Don't Stop Me Now! Identification and Prediction of Unnecessary Volatility Interruptions

Abstract

Exchange operators use circuit breakers like volatility interruptions to prevent transitory or error-induced price shocks. However, these mechanisms can impede market efficiency if triggered by legitimate price changes due to new information. We introduce a clustering approach to identify unnecessary volatility interruptions that are triggered within persistent price trends, thereby delaying price discovery. Our findings indicate that such interruptions are more likely to occur when liquidity and order book activity are high and relevant news is present. To improve market design, we propose a deep learning model that can predict unnecessary interruptions based on pre-interruption market and news data.

Keywords: Circuit Breakers, Volatility Interruptions, Market Quality, Price Discovery, Market Design, Machine Learning

JEL: G14, G18

1. Introduction

Securities markets often face sudden, unsubstantiated price changes due to liquidity shocks, erroneous orders, or uncertainty about the impact of unanticipated company-specific or macroeconomic news. The emergence of algorithmic and high-frequency trading with fully automated order submissions and cancellations at high speed (O'Hara, 2015) has increased the occurrence and risk of short-term liquidity imbalances and large transitory price changes (U.S. Securities and Exchange Commission, 2016). These heightened volatility risks posed by modern securities markets require market operators and regulators to consider measures to mitigate them. Therefore, exchanges around the globe use circuit breakers, i.e., market safeguards such as trading halts or volatility interruptions, which pause or slow down trading in the event of significant price changes. These safeguards aim to ensure price continuity and the proper functioning of fully electronic securities markets by preventing transitory or error-induced price shocks.

Circuit breakers were first introduced on major stock exchanges in the late 1980s (e.g., Lee et al., 1994; Lauterbach and Ben-Zion, 1993) and have been increasingly adopted in securities markets worldwide. According to a 2016 survey among exchange operators (Gomber et al., 2016), 86% of the responding trading venues employ circuit breakers to ensure investor protection and market stability, up from 60% reported in a similar survey conducted by the World Federation of Exchanges in 2008 (World Federation of Exchanges, 2008). Moreover, regulators in most jurisdictions mandate the use of circuit breakers.[1] However, their effectiveness remains a subject of ongoing debate. Critics argue that circuit breakers unnecessarily disrupt the regular trading process, potentially hindering market efficiency by delaying price discovery (Fama, 1988) and causing volatility spillovers across time and markets (Subrahmanyam, 1994).

The main issue with the design of current circuit breakers is their reliance on simplistic rule-based mechanisms, which trigger interruptions as soon as the price of a security or index crosses a pre-defined threshold. Thus, these mechanisms do not differentiate between legitimate market movements driven by fundamental information and erratic price jumps caused by liquidity shocks, erroneous orders, misconfigured trading algorithms, or fake news.[2] As a result, existing circuit breakers are characterized by a trade-off between their protective role in maintaining market stability and their potential adverse effects on market quality (Hautsch and Horvath, 2019).

The issue of simplistic circuit breaker rules leading to unnecessary trading interruptions becomes even more critical when considering the frequency of these events. Rather than being rare events, circuit breakers are frequently triggered, even in the

[1] For example, the U.S. Securities and Exchange Commission mandates a trading pause for individual stocks through the Limit Up/Limit Down mechanism in the event of sudden price swings, and a market-wide circuit breaker in the form of a trading halt in case of a severe decline in the S&P 500 Index (U.S. Securities and Exchange Commission, 2012). Similarly, the European financial market regulation MiFID II mandates trading venues to have mechanisms in place to halt or restrict trading in case of significant price movements (European Commission, 2014).

[2] See Bongaerts et al. (2024) for a related discussion on the issues arising from the simplistic and backward-looking nature of current circuit breaker designs.

most liquid stocks.[3] For instance, the European Securities and Markets Authority (ESMA) reported a peak of more than 3,000 circuit breaker events within a single week in March 2020 due to market turmoil related to COVID-19 (European Securities and Markets Authority, 2022). In our sample of German blue chip stocks, we observe an average of 3.15 interruptions per day and nearly 100 interruptions per stock over the observation period of almost five years. Consequently, identifying and avoiding unnecessary interruptions could significantly enhance the effectiveness of circuit breakers and improve overall market quality.

To address this issue and to contribute to solving the trade-off between market stability and the negative impact of unnecessary interruptions, this paper introduces a novel cluster-based approach utilizing volatility interruptions, the prevalent circuit breaker mechanism in Europe. This method allows for the identification of those volatility interruptions that are triggered during persistent price trends, where they merely delay the incorporation of new information into prices. To the best of our knowledge, this is the first framework designed to detect such unnecessary volatility interruptions. Our study is based on a dataset of 3,899 historic volatility interruptions in the 40 stocks of the benchmark index DAX40 from April 2019 to December 2023. The dataset includes high-frequency order book, trade, and news data. By combining an autoencoder model that consolidates order book information before and after each interruption with a k-means algorithm, we classify 39.8% of the observed volatility interruptions as unnecessary, i.e., interruptions that simply delay the price discovery process.

To examine the drivers of unnecessary interruptions, an analysis utilizing probit regressions reveals that unnecessary volatility interruptions are more likely to occur during periods of high liquidity and volatility, the presence of relevant news, and increased order book activity. Economically, these conditions suggest that the price changes leading to the interruption are likely driven by shifts in value expectations and corresponding trading activity. In such cases, the rule-based nature of current circuit breaker mechanisms unnecessarily interferes with the trading process. Moreover, our results suggest that volatility interruptions are more likely to be unnecessary when the last price before the interruption is near the pre-defined threshold that triggers the circuit breaker. When prices approach these thresholds due to sustained movements throughout the trading day, even minor price fluctuations from normal trading activity can activate the circuit breaker, causing unwarranted market disruptions.

The probit regression model also serves as a benchmark for predicting unnecessary interruptions based on ex-ante order book, trade, and news information. However, the predictive performance of this model is only marginally better than random guessing. This suggests that more granular, non-linear dynamics in price and order book behavior must be incorporated to predict unnecessary interruptions in advance.

Based on these findings, we develop a deep learning model that significantly enhances prediction capabilities by effectively capturing complex spatial and temporal

[3]For an overview of the number of circuit breaker events in selected empirical studies, see Bongaerts et al. (2024), Table 2.

relationships among input features. Unlike probit regression models, it eliminates the need to aggregate input time series data and pre-select relevant variables, as it is capable of processing raw data and autonomously identifying useful features. The model combines convolutional neural networks (CNNs), an inception module, and long short-term memory (LSTM) layers. The CNN component extracts features from individual order book snapshots, while the inception module focuses on time-wise convolutions, summarizing features over time. These features are then processed by the LSTM layer, capturing temporal changes essential for the classification task. Using this advanced model, we are able to predict unnecessary volatility interruptions with a precision between 60% and 100%, depending on the sensitivity to false negatives.

Our findings contribute to two strands of literature. First, our paper adds to the extensive research stream that discusses and analyzes the effectiveness of circuit breakers from both theoretical and empirical perspectives.[4] Circuit breakers can help to "cool down" markets and reduce volatility by providing market participants with time to reassess their trading strategies, inventories, and the impact of news (Ma et al., 1989). However, these safeguards also interfere with trading and market liquidity, making it difficult for liquidity providers to manage their inventories (Lauterbach and Ben-Zion, 1993), delaying price discovery (Lehmann, 2019), and causing volatility spillovers to other markets and subsequent trading periods (Subrahmanyam, 1994). Despite these drawbacks, circuit breakers seem necessary to prevent erroneous price jumps in today's fully electronic securities markets (Subrahmanyam, 2013), where order submissions, executions, and price determination occur autonomously at millisecond frequency (O'Hara, 2015). Empirical studies reach contradictory conclusions regarding the effectiveness of circuit breakers in reducing volatility, although they generally agree on their harmful effects on liquidity and price discovery (e.g., Abad and Pascual, 2010; Hautsch and Horvath, 2019; Kim and Rhee, 1997).[5] Based on an in-depth analysis of trading interruptions at Nasdaq, Hautsch and Horvath (2019) conclude that there is a trade-off between the protective role of trading interruptions and their potentially adverse effects on volatility, liquidity, and price efficiency. Variations in distributions across different observation periods and datasets, along with differences in the design of the safeguards (Clapham et al., 2017), may explain the differing conclusions regarding their effectiveness in reducing volatility.

The study most closely related to ours is the model by Bongaerts et al. (2024), which demonstrates that properly calibrated circuit breakers can prevent market runs by curbing excessive trading. Like us, the authors argue that the current simplistic, price-triggered circuit breaker mechanisms fail to differentiate between legitimate liquidity demand and inefficient excessive trading. To enhance the current market design, they propose a forward-looking circuit breaker that becomes increasingly restrictive as the expected welfare losses from market runs increase.

[4]For an overview of the literature, refer to the surveys by Abad and Pascual (2013) and Sifat and Mohamad (2019).

[5]Exhibit 17.2 in the survey by Abad and Pascual (2013) provides a systematic comparison of empirical studies examining whether circuit breakers reduce volatility, improve price discovery, or interfere with liquidity and the trading process.

We contribute to the literature on circuit breakers by demonstrating that it is possible to identify and avoid unnecessary interruptions. This can mitigate the adverse effects of circuit breakers and potentially resolve the trade-off between their protective role and their negative impact on market quality. To the best of our knowledge, no empirical study has yet attempted to differentiate between the circumstances in which circuit breakers are triggered, although this is crucial in determining whether an interruption is necessary.

Second, we contribute to the emerging body of literature that applies machine learning techniques to market microstructure research. For instance, Easley et al. (2021) utilize machine learning methods to predict future levels of liquidity, volatility, and other critical variables for market participants and researchers. Similarly, Kwan et al. (2021) employ reinforcement learning to investigate the price discovery process, while Sirignano and Cont (2019) leverage deep learning to forecast the direction of price movements based on historical limit order book (LOB) data. Our study extends the application of machine learning in market microstructure by exploring its use in analyzing circuit breakers and its potential to enhance the design and effectiveness of these mechanisms.

