting point is entirely justified, since one of the main criteria for media is that the media is not inherently biased but must be objective. Only many years later that economic and political news was found to be often biased and targeted at the consumers of media outlets (Mullainathan & Shleifer, 2005). Previous studies found that media coverage effectively provides investors with additional information and reduces information asymmetry, affecting asset prices (Niederhoffer, 1971; Johnson et al., 2005; Bushee et al., 2010).

Most studies have linked the impact of news published in the media to changes in stock prices. Since Niederhofer’s (1971) paper, news sentiment analysis has been a popular area of research. His research demonstrated that price movements tend to move in the same direction on the first and second days, and that the market tends to overreact after horrific world events. Shiller (1984) found that when many market participants consider the sentiment expressed in the media, have similar irrational expectations, and follow each other (herd effect), this affects stock prices.

Tetlock (2007) conducted the following influential study, who examined the possible relationship between media and stock prices using the Wall Street Journal, finding that high pessimism causes market prices to decline. Later, Tetlock et al. (2008) used a bag-of-words model to assess whether corporate financial news can predict stock returns. The results showed that negative wording of corporate news predicted underreaction to the information. Fang and Peress (2009) examined the relationship between media coverage and expected stock returns. They found that stocks without media coverage earn more than those with media coverage, with the results particularly striking for small stocks, which have high ownership, low analyst followings, and often higher volatility. Research from the last decade tends to contradict this - the more media attention a particular stock receives or the more the financial media analyses what is happening in the financial market, the greater the trading volume or stock price reactions (Alanyali et al., 2013; Dougal et al., 2012; Pinnuck, 2014). Research into social media's impact has also increasingly emerged. For example, Lachana and Schröder (2023) compared sentiment from the Wall Street Journal and the social media platform Seeking Alpha and found that social media sentiment is more strongly related to short-term changes in stock prices. From this I arrive at the first hypothesis:

***$H_1$*:** *News sentiment has a statistically significant effect on stock price returns.*

Recent improvements in computing power and machine learning applications in text analysis have led to an exponential growth in research volumes over the past decade (Ligthart et al., 2021). Boudoukh et al (2013) found that when news is correctly identified, there is significantly more evidence of a strong relationship between stock price changes and information. Several studies have shown that negative news has a greater impact and predictive power on stock prices than positive news (Tetlock, 2007; Engelberg & Parsons, 2011). This is further supported by Uhl (2014), who analysed Reuters news and confirmed that negative sentiment has a stronger influence on market movements. Similarly, Liu et al. (2023) found that while company-specific positive news can increase trading activity, negative-toned news tends to have greater predictive power regarding future stock price performance. From this I arrive at the second hypothesis:

***$H_2$:*** *Negative news sentiment has a significantly greater effect on stock price performance than positive sentiment.*

As a result, previous studies have identified several relationships between news and stock prices over time, which is also the basis of this work. This study aims to provide an initial overview and analysis of the relationship between news published in the Estonian media and developments on the Tallinn Stock Exchange. Despite the growing role of media influence in shaping investor expectations and behavior, no research currently focuses on what is happening in the Estonian market. By filling this gap, the study helps better understand how media sentiment can influence local investor decisions and market movements in a small market.

# 3. Research design

## 3.1 Data collection and sample

This study uses a media-based quantitative methodology that combines sentiment analysis with the ARMAX model to investigate the relationship between media sentiment and stock market movements. This approach is well-suited to identify broad patterns of investor behavior over time based on publicly available data. Compared to surveys or interviews, which can often be biased, have limited capacity, and are not competent to capture real-time reactions, sentiment-based modeling allows for systematic and iterative analysis of real-world market-related content (Tetlock, 2007; Nassirtoussi et al., 2014).

Previous studies have shown that media sentiment measures are reliable indicators of investor sentiment and attention, which in turn affect asset pricing and market performance (Liu & Yu, 2020; Zhang & Wang, 2020). Accordingly, this method is theoretically and empirically acceptable for researching news sentiment and stock market dynamics.

The study analyses 12 companies listed on the Tallinn Stock Exchange (Table 1). The sample includes the most traded shares on the Tallinn Stock Exchange, accounting for transactions, trading volume, and turnover as components of the share. The most important factor is trading volume, which indicates interest in the share and indicates whether the share is in demand or not. The sample excluded shares that were only listed on the stock exchange at the end of the research period or had such a low trading volume that measuring the impact of the news was not justified.

Based on the sample, the companies can be divided into two groups. According to the company reports available on the Nasdaq Baltic Stock Exchange website, the most traded shares are Tallink Grupp, Tallinna Sadam, Coop Pank, Enefit Green, LHV Group, Nordecon, and TKM Grupp, whose annual trading volume exceeds the threshold of 15 million euros per period. The trading volume of other companies, Ekspress Grupp, Merko Ehitus, Arco Vara, Harju Elekter, and Tallinna Vesi, is below 10 million euros (Table 1).

TABLE 1. Tallinn Stock Exchange company's volume turnover and trades

| | Company | Ticker symbol | Volume | Turnover | Trades |
|---|---|---|---|---|---|
| 1 | AS Tallinna Vesi | (TVE1T.TL) | 3 992 785 | 50 784 375 | 89 234 |
| 2 | AS Harju Elekter | (HAE1T.TL) | 5 825 054 | 34 592 052 | 106 920 |
| 3 | Arco Vara AS | (ARC1T.TL) | 6 230 662 | 11 454 473 | 59 160 |
| 4 | AS Merko Ehitus | (MRK1T.TL) | 7 957 805 | 96 072 066 | 146 609 |
| 5 | AS Ekspress Grupp | (EEG1T.TL) | 8 370 992 | 8 826 813 | 36 977 |
| 6 | TKM Grupp AS | (TKM1T.TL) | 15 562 718 | 145 072 662 | 245 634 |
| 7 | Nordecon AS | (NCN1T.TL) | 20 662 286 | 2 253 519 | 94 183 |
| 8 | AS LHV Group | (LHV1T.TL) | 31 295 815 | 288 007 079 | 344 037 |
| 9 | Enefit Green AS | (EGR1T.TL) | 62 593 577 | 251 378 200 | 423 458 |
| 10 | Coop Pank AS | (CPA1T.TL) | 69 692 301 | 162 771 152 | 360 843 |
| 11 | AS Tallinna Sadam | (TSM1T.TL) | 87 314 392 | 151 248 215 | 255 114 |
| 12 | AS Tallink Grupp | (TAL1T.TL) | 301 643 270 | 208 637 044 | 368 773 |
| | **Total** | | **621 141 657** | **1 411 097 650** | **2 530 942** |

*Note: Values are ordered from smallest to largest by trading volume. Source: The Nasdaq Tallinn website; created by the author.*

All analyzed articles come from Äripäev, Estonia's most popular and only newspaper specializing in economic news. Äripäev's web portal publishes news daily, but articles related to stock market trading are published mainly on weekdays. Articles published on weekends are postponed until Monday of the new week or until the first trading day of the market's open business day. News published on holidays is also postponed to the next business day when the market is open. For news published on a business day, the time of publication of the news was also monitored to coincide with the time when the stock market is open for trading.

The author used the Äripäev news portal as the primary source to select the articles. The first step was to define the search period, which spanned from January 1, 2019, to December 31, 2023. This means that the news sample spans a five-year period to observe longer-term market trends and to identify greater commonalities. The main criterion for inclusion was that the primary focus of the news article was the company under analysis; such articles constitute the majority of the sample. In rare cases, the sample also includes opinion pieces, investment recommendations, comparative articles on listed companies (by sector), and stock market summaries. The latter's selection was based on whether the company was mentioned in the headline/introductory paragraph.

An exception was made for companies not listed on the stock exchange as of January 1, 2019, such as Coop Pank and Enefit Green. For these companies, articles were selected based on their listing date to avoid the influence of pre-listing information, which would also lack relevant market statistics for analysis. Additionally, the author excluded all news related solely to listed company subsidiaries, which is an important distinction contributing to more accurate results. The final sample consists of 1,930 articles (Table 2).

The author used R as the main programming environment for collecting and structuring the textual data, as it supports extensive text analysis and web scraping capabilities. The packages used included *RSelenium* for automated content extraction from the news website and *rvest* for parsing static HTML content. Similar tools have been effectively applied in previous sentiment-based market studies (Nassirtoussi et al., 2014). The final selection of articles was carried out manually based on relevance. Extracted content was analyzed using HTML element identifiers such as headlines and the 'div class' containing the article text.

TABLE 2. Sample of Articles

| | Company | 2019 | 2020 | 2021 | 2022 | 2023 | Total |
|---|---|---|---|---|---|---|---|
| 1 | AS Harju Elekter | 9 | 12 | 19 | 10 | 12 | 62 |
| 2 | Arco Vara AS | 8 | 15 | 34 | 13 | 7 | 77 |
| 3 | TKM Grupp AS | 22 | 25 | 15 | 6 | 18 | 86 |
| 4 | AS Ekspress Grupp | 13 | 19 | 22 | 22 | 16 | 92 |
| 5 | AS Tallinna Vesi | 33 | 16 | 22 | 14 | 15 | 100 |
| 6 | AS Tallinna Sadam | 17 | 31 | 18 | 21 | 19 | 106 |
| 7 | Nordecon AS | 29 | 27 | 31 | 17 | 24 | 128 |
| 8 | AS Merko Ehitus | 30 | 26 | 31 | 15 | 27 | 129 |
| 9 | Enefit Green AS | 0 | 0 | 31 | 60 | 74 | 165 |
| 10 | Coop Pank AS | 16 | 47 | 80 | 53 | 39 | 235 |
| 11 | AS Tallink Grupp | 43 | 113 | 47 | 35 | 45 | 283 |
| 12 | AS LHV Group | 83 | 108 | 115 | 83 | 78 | 467 |
| | **Total** | 303 | 439 | 465 | 349 | 374 | **1930** |

*Note: Enefit Green AS articles start from 2021 because trading of the company's shares on the stock exchange began in October 2021. Created by the author.*

## 3.2 Data processing and sentiment analysis

After scraping the news from the website, the author prepared the texts for analysis. Initial attempts to perform sentiment analysis using Estonian-language packages yielded unsatisfactory results, with sentiment scores clustering heavily around zero. Therefore, drawing on previous similar studies, the author opted to translate the news texts into English. Translation is widely regarded as an acceptable and practical method for incorporating non-English media sources into content analysis, especially when the working language of the study is English (Krippendorff, 2018). This study applies sentiment analysis, which necessitates the structuring and linguistic normalization of textual data. As most sentiment analysis tools and language resources are developed for English, it is standard practice to translate non-English texts, such as news articles, into English and to perform sentiment analysis on the translated versions. This approach has been successfully employed in earlier studies. For instance, Lin et al. (2022) translated multilingual social media posts into English and demonstrated that the combination of machine translation and deep learning improved the predictability of stock market movements. In this study, the translation was carried out using the TartuNLP neural machine translation system developed at the University of Tartu.
The word cloud in Figure 1 reflects the frequency of words in the headlines and main text of the analyzed articles - the larger the word, the higher the frequency of occurrence.


FIGURE 1. Word cloud constructed from the complete corpus of all news articles.
*Note: Created by the author.*

For sentiment extraction, the author pre-processed the texts by removing stop words, stripping punctuation, and converting all characters to lowercase. As previously mentioned, sentiment analysis is used to determine whether the tone of the text is positive or negative. The Python Natural Language Toolkit (NLTK) dictionary package was employed to compute sentiment scores. This tool assigns scores to individual words based on their emotional polarity, which are then aggregated into an overall sentiment score for the text. These scores were normalized onto a scale ranging from -1 to 1. Sentiment scores were subsequently linked with daily stock return data, calculated using the logarithmic difference between opening and closing prices, and retrieved from Yahoo Finance.

$$r_t = log(P_t/P_{t-1}) = log(P_t) - log(P_{t-1})$$

The following table highlights the 10 news stories with the most positive and 10 with the most negative news sentiment (Table 3). Articles with the highest positive sentiment scores are predominantly associated with increases in company revenues and profits. At the sectoral level, strong positive sentiment was observed in stories highlighting improved performance in specific industries, such as real estate and banking. News articles focusing on corporate leadership also received high positive sentiment scores, for example, when a company CEO was awarded the title of 'Executive of the Year'.