Overall, our findings can enhance circuit breaker mechanisms by reducing the number of unnecessary interruptions. This can mitigate potential adverse effects of these mechanisms on market quality (Hautsch and Horvath, 2019; Subrahmanyam, 1994) and ultimately improve market efficiency. Additionally, our findings can inform discussions with both market operators and regulatory authorities regarding the design and rules of circuit breakers.

The remainder of the paper is organized as follows: Section 2 provides an overview of the institutional background, details on the dataset, and key descriptive statistics. Section 3 outlines our methodology for clustering volatility interruptions and presents the corresponding results. In Section 4, we describe the regression-based approach for predicting unnecessary volatility interruptions and examine their economic drivers. Section 5 introduces a more advanced deep learning model for predicting unnecessary interruptions and evaluates its performance. Section 6 covers robustness checks and discusses limitations. Finally, Section 7 concludes.

2. Institutional background & data

2.1. Volatility interruptions on Xetra

Our analyses are based on volatility interruptions, which are the common type of circuit breakers in European securities markets (Gomber et al., 2016). Instead of leading to a complete trading halt, volatility interruptions temporarily switch the trading phase from continuous trading to an unscheduled call auction in individual stocks once they are triggered. This is similar to the Limit Up/Limit Down mechanism in the U.S., suggesting that our results are transferable to other markets. Our sample of volatility interruptions represents data from the German trading venue Xetra. Xetra, operated by Deutsche Börse in Frankfurt, is a fully automated, order-driven trading system where buy and sell orders are matched based on price-time priority within a standard open LOB. The system provides continuous trading for the immediate execution of orders and scheduled auctions at specific times during the trading day. These auctions determine opening and closing prices by pooling liquidity and matching orders at a single clearing price. In contrast to other European markets, Xetra also has a scheduled auction in the middle of the trading day (hereafter referred to as intraday auction). Xetra is the most liquid market and reference market for German equities and ETFs.

Like most other European markets, Xetra employs volatility interruptions. A volatility interruption is a rule-based circuit breaker that is triggered during periods of unusually high volatility, utilizing a dual price corridor mechanism as depicted in Figure 1. Once a potential execution price exceeds either a dynamic range around the last traded price or a static range around the last auction price, continuous trading is paused, and trading switches to a non-scheduled call auction, i.e., a volatility interruption. The width of these price ranges is determined by the exchange operator on a stock-specific basis, taking into account each stock's historical volatility.[6] While the exact price limits are not disclosed by the market operator, they can be estimated fairly accurately by reverse-engineering them based on historical data at least for the static price range. Volatility interruptions last 2 minutes plus a random auction end of up to 30 seconds to avoid market manipulation. After the auction clearing price is determined, trading resumes to continuous trading (Deutsche Boerse Group, 2024).[7]

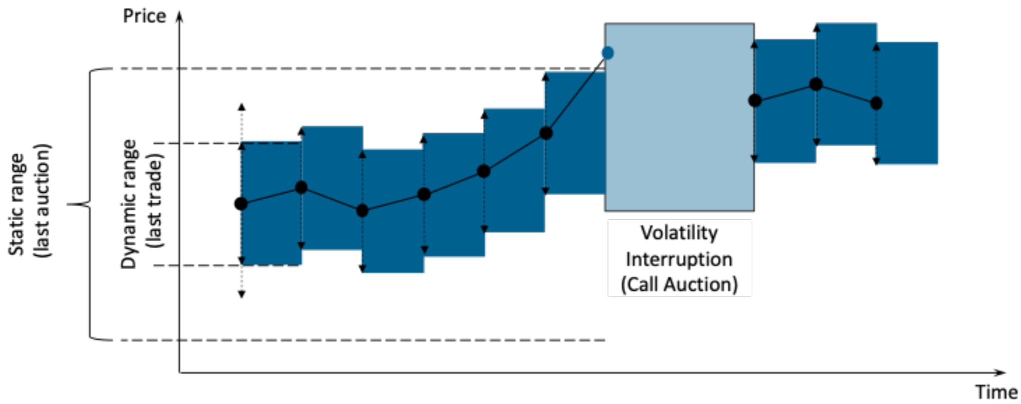

Figure 1: Xetra price ranges for volatility interruptions
The static price range represents a symmetric corridor around the last auction price. The dynamic price range represents a symmetric corridor around the last trade price. If the next potential price is outside one or both of the two corridors, a volatility interruption is triggered, i.e., trading switches from continuous trading to a call auction.

[6]In anticipation of events likely to cause heightened volatility (e.g., the day following the UK's referendum to leave the European Union), the market operator can declare a "fast market" condition, which results in doubling the price ranges that trigger volatility interruptions.

[7]In the rare event when the auction price at the end of a volatility interruption is outside the doubled price ranges, the volatility interruption is extended and concluded manually by the market operator after consulting with the market participants who triggered the interruption.

2.2. Data

For our analysis, we use volatility interruptions triggered in stocks of the DAX40, Germany's leading stock index that includes the largest German companies based on market capitalisation. Our dataset covers the period from 2019[8] to 2023, taking into account the DAX40 constituents as of the end of 2023 (see also Figure 10 in Appendix A for a list of all included DAX constituents). In total, we observe 4,933 triggered interruptions for these stocks in this period. We collect 10 minutes of order book, trade, and message data before and after each interruption from the Deutsche Börse A7 market data platform, creating a 20-minute observation period per event. To ensure a complete 20-minute observation window for each event, we exclude 851 interruptions whose 10 minutes pre- or post-interruption period overlaps with scheduled auctions (i.e., opening, intraday, or closing auction) and 183 that overlap with other volatility interruptions in the same stock.[9] This cleaning step results in a final dataset which comprises 3,899 volatility interruptions. Details on the quantity of removed observations for each cleaning step are depicted in Table 1.

	Number of volatility interruptions
Total number of volatility interruptions	4,933
Overlapping with scheduled auctions	851
Overlapping with other volatility interruptions	183
Final dataset	3,899
Percentage of the sample	79%

Table 1: Cleaning process and dataset summary

In our empirical study, we utilize Xetra market data with nanosecond granularity, including limit prices and quantities for the top ten levels of the order book, trade executions with corresponding prices and quantities, and order messages detailing submissions, cancellations, and modifications. Furthermore, we incorporate news data from RavenPack, which provides the number of news related to a stock around a volatility interruption together with the relevance and sentiment scores associated with each news item. This dataset from two different sources gives us a comprehensive view of both market microstructure and external informational influences.

We resample the dataset into time series with a one-second frequency to synchronize array length and time steps of the model inputs. Before deploying machine learning techniques - specifically clustering and deep learning for prediction - we apply min-max normalization on all features, ensuring each sample is standardized on a per-sample basis.

[8] Our observation periods starts with April 2019 as earlier data is not available on the Deutsche Börse A7 market data platform.

[9] With these two filtering steps we also exclude the 159 instances of extended volatility interruptions that occurred during our observation period.

We use the entire dataset of volatility interruptions for the clustering process. For the prediction task, we split the dataset into a training period from April 2019 to November 2022, a validation period in December 2022, and a test period from January to December 2023[10]. For individual volatility interruptions, we further distinguish between an ex-ante "pre-interruption" phase and an ex-post "post-interruption" phase. The selected observation windows of 10 minutes for each phase are short enough to reflect the market's high-frequency nature, yet sufficiently long to capture gradual, longer-term movements that drive the price to the static triggering threshold.

2.3. Descriptive statistics

Descriptive statistics on the occurrence of volatility interruptions in our dataset are provided in Table 2. During the observation period, DAX40 stocks on Xetra experienced an average of 3.15 volatility interruptions per trading day. However, there are significant outliers, mainly due to market-wide events that triggered numerous interruptions. For instance, the highest number of volatility interruptions on a single trading day was 152. Thus, volatility interruptions are not rare events, even among the most liquid German stocks, but occur frequently, underscoring the importance of identifying and minimizing unnecessary interruptions of trading and price discovery. With respect to variations across stocks, Zalando SE experienced the most volatility interruptions (268), while Daimler Truck Holding AG had only 14 interruptions between 2019 and 2023. A detailed breakdown of the number of interruptions per stock during our observation period is provided in Figure 10 in Appendix A.

	mean	median	min	max	std
Volatility interruptions per day	3.15	1.00	0.00	152.00	7.90
Volatility interruptions per stock	97.48	78.00	14.00	268.00	64.90
Duration [sec]	135.05	135.08	120.00	150.00	8.66

Table 2: Descriptive statistics on the observed volatility interruptions in our sample

On average, volatility interruptions on Xetra lasted 135.05 seconds, closely aligning with the expected duration of 135 seconds, which includes a 120-second auction phase followed by a random end of up to 30 seconds.

The histogram of volatility interruptions, as illustrated in Figure 2, reveals that these interruptions are occurring regularly on Xetra. It shows significant concentration of volatility interruptions during two global events between 2019 and 2023: the COVID-19 pandemic outbreak around March 2020 and the Russian invasion of Ukraine in late February 2022. Both events led to major disruptions and economic turmoil across various sectors, resulting in a high number of volatility interruptions during these periods. The frequency of volatility interruptions thereby correlates with overall market volatility, as indicated by the volatility index VIX.

[10] We split the data into training, validation and test set chronologically rather than randomly, as this approach reflects how a market operator, as a potential adopter of the mechanism, would apply the prediction model in practice.

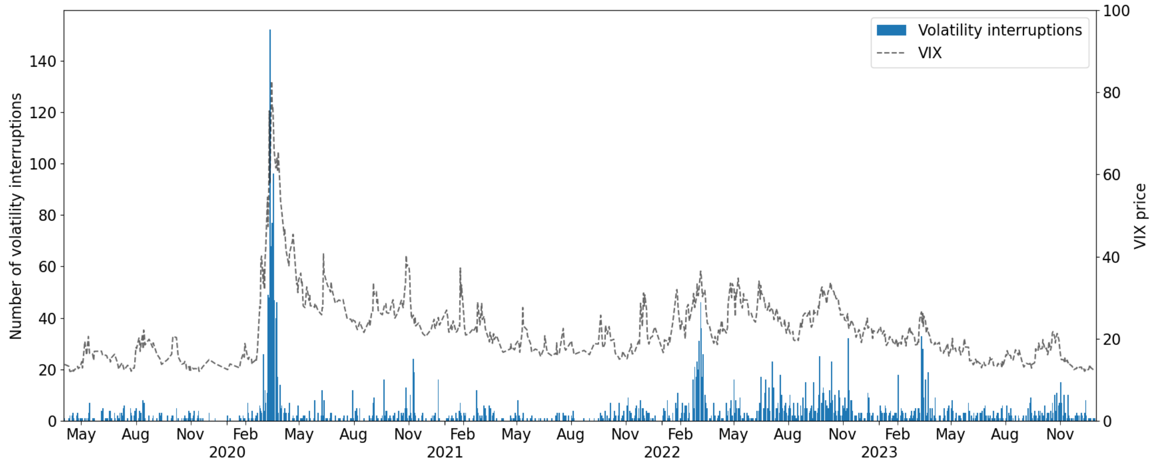

Figure 2: Histogram of volatility interruptions during the observation period and development of the VIX
The histogram in blue shows the total number of volatility interruptions in DAX40 constituents on Xetra for each trading day of the considered time period from 2019 to 2023. The dashed line provides the price development of the volatility index VIX over the same period.

Figure 3 illustrates the evolution of market quality and trading activity around volatility interruptions in our dataset. The figure shows that liquidity, as measured by the average relative spread, remains relatively stable before the start of a volatility interruption. However, once continuous trading resumes following the auction phase of the interruption, the spread increases by an average of 52% compared to its level immediately prior to the interruption. Similarly, liquidity in terms of depth, measured by the euro volume quoted across the first five levels of the order book (level 5 depth), remains relatively stable, with a slight dip just before the interruption, but then declines by an average of 56% after the interruption. During the 10-minute observation window following the interruption, liquidity gradually improves in both breadth and depth, though it does not fully recover to pre-interruption levels even after 10 minutes. Consequently, volatility interruptions not only delay price discovery but are also associated with reduced liquidity, leading to higher trading costs for market participants. Also from this perspective, minimizing unnecessary interruptions would be beneficial for overall market quality.

As expected, volatility increases just before a volatility interruption is triggered. After the interruption, it spikes sharply in the first few seconds before gradually settling back to pre-interruption levels over the course of approximately 30 seconds. These descriptive statistics align with the findings of the majority of empirical studies (e.g., Hautsch and Horvath, 2019), which also report increased volatility and reduced liquidity following trading interruptions triggered by circuit breakers.

Trading activity, measured by the number of trades and trading volume, experiences sharp spikes both immediately before a volatility interruption is triggered and immediately after the unscheduled auction phase concludes. Apart from these two spikes, trading is rather stable around volatility interruptions, with a slight increase

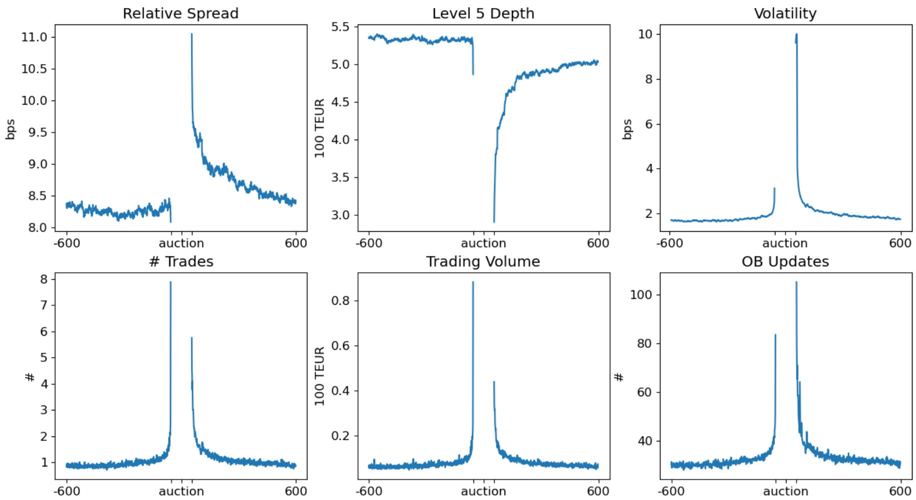

Figure 3: Average liquidity, trading activity, order book activity, and volatility around volatility interruptions
This figure plots the average relative spread, depth on the top five order book levels, midpoint return volatility, number of trades, trading volume, and number of order book updates in each of the 600 seconds of the continuous trading phase before and after a volatility interruption (i.e., excluding the volume executed during the auction phase of the interruption). Since the Xetra order book is not visible during the auction phase triggered by the volatility interruption, these metrics cannot be calculated for that period.

in activity in the seconds following an interruption. A similar pattern is observed in order book activity. The spike in trading activity just before a volatility interruption may be driven by market participants responding to new information or reassessing the stock's expected value. It could also result from traders anticipating the interruption and rushing to execute their trades before it takes effect, consistent with the "magnet effect" hypothesis Subrahmanyam (1994). In contrast, the heightened trading activity following the interruption likely stems from participants seeking to execute their intended trades, manage their inventories, and update their orders and quotes—actions they were either unable or unwilling to perform during the auction phase of the volatility interruption.

3. Clustering and labeling of volatility interruptions

3.1. Methodology

To systematically identify unnecessary volatility interruptions, we analyze historical volatility interruptions by clustering them based on comprehensive LOB data in the 10-minute windows preceding and following each interruption. By using the full set of information contained within the LOB, i.e., limits and volumes of the first ten levels at the bid and ask side of the order book, we ensure that only those volatility interruptions characterized by similar market dynamics both before and after the event are grouped together. This approach avoids reliance on a limited set of handcrafted features from the LOB (e.g., spread or depth measures), offering a more robust categorization of different types of volatility interruptions.

Based on this differentiation of various types of historically observed volatility interruptions, we are able to investigate whether certain clusters exhibit a persistent price trend in both the pre- and post-interruption phase. If such trends are present, it would suggest that these interruptions may have merely delayed the natural price formation process, rather than serving as a corrective function for unreasonable price jumps.

To implement this categorization, we utilize an unsupervised clustering algorithm. We apply the k-means clustering method due to its simplicity and computational efficiency. Given the high dimensionality and temporal dependencies inherent in LOB time series, we first employ an autoencoder model to reduce the data to a lower-dimensional set of informative features. This dimensionality reduction ensures that the input to the k-means clustering algorithm consists of time-independent, meaningful features that effectively capture the full range of market dynamics surrounding each volatility interruption.

We base our autoencoder architecture on the DeepLOB model proposed by Zhang et al. (2019), which has demonstrated efficacy in automatically extracting significant features from LOB data. The encoder component of our model replicates the DeepLOB model, while the decoder component is an exact inversion of the encoder. A detailed illustration of the architecture for both the encoder and decoder is provided in Appendix B, in Figure 13 and Figure 14, respectively. We constrain the latent space between the encoder and decoder to a size of 64[11], ensuring the model extracts only meaningful information from the LOB data. After training the model, we further utilize only the encoder to transform the data for the k-means clustering.

The effectiveness of our clustering results is highly dependent on the choice of k, the number of clusters. To identify the optimal k, we employ the elbow method, a widely-used technique that sets the value of k where the marginal improvement from adding an additional cluster becomes insignificant (Thorndike, 1953; Syakur et al., 2018). This is typically identified as a "bend" or "elbow" in the plot of the sum of squared distances between each sample and its corresponding cluster center for different k. The cluster center, in this context, is defined as the average of all sample midpoints within a given cluster.

While our clustering algorithm is primarily based on feature maps derived from the LOB data, the elbow method is applied specifically to the midpoint data as we aim to analyze the price determination around volatility interruptions in order to identify unnecessary interruptions. This dual approach ensures that the clustering model optimally represents both overall market dynamics and the price formation process around volatility interruptions.

Finally, we classify clusters where volatility interruptions clearly delay the price formation process, i.e., postpone the continuation of a persistent price trend where the direction of the trend before and after the interruption remains consistent, as unnecessary interruptions. In these cases, price changes leading to the volatility

[11] We opted for 64 as we evaluated various other parameter choices (i.e. 512, 256, 128, 64, 32) and 64 was the smallest size with reasonable results as sizes < 64 would result in a much higher loss.

interruption were obviously not caused by irrational or transitory price jumps resulting from liquidity shocks or erroneous orders but reflected new information and corresponding changes in expectations. Here, the volatility interruption defers the incorporation of new information into prices and unnecessarily interrupts the trading process with all negative implications for price discovery and liquidity. For instance, if the midpoint exhibits an upward trend that triggers an interruption by exceeding the upper threshold, and this trend continues in a similar manner after the interruption, it indicates that the market agreed on the upward movement. In such cases, the interruption merely delays the ongoing price formation process and is therefore consequently classified as unnecessary.

We introduce a binary target variable for the different types of volatility interruptions where 1 represents an unnecessary interruption and 0 all other interruptions.

3.2. Results

The elbow method shows that a clustering algorithm with $k = 12$ provides an optimal balance. This configuration introduces a sufficient number of clusters to capture the individual dynamics of samples around volatility interruptions, while avoiding an excessive number of clusters that could lead to redundancy. Figure 4 illustrates the inertia values for clustering models with varying k. The "elbow" point, indicating where the marginal gain of adding more clusters diminishes, is observed at $k = 12$.