Negative news stories were primarily related to company losses and poor financial results. The analysis period also partially overlapped with the COVID-19 pandemic, which had a noticeable negative impact, particularly evident in the context of the shipping sector. The negative sentiment scores also include fraud-related news, legal disputes, and controversies involving listed companies.

TABLE 3. The most positive and negative news sentiments

| Date | Sentiment | Company |
|---|---|---|
| 08.02.2022 | 0,9600 | LHV Group |
| 05.11.2021 | 0,9590 | Enefit Green |
| 10.10.2019 | 0,9570 | TKM Grupp |
| 09.04.2021 | 0,9520 | Merko Ehitus |
| 09.04.2021 | 0,9520 | Nordecon |
| 28.10.2022 | 0,9520 | Tallink Grupp |
| 31.10.2023 | 0,9490 | Ekspress Grupp |
| 04.02.2020 | 0,9440 | LHV Group |
| 05.05.2023 | 0,9430 | Enefit Green |
| 25.08.2021 | 0,9420 | Coop Pank |

| Date | Sentiment | Company |
|---|---|---|
| 16.10.2023 | -0,9460 | LHV Group |
| 30.04.2021 | -0,9210 | Ekspress Grupp |
| 13.01.2022 | -0,9140 | Ekspress Grupp |
| 26.09.2023 | -0,9120 | Enefit Green |
| 10.03.2021 | -0,9100 | Tallink Grupp |
| 23.03.2020 | -0,9060 | Tallink Grupp |
| 19.07.2019 | -0,8930 | Tallinna Vesi |
| 14.02.2019 | -0,8720 | Arco Vara |
| 29.10.2020 | -0,8640 | Arco Vara |
| 02.03.2021 | -0,8600 | Tallink Grupp |

*Note: This table shows the 10 highest and lowest sentiments in news published between 01.01.2019 and 31.12.2023. It also shows the company's name. Authors' calculations.*

## 3.3 Approach to analysis and ARMAX model

The author decided to find a connection between news sentiments and stock returns based on existing stock price movement values. Such predictions are mostly made with different econometric models. The best solution seemed to be to use the Autoregressive moving average model. More specifically, the ARMAX model is a further development of ARMA. The author of the thesis used to implement an autoregressive-moving-average with exogenous inputs (ARMAX). The ARMAX model can be used to analyze how exogenous inputs (news sentiment) affect the dependent variable (stock return). In addition, the model also considers the time series patterns of stock returns, considering the past value of the stock and random noise. Thus, ARMAX is a suitable tool to investigate to what extent and with what time lag news sentiment affects stock returns. The AR component of the ARMAX model ($p$) shows how the previous day's performance affects the current day. MA component ($q$) shows how the effect of noise from previous random errors is transmitted to the value of the dependent variable. Exogenous inputs (*x*) measure the effect of an exogenous input (news sentiment) on the dependent variable (stock return). The last one is the error component for a model unexplained part. Taking as a main example such an early similar paper, Puzanova and Eratalay (2021) and Lapitskaya, Eratalay, and Sharma (2022), the following ARMAX model is constructed:

$$Y_t = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \gamma_1 D_t^+ + \gamma_2 D_t^- + \varepsilon_t$$

where Yt – log-return, Xt, Xt−1, Xt−2 – sentiment score on the same day and with day lags, Dt+ and Dt− – positive and negative outlier's dummies, εt – model residue, ϕi – AR component coefficients, θj – MA component coefficients, α – intercept, p, and q – orders of the AR and MA.

The empirical analysis is based on ARMAX(p,d,q) models, where exogenous variables related to news sentiment (xreg) were added sequentially. First, a baseline model was fitted, which included only dummy variables for extreme returns. The best ARMAX order was selected based on the Akaike Information Criterion (AIC), avoiding bias in comparing models.

Outlier cutoffs were identified based on the Z-scores of daily logarithmic returns, comparing returns to the mean and standard deviation of the entire sample. Several thresholds (±2.0; ±2.5; ±3.0) were tested to find the optimal balance between model stability and sensitivity (Table 4). The best fit was achieved with a threshold of ±2, which marked approximately 5% of the data points as outliers – this improved the level of autocorrelation of the residuals, lowered the AIC and SSE indicators, and improved the model's ability to describe stock price dynamics. Higher threshold values ignored important market shocks and degraded the model accuracy.

TABLE 4. AIC values of ARMAX models under different z-score cutoffs

| z-score | Index t | Index t-1 |
|---|---|---|
| 2 | 2807.535 | 2136.941 |
| 2,5 | 2955.302 | 2277.036 |
| 3 | 3059.384 | 2403.221 |

*Note: Authors' calculations.*

Four models were tested to assess the impact of sentiment: a baseline model, a model with general sentiment, a model with a distinction between positive and negative sentiment, and a combined model that included outliers and allowed testing for asymmetric effects. All models included dummy variables pos_outlier and neg_outlier. Sentiment variables were included with up to two-day lags (lags 0–2) and the fit of the models was compared based on AIC values, regression coefficients, and Ljung-Box residual diagnostics.

# 4. Results

## 4.1 Daily results of LHV Group LHV

This paper presents the indicators of LHV Group, as the most comprehensive data for analysis was available for this company. A brief overview of other companies is at the end of the chapter.

Four successively expanding ARMAX models were estimated for daily data, all using the ARMAX (1,0,1) base model. The fit of each model was assessed using AIC. In addition, the statistical significance and direction of the coefficients of the sentiment variables were interpreted. The returns data are stationary (p=0.01) based on the augmented Dickey-Fuller test. Residual autocorrelation analysis (Ljung-Box test with p-value 0.15) confirms that the model is statistically well-specified, the residuals are not autocorrelated, and the model is reliable.

The base model ARMAX (1,0,1) (Table 5) included two autoregressive components (AR and MA) and two outlier dummy variables representing extreme positive and negative returns.

$$R_t=\mu+\phi_1 R_{t-1}+\theta_1 \varepsilon_{t-1}+\gamma_1 \cdot pos\_outlier_t+\gamma_2 \cdot neg\_outlier_t+\varepsilon_t$$

The coefficients of AR(1) and MA(1) were not statistically significant (p > 0.1), suggesting that the dependence on the previous return and its shock is marginal. The base model serves as a reference model to assess the effects of adding sentiment variables.

In Model 1.1 (Table 5), the base model was extended by adding sentiment variables: general sentiment (x0) and its lagged values after one and two days (xlag1, xlag2).

$$r_t = \alpha + \beta_0 x_{0,t} + \beta_1 x_{lag1,t} + \beta_2 x_{lag2,t} + \phi_1 r_{t-1} + \theta_1 \varepsilon_{t-1} + \sum_{j=1}^{k} \gamma_j D_{j,t} + \varepsilon_t$$

General sentiment on the same day (x0) was statistically significant ($p < 0.01$), indicating a clear positive relationship between sentiment and same-day returns. The lagged variables were not significant. The autoregressive component AR(1) of the model and the moving average component MA(1) were both statistically highly significant ($p < 0.01$). A negative and high AR(1) value indicates a strong mean reversion effect in returns. This means that if there was a positive deviation in the previous period (i.e., the return was higher than the model predicted), it is highly likely to be followed by a negative correction. Such dynamics are consistent with a weaker form of the efficient market hypothesis, according to which prices tend to compensate for short-term over- or under-reactions. At the same time, a positive MA(1) value suggests that return shocks are transmitted in the short term. For example, if positive sentiment causes an unexpected price increase, this effect may continue into the next period before the market smooths it out with a correction. This can be interpreted as short-term inertia or momentum that sentiment and news can temporarily add to market behavior. Combined, these results suggest that the price behavior of LHV shares is characterized by a two-phase reaction: first, a slight continuation (positive MA) due to news or sentiment, followed by an intense and rapid normalization (negative AR). The model fit improved. This suggests that sentiment on the same day explains some of the return variation that the base model did not capture.

$$R_t = \mu + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \gamma_1 \cdot \mathrm{pos\_outlier}_t + \gamma_2 \cdot \mathrm{neg\_outlier}_t + \sum_{i=0}^{2} \beta_i^{(+)} \cdot \mathrm{pos}_{t-i} + \sum_{i=0}^{2} \beta_i^{(-)} \cdot \mathrm{neg}_{t-i} + \varepsilon_t$$

In Model 1.2 (Table 5, Figure 2), sentiment was separated into positive (pos0, pos1, pos2) and negative (neg0, neg1, neg2), each with a lag of up to two days.

The results show that both positive sentiment on the same day (pos0) and negative sentiment on the same day (neg0) were statistically significant ($p < 0.01$). Their coefficients suggested that both good and bad news immediately impact returns, with the negative impact being more substantial. The lagged sentiment variables (pos1, pos2, neg1, neg2) were not statistically significant, although neg1 and neg2 showed a moderate positive trend. The AR(1) and MA(1) coefficients are similar to the previous models. The model fit was also strong. Although the AIC was slightly higher than in model 1.1, the effect size and statistical significance suggest that distinguishing sentiments reveals an asymmetric effect – negative news has a more substantial impact than positive news.

In Model 1.3 (Table 5), general sentiments (x0, xlag1, xlag2) and negative sentiment (neg0, neg1, neg2) were combined to investigate whether distinguishing negative information from general tone could better explain returns.

$$R_t = \mu + \phi_1 R_{t-1} + \theta_1 \varepsilon_{t-1} + \gamma_1 \cdot \text{pos\_outlier}_t + \gamma_2 \cdot \text{neg\_outlier}_t + \sum_{i=0}^{2} \beta_i \cdot x_{t-i} + \sum_{i=0}^{2} \delta_i \cdot \text{neg}_{t-i} + \varepsilon_t$$

In addition to the general market tone, negative information can have an independent and stronger impact on returns. Therefore, distinguishing negative sentiment from general tone can indeed improve the explanatory power of the model, as negative news tends to affect investors more than neutral or positive news, which is also consistent with previous empirical findings. General sentiment on the same day (x0) was again statistically significant ($p < 0.01$), confirming an immediate effect. However, lagged sentiment (xlag1, xlag2) was not statistically significant. The negative sentiment variables (neg0, neg1, neg2) were not statistically significant, although neg1 and neg2 indicated possible lagged effects. This may suggest that negative sentiment no longer has an independent statistical effect once the general tone is considered. AR(1) and MA(1) were again similar to the previous. The model fit was equivalent to model 1.3. The results suggest that while the sentiment lag has some effect, the overall tone explains the majority of the return movement on the same day.

Daily data analysis showed that including sentiment variables improved the fit of the ARMAX models. The most influential factor was the general sentiment on the same day, which was statistically significant in all extended models. When sentiment was divided into positive and negative, negative sentiment substantially impacted stock returns. The results suggest an asymmetric market response: bad news affects returns more and more immediately than good news.