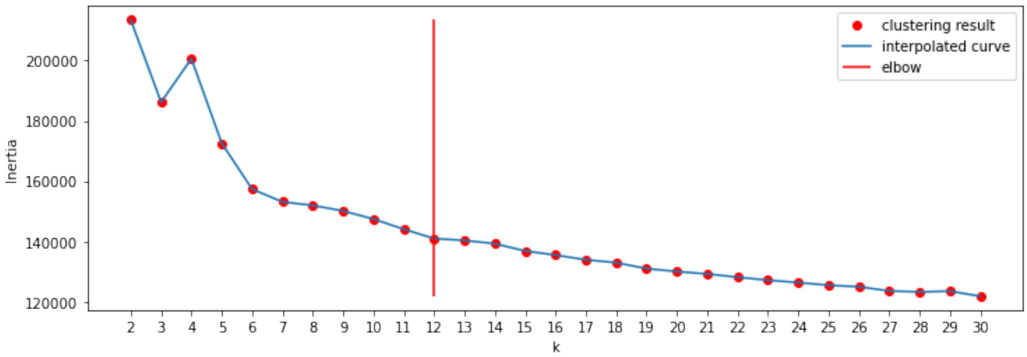

Figure 4: Inertia values at different k
The elbow is determined as the point where the second derivative is maximized in absolute terms. For this calculation every $k < 6$ is ignored as these values for k are not able to capture all complex dynamics happening around a volatility interruption.

The clustering results from the k-means model are presented in Figure 5 and Figure 7. These plots show the midpoint development of each volatility interruption within the clusters, both before and after the volatility interruption, along with the average midpoint trajectory across all observations in each cluster. The interruption is marked at $t = 0$, with $t = -600$ corresponding to 10 minutes prior to and $t = 600$ corresponding to 10 minutes after the interruption, given that the data is sampled at one-second intervals.

Figure 5 displays the first four clusters, which all exhibit a distinct pattern of volatility interruptions delaying the price formation process. Here, the continuous upward or downward trends observed before the interruption persist into the post-interruption phase. The auction phase of the volatility interruption appears to be ineffective in establishing an efficient price for observations in these clusters and simply interrupts a price trend, thereby further hindering prompt price formation. In these cases, the volatility interruption does not help to avoid welfare losses resulting from unsubstantiated price jumps but only comes with the costs of delayed price discovery, unrealized gains from trade, and worse market quality after the interruption. Based on this clustering approach, we find that 39.8% of all samples are classified as unnecessary volatility interruptions.[12]

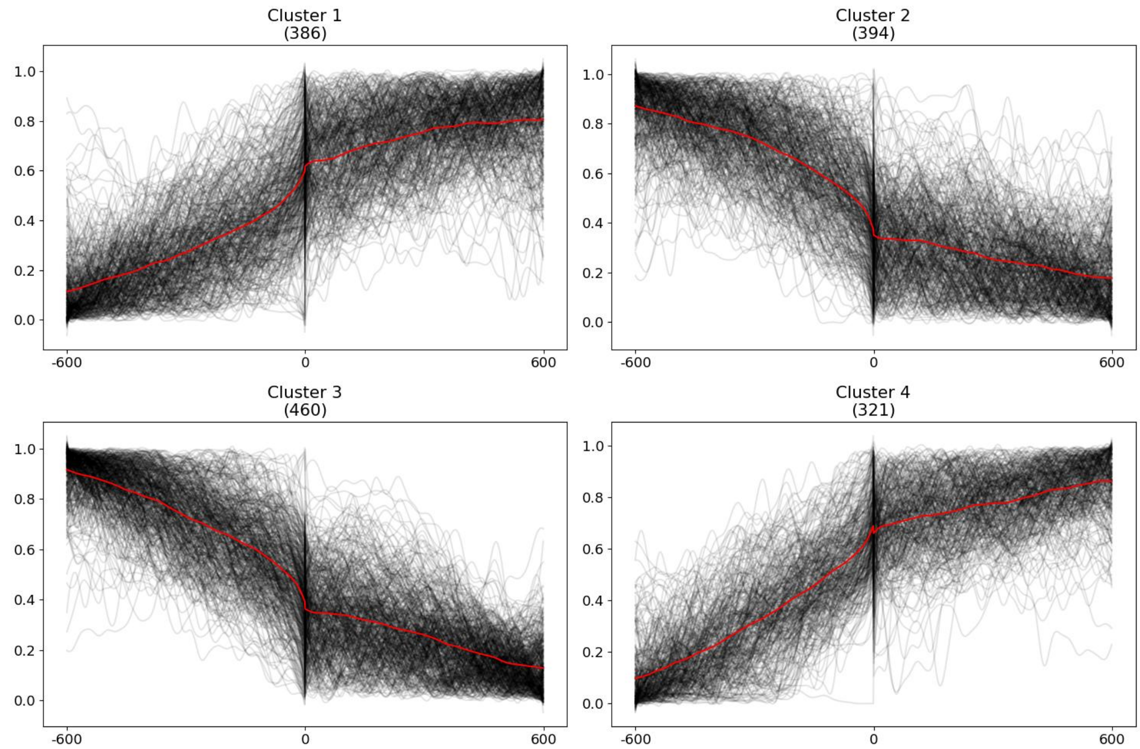

Figure 5: Clustering results for the first four clusters identified as unnecessary interruptions
The plot shows the midpoint trends (in black) for the volatility interruptions in the first four clusters identified as unnecessary interruptions as well as the cluster average (in red). The auction phase at $t = 0$ is omitted as no continuous price determination happens during this period. For a better visualization of the samples, the midpoint trend for both the pre-interruption period ($t \in [-600, -1]$) and post-interruption period ($t \in [1, 600]$) is smoothed by applying the Savitzky-Golay filter (Savitzky and Golay, 1964) with a window length of 31 and a polynomial order of 2. The number of observations in each cluster is given in parentheses under each subheading. The midpoint on the y-axis is bound between 0 and 1 due to the min-max scaling.

[12] Because the data was split chronologically into training, validation, and test sets (see Section 2.2), the imbalance slightly varies across these subsets. Specifically, in the training set, 39.1% of volatility interruptions are labeled as unnecessary, compared to 46.2% in the validation set and 44.7% in the test set.

To illustrate volatility interruptions identified as unnecessary, Figure 6 depicts the midpoint price development for exemplary stocks where an unnecessary volatility interruption occurred on a specific day.[13] In each example, the volatility interruption disrupts a persistent price trend that began shortly after the release of significant ad-hoc news related to the respective stock. Consequently, the price changes delayed by the interruption were likely driven by this new information and corresponding shifts in value expectations. For the stocks shown in Panels A, B, and C, the news was positive, resulting in a sustained price increase (a significant upward revision in long-term financial targets for Infineon Technologies, a major share repurchase program by Deutsche Post DHL Group, and better-than-expected quarterly results for Rheinmetall). In contrast, Panel D features negative news for Deutsche Bank (weaker performance in the credit and derivatives business), leading to a sustained price decline.

All other remaining clusters - not considered as unnecessary - are visualized in Figure 7. These clusters do not show a delayed price formation as observed in the earlier clusters and are therefore not labeled as clearly unnecessary.

The clusters depicted in Figure 7 demonstrate a positive contribution of volatility interruptions to market stability. Volatility interruptions appear effective in "cooling down" the market during periods of extreme price movements, providing market participants with time to reassess available information and determine whether the current price trend is justified as described by Ma et al. (1989). This behavior is particularly evident in Clusters 6, 7, and 12, where the extreme price trends observed before the interruption are almost entirely reversed in the post-interruption phase. This suggests that, during the pre-interruption phase, the market may have either been uncertain about the true valuation of an asset, overreacted to certain news, or experienced substantial market impact from the execution of a large order. In these cases, the interruption phase allowed for a reassessment, showing that the volatility interruption served as a valuable safeguard and contributed directly to price discovery.

Clusters 5, 8, 10, and 11 display a similar behavior, though without the trend-reversing effect observed in the other clusters. Instead, these clusters exhibit a trend-breaking effect, where the price trend stabilizes around the auction price level in the post-interruption phase. This suggests that the volatility interruption effectively contributed to the price formation process, as prices tend to remain close to the price level determined by the auction mechanism of the volatility interruption. Although Cluster 9 also shows a continuous price trend before and after the interruption, it is not classified as a cluster of unnecessary volatility interruptions due to the high level of dispersion and noise within the cluster. In particular, we observe both upward and downward movements of individual samples in the pre-interruption phase, contradicting the interpretation of the cluster center as a stable, one-directed price trend. As

[13] Since all volatility interruptions in this exemplary sample occurred in the afternoon, we only plot the period from the resumption of continuous trading after the intraday auction to market close. No additional volatility interruptions or ad-hoc disclosures occurred for these stocks on the respective dates. Ad-hoc news reports are included in the news data from RavenPack used in this study.

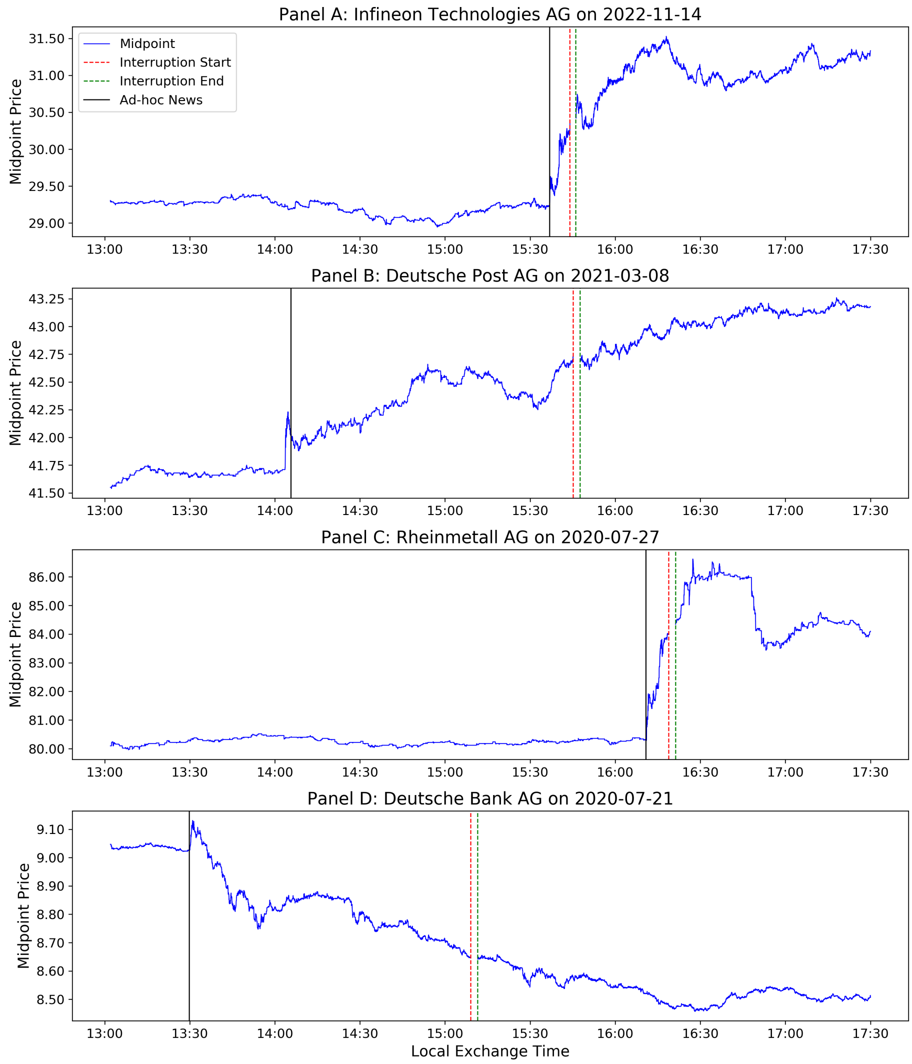

Figure 6: Examples of volatility interruptions identified as unnecessary
Panels A, B, C, and D show the midpoint price development (y-axis) in blue for Infineon Technologies, Deutsche Post DHL Group, Rheinmetall, and Deutsche Bank, respectively, on days experiencing unnecessary volatility interruptions. The start of the volatility interruption is marked in red, while its conclusion and subsequent return to continuous trading is indicated in green. The black vertical line denotes significant ad-hoc news releases. The x-axis displays time in local exchange time.

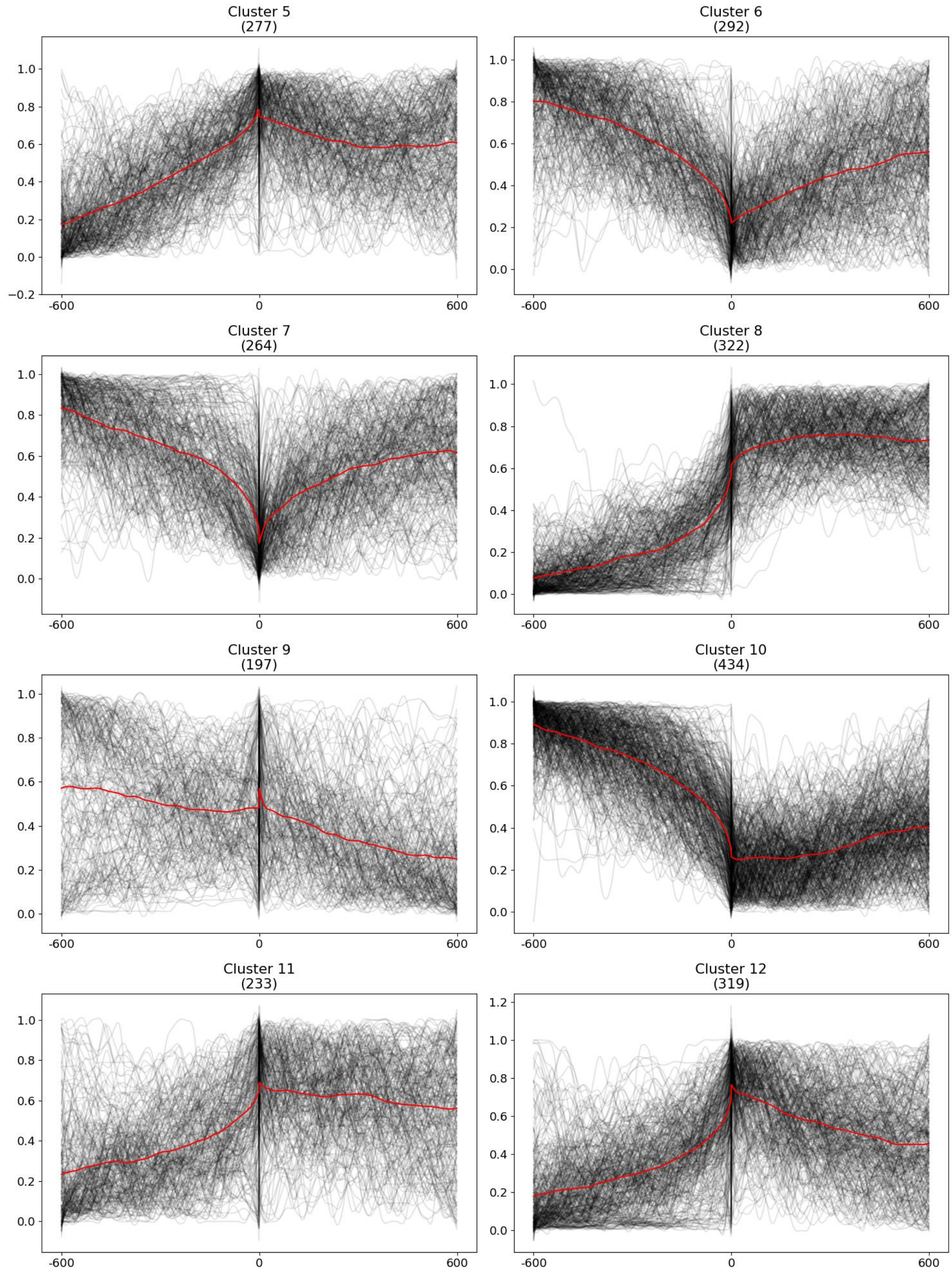

Figure 7: Clustering results for the remaining clusters not identified as unnecessary interruptions
The plot shows the midpoint trends (in black) for the volatility interruptions in the eight remaining clusters not identified as unnecessary interruptions as well as the cluster average (in red). The auction phase at $t = 0$ is omitted as no continuous price determination happens during this period. For a better visualization of the samples, the midpoint trend for both the pre-interruption period ($t \in [-600, -1]$) and post-interruption period ($t \in [1, 600]$) is smoothed by applying the Savitzky-Golay filter (Savitzky and Golay, 1964) with a window length of 31 and a polynomial order of 2. The number of observations in each cluster is given in parentheses under each subheading. The midpoint on the y-axis is bound between 0 and 1 due to the min-max scaling.

we apply the unnecessary label conservatively, we exclude samples from this cluster from our classification of price-delaying volatility interruptions.

Consequently, the clusters shown in Figure 7 do not exhibit a price-delaying effect but rather support overall price discovery during volatile market phases and are therefore not considered as unnecessary volatility interruptions. Figure 11 in Appendix A presents a histogram of the number of samples assigned to each cluster, showing a relatively uniform distribution across all clusters, with cluster 9 consisting of a comparatively smaller number of samples.

4. Prediction and drivers of unnecessary volatility interruptions - The regression approach

4.1. Methodology

Building on the method developed in the previous section, we are able to identify cases of unnecessary volatility interruptions. To improve the effectiveness of circuit breakers and reduce the frequency of unnecessary interruptions, we aim to assess whether it is possible to predict the necessity of a volatility interruption using only pre-interruption (ex-ante) information. This predictive capability would enhance the effectiveness of circuit breakers, preventing disruptions of price trends that are justified, e.g., due to company-specific or macroeconomic news. Moreover, we analyze the market conditions that lead to unnecessary volatility interruptions.

For this purpose, we employ benchmark probit regression models to investigate the predictability of unnecessary volatility interruptions and to reveal their economic drivers. These models apply different sets of independent variables to explain the binary dependent variable, which indicates whether an interruption was identified as unnecessary. Given that our dependent variable is binary, probit regression is an appropriate choice. It effectively models binary outcomes by producing a continuous output that represents the probability of the dependent variable belonging to the target class (here: unnecessary volatility interruptions).

Our first model focuses on the *market quality* conditions prior to the interruption, with the goal of determining whether liquidity, volatility, or trading activity related factors influence the likelihood of an interruption being unnecessary. Therefore, we incorporate the average relative spread, order book depth, trading volume, midpoint volatility and number of order book messages prior to each volatility interruption in this analysis.

Factors beyond market data, such as news events, are often associated with price adjustments. Therefore, we introduce a second probit regression model that incorporates *news and contextual factors* to account for macroeconomic conditions prior to the volatility interruption. Inputs for this model include the number and relevance of recent news articles related to the issuer of the affected stock, the frequency of volatility interruptions across all stocks in the respective market (Xetra) within the preceding hour[14], the fast market indicator[15], and the proximity of the current price to the estimated static triggering threshold[16].

A comprehensive overview of the variables used in both models, including their calculations and a brief explanation of their informational content, is provided in Table 5 in Appendix A.

To address the time-dependent nature of our input data, which covers 10 minutes before each volatility interruption, we use an exponentially decaying average (or sum for count- and volume-based variables such as the number of order book messages and trading volume). This method assigns progressively less weight to data points further from the interruption, ensuring that more recent information has a greater influence on the model's predictions. By transforming the time-series data into one time-independent measure for each interruption, this approach makes it compatible with the probit regression framework.