TABLE 5. LHV Group ARMAX (1,0,1) daily based values

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| ***ar1*** | 0.08 | -0.86*** | -0.86*** | 0.00 | 0.20 | -0.86*** |
| ***ma1*** | -0.12 | 0.80*** | 0.81*** | 0.10 | -0.23 | 0.81*** |
| ***intercept*** | 0.01 | 0.01 | 0.02 | -0.04** | -0.04** | 0.02 |
| ***pos_outlier*** | 4.16*** | 3.99*** | 4.01*** | 2.33*** | 2.33*** | 4.01*** |
| ***neg_outlier*** | -4.42*** | -4.24*** | -4.24*** | -1.62*** | -1.61*** | -4.24*** |
| ***x0*** | NA | 0.36*** | NA | 0.25*** | NA | 0.32*** |
| ***xlag1*** | NA | -0.04 | NA | -0.05 | NA | -0.01 |
| ***xlag2*** | NA | 0.02 | NA | 0.05 | NA | 0.00 |
| ***index t*** | NA | NA | NA | 0.98*** | 0.98*** | NA |
| ***index t-1*** | NA | NA | NA | -0.06*** | -0.06*** | NA |
| ***pos0*** | NA | NA | 0.32*** | NA | 0.25*** | NA |
| ***pos1*** | NA | NA | -0.01 | NA | -0.06 | NA |
| ***pos2*** | NA | NA | 0.00 | NA | 0.07 | NA |
| ***neg0*** | NA | NA | 0.63** | NA | 0.63*** | 0.31 |
| ***neg1*** | NA | NA | -0.26 | NA | 0.04 | -0.25 |
| ***neg2*** | NA | NA | 0.16 | NA | -0.14 | 0.16 |
| **Log-Likelihood** | -2297.37 | -2288.76 | -2287.95 | 1649.26 | -1648.66 | -2287.95 |
| **AIC** | 4606.73 | 4595.53 | 4599.89 | 3320.51 | 3325.32 | 4599.89 |
| **BIC** | 4639.78 | 4645.11 | 4666.00 | 3381.11 | 3402.44 | 4666.00 |
| **SSE** | 1225.95 | 1313.59 | 1312.41 | 651.52 | 651.09 | 1312.41 |

*Standard errors are given in parentheses. [*], [**], [***] show significance at 0.1, 0.05, and 0.01 significance levels, respectively–source: Authors' calculations.*

To assess the model's statistical robustness, the results obtained using OLS and HAC (Newey-West) standard errors were compared. Although the HAC standard errors were generally larger, statistically significant variables (including sentiment t, stock market index, and outlier dummies) retained their significance even with robust estimates. This confirms that the model conclusions do not depend strictly on the fulfillment of the classical OLS assumptions and are robust even in the case of possible autocorrelation or heteroscedasticity of the residuals.

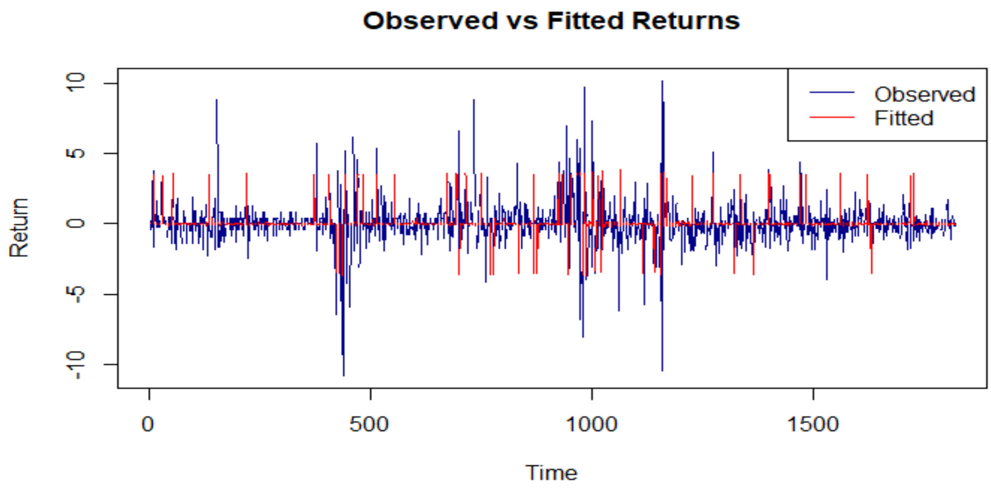


FIGURE 2. Observed and fitted daily returns of LHV Group Pos/Neg model (1.2).
*Note: The y-axis shows daily log returns, which approximate percentage price changes. Values above zero indicate positive returns, while values below zero indicate negative returns. The x-axis represents time in days. Created by the author.*

Figure 1 shows the fitted and observed daily stock returns for model 1.3. The fitted returns broadly follow the actual values, especially in the middle. The model remains modest in the case of extreme movements, which is typical of linear ARMAX approaches. However, the graph illustrates that the overall fit and directionality of the model improved when sentiment was included.

In this article, the determinants of LHV share returns were estimated using ARMAX models, which included message sentiment and the Tallinn Stock Exchange Index as exogenous variables. The impact of the stock exchange index was considered separately, both in the same period (Index at time t) and with a lag (Index at time t–1), to analyze the potential lagged market influence. To model extreme return values, dummy variables were constructed based on z-score cutoffs (2, 2.5, and 3). The Akaike Information Criterion (AIC) results were compared to identify the best-fitting model. The empirical results showed that the best model was achieved at a z-score threshold 2 when current and lagged stock exchange index values were included as exogenous inputs. This suggests that moderately defined extreme return days and the lag component of the market trend significantly increase the model's explanatory power. The found reference effect may reflect the inertia of market participants' reactions or the late reflection of news in the share price.

The results of the ARMAX modelling process for LHV Group confirm that including news sentiment improves explanatory power compared to the baseline ARMAX (1,0,1) model. General sentiment on the same day (x0) was consistently statistically significant ($p < 0.01$) across all extended models, suggesting a clear contemporaneous relationship between sentiment tone and stock returns. This supports the **first hypothesis**, which proposed that news sentiment statistically affects stock price performance. When sentiment was disaggregated into positive and negative components (Model 1.2), both pos0 and neg0 were statistically significant ($p < 0.01$), with the coefficient for negative sentiment being larger in magnitude. This indicates that both good and bad news influence returns immediately, but the effect of negative sentiment is more substantial. Although the AIC value for this model was slightly higher than for Model 1.1, the size and significance of the coefficients support the **second hypothesis**, suggesting an asymmetric market response: negative sentiment has a greater impact on daily stock returns than positive sentiment. Model 1.3, which combined general sentiment with specific negative sentiment variables, showed that general sentiment remained significant while the individual effect of lagged negative sentiment diminished. This suggests that much of the return variation related to sentiment is already captured by the general tone of news, especially on the same day. Taken together, the results point to a two-phase price reaction pattern for LHV Group: an immediate reaction driven by general sentiment and, to a lesser extent, a possible delayed response to negative news. These findings are consistent with theories of limited investor attention and asymmetric information processing, and they reinforce the conclusion that sentiment, particularly negative sentiment, plays a statistically and economically meaningful role in short-term price formation.

The results of the daily data analysis of other companies presented in the appendix (Appendix A) largely confirm the main results for the LHV share. A statistically significant short-term effect of sentiment was evident, especially when news was released. Models that included sentiment (especially negative sentiment) showed a better fit than the baseline model. However, there were also differences between companies. For three companies, including sentiment variables, did not improve the fit of the models (AIC/SSE) compared to the baseline model. This suggests that not all stocks respond to news signals in the same way - possible reasons range from company visibility and the intensity of the news cycle to differences in investor behavior.

Including the Tallinn Stock Exchange Index as an exogenous variable provided valuable additional information. This suggests that the index helps explain the entire price movement. and may mean that sentiment and the index partly reflect simultaneous market mechanisms, where the news flow and the general stock market sentiment move hand in hand. However, it is noteworthy that many news reports noted that the rise/fall of the stock of company X or Y contributed to the rise/fall of the stock exchange index. Secondly, it is worth noting that the effect of positive sentiment was significant in some cases, contrary to the dominant negative focus of LHV results. This suggests that the market is more sensitive to positive news flows for some companies. The sentiment effect is more pronounced for more companies than a slightly longer-term effect. It also became apparent when analyzing other companies' results that negative news's impact is longer-term, and positive sentiment dissipates faster.

## 4.2 Weekly results of LHV Group

Due to a substantial amount of missing values in the dataset, the author also tested models using weekly aggregated data. However, the weekly data results were less conclusive than those derived from daily data (Table 6). This suggests that the impact of news sentiment on stock price performance may be more detectable at a higher temporal resolution. Thus, the effect of sentiment variables on LHV Group share performance was estimated using four ARMAX models (with MA components), also using weekly data. All models also included exceptional dummy variables and were estimated using log-likelihood, AIC, BIC, and SSE.

A comparison of regression models based on weekly data showed that sentiment-based models improved the explanatory power of the underlying structure. Still, increasing model complexity did not always lead to a better fit. The underlying model, which included only MA components, intercept, and extreme returns dummy variables, achieved the lowest AIC and BIC values, indicating optimal simplicity and a small number of parameters. However, the log-likelihood and sum of squares of residuals of this model were higher, indicating limited explanatory power. Adding sentiment (model "Sentiment") significantly improved the model's fit based on log-likelihood and SSE indicators, with same-week sentiment being positive and statistically significant, indicating that the general sentiment of market participants affects stock returns. However, there were also lag effects, suggesting a possible lag in market adjustment. The "Pos/Neg" model, where sentiment was split into positive and negative components, did have some interesting individual effects (the adverse sentiment lag effect) (Figure 3). Still, the overall model quality did not improve compared to the general sentiment model. The even more complex "Mixed" model combined all sentiment components and showed signs of collinearity and decreased statistical efficiency, suggesting excessive parameterization without additional explanatory value. In all models, the positive and negative outlier dummies remained statistically highly significant and largely persistent, confirming their role in modeling extreme movements in returns even in weekly data.

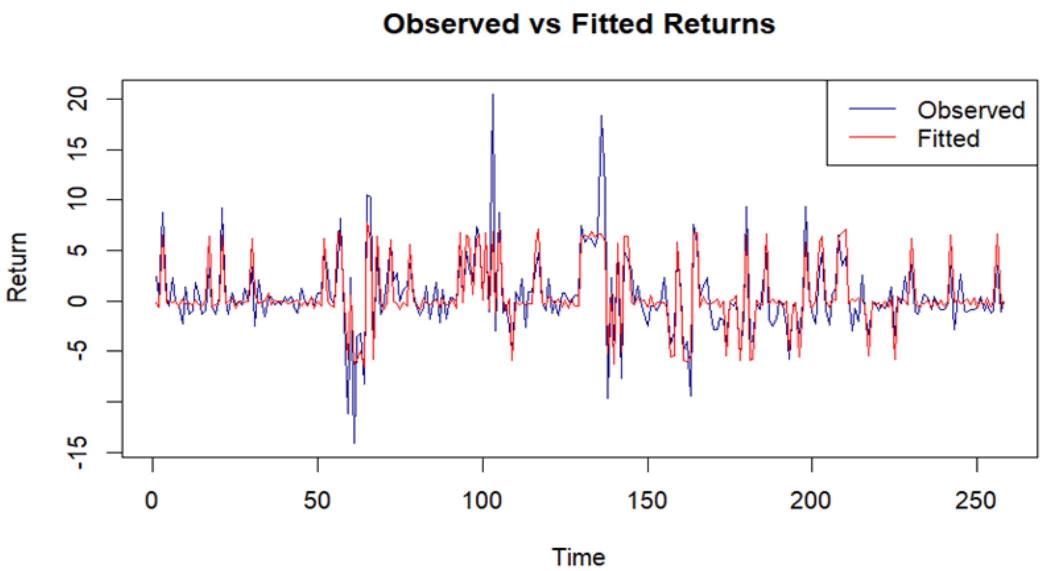


FIGURE 3. Observed and fitted weekly returns of LHV Group Pos/Neg model (1.2)
*Note: The y-axis shows weekly log returns, which approximate percentage price changes. Values above zero indicate positive returns, while those below zero indicate negative returns. The x-axis represents consecutive weekly periods. Created by the author.*

TABLE 6. LHV Group ARMAX (0,0,4) weekly based values

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| ***ar1*** | 0.30 | 0.31 | 0.31 | -0.07 | -0.06 | 0.31 |
| ***ar2*** | 0.47 | 0.46 | 0.47 | -0.00 | -0.00 | 0.47 |
| ***ma1*** | -0.26 | -0.26 | -0.26 | 0.15** | 0.14** | -0.26 |
| ***ma2*** | -0.52 | -0.53 | -0.53 | -0.16** | -0.16** | -0.53 |
| ***intercept*** | -0.10 | -0.39 | -0.45 | -0.35** | -0.40* | -0.45 |
| ***pos_outlier*** | 6.22*** | 6.18*** | 6.18*** | 4.60*** | 4.60*** | 6.18*** |
| ***neg_outlier*** | -4.84*** | -4.94*** | -4.93*** | -2.33*** | -2.32*** | -4.93*** |
| ***x0*** | NA | 0.48 | NA | 0.35 | NA | 0.51 |
| ***xlag1*** | NA | -0.14 | NA | -0.24 | NA | -0.14 |
| ***xlag2*** | NA | 0.68* | NA | 0.47* | NA | 0.77* |
| ***index t*** | NA | NA | NA | 6.03*** | 6.06*** | NA |
| ***index t-1*** | NA | NA | NA | -0.30 | -0.31 | NA |
| ***pos0*** | NA | NA | 0.51 | NA | 0.39 | NA |
| ***pos1*** | NA | NA | -0.14 | NA | -0.32 | NA |
| ***pos2*** | NA | NA | 0.77* | NA | 0.60* | NA |
| ***neg0*** | NA | NA | 0.33 | NA | 0.28 | -0.18 |
| ***neg1*** | NA | NA | -0.23 | NA | 0.26 | -0.10 |
| ***neg2*** | NA | NA | -0.02 | NA | -0.61 | -0.80 |
| **Log-Likelihood** | -553.13 | -550.65 | -550.52 | -476.71 | -476.09 | -550.52 |
| **AIC** | 1122.27 | 1123.29 | 1129.04 | 979.42 | 984.18 | 1129.04 |
| **BIC** | 1150.69 | 1162.37 | 1178.78 | 1025.61 | 1041.03 | 1178.78 |
| **SSE** | 1099.83 | 1078.80 | 1077.76 | 607.81 | 604.91 | 1077.76 |

*Standard errors are given in parentheses. [*], [**], [***] show significance at 0.1, 0.05, and 0.01 significance levels, respectively–source: Authors' calculations.*

The results obtained from weekly data for all analyzed companies were much weaker and less consistent than those obtained from daily data. For all models considered (baseline, sentiment, positive/negative, mixed model), the effects of sentiment variables mostly were statistically insignificant, and the regression coefficients were small or unstable.

First, the effect of sentiment is not systematically evident. In some cases, there was a statistically significant effect on sentiment or its lag. However, these effects were patchy; for example, they were only in the second week lag (2nd lag), without a logical pattern from the short to medium term lag. Such isolated effects, which do not repeat across models or companies, are more likely to be random or data-sensitive than to indicate a reliable pattern. Second, distinguishing sentiment into positive and negative did not improve the overall model fit (AIC, SSE) or increase the significance of the coefficients compared to the usual general sentiment. Also, the Tallinn Stock Exchange Index did not improve the quality of the models, although in some cases it improved the distribution of residuals and reduced autocorrelation.

In summary, these companies' findings suggest that it is more difficult to assess the impact of sentiment based on weekly data: the effect may dissipate as the data is aggregated or weaken as news hits the market in the first few days. However, some models showed a slight and immediate response to positive sentiment, suggesting that certain types of news can trigger immediate reactions, but these results were inconsistent.

# 5. Discussion

This work aimed to investigate whether news sentiment affects the performance of Tallinn Stock Exchange stocks and how important the distinction between positive and negative information is. The analysis was conducted using ARMAX-type models, where sentiment variables were added as exogenous factors. The model mainly focused on analyzing daily data. In addition, it also tested the aggregation of data into weeks.

Daily data analysis revealed that sentiment has a clear and direct effect on share performance, especially in the case of negative sentiment. The best fit was provided by the model divided into positive and negative news (1.2) and the combined model (model 1.3), which included both general and negative sentiment. Negative information had a more substantial, statistically significant effect than positive information, which aligns with the psychological and financial theory position, according to which the market reacts to negative information faster and more intensively. Negative news often substantially impacts market behavior because it creates higher risk and uncertainty perceptions, and triggers faster and more extensive investment decisions. This is confirmed by both behavioral theories (negativity bias, loss aversion) and empirical results.

The results of this paper show that improving the baseline model with media sentiment improves model fit and has a rapid and short-term impact on stock returns on the same day. This supports the notion that news sentiment contains information that immediately influences market participants' expectations and behavior.

Today's financial markets process information at a rapid pace: news is disseminated in real time, and both professional investors and algorithmic trading systems react to its tone immediately. General sentiment is a mood barometer that summarizes the emotional tone of financial and macro news. Thus, the market does not always need a fundamental shock; a change in general sentiment can trigger an immediate price adjustment.

More recent studies support this. For example, Li et al. (2023) and Deng et al. (2018) show that text-based sentiment (based on both news and social media) affects stock returns on the same day and improves forecasting models. Tetlock (2007) also showed that media tone (pessimism index) affects market returns on the following or same day, even if fundamental information does not change. This suggests that sentiment is part of the market's perceived information, which is evaluated and considered immediately.

The article is also consistent with Puzanova and Eratalay’s (2021) findings on real estate news sentiment on stock returns, confirming that news sentiment affects stock returns on the same day. It was also observed that the positive effect was more diffuse over time, while general or negative sentiment had an immediate effect, suggesting the primacy of price reactions in sentiment perception.

This knowledge is also common in previous literature: negative sentiment has a more substantial effect (Tetlock, 2007; Uhl, 2014). This thesis aligns with previous work that has also found that media sentiment has a short-term impact on market events. That negative news has a greater impact on market events than positive news. One possible reason may be that people fear and perceive losses more than gains, which is consistent with prospect theory (Kahneman & Tversky, 1979). People have a cognitive tendency to pay more attention to negative news, as well as to react to it more quickly and to consider it more credible. Negative news encourages more herd behavior (Gross & Tsetlin, 2004). Negative news also makes investors act more, such as selling or revaluing, while investors react less positive news, which tends to simply “confirm” the knowledge.

The sentiment effect was generally weaker and less statistically significant based on weekly data. However, some effects appeared with a lag. This suggests that the sentiment effect is short-term and mainly perceptible only on a daily basis. When daily fluctuations are averaged over a week, information about when precisely the sentiment effect occurred is lost (whether the effect occurred on Monday or Friday). As a result, the reaction to the "average" is spread out, causing the statistical relationship to dissipate. This is supported by Engelberg & Parsons (2011), who emphasize that the effect of sentiment is more substantial at the daily level and decreases when information is aggregated over more extended periods. Sentiment can be short-lived and quickly dissipated by nature. The emotional charge of news affects the market immediately, but its effect can fade within a few days, so weekly price movements no longer reflect it. Da, Engelberg & Gao (2011) show also that sentiment is strongest within 1–2 days after the event. Yoon & Takahashi (2025) found that small-cap stocks respond quickly to sentiment, but the effect dissipates within a few days, which could also be characteristic of the Tallinn Stock Exchange.

If sentiment is reacted to immediately, within a week, the market has already fully priced in the information, resulting in a weak correlation "in hindsight". Lagged effects may indicate secondary processing, for example, for assets or news with less attention. Weekly data means fewer observations, which in a statistical model means less power to find significant relationships. Ordinary effects may also remain statistically insignificant, even though they exist.

Including the index in the model helped better control for the broader underlying market trend and allowed us to distinguish company-specific effects from broader market dynamics. The inclusion of the index improved the result by showing that the returns of shares were closely related to the general market trend – the share moved in line with the general market rise and uptrend. This confirms that the sentiment effect does not operate in a vacuum but largely corresponds to the market background. However, the opposite effect cannot be ruled out. For example, LHV shares are one of the most traded shares on the Tallinn Stock Exchange, and the movement of their share price can potentially also lead to a change in the general market index. This is supported by the fact that some news items included, among other things, the text: "The rise/fall of X stock contributed to the rise/fall of the Tallinn Stock Exchange index." It also shows that the stock exchange index has significant explanatory power in determining short-term returns. The media effect is more potent in smaller, less analyzed stocks – the index helps to distinguish the mixing of market and sentiment (Engelberg & Parsons, 2011).

In conclusion, news sentiment, especially negative news, impacts the performance of the Tallinn Stock Exchange. However, this impact is short-term and evident in daily data. In practice, investors and traders could consider the tone of the news feed when making decisions, especially for short-term positions. A limitation of the study is that sentiment is measured using automatic text analysis methods, which may miss some of the context. The models used a reliable ARMAX approach, which may not capture nonlinear effects. A more sector-specific analysis of the news could be tested in the future. Certain limitations may also arise from the translated texts, as many of the original Estonian texts stood out for their use of figurative language. It would also be worth treating outliers with a higher z-score cutoff so that the dataset considers more exceptions and tries to explain them better. In the future, a more sector-specific analysis of news could be tested. In addition, it would be worth including more articles in the future, because at the moment, news from only one news portal was used, which resulted in a sample size that was not very large in terms of companies. In addition, considering the implementation of the GARCH model in the future would allow studying how sentiment affects not only the return on a stock but also its volatility and the extent of price fluctuations.

# 6. Conclusions

Predicting what will happen on the stock market is the dream of all profit-maximizing investors. Throughout history, various factors have been seen as the causes of stock market price movements. The attitude towards understanding stock market price movements has been different. The most important and main starting point of the last century was that markets function efficiently, according to which investors behave rationally, and what happens in the markets can be explained on the basis of fundamental analysis. In the later period, a clear understanding has developed that in addition to these factors, investor behavior is also influenced by psychological aspects, emotions, herd behavior, heuristics, cognition, and media coverage. There has been much speculation about the role of the media in the formation of stock prices. It has increasingly been found that the media plays an important role in price movements, and both the influence of the media and the amount of research on this topic have grown over time. The greater volume and faster and wider distribution of information have contributed to the growth of the impact of the media. Social media has opened up entirely new research opportunities.

The aim of this paper was to investigate the relationship between news sentiment and stock price movements on the Tallinn Stock Exchange and to find confirmation of the hypotheses that news sentiment affects stock price movements and that the effect of negative sentiment is more substantial than positive sentiment. The paper analyzed news from 12 Tallinn Stock Exchange companies in 2019–2023. The sample consisted of 1,930 news items, all of which originated from the newspaper Äripäev and were originally in Estonian. The analysis used news headlines and introductory paragraph translated into English for better analysis, which is a common practice for non-English news. The NLTK dictionary package was used to determine the news sentiment, which helped define the sentiment of the news by giving it a score between 1 and -1.

The results of the analysis showed that there is a relationship between news sentiment and stock performance. Adding sentiment to the base model and dividing the data into positive and negative led to results that showed that the effect of sentiment is manifested on the same day, and that negative news affects market events more than positive news. The suitability of the models was assessed based on AIC values. The improvements in the AIC values of the models after sentiment and division into positive/negative were not large but were still statistically significant and improved the model. The outliers that emerged prompted to test whether adding a general market index to the model would help further improve the explanation of what is happening in the market, which was also confirmed. Analysis of data aggregated by week gave weaker and less significant results. No precise data pattern emerged across the analyzed companies, and in some cases, the model detected the effect after a long delay.

From the daily results, it can be concluded that although the effect of news on shares on the Tallinn Stock Exchange is not significantly strong, the relationships that have emerged are still statistically significant. The inclusion of news sentiment in prices indicates a short-term effect of sentiment. The emergence of a more substantial impact of negative news can be explained by the fact that it is inherent in human nature to perceive losses more painfully than equivalent gains. Negative news also makes investors take more action, since positive news is taken as confirmation of what is already known. Due to the fact that the sample of the work was not very large, broader generalizations cannot be made in the context of the Tallinn Stock Exchange, but the approach to the topic raised new research questions for assessing the relationship between news and stock prices. The work could be further developed by including volatility in the stock price formation model and compiling a larger sample.