Following this approach, we set up the first probit regression - the *market quality* model - using the following equation:

$$\begin{aligned} \Pr\left(y_i = 1 | X_i\right) = \Phi\,(&\alpha + \beta_1 \cdot rel\_spread_i + \beta_2 \cdot level\_5\_depth_i \\ &+\beta_3 \cdot trade\_volume_i + \beta_4 \cdot midpoint\_return\_vola_i \\ &+\beta_5 \cdot message\_count_i) \end{aligned} \tag{1}$$

where $i \in \{1, 2, ..., 3899\}$ is the index for each observation, y is the binary target, Φ is the cumulative normal distribution and $X_i = \{rel\_spread_i$, $level\_5\_depth_i$, $trade\_volume_i$, $midpoint\_return\_vola_i$, $message\_count_i\}$ is the list of independent variables.

[14]The inclusion of the number of volatility interruptions across all DAX40 stocks within the hour before the interruption is designed to capture the influence of broader market events, such as the COVID-19 pandemic or the Russian invasion of Ukraine, which may trigger parallel price adjustments across multiple stocks.

[15]The Xetra market supervision department can define a so called fast market for all stocks for a specific day if unusually high volatility is anticipated. This could be the case if major macroeconomic events or crucial announcements (e.g., by central banks, governments, or companies) are expected or occurred before market opening, which affect the entire market. In such a case, the price corridors for the triggering of volatility interruptions are doubled.

[16]Since the actual price thresholds of volatility interruptions on Xetra are not publicly disclosed, we reverse-engineer the static threshold for each stock by observing the maximum price deviation from the last auction price that did not trigger an interruption. Given that the market operator may adjust these thresholds in response to overall market volatility, especially during volatile periods such as the COVID-19 pandemic, we independently estimate the static price range for each stock and month to account for potential adjustments. Figure 9 in Appendix A shows the resulting approximated static price ranges. The distance of the last price prior to the interruption to the static threshold offers insights into whether a significant price jump triggered the volatility interruption or if only minor price changes were sufficient to initiate it. This distinction can provide valuable information for differentiating between unnecessary and relevant interruptions.

The second regression - the *news and contextual factors* model - is set up by the following equation:

$$\begin{aligned}\Pr\left(y_i = 1|X_i\right) = \Phi\,(&\alpha + \beta_1 \cdot news\_count_i + \beta_2 \cdot news\_relevance_i \\ &+\beta_3 \cdot news\_count_i \times news\_relevance_i \\ &+\beta_4 \cdot vola\_interruptions\_market_i \\ &+\beta_5 \cdot fast\_market\_dummy_i \\ &+\beta_6 \cdot distance\_to\_static\_barrier_i)\end{aligned} \tag{2}$$

where $X_i = \{news\_count_i,\ news\_relevance_i,\ vola\_interruptions\_market_i,$ $fast\_market\_dummy_i,\ distance\_to\_static\_barrier_i\}$ is the list of all independent variables in the *news and contextual factors* model.

As a third model, we introduce a comprehensive regression model that incorporates all independent variables used in the previously discussed models. This *full* model aims to integrate a broad range of factors, including market quality metrics, news, and other contextual variables, to predict the likelihood of an unnecessary volatility interruption ex-ante. By encompassing all relevant factors, this *full* model is anticipated to outperform the other models in the prediction task, offering a more robust and accurate assessment.

To evaluate each model's capability in accurately predicting the probability of a volatility interruption being unnecessary using only ex-ante information, we apply common classification metrics to the models' outputs such as accuracy, precision, and recall: Accuracy measures the overall agreement between predicted labels and true labels. Precision for a specific class measures the proportion of true positives among all predicted positives, reflecting the model's likelihood of correctly predicting the target class (here: unnecessary volatility interruptions). Recall measures the proportion of true positives identified out of all actual positives in the target class, indicating the model's ability to capture all relevant instances of unnecessary volatility interruptions. Following standard practice in evaluating prediction models, we conduct an out-of-sample evaluation.

The primary goal of this study is the accurate identification of unnecessary volatility interruptions to minimize these interruptions and mitigate their negative effects. Therefore, an ideal classification model should maximize both precision and recall when classifying unnecessary interruptions. High precision ensures that when the model predicts an interruption as unnecessary, it is highly likely to be correct. High recall ensures that the model identifies as many unnecessary interruptions as possible. However, prediction models regularly face a trade-off between precision and recall. For the task at hand, i.e., improving current circuit breaker mechanisms, precision is the key metric to optimize as a high precision ensures that the model predicts an unnecessary interruption only when the prediction is highly likely to be correct, thereby minimizing the risk of false positives. In a real-world implementation, it is less critical to capture every unnecessary interruption (recall), as each correctly identified and avoided interruption would already improve market efficiency relative to the status quo. However, false positives could disrupt price continuity and compromise overall market stability. Therefore, the model's precision must be prioritized to reduce the risk of misclassifying necessary interruptions as unnecessary, which is crucial to prevent welfare losses caused by large transitory price swings due to, e.g., short-term liquidity crashes, erroneous trading algorithms, or similar events.

The dataset is split into training and test subsets as described in Section 2. Specifically, the training set is created by combining the training and validation sets. This split is performed in a chronological manner, with the training subset containing samples from 2019 to the end of 2022, and the test subset containing the remaining samples from the year 2023.

4.2. Results

Table 3 presents the outcomes of the proposed regression models. The coefficients in this table represent marginal effects, indicating the absolute increase (or decrease) in the probability of the dependent variable for each explanatory variable. The classification metrics at the end of Table 3 reveal that the first two models - the *market quality* and the *news and contextual factors* model - exhibit low predictive power, with accuracy scores below 56%. Such low accuracy indicates that these models perform similar to a naive random guessing approach always predicting the majority class of the imbalanced test sample, where 55.3% of the observations are not deemed unnecessary.

Furthermore, the highest achieved precision scores for these models remain below 55%, underscoring their limited effectiveness in correctly identifying true positives, i.e., unnecessary interruptions. The recall score for the first model reflects similarly poor performance, with a maximum of only 57%. However, for the second model the recall score is quite decent at nearly 76%. The higher recall is likely a result of the model's tendency to predict that a volatility interruption will be unnecessary in most cases, leading to improved recall at the expense of precision and overall accuracy.

In summary, neither model demonstrates the ability to reliably predict, in an ex-ante manner, whether a volatility interruption is unnecessary. These findings suggest that the models may require further refinement or alternative approaches to improve predictive accuracy.

In addition to serving as a benchmark prediction model, the probit regressions enable identifying the economic conditions under which unnecessary volatility interruptions occur. In the first probit model, we do not observe significant results related to volatility, suggesting that unnecessary volatility interruptions can occur during both periods of high volatility and relatively calm phases. Thus, unnecessary interruptions appear to be influenced more by factors other than volatility. With respect to liquidity, our findings reveal that high levels of liquidity prior to an interruption significantly increase the likelihood of the volatility interruption being unnecessary. Both a narrower bid-ask spread and larger depth at the top order book levels are associated with a higher probability of the volatility interruption being unnecessary. In such cases, the cause behind the price change triggering the interruption was not a short-term liquidity shock, which is one reason why circuit breakers are in place. Moreover, unnecessary interruptions are associated with high levels of order book activity, suggesting that market participants might incorporate new information in

Table 3: Marginal effects of probit regression models

This table shows the marginal effects of the proposed probit models. Marginal effects describe the absolute change in the output probability given a change in the independent variables. The relative spread and midpoint return volatility are given in basis points (bps). The level 5 depth and the trade volume are given in hundred thousand Euro. The message count is given in thousands.

	Market Quality	News and Contextual Factors	Full Model
rel_spread	-0.0045**		-0.0033*
	(0.0020)		(0.0019)
level_5_depth	0.0052**		0.0055**
	(0.0026)		(0.0022)
trade_volume	-0.0036***		-0.0036***
	(0.0008)		(0.0007)
midpoint_return_vola	0.0092		0.0153*
	(0.0100)		(0.0091)
message_count	0.0082**		0.0097***
	(0.0040)		(0.0036)
news_count		-0.0038	-0.0033
		(0.0034)	(0.0035)
news_relevance		-0.1114***	-0.0954***
		(0.0358)	(0.0360)
news_relevance_interaction		0.0132**	0.0118**
		(0.0052)	(0.0052)
vola_interruptions_market		-0.0020*	-0.0019
		(0.0012)	(0.0013)
fast_market_dummy		-0.0011***	-0.0010**
		(0.0004)	(0.0004)
distance_to_static_barrier		0.1401***	0.1129**
		(0.0534)	(0.0541)
Pseudo R2	0.007	0.008	0.013
Max. Accuracy	0.557	0.551	0.511
Max. Precision	0.546	0.463	0.553
Max. Recall	0.574	0.760	0.745

Standard errors in parentheses.

* $p < .1$, ** $p < .05$, ***$p < .01$

Maximum accuracy and maximum precision evaluated using multiple thresholds ranging from 0.4 to 0.8 (in 0.05 steps) to turn the continuous probability output of the model into a classification of 0 and 1.

their resting orders leading to many order modifications (or cancellations and subsequent submissions with changed volumes and prices). Conversely, lower trading volumes are also a relevant factor contributing to unnecessary volatility interruptions. High order book activity and liquidity combined with low trading volume may suggest increased participation by algorithmic or high-frequency traders and reduced activity from institutional investors in these market conditions. Algorithmic and high-frequency trading is typically characterized by a high number of messages per trade and smaller trade sizes (Friederich and Payne, 2015), whereas institutional investors tend to execute larger volumes. However, our dataset does not reveal different types of market participants.