## References

1. Ahern, K. R., & Sosyura, D. (2014). Rumor has it: Sensationalism in financial media. The Review of Financial Studies, 28(7), 2050–2093. http://dx.doi.org/10.2139/ssrn.2264468
2. Alanyali, M., Moat, H. S., Curme, C., Preis, T., & Stanley, H. E. (2013). Quantifying the relationship between financial news and the stock market. Scientific Reports, 3, 3578. https://doi.org/10.1038/srep03578
3. Barakat, M., & Barakat, H. (2024). Does the overreaction hypothesis stand the test of time across comparable developed markets: A tale of two crises. Journal of International Finance and Economics, 24, 24–52. https://doi.org/10.18374/JIFE-24-1.3
4. Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies, 21(2), 785–818. https://doi.org/10.1093/rfs/hhm079
5. Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. Handbook of the Economics of Finance, 1, 1053–1128.
6. Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2018). Extrapolation and bubbles. The Journal of Financial Economics, 129(2), 203–227. https://doi.org/10.1016/j.jfineco.2018.04.008
7. Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. Journal of Financial Economics, 49(3), 307–343.
8. Bernstein, E. S. (2012). The Transparency Paradox: A Role for Privacy in Organizational Learning and Operational Control. Administrative Science Quarterly, 57(2), 181-216. https://doi.org/10.1177/0001839212453028
9. Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2013). Which news moves stock prices? A textual analysis. The Review of Financial Studies, 26(12), 3481–3520. https://doi.org/10.1093/rfs/hht027
10. Bushee, B. J., & Miller, G. S. (2012). Investor relations, firm visibility, and investor following. The Accounting Review, 87(3), 867–897. https://doi.org/10.2308/accr-10211

11. Bushee, B. J., Core, J. E., Guay, W., & Hamm, S. J. W. (2010). The Role of the Business Press as an Information Intermediary. Journal of Accounting Research, 48(1), 1–19. http://www.jstor.org/stable/40389239
12. Brunswick Group. (2023). *2023 Digital Investor Survey*. Retrieved from https://www.brunswickgroup.com/perspectives/digital-investor-survey/
13. Byrne, A., & Brooks, M. (2008). Behavioral finance: Theories and evidence. Financial Services Review, 17(3), 261–277.
14. Campbell, J. Y., & Shiller, R. J. (1988). Stock prices, earnings, and expected dividends. The Journal of Finance, 43(3), 661–676. https://doi.org/10.1111/j.1540-6261.1988.tb04598.x
15. Chapman, C., & Roberts, S. (2022). Loss aversion and overconfidence in U.S. equity markets. Review of Financial Studies, 35(4), 1620–1650. https://doi.org/10.1093/rfs/hhac002
16. Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2021). Wisdom of crowds: The value of stock opinions transmitted through social media. The Review of Financial Studies, 27(5), 1367–1403. https://doi.org/10.1093/rfs/hhaa051
17. Chowdhury, S. G., Routh, S., & Chakrabarti, S. (2014). News Analytics and Sentiment Analysis to Predict Stock Price Trends.
18. Clarkson, P. M., Kao, J. L., & Richardson, G. D. (2006). Evidence that management discussion and analysis (MD&A) is a part of a firm's overall disclosure package. Contemporary Accounting Research, 23(1), 301–331. https://doi.org/10.1506/2H6Q-1H68-RC64-H0P5
19. Collins, W. T. (2020). Noise traders, fintech, and equity market volatility (Undergraduate honors thesis). Georgia Southern University. https://digitalcommons.georgiasouthern.edu/honors-theses/544
20. Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. The Journal of Finance, 66(4), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x
21. Davis, A. (2006). Media effects and the active elite audience: A study of communications in the city of London. European Journal of Communication, 21(3), 303–321. https://doi.org/10.1177/0267323106066635
22. De Bondt, W. F., & Thaler, R. H. (1985). Does the stock market overreact? The Journal of Finance, 40(3), 793–805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x

23. De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. Journal of Political Economy, 98(4), 703–738. https://doi.org/10.1086/261703
24. Deng, S., Mitsubuchi, T., Shioda, K., Shimada, T., & Sakurai, A. (2018). Combining technical analysis with sentiment analysis for stock price prediction. IEICE Transactions on Information and Systems, E101.D(4), 901–910. https://doi.org/10.1587/transinf.2017EDP7295
25. Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. Academy of Management Review, 20(1), 65–91. https://doi.org/10.5465/amr.1995.9503271992
26. Dong, X., Xu, S., Liu, J., & Tsai, F. S. (2022). Does media sentiment affect stock prices? Evidence from China's STAR market. Frontiers in Psychology, 13, 1040171. https://doi.org/10.3389/fpsyg.2022.1040171
27. Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. Review of Financial Studies, 25(3), 639–679. https://doi.org/10.1093/rfs/hhr133
28. Elhussein, N. H. A., & Abdelgadir, J. N. A. (2020). Behavioral Bias in Individual Investment Decisions: Is It a Common Phenomenon in Stock Markets? International Journal of Financial Research, 11(6), 25-36. https://doi.org/10.5430/ijfr.v11n6p25
29. Fama, E. F. (1965). The Behavior of Stock-Market Prices. The Journal of Business, 38(1), 34–105. http://www.jstor.org/stable/2350752
30. Fama, E. F. (1970). Efficient Capital Markets: Review of Theory and Empirical Work. Journal of Finance, 25(2), 383–417.
31. Fama, E. F. (1997). Market Efficiency, Long-Term Returns, and Behavioral Finance. Available at SSRN: https://ssrn.com/abstract=15108 or http://dx.doi.org/10.2139/ssrn.15108
32. Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. The Journal of Finance, 64(5), 2023–2052. https://doi.org/10.1111/j.1540-6261.2009.01493.x
33. Fromlet, H. (2001). Behavioral finance: Theory and practical application. Business Economics, 36(3), 63–69.

34. Guldiken, O., Tupper, C., Nair, A., & Yu, H. (2017). The impact of media coverage on IPO stock performance. Journal of Business Research, 72, 24-32.
35. Gupta, E., Preetibedi, & Poonamlakra. (2014). Efficient Market Hypothesis V/S Behavioural Finance. IOSR Journal of Business and Management, 16(4), 56–60. http://iosrjournals.org/iosr-jbm/papers/Vol16-issue4/Version4/H016445660.pdf
36. Hautsch, N., & Groß-Klußmann, A. (2011). When machines read the news: Using automated text analytics to quantify high-frequency news-implied market reactions. Journal of Empirical Finance, 18, 321–340. https://doi.org/10.1016/j.jempfin.2010.11.009
37. Hermida, A., Fletcher, F. J., Korell, D., & Logan, D. (2012). SHARE, LIKE, RECOMMEND. Journalism Studies, 13, 815–824. https://www.tandfonline.com/doi/full/10.1080/15427560.2012.653020?src=recsys
38. Hervé, F., Zouaoui, M., & Belvaux, B. (2019). Noise traders and smart money: Evidence from online searches. Economic Modelling, 83, 141–149. https://doi.org/10.1016/j.econmod.2019.03.008
39. Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance, 48(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x
40. Joanna Strycharz, J., Strauss, N., & Trilling, D. (2018). The Role of Media Coverage in Explaining Stock Market Fluctuations: Insights for Strategic Financial Communication. International Journal of Strategic Communication, 12(1), 67–85. https://doi.org/10.1080/1553118X.2017.1378220
41. Johnson, W. C., Kang, H., & Yi, S. (2005). The value of media in financial markets. Working Paper, University of Iowa. (Kui sul on täpne väljaande nimi või DOI, saan seda täiendada.)
42. Kahneman, D., & Tversky, A. (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453–458. https://doi.org/10.1126/science.7455683
43. Karlsson, M. (2011). The immediacy of online news, the visibility of journalistic processes, and a restructuring of journalistic authority. Journalism, 12, 279–295. https://doi.org/10.1177/1464884910388223

44. Kohut, A., Doherty, C., Dimock, M., & Keeter, S. (2010). Americans spend more time following the news. Pew Research Center.
45. Kräussl, R., & Mirgorodskaya, E. (2014). The role of sentiment in European financial markets. Applied Economics Letters, 21(16), 1157–1160. https://doi.org/10.1080/13504851.2014.910173
46. Krippendorff, K. (2018). Content analysis: An introduction to its methodology (4th ed.). SAGE Publications.
47. Lachana, R., & Schröder, M. (2023). Financial news and social media: Who is more informative? Finance Research Letters, 54, 103738. https://doi.org/10.1016/j.frl.2023.103738
48. Lapitskaya, D., Eratalay, M. H., & Sharma, D. (2022). Investor sentiment, media coverage, and stock return dynamics: Evidence from an emerging market. Journal of Behavioral and Experimental Finance, 34, 100726. https://doi.org/10.1016/j.jbef.2022.100726
49. Lehavy, R., & Sloan, R. G. (2008). Investor recognition and stock returns. Review of Accounting Studies, 13(2–3), 327–361. https://doi.org/10.1007/s11142-007-9049-4
50. Li, T., Urquhart, A., & Zhang, Y. (2022). Investor behavior and cryptocurrency markets. Journal of International Financial Markets, Institutions and Money, 76, 101496. https://doi.org/10.1016/j.intfin.2021.101496
51. Li, X., Shen, H., Zhang, Y., & Liu, B. (2023). News and social media sentiment for stock return prediction: A deep learning approach. Expert Systems with Applications, 213, 119157. https://doi.org/10.1016/j.eswa.2022.119157
52. Ligthart, M., van Ooijen, R., & van der Wel, M. (2021). The rise of textual analysis in finance. Journal of Economic Surveys, 35(5), 1486–1517. https://doi.org/10.1111/joes.12443
53. Lill, A. (2025, January 12). *Andrei Zevakin: Tallinna börs on liiga aeglane*. Äripäev. Retrieved May 20, 2025, from https://www.aripaev.ee/borsiuudised/2025/01/12/andrei-zevakin-tallinna-bors-on-liiga-aeglane
54. Lin, Y., He, Y., Evgeniou, T., & Jin, W. (2022). Multilingual social media sentiment analysis and stock market prediction using machine translation and

deep learning. Decision Support Systems, 156, 113738. https://doi.org/10.1016/j.dss.2022.113738

55. Liu, J.-X., Leu, J.-S., & Holst, S. (2023). Stock price movement prediction based on Stocktwits investor sentiment using FinBERT and ensemble SVM. PeerJ Computer Science, 9, e1403. https://doi.org/10.7717/peerj-cs.1403
56. Liu, Z., & Yu, W. (2020). Media sentiment and stock returns: Evidence from the Chinese stock market. Pacific-Basin Finance Journal, 61, 101332. https://doi.org/10.1016/j.pacfin.2020.101332
57. Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x
58. Madaan, G., & Singh, S. (2019). An analysis of behavioral biases in investment decision-making. International Journal of Financial Research, 10(4), 55–67. https://doi.org/10.5430/ijfr.v10n4p55
59. Madaan, M., & Singh, S. (2019). Behavioural finance: A review of theory and literature. International Journal of Economics and Financial Issues, 9(1), 80–86. https://doi.org/10.32479/ijefi.7382
60. Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. Journal of Economic Perspectives, 17(1), 59-82. https://doi.org/10.1257/089533003321164958
61. Malkiel, B. G. (2005). Reflections on the efficient market hypothesis: 30 years later. Financial Review, 40(1), 1–9. https://doi.org/10.1111/j.0732-8516.2005.00090.x
62. Mehta, P., & Pandya, D. (2020). A Review On Sentiment Analysis Methodologies, Practices, And Applications. International Journal of Scientific & Technology Research, 9, 601-609.
63. Mullainathan, S., & Shleifer, A. (2005). The market for news. American Economic Review, 95(4), 1031–1053. https://doi.org/10.1257/0002828054825619
64. Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. Expert Systems with Applications, 41(16), 7653–7670. https://doi.org/10.1016/j.eswa.2014.06.009