Based on the second regression model, our analysis of the impact of news and contextual factors reveals that the presence of relevant news, as indicated by the interaction term between news count and relevance score, significantly increases the likelihood of a volatility interruption being unnecessary. This finding supports the hypothesis that unnecessary interruptions may disrupt the ongoing price discovery process, thereby delaying the integration of new information into stock prices. Additionally, the proximity of the current price to the static price limit significantly affects the classification of the volatility interruption; the closer the price is to this limit, the more likely the interruption is unnecessary. This observation suggests that the existing simplistic rule-based volatility interruption mechanism fails to accurately assess market conditions leading up to the interruption. The mechanism triggers an interruption as soon as the next possible price exceeds the price range, irrespective of whether this breach results from a large, sudden price jump or gradual price adjustments during normal trading conditions if the price is already close to the threshold. Furthermore, significant factors that reduce the likelihood of an unnecessary volatility interruption include the activation of the fast market indicator and the number of triggered volatility interruptions in all DAX40 constituents in the previous hour. When market operators anticipate high volatility, fast market is activated, resulting in doubled price ranges. If doubled ranges are breached, the likelihood of the interruption being necessary increases. Similarly, if numerous volatility interruptions across all stocks have already been triggered in the past hour, it indicates a volatile market, making further interruptions more justified to stabilize the market using safeguards.

The results from the *full* model, which combines both proposed models, demonstrate that the observed effects remain consistent, highlighting the robustness of the findings. Additionally, the coefficient for volatility becomes significant at the 10%-level, suggesting that higher volatility also increases the probability of an interruption being unnecessary. One possible explanation is that volatility interruptions are more likely to be necessary if they are triggered in relatively calm market phases with lower levels of volatility in the pre-interruption period, e.g., due an erroneous or very large order. Similar to the first two models, the prediction accuracy of the *full* model with a maximum precision of 55% is still not sufficient although it includes both market quality and context-related variables. In terms of accuracy, it even performs slightly worse than a naive random guessing approach.

Overall, unnecessary volatility interruptions are more likely to occur during active price formation processes, characterized by increased order book activity and the presence of relevant news. Despite high levels of liquidity, trading volumes are moderate. In such conditions, liquidity providers are more likely to continue quoting rather than exiting the order book despite potentially higher levels of volatility. These market conditions suggest that current price changes are supported by market participants, and volatility interruptions in such situations unnecessarily disrupt ongoing price discovery processes. Furthermore, triggering interruptions when prices are close to the static price limit is more likely to be unnecessary, highlighting the inherent problem of simplistic, rule-based circuit breakers.

However, our analysis shows that predicting unnecessary volatility interruptions ex-ante based on our probit regressions is neither feasible nor adequate for improving the circuit breaker mechanism. While the probit regressions offer insights (explainability) into the factors driving different types of volatility interruptions, the complexity of the market conditions in which these interruptions occur goes beyond the capacity of such models to capture dynamic interactions between various factors.

Therefore, the next section proposes a comprehensive deep learning model designed to capture and model the complex relationships between multiple market quality and contextual factors.

5. Prediction of unnecessary volatility interruptions - The deep learning approach

5.1. Methodology

Building on the results presented in Section 4, we propose a deep learning model designed to predict the likelihood of a volatility interruption being unnecessary using only information available prior to its triggering. Unlike the models employed in the previous section, our deep learning approach can effectively model complex spatial and temporal non-linear relationships among various input factors. This capability eliminates the need to aggregate input time series into a single value, as it is required for the probit regressions using the exponential mean. Moreover, it removes the necessity of pre-selecting potentially relevant independent variables, as the model's architecture inherently identifies and extracts useful features.

Consequently, our deep learning model can process raw information directly without relying on pre-calculated measures. This approach enables us to leverage large volumes of unprocessed data preceding each volatility interruption, capturing the full spectrum of dynamics and relationships among various factors that influence whether the interruption would be unnecessary.

The classification model incorporates two distinct data streams: LOB data and time-dependent news and contextual information. The following subsections describe the deep learning architecture:

The LOB data stream

The LOB data stream comprises 600 snapshots (equivalent to 10 minutes) of LOB data, each containing ten levels of price and quantity information for both the bid and ask sides. The model structure for the LOB data stream is inspired by the architecture of DeepLOB (Zhang et al., 2019). The initial part of the architecture features a CNN comprising multiple neurons responsible for convolution operations. The CNN aims to extract useful features from single LOB snapshots, potentially encompassing convolutions of bid and ask-side price levels or prices and quantities. This distinguishes it from the previous approach using probit regression (Section 4), as it autonomously extracts useful features by design from the order book data, removing the need for pre-calculated and time-aggregated measures. Consequently, it should lead to more meaningful features to be extracted and later used since the model itself employs the features needed for successful classification or dimensionality reduction.

Subsequent to the feature extraction by the CNN layers, the architecture employs an inception module, focusing on time-wise convolution rather than convolution within a single orderbook snapshot. This module is designed to extract features based on the ones previously computed over a specific time frame, summarizing them into one value. For instance, it could calculate the maximum spread over the last five time steps. The extracted features are then fed into an LSTM layer, which captures temporal changes in these features essential for the time-dependent characteristics of the stream of LOB data. Two key modifications have been made to the original DeepLOB architecture of Zhang et al. (2019): the dropout layer's activation probability is reduced to 0.02, which has demonstrated improved performance. Additionally, the architecture is truncated after the final LSTM layer to produce a feature vector for concatenation with the news and contextual data stream.

The news and contextual data stream

Analogous to the regression-based *news and contextual factors* model in Section 4, the news and contextual data stream in the deep learning model is designed to capture information that is not directly observable in market data but may influence whether an interruption is deemed unnecessary. This stream covers the same types of information as those used in the second probit regression in Section 4. However, instead of aggregating this information, as it was done for the regressions, we utilize the raw time series data as input for the deep learning model.

Given the deep learning model's higher expressive power and its ability to autonomously extract relevant features for prediction, we include a broader array of information in the news and contextual data stream, beyond the pre-selected measures used earlier. This enables the incorporation of a wide range of contextual factors during the model training process, aiming to capture as many potential dynamics and relationships as possible to enhance prediction accuracy.

In addition to the factors used in the regression model, we include the number of triggered volatility interruptions for the affected stock, aiming to capture stock-specific high-volatility dynamics. Our hypothesis is that multiple volatility interruptions in the same stock may indicate unnecessary interruptions triggered by a substantial but relevant change in the fundamental value of the stock. We also incorporate the number of volatility interruptions at both the stock and market level over the past 24 hours to account for long-term dynamics. To enrich the trade data, we add the number of executed trades and their average trade volume. In terms of news data, we include the sentiment score of each news item and the similarity of new information to past news, which helps to determine whether the news is recent or if its content may already be reflected in prices. Furthermore, we include order message data, encompassing the number of order submissions, cancellations, and modifications, to provide a detailed view of activities within each order book update. Lastly, we factor in the current minute of the observation, recognizing that key announcements by central banks or similar entities often occur at specific times, such as on the quarter, half, or full hour.[17]

In summary, the news and contextual data stream includes the following features: the number of past volatility interruptions in both the market and the individual stock in the last hour and in the last 24 hours, a dummy variable for the fast market indicator, trade data (including the number of trades, trade volume, and average volume per trade), news data (including the number of news items, sentiment, relevance, and similarity), the number of aggregated order book snapshots for each time interval, the relative distance to the approximated static barrier, message data (including the number of order submissions, cancellations, and modifications) and the current minute of the observation. A detailed overview of these features, their calculations, and their informational content is provided in Table 5 in Appendix A.

The news and contextual data stream is first processed through an LSTM layer to capture temporal dependencies and extract relevant features. This is then followed by a fully connected layer. Unlike the LOB channel, we avoid using a CNN layer for the news and contextual data channel, as there is no rationale to assume meaningful spatial interactions between individual features, such as news sentiment and past volatility interruptions.

The combined model

The combined model architecture integrates the outputs from both data streams. The LOB data channel's feature vector is concatenated with the feature vector from the news and contextual data channel. The resulting combined feature map is then passed through a final fully connected layer, which outputs the probability of the volatility interruption being unnecessary. A detailed illustration of the entire model architecture is provided in Figure 15 in Appendix B.

The deep learning approach is expected to outperform the regression models proposed in Section 4 as it leverages not only pre-calculated features but also raw time series data from the market, news, and other contextual factors. Moreover, it enables the model to automatically extract relevant features for the classification task, potentially leading to more accurate predictions. However, a drawback of this approach is the loss of explainability, as the model's internal operations become more complex and less transparent.

[17] For example, the European Central Bank's monetary policy decisions are published in a press release at 14:15 CET (equals local time for the trading venue Xetra) followed by a press conference starting at 14:45 CET. See https://www.ecb.europa.eu/press/govcdec/mopo/html/index.en.html.

Model training and evaluation

For training the model, the dataset was split into training, validation, and test subsets as described in Section 2. The split was conducted chronologically: the training subset contains samples from April 2019 to November 2022 and the validation subset contains samples from December 2022, while the test subset consists of samples from the year 2023. Due to the imbalanced distribution of the class labels, the training dataset was oversampled using the Synthetic Minority Oversampling Technique (SMOTE) (Chawla et al., 2002). Oversampling ensures a more balanced training dataset, leading to a more effective training process and potentially better classification performance. The training set was exclusively used for model training. During the training process, the validation set was used to monitor loss and accuracy metrics on a small out-of-sample subset. This monitoring was essential to detect potential overfitting. If overfitting was observed, the training process would have been interrupted and terminated. The test subset, which was never used in the training process to prevent information leakage, was utilized to evaluate the model's performance using the standard classification metrics accuracy, recall, and precision as already discussed in Section 4. Additionally, we consider the F_β-score to assess the balance between precision and recall.

As with the regression-based approach, the precise identification of unnecessary volatility interruptions is prioritized over general model accuracy. It is more important for the model to detect unnecessary interruptions with high precision to avoid false positives, even if this comes at the expense of lower overall accuracy. Therefore, also the deep learning model is optimized to maximize precision when identifying unnecessary interruptions.

5.2. Results

The model training converged after 128 iterations. The evolution of the loss function during the training process is documented in Figure 12 in Appendix A. The main objective of this study is to achieve an optimal balance between high precision and recall for predicting unnecessary volatility interruptions, with the primary emphasis on precision to avoid missclassified but relevant interruptions. This balance can be adjusted by testing different cut-off values[18] for the predicted probabilities that indicate unnecessary interruptions.

Figure 8 presents the precision-recall curve, which illustrates the trade-off between precision and recall at different thresholds. This curve is instrumental in identifying an optimal threshold for the underlying prediction task. It becomes evident that thresholds ranging from 0.6 to 0.75 appear to be suitable given the classification task at hand. While those above 0.6 lead to sufficiently high precision scores, thresholds above 0.75 do not further increase precision but only lead to a decrease in recall.

[18] As the output of the model is a continuous value (i.e. the probability of the sample being in the target class) the value needs to be transformed into a binary value representing the class which is done by defining a cut-off value. Every output higher than the cut-off value is considered as part of the target class, in this case, an unnecessary volatility interruption.

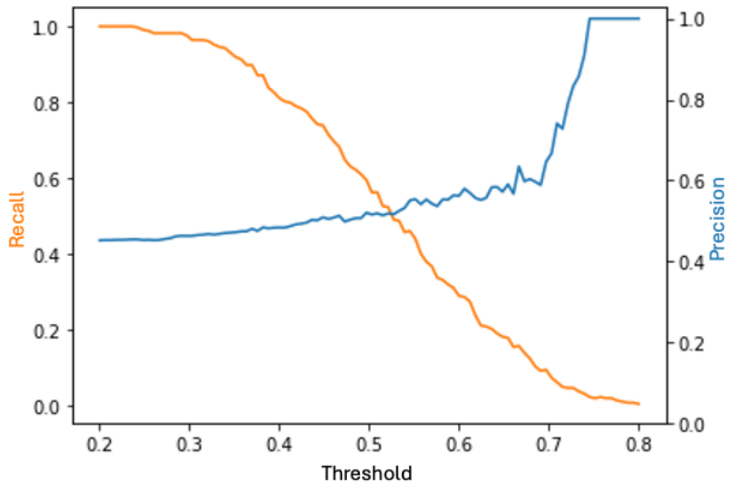

Figure 8: Precision-recall curve of model outputs at different thresholds for labeling as unnecessary

To analyze the impact of the different thresholds within this range on model performance more closely, Table 4 presents the average precision and recall scores over 100 repeated predictions for thresholds ranging from 0.6 to 0.75. The results indicate that using a threshold of 0.7 or higher yields high precision scores exceeding 65% in identifying unnecessary volatility interruptions. However, this comes at the cost of lower recall scores, with only a relatively small proportion (2% to 8%) of all actually unnecessary interruptions being detected. Conversely, implementing the model with lower thresholds results in significantly higher recall scores. For instance, at a threshold of 0.6, the recall is nearly 30%, indicating that nearly one-third of all unnecessary interruptions are correctly identified. However, the precision at this level is lower, just under 60%, indicating that approximately 40% of the interruptions labeled as unnecessary are actually not part of that class. Larger thresholds are also supported by the F_β-score, which summarizes a model's predictive performance by balancing precision and recall. With a β of 0.1 (0.2), ten (five) times more weight is placed on achieving high precision in predicting unnecessary volatility interruptions, thereby minimizing false positives. This means avoiding predictions of unnecessary interruptions when they are in fact justified. For a β of 0.1, the optimal balance between precision and recall is achieved at a threshold of 0.75, whereas for a β of 0.2, the score reaches its peak at a threshold of 0.65.

Given these findings, thresholds between 0.65 and 0.75 appear to offer a balanced performance in accurately identifying unnecessary volatility interruptions. Using these thresholds, a precision between 60% and 100% can be achieved, offering the potential to prevent the triggering of a volatility interruption when there is a high likelihood that this interruption is unnecessary. Such an approach would minimize delays in price determination and enhance overall market quality by mitigating the negative consequences of volatility interruptions. Our approach can either be fully integrated into a circuit breaker mechanism or serve as a decision support system for market operators to shorten the auction phase if the probability of an unnecessary interruption increases as more data becomes available during the auction. When

Table 4: Classification metrics of the deep learning prediction model

Threshold	Accuracy	Precision	Recall	$F_{\beta=0.1}$	$F_{\beta=0.2}$
0.600	0.580 (0.01)	0.561 (0.01)	0.293 (0.01)	0.555 (0.01)	0.541 (0.01)
0.625	0.574 (0.01)	0.564 (0.01)	0.221 (0.01)	0.555 (0.01)	0.532 (0.01)
0.650	0.579 (0.01)	0.597 (0.01)	0.183 (0.00)	0.584 (0.01)	0.549 (0.01)
0.675	0.575 (0.02)	0.619 (0.02)	0.134 (0.01)	0.598 (0.02)	0.544 (0.01)
0.700	0.569 (0.03)	0.659 (0.03)	0.080 (0.00)	0.615 (0.03)	0.515 (0.02)
0.725	0.567 (0.04)	0.803 (0.04)	0.045 (0.00)	0.687 (0.03)	0.487 (0.02)
0.750	0.563 (0.00)	1.000 (0.00)	0.024 (0.00)	0.708 (0.02)	0.385 (0.03)

Classification scores in this table are obtained by predicting the test dataset 100 times, where the minority class of unnecessary interruptions represents 44.7% of the observations in the test sample. The table reports the average classification score as well as the corresponding standard deviation in parentheses. We report results for $\beta \in \{0.1, 0.2\}$ for the F_β-score to emphasize the importance of predicting unnecessary volatility interruptions with a high precision to avoid false positives, i.e., predicting an interruption to be unnecessary although it is actually relevant.

implementing this approach in real-world exchange systems, the optimal threshold selection should depend on the level of automation relative to human intervention. In a largely automated environment, a more restrictive mechanism (i.e., higher thresholds) is preferable to minimize false positives. In contrast, when the mechanism is used as a decision support system with a higher level of human supervision, a less restrictive setting may be more suitable tolerating more false positives, as final decisions are reviewed manually by the market supervision team. Accordingly, in settings with high automation, using an F_β-score with a lower β could help identify a suitable threshold, as this is emphasizing precision and stressing the sensitivity to false positives. Conversely, in lower automation settings, a higher β may be more appropriate, as it places greater emphasis on recall, reducing the risk of missed detections.

Therefore, the model can be tailored to the to the respective market operator's needs to define thresholds that satisfy his requirements on avoiding false positives versus detecting as many unnecessary volatility interruptions as possible. For decision support systems used by market operators who manually supervise volatility interruptions, a less restrictive threshold may be preferable. On the other hand, for automatic mechanisms aimed at canceling unnecessary volatility interruptions, a more restrictive threshold would be more appropriate.

6. Limitations and robustness tests

While this study offers valuable insights into improving the effectiveness of circuit breaker mechanisms in securities markets, certain limitations should be acknowledged when interpreting and applying our findings.

Our analysis primarily examines a specific implementation of circuit breakers — volatility interruptions. Although volatility interruptions are a common safeguard across European stock exchanges, they are less frequently used in other parts of the world (Gomber et al., 2016). In the U.S., circuit breakers are typically implemented as trading halts, with the Limit Up/Limit Down mechanism governing single-stock trading halts (U.S. Securities and Exchange Commission, 2012). Despite the differences between volatility interruptions and trading halts, both types of circuit breakers are triggered based on pre-determined price thresholds. Our methodology is broadly applicable to rule-based circuit breakers in general, making our results relevant to other mechanisms, including the U.S. Limit Up/Limit Down mechanism. Nonetheless, further comprehensive analyses should be conducted before generalizing our findings to other circuit breaker implementations.

Regarding our dataset, another limitation is the absence of information on the specific orders that triggered the volatility interruptions. This data would provide crucial insights into whether these orders were erroneously sent to the exchange or were reasonable. With access to the triggering order, we could directly identify misconfigured or error-induced orders for which a volatility interruption is necessary. Additionally, such data could enhance our prediction models by providing valuable information to better identify unnecessary interruptions. Unfortunately, the triggering order is not included in public market data feeds. However, as the market operator has knowledge of this order, the inclusion of this information into our proposed approach will likely further increase its performance.

To ensure the robustness of our findings and to account for potential alternative explanations for unnecessary volatility interruptions, we conducted several robustness tests based on the probit regression in Section 4, with results detailed in Appendix C. Table 6 presents the results of a regression analysis that includes stock fixed effects to control for any stock-specific factors that might influence our results. We included dummy variables for each stock, and the coefficients for these stock dummies were found to be insignificant. All other effects remained consistent with our original analysis, suggesting that stock-specific factors do not influence our main results.

As highlighted in Figure 2 and described in Section 2, our dataset was significantly affected by the COVID-19 pandemic. To account for potential pandemic-related effects, we conducted a robustness test incorporating a COVID-19 dummy variable for all samples triggered between February and May 2020. Table 7 presents the results of this analysis, showing an insignificant coefficient for the COVID-19 dummy variable in the *full* model, while all other effects remained consistent with our main results. This suggests that our approach is applicable across periods of both high and low volatility and is not influenced by market-wide disruptions.

Lastly, we tested for potential yearly effects by conducting a regression analysis that included dummy variables for each year within our observation period. The results, depicted in Table 8, revealed no significant coefficients for the yearly dummy variables, and no deviations were observed compared to our main results. This indicates that our findings are stable across different years and are not influenced by year-specific factors.

7. Conclusion

Our study addresses a significant gap in the literature on circuit breakers as it demonstrates that it is possible to identify and avoid unnecessary interruptions. Interruptions caused by circuit breakers can negatively impact market quality by delaying price discovery and disrupting liquidity, leading to a trade-off between their protective role and adverse effects on market quality (Hautsch and Horvath, 2019). Our research presents a novel