65. Niederhoffer, V. (1971). The analysis of world events and stock prices. The Journal of Business, 44(2), 193–219. https://doi.org/10.1086/295358
66. Nofsinger, J. R. (2005). The Psychology of Investing (2nd ed.). Prentice Hall.
67. Nofsinger, J. R. (2017). The Psychology of Investing (6th ed.). New York: Routledge.
68. Oberlechner, T., & Hocking, S. (2004). Information sources, news, and rumors in financial markets: Insights into the foreign exchange market. Journal of Economic Psychology, 25(3), 407–424. https://doi.org/10.1016/S0167-4870(03)00011-5
69. Peterson, R. L. (2002). Investing with the trend: A rules-based approach to money management. John Wiley & Sons.
70. Pinnuck, M. (2014). A review of the impact of media on firm value. Journal of Accounting Literature, 33, 36–67. https://doi.org/10.1016/j.acclit.2014.02.002
71. Puzanova, N., & Eratalay, M. H. (2021). The impact of real estate news sentiment on stock returns. Journal of Risk and Financial Management, 14(7), 303. https://doi.org/10.3390/jrfm14070303
72. Ramiah, V., Xu, X., & Moosa, I. A. (2015). Neoclassical finance, behavioral finance, and noise traders: A review and assessment of the literature. International Review of Financial Analysis, 41, 89–100. https://doi.org/10.1016/j.irfa.2015.07.012
73. Read, C. (2013). Understanding insider trading laws: A guide to the legal and ethical issues. New York: Financial Times Press.
74. Richard, C., & Peterson, R. L. (2006). Affect and financial decision-making: How neuroscience can inform market participants. Journal of Behavioral Finance, 7(2), 70–78. https://doi.org/10.1207/s15427579jpfm0702_4
75. Schmidt, H. (2020). Buy on the rumor, sell on the news: Market reactions to media-based signals. Journal of Behavioral and Experimental Finance, 28, 100407. https://doi.org/10.1016/j.jbef.2020.100407
76. Schumaker, R. P., Zhang, Y., Huang, C.-N., & Chen, H. (2012). Evaluating sentiment in financial news articles. Decision Support Systems, 53(3), 458–464. https://doi.org/10.1016/j.dss.2012.03.001
77. Shiller, R. J. (1984). Stock prices and social dynamics. Brookings Papers on Economic Activity, 1984(2), 457–498. https://doi.org/10.2307/2534436

78. Shiller, R. J. (2005). Irrational Exuberance (2nd ed.). Princeton University Press.
79. Shiller, R. J. (2017). Narrative Economics: How Stories Go Viral and Drive Major Economic Events. Princeton University Press.
80. Siering, M. (2012). Investigating the Impact of Media Sentiment and Investor Attention on Financial Markets. Enterprise Applications and Services in the Finance Industry.
81. Simon, H. A. (1955). A behavioral model of rational choice. The Quarterly Journal of Economics, 69(1), 99–118.
82. Singh, A. K., & Sharma, R. (2024). Thematic review and discussion of research on herd behavior in financial markets. SAGE Open, 14(2). https://doi.org/10.1177/21582440241239221
83. Statman, M. (2014). Behavioral finance: Finance with normal people. Borsa Istanbul Review, 14. https://doi.org/10.1016/j.bir.2014.03.001
84. Tan, M., Li, X., & Zhou, Y. (2023). Rumors, information asymmetry, and stock price manipulation. Journal of Financial Markets, 67, 100806. https://doi.org/10.1016/j.finmar.2023.100806
85. Tetlock, P. C. (2010). Does Public Financial News Resolve Asymmetric Information? The Review of Financial Studies, 23(9), 3520–3557. https://doi.org/10.1093/rfs/hhq052
86. Tetlock, P. C. (2011). All the News That is Fit to Reprint: Do Investors React to Stale Information? The Review of Financial Studies, 24(5), 1481–1512. http://www.jstor.org/stable/20869278
87. Tetlock, P. C. (2014). Information transmission in finance. Annual Review of Financial Economics, 6, 365–390. https://doi.org/10.1146/annurev-financial-110613-034217
88. Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. The Journal of Finance, 63(3), 1437–1467. https://doi.org/10.1111/j.1540-6261.2008.01362.x
89. Thaler, R. H. (Ed.). (2005). Advances in Behavioral Finance (Vol. II). Princeton University Press. http://www.jstor.org/stable/j.ctt1j1nwfj
90. Thompson, R. (2013). A comparative analysis of insider trading regulation and its impact on financial market trust. Journal of Financial Regulation and Compliance, 21(4), 357–370. https://doi.org/10.1108/JFRC-01-2013-0005

91. Ţiţan, A. (2015). The Efficient Market Hypothesis: Review of Specialized Literature and Empirical Research. Procedia Economics and Finance, 32, 442–449. https://doi.org/10.1016/S2212-5671(15)01416-1
92. Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9
93. Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291. https://doi.org/10.2307/1914185
94. Uhl, M. W. (2014). Reuters Sentiment and Stock Returns. Journal of Behavioral Finance, 15(4), 287–298. https://doi.org/10.1080/15427560.2014.967852
95. Vandolder, D. (2023). Behavioral risk profiling: Measuring loss aversion of individual investors. Journal of Behavioral and Experimental Finance, 40, 100867. https://doi.org/10.1016/j.jbef.2023.100867
96. Wan, X., Wang, Q., Zhang, X., & Li, J. (2020). Predicting stock price movement using social media and news sentiment. Multimedia Tools and Applications, 79, 21603–21624. https://doi.org/10.1007/s11042-020-08764-9
97. Wang, Y., & Liu, J. (2023). A new perspective on how investor sentiment affects herding in cryptocurrency markets. Finance Research Letters, 49, 102800. https://doi.org/10.1016/j.frl.2022.102800
98. Yang, H., Kuang, K., & Zhang, J. E. (2017). Media coverage and investor sentiment: Evidence from the stock market in China. Journal of Behavioral Finance, 18(4), 390–403. https://doi.org/10.1080/15427560.2017.1344677
99. Yang, J., Song, D., & Wu, W. (2022). News sentiment and stock returns: Evidence from machine learning and natural language processing. Finance Research Letters, 46, 102281. https://doi.org/10.1016/j.frl.2021.102281
100. Yoon, S., & Takahashi, H. (2025). Investor sentiment and stock return dynamics of small-cap firms. Journal of Behavioral Finance, Forthcoming.
101. Zhang, X., & Wang, X. (2020). Investor attention, media sentiment and stock return: Evidence from the Chinese market. Physica A: Statistical Mechanics and its Applications, 542, 123496. https://doi.org/10.1016/j.physa.2019.123496

102. Zhao, X., Zuo, L., & Zhang, W. (2010). The effect of rumor propagation on stock prices: Evidence from the Chinese stock market. China Economic Review, 21(4), 645–657. https://doi.org/10.1016/j.chieco.2010.05.007

## APPENDIX A - Estimation results presented in detail by equations

**Arco Vara ARMAX (1,0,4)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ar1* | 0.63*** | 0.61** | 0.54* | 0.72*** | 0.70*** | 0.54* |
| *ma1* | -0.69*** | -0.67*** | -0.61** | -0.80*** | -0.78*** | -0.61** |
| *ma2* | 0.01 | 0.01 | 0.00 | 0.02 | 0.02 | 0.00 |
| *ma3* | -0.01 | -0.01 | -0.01 | -0.01 | -0.02 | -0.01 |
| *ma4* | 0.06** | 0.06** | 0.06** | 0.06*** | 0.07*** | 0.06** |
| *intercept* | 0.01 | 0.01 | 0.00 | -0.00 | -0.01 | 0.00 |
| *pos_outlier* | 3.96*** | 3.94*** | 3.91*** | 3.85*** | 3.82*** | 3.91*** |
| *neg_outlier* | -3.98*** | -3.98*** | -4.00*** | -3.80*** | -3.82*** | -4.00*** |
| *x0* | NA | 0.42* | NA | 0.03 | NA | 0.58** |
| *xlag1* | NA | 0.04 | NA | 0.20*** | NA | 0.37 |
| *xlag2* | NA | 0.00 | NA | 0.08*** | NA | -0.15 |
| *index t* | NA | NA | NA | 0.35 | NA | NA |
| *index t-1* | NA | NA | NA | 0.06 | NA | NA |
| *pos0* | NA | NA | 0.58** | NA | 0.52* | NA |
| *pos1* | NA | NA | 0.37 | NA | 0.36 | NA |
| *pos2* | NA | NA | -0.15 | NA | -0.09 | NA |
| *neg0* | NA | NA | 0.12 | NA | 0.03 | -0.47 |
| *neg1* | NA | NA | -0.84* | NA | -0.76* | -1.21** |
| *neg2* | NA | NA | 0.48 | NA | 0.42 | 0.63 |
| **Log-Likelihood** | -2562.55 | -2560.80 | -2557.18 | -2533.87 | -2530.59 | -2557.18 |
| **AIC** | 5143.10 | 5145.60 | 5144.35 | 5095.75 | 5095.17 | 5144.35 |
| **BIC** | 5192.68 | 5211.70 | 5226.98 | 5172.87 | 5188.82 | 5226.98 |
| **SSE** | 1773.50 | 1770.10 | 1763.08 | 1718.60 | 1712.42 | 1763.08 |

**Coop Pank ARMAX (3,0,2)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ar1* | -1.00*** | -0.59*** | -1.55*** | -1.16*** | -1.53*** | -1.55*** |
| *ar2* | -1.04*** | -1.01*** | -1.02*** | -0.61 | -0.98*** | -1.02*** |
| *ar3* | -0.05** | -0.05** | -0.03 | 0.03 | -0.01 | -0.03 |
| *ma1* | 0.95*** | 0.54*** | 1.50*** | 1.14*** | 1.51*** | 1.50*** |
| *ma2* | 1.00*** | 1.00*** | 0.95*** | 0.61* | 0.95*** | 0.95*** |
| *intercept* | -0.02 | -0.03 | -0.03 | -0.04** | -0.04* | -0.03 |
| *pos_outlier* | 3.54*** | 3.50*** | 3.49*** | 2.95*** | 2.95*** | 3.49*** |
| *neg_outlier* | -4.08*** | -4.10*** | -4.14*** | -2.96*** | -2.96*** | -4.14*** |
| *x0* | NA | 0.30*** | NA | 0.25** | NA | 0.33*** |
| *xlag1* | NA | -0.02 | NA | 0.06 | NA | -0.05 |
| *xlag2* | NA | -0.02 | NA | -0.04 | NA | -0.03 |
| *index t* | NA | NA | NA | 0.54*** | 0.54*** | NA |
| *index t-1* | NA | NA | NA | -0.01 | -0.00 | NA |
| *pos0* | NA | NA | 0.33*** | NA | 0.26** | NA |
| *pos1* | NA | NA | -0.05 | NA | 0.03 | NA |
| *pos2* | NA | NA | -0.03 | NA | -0.09 | NA |
| *neg0* | NA | NA | -0.31 | NA | 0.11 | -0.64 |
| *neg1* | NA | NA | -0.18 | NA | 0.45 | -0.12 |
| *neg2* | NA | NA | 0.94* | NA | 0.93* | 0.97* |
| **Log-Likelihood** | -2252.49 | -2246.26 | -2244.30 | -2104.78 | -2095.59 | -2244.30 |
| **AIC** | 4522.99 | 4516.52 | 4518.59 | 4237.57 | 4225.18 | 4518.59 |
| **BIC** | 4572.57 | 4582.63 | 4601.23 | 4314.69 | 4318.83 | 4601.23 |
| **SSE** | 1261.36 | 1250.50 | 1250.96 | 1073.61 | 1062.72 | 1250.96 |

**Ekspress Grupp ARMAX (0,0,2)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ma1* | -0.07*** | -0.07*** | -0.07*** | -0.07*** | -0.07*** | -0.07*** |
| *ma2* | -0.05* | -0.04* | -0.04* | -0.05** | -0.05* | -0.04* |
| *intercept* | -0.02 | -0.02 | -0.04* | -0.03* | -0.04** | -0.04* |
| *pos_outlier* | 4.03*** | 4.01*** | 3.98*** | 3.87*** | 3.84*** | 3.98*** |
| *neg_outlier* | -3.62*** | -3.62*** | -3.62*** | -3.41*** | -3.42*** | -3.62*** |
| *x0* | NA | 0.54*** | NA | 0.58*** | NA | 0.96*** |
| *xlag1* | NA | 0.13 | NA | 0.16 | NA | 0.18 |
| *xlag2* | NA | -0.21 | NA | -0.20 | NA | -0.19 |
| *index t* | NA | NA | NA | 0.27*** | 0.27*** | NA |
| *index t-1* | NA | NA | NA | -0.01 | -0.01 | NA |
| *pos0* | NA | NA | 0.96*** | NA | 0.97*** | NA |
| *pos1* | NA | NA | 0.18 | NA | 0.17 | NA |
| *pos2* | NA | NA | -0.19 | NA | -0.19 | NA |
| *neg0* | NA | NA | -0.74* | NA | -0.64 | -1.69*** |
| *neg1* | NA | NA | -0.22 | NA | -0.05 | -0.40 |
| *neg2* | NA | NA | -0.26 | NA | -0.20 | -0.07 |
| **Log-Likelihood** | -2476.27 | -2472.00 | -2465.05 | -2429.44 | -2423.20 | -2465.05 |
| **AIC** | 4964.53 | 4961.99 | 4954.11 | 4880.87 | 4874.40 | 4954.11 |
| **BIC** | 4997.58 | 5011.57 | 5020.21 | 4941.47 | 4951.53 | 5020.21 |
| **SSE** | 1613.42 | 1605.88 | 1593.70 | 1532.66 | 1522.22 | 1593.70 |

**Harju Elekter ARMAX (5,0,0)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ar1* | 0.08*** | 0.08*** | 0.08*** | NA | 0.08*** | 0.08*** |
| *ar2* | 0.02 | 0.03 | 0.03 | NA | 0.03 | 0.03 |
| *ar3* | 0.02 | 0.02 | 0.02 | NA | 0.02 | 0.02 |
| *ar4* | -0.03 | -0.03 | -0.03 | NA | -0.05* | -0.03 |
| *ar5* | -0.11*** | -0.11*** | -0.11*** | NA | -0.13*** | -0.11*** |
| *intercept* | -0.02 | -0.02 | -0.02 | -0.03 | -0.04 | -0.02 |
| *pos_outlier* | 4.06*** | 4.06*** | 4.06*** | 3.78*** | 3.76*** | 4.06*** |
| *neg_outlier* | -3.62*** | -3.62*** | -3.60*** | -3.19*** | -3.09*** | -3.60*** |
| *x0* | NA | -0.02 | NA | 0.09 | NA | -0.11 |
| *xlag1* | NA | -0.17 | NA | -0.15 | NA | -0.11 |
| *xlag2* | NA | 0.23 | NA | 0.21 | NA | 0.30 |
| *index t* | NA | NA | NA | 0.42*** | 0.42*** | NA |
| *index t-1* | NA | NA | NA | -0.01 | 0.00 | NA |
| *pos0* | NA | NA | -0.11 | NA | -0.03 | NA |
| *pos1* | NA | NA | -0.11 | NA | -0.10 | NA |
| *pos2* | NA | NA | 0.30 | NA | 0.24 | NA |
| *neg0* | NA | NA | 0.76 | NA | 0.84 | 0.87 |
| *neg1* | NA | NA | -0.73 | NA | -0.66 | -0.62 |
| *neg2* | NA | NA | -0.49 | NA | -0.25 | -0.79 |
| **Log-Likelihood** | -2541.73 | -2540.99 | -2539.53 | -2468.89 | -2450.64 | -2539.53 |
| **AIC** | 5101.46 | 5105.97 | 5109.05 | 4959.78 | 4935.28 | 5109.05 |
| **BIC** | 5151.04 | 5172.08 | 5191.68 | 5020.37 | 5028.93 | 5191.68 |
| **SSE** | 1733.42 | 1732.01 | 1729.23 | 1600.42 | 1568.64 | 1729.23 |

**TKM Grupp ARMAX (2,0,2)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| ***ar1*** | 0.37*** | -1.03*** | -1.03*** | -1.03*** | -1.58*** | -1.03*** |
| ***ar2*** | -0.97*** | -0.98*** | -0.98*** | -0.96*** | -0.94*** | -0.98*** |
| ***ma1*** | -0.37*** | 1.03*** | 1.03*** | 1.02*** | 1.56*** | 1.03*** |
| ***ma2*** | 0.98*** | 0.99*** | 0.99*** | 0.98*** | 0.91*** | 0.99*** |
| ***intercept*** | 0.03** | 0.02* | 0.01 | 0.01 | 0.00 | 0.02 |
| ***pos_outlier*** | 2.95*** | 2.96*** | 2.90*** | 2.43*** | 2.38*** | 2.90*** |
| ***neg_outlier*** | -3.89*** | -3.92*** | -3.91*** | -2.97*** | -2.98*** | -3.92*** |
| ***x0*** | NA | 0.05 | NA | 0.08 | NA | 0.29** |
| ***xlag1*** | NA | 0.30** | NA | 0.28** | NA | 0.31** |
| ***xlag2*** | NA | -0.05 | NA | -0.02 | NA | -0.06 |
| ***index t*** | NA | NA | NA | 0.34*** | 0.34*** | NA |
| ***index t-1*** | NA | NA | NA | -0.00 | 0.00 | NA |
| ***pos0*** | NA | NA | 0.29** | NA | 0.23** | NA |
| ***pos1*** | NA | NA | 0.32** | NA | 0.31*** | NA |
| ***pos2*** | NA | NA | -0.06 | NA | -0.05 | NA |
| ***neg0*** | NA | NA | -1.45*** | NA | -0.81*** | -1.75*** |
| ***neg1*** | NA | NA | 0.08 | NA | 0.04 | -0.22 |
| ***neg2*** | NA | NA | 0.01 | NA | 0.09 | 0.06 |
| **Log-Likelihood** | -1482.33 | -1476.07 | -1463.93 | -1295.55 | -1290.90 | -1463.93 |
| **AIC** | 2980.67 | 2974.14 | 2955.85 | 2617.10 | 2613.80 | 2955.87 |
| **BIC** | 3024.74 | 3034.74 | 3032.97 | 2688.71 | 2701.94 | 3032.99 |
| **SSE** | 542.47 | 538.51 | 531.35 | 441.99 | 439.80 | 531.35 |

**Enefit Green ARMAX (5,1,0)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| ***ar1*** | -0.88*** | -0.87*** | -0.88*** | -0.83*** | -0.84*** | -0.88*** |
| ***ar2*** | -0.76*** | -0.76*** | -0.76*** | -0.72*** | -0.73*** | -0.76*** |
| ***ar3*** | -0.50*** | -0.50*** | -0.50*** | -0.51*** | -0.52*** | -0.50*** |
| ***ar4*** | -0.37*** | -0.38*** | -0.37*** | -0.37*** | -0.36*** | -0.37*** |
| ***ar5*** | -0.15*** | -0.16*** | -0.15*** | -0.14*** | -0.14*** | -0.15*** |
| ***pos_outlier*** | 3.12*** | 3.07*** | 3.07*** | 2.20*** | 2.20*** | 3.07*** |
| ***neg_outlier*** | -3.03*** | -3.03*** | -3.00*** | -2.03*** | -1.96*** | -3.00*** |
| ***x0*** | NA | 0.30*** | NA | 0.38*** | NA | 0.26*** |
| ***xlag1*** | NA | -0.07 | NA | -0.05 | NA | -0.08 |
| ***xlag2*** | NA | 0.07 | NA | -0.01 | NA | 0.07 |
| ***index t*** | NA | NA | NA | 0.71*** | 0.72*** | NA |
| ***index t-1*** | NA | NA | NA | 0.00 | 0.01 | NA |
| ***pos0*** | NA | NA | 0.26*** | NA | 0.26*** | NA |
| ***pos1*** | NA | NA | -0.08 | NA | -0.09 | NA |
| ***pos2*** | NA | NA | 0.07 | NA | -0.02 | NA |
| ***neg0*** | NA | NA | 0.61*** | NA | 1.17*** | 0.36 |
| ***neg1*** | NA | NA | 0.05 | NA | 0.35* | 0.13 |
| ***neg2*** | NA | NA | 0.06 | NA | 0.14 | -0.01 |
| **Log-Likelihood** | -1181.11 | -1172.35 | -1171.16 | -806.52 | -794.21 | -1171.16 |
| **AIC** | 2378.22 | 2366.70 | 2370.32 | 1639.04 | 1620.41 | 2370.32 |
| **BIC** | 2422.29 | 2427.29 | 2447.43 | 1710.65 | 1708.54 | 2447.43 |
| **SSE** | 389.75 | 386.03 | 385.52 | 258.42 | 254.95 | 385.52 |

**Merko Ehitus ARMAX (0,0,0)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *intercept* | 0.03** | 0.03 | 0.03* | 0.01 | 0.02 | 0.03* |
| *pos_outlier* | 3.24*** | 3.20*** | 3.21*** | 2.73*** | 2.74*** | 3.21*** |
| *neg_outlier* | -3.86*** | -3.85*** | -3.86*** | -3.02*** | -3.02*** | -3.86*** |
| *x0* | NA | 0.36** | NA | 0.31** | NA | 0.39*** |
| *xlag1* | NA | 0.12 | NA | 0.20 | NA | 0.03 |
| *xlag2* | NA | -0.09 | NA | -0.10 | NA | -0.17 |
| *index t* | NA | NA | NA | 0.42*** | 0.42*** | NA |
| *index t-1* | NA | NA | NA | -0.03 | -0.03* | NA |
| *pos0* | NA | NA | 0.39*** | NA | 0.33** | NA |
| *pos1* | NA | NA | 0.03 | NA | 0.11 | NA |
| *pos2* | NA | NA | -0.17 | NA | -0.15 | NA |
| *neg0* | NA | NA | -0.13 | NA | 0.02 | -0.53 |
| *neg1* | NA | NA | 1.15** | NA | 1.44*** | 1.11* |
| *neg2* | NA | NA | 1.00* | NA | 0.68 | 1.17** |
| **Log-Likelihood** | -1890.19 | -1886.43 | -1881.69 | -1720.70 | -1715.63 | -1881.69 |
| **AIC** | 3788.38 | 3786.86 | 3783.38 | 3459.41 | 3455.25 | 3783.38 |
| **BIC** | 3810.42 | 3825.42 | 3838.47 | 3508.98 | 3521.36 | 3838.47 |
| **SSE** | 848.51 | 845.02 | 840.64 | 704.61 | 700.70 | 840.64 |

**Nordecon ARMAX (0,0,0)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *intercept* | -0.03 | -0.03 | -0.03 | -0.05** | -0.04** | -0.03 |
| *pos_outlier* | 3.46*** | 3.40*** | 3.41*** | 3.15*** | 3.16*** | 3.41*** |
| *neg_outlier* | -3.43*** | -3.42*** | -3.41*** | -3.09*** | -3.07*** | -3.41*** |
| *x0* | NA | 0.60*** | NA | 0.65*** | NA | 0.56*** |
| *xlag1* | NA | 0.32* | NA | 0.36** | NA | 0.21 |
| *xlag2* | NA | -0.27 | NA | -0.25 | NA | -0.29 |
| *index t* | NA | NA | NA | 0.34*** | 0.35*** | NA |
| *index t-1* | NA | NA | NA | 0.03 | 0.03 | NA |
| *pos0* | NA | NA | 0.56*** | NA | 0.55*** | NA |
| *pos1* | NA | NA | 0.21 | NA | 0.23 | NA |
| *pos2* | NA | NA | -0.29 | NA | -0.29 | NA |
| *neg0* | NA | NA | 0.72* | NA | 1.04** | 0.16 |
| *neg1* | NA | NA | 0.78* | NA | 0.82** | 0.58 |
| *neg2* | NA | NA | -0.23 | NA | -0.14 | 0.06 |
| **Log-Likelihood** | -2295.06 | -2287.11 | -2286.25 | -2204.61 | -2202.95 | -2286.25 |
| **AIC** | 4598.11 | 4588.21 | 4592.49 | 4427.21 | 4429.91 | 4592.49 |
| **BIC** | 4620.15 | 4626.77 | 4647.58 | 4476.79 | 4496.01 | 4647.58 |
| **SSE** | 1322.68 | 1311.21 | 1309.97 | 1197.80 | 1195.63 | 1309.97 |

**Tallink Grupp ARMAX (5,1,0)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ar1* | -0.78*** | -0.78*** | -0.78*** | -0.81*** | -0.81*** | -0.78*** |
| *ar2* | -0.65*** | -0.65*** | -0.65*** | -0.66*** | -0.66*** | -0.65*** |
| *ar3* | -0.53*** | -0.53*** | -0.53*** | -0.51*** | -0.51*** | -0.53*** |
| *ar4* | -0.28*** | -0.28*** | -0.28*** | -0.32*** | -0.32*** | -0.28*** |
| *ar5* | -0.16*** | -0.16*** | -0.16*** | -0.18*** | -0.18*** | -0.16*** |
| *pos_outlier* | 3.17*** | 3.13*** | 3.15*** | 2.51*** | 2.51*** | 3.15*** |
| *neg_outlier* | -3.48*** | -3.46*** | -3.45*** | -2.42*** | -2.42*** | -3.45*** |
| *x0* | NA | 0.39*** | NA | 0.26*** | NA | 0.28** |
| *xlag1* | NA | -0.01 | NA | 0.07 | NA | -0.07 |
| *xlag2* | NA | -0.07 | NA | -0.01 | NA | -0.08 |
| *index t* | NA | NA | NA | 0.61*** | 0.61*** | NA |
| *index t-1* | NA | NA | NA | -0.02 | -0.02 | NA |
| *pos0* | NA | NA | 0.28** | NA | 0.25** | NA |
| *pos1* | NA | NA | -0.07 | NA | 0.03 | NA |
| *pos2* | NA | NA | -0.09 | NA | -0.02 | NA |
| *neg0* | NA | NA | 0.62*** | NA | 0.29* | 0.34 |
| *neg1* | NA | NA | 0.12 | NA | 0.16 | 0.20 |
| *neg2* | NA | NA | -0.02 | NA | 0.01 | 0.06 |
| **Log-Likelihood** | -2334.90 | -2328.55 | -2327.28 | -2075.14 | -2074.93 | -2327.28 |
| **AIC** | 4685.80 | 4679.09 | 4682.55 | 4176.29 | 4181.85 | 4682.55 |
| **BIC** | 4729.86 | 4739.68 | 4759.67 | 4247.89 | 4269.98 | 4759.67 |
| **SSE** | 1382.17 | 1372.57 | 1370.66 | 1039.43 | 1039.18 | 1370.66 |

**Tallinna Sadam ARMAX (2,0,2)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ar1* | -0.05 | 0.00 | -0.14* | 1.22*** | 1.21*** | -0.14* |
| *ar2* | 0.73*** | 0.69*** | 0.82*** | -0.79*** | -0.77*** | 0.82*** |
| *ma1* | 0.08 | 0.03 | 0.19** | -1.23*** | -1.21*** | 0.19** |
| *ma2* | -0.67*** | -0.63*** | -0.76*** | 0.83*** | 0.81*** | -0.76*** |
| *intercept* | -0.01 | -0.01 | -0.01 | -0.02** | -0.03** | -0.01 |
| *pos_outlier* | 3.23*** | 3.19*** | 3.19*** | 2.65*** | 2.66*** | 3.19*** |
| *neg_outlier* | -3.75*** | -3.73*** | -3.67*** | -3.12*** | -3.11*** | -3.67*** |
| *x0* | NA | 0.33*** | NA | 0.23** | NA | 0.12 |
| *xlag1* | NA | 0.00 | NA | -0.01 | NA | 0.11 |
| *xlag2* | NA | -0.00 | NA | 0.00 | NA | 0.04 |
| *index t* | NA | NA | NA | 0.32*** | 0.31*** | NA |
| *index t-1* | NA | NA | NA | 0.04*** | 0.04*** | NA |
| *pos0* | NA | NA | 0.12 | NA | *0.14* | NA |
| *pos1* | NA | NA | 0.11 | NA | *0.07* | NA |
| *pos2* | NA | NA | 0.04 | NA | *0.04* | NA |
| *neg0* | NA | NA | 1.06*** | NA | 0.52*** | 0.94*** |
| *neg1* | NA | NA | -0.29 | NA | -0.23 | -0.40 |
| *neg2* | NA | NA | -0.18 | NA | -0.14 | -0.23 |
| **Log-Likelihood** | -1339.76 | -1334.80 | -1326.12 | -1125.98 | -1123.40 | -1326.12 |
| **AIC** | 2695.51 | 2691.61 | 2680.25 | 2277.96 | 2278.80 | 2680.25 |
| **BIC** | 2739.58 | 2752.20 | 2757.37 | 2349.57 | 2366.94 | 2757.37 |
| **SSE** | 464.02 | 461.51 | 457.13 | 367.06 | 366.02 | 457.13 |

**Tallinna Vesi ARMAX (0,0,2)**

| | Base | Sentiment | Pos/Neg | Sentiment + Index | Pos/Neg + Index | Mixed |
|---|---|---|---|---|---|---|
| *ma1* | -0.01 | -0.01 | -0.01 | -0.02 | -0.02 | -0.01 |
| *ma2* | 0.04 | 0.04 | 0.04 | 0.05* | 0.04* | 0.04 |
| *intercept* | 0.02 | 0.01 | 0.01 | 0.00 | 0.01 | 0.01 |
| *pos_outlier* | 3.18*** | 3.18*** | 3.17*** | 2.93*** | 2.93*** | 3.17*** |
| *neg_outlier* | -3.35*** | -3.35*** | -3.35*** | -2.94*** | -2.94*** | -3.35*** |
| *x0* | NA | 0.04 | NA | 0.03 | NA | 0.04 |
| *xlag1* | NA | 0.21* | NA | 0.20* | NA | 0.16 |
| *xlag2* | NA | 0.07 | NA | 0.08 | NA | 0.06 |
| *index t* | NA | NA | NA | 0.22*** | 0.22*** | NA |
| *index t-1* | NA | NA | NA | 0.01 | 0.01 | NA |
| *pos0* | NA | NA | 0.04 | NA | 0.04 | NA |
| *pos1* | NA | NA | 0.16 | NA | 0.16 | NA |
| *pos2* | NA | NA | 0.06 | NA | 0.06 | NA |
| *neg0* | NA | NA | 0.03 | NA | 0.02 | -0.01 |
| *neg1* | NA | NA | 0.41 | NA | 0.38 | 0.25 |
| *neg2* | NA | NA | 0.13 | NA | 0.15 | 0.08 |
| **Log-Likelihood** | -1417.84 | -1415.98 | -1415.60 | -1325.37 | -1324.99 | -1415.60 |
| **AIC** | 2847.68 | 2849.97 | 2855.20 | 2672.75 | 2677.99 | 2855.20 |
| **BIC** | 2880.73 | 2899.55 | 2921.31 | 2733.34 | 2755.11 | 2921.31 |
| **SSE** | 505.50 | 504.48 | 504.27 | 456.76 | 456.57 | 504.27 |

## Kokkuvõte

Aktsiaturul toimuva ennustamine on kõikide kasumit maksimeerivate investorite unistus. Läbi aegade on aktsiaturu hinnaliikumise põhjustena nähtud erinevaid tegureid. Eelmise sajandi tähtsaim ja peamine lähtepunkt oli, et turud toimivad efektiivselt, mille kohaselt käitub investor ratsionaalselt ja turgudel toimuv on selgitatav fundamentaalanalüüsi baasil. Hilisemal perioodil on kujunenud selge arusaam, et lisaks neile teguritele mõjutavad investori käitumist veel psühholoogilised aspektid, emotsioonid, karjakäitumine, heuristikad, kognitiivsus, aga ka meediakajastus. Meedia rolli üle aktsiahindade kujunemises on palju spekuleeritud. Üha enam on leitud, et meedia mängib hinnaliikumises olulist rolli. Nii meedia mõju kui uurimuste hulk sel teemal on ajas kasvanud. Meedia mõju kasvule on kaasa aidanud informatsiooni suurem maht ning kiirem ja laiem levik. Hoopis uued uurimisvõimalused on seejuures avanud sotsiaalmeedia.

Käesoleva töö eesmärk oli uurida uudiste meeleolu ja aktsiahindade liikumise vahelist seost Tallinna börsil ja leida kinnitust püstitatud hüpoteesidele, et uudise sentiment mõjutab aktsiahindade liikumist ning negatiivse meelsuse mõju on tugevam kui positiivse meelsuse mõju. Töös analüüsiti 12 Tallinna börsi ettevõtte uudiseid aastatel 2019–2023. Valim koosnes 1930 uudisest, mis kõik pärinesid ajalehest Äripäev ja olid algselt eesti keeles. Analüüsiks kasutati uudiste pealkirju ja juhtlõiku, need tõlgiti paremate analüüspakettide kasutamiseks inglise keelde, mis on mitte-ingliskeelsete uudiste puhul levinud praktika. Uudise meeleolu ehk sentimendi määramiseks kasutati Natural language toolkit (NLTK) sõnastiku paketti, mis aitas määrata uudise meelsuse andes sellele skoori vahemikus 1 kuni -1.

Analüüsi tulemused näitasid, et uudiste meelsuse ja aktsiate tootluse vahel on seos. Meelsuse lisamine baasmudelile ja andmete jagamine positiivseteks ja negatiivseteks meelsusteks viis tulemuseni, mis näitas, et meeleolu mõju avaldub samal päeval, olles lühiajaline ning negatiivsed uudised mõjutavad turusündmusi rohkem kui positiivsed uudised. Mudelite sobivust hinnati Akaike information criterion (AIC) väärtuste põhjal. Mudelite AIC väärtuste paranemised ei olnud meelsuse ja positiivseteks/negatiivseteks jagunemise järel suured, ent siiski statistiliselt olulised ja parandasid mudelit. Esile kerkinud erindid (outliers) ajendasid proovima, kas üldise turuindeksi lisamine mudelisse aitab turul toimuva selgitamist veelgi parandada, mis leidis ka kinnitust. Nädalakaupa koondatud andmete analüüs andis nõrgemaid ja vähemolulisemaid tulemusi. Analüüsitud ettevõtete lõikes ei kujunenud välja selget andmete mustrit, kohati tuvastas mudel mõju pika viivituse järel.

Päevastest tulemustest saab järeldada, et kuigi uudiste mõju Tallinna börsi aktsiatele ei ole märkimisväärselt tugev, on ilmnenud seosed siiski statistiliselt olulised. Uudiste meelsuse hindadesse lülitumine viitab meeleolu lühiajalisele mõjule. Negatiivsete uudiste tugevama mõju esilekerkimine on põhjendatav sellega, et inimeste loomusele on omane tajuda kaotusi valusamalt kui samaväärseid võite. Ühtlasi panevad negatiivsed uudised investorit rohkem tegutsema, kuivõrd positiivseid uudiseid võetakse kui kinnitust juba teadaolevale. Töö tulemused näitasid, et ka väikesel ja aeglaseks peetaval turul on võimalik leida seoseid uudiste meelsuse ja aktsiahindade liikumise vahel. Tööle pakuks edasiarendust volatiilsuse kaasamine aktsiahindade kujunemise mudelisse ning suurema valimi koostamine.

**Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks**

Mina, **Kristin Lass**,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose Stock price reactions to news: the case of the Tallinn Stock Exchange,
mille juhendajad on **Kristjan Pulk** ja **Mustafa Hakan Eratalay**
reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

*Kristin Lass*
*21.05.2025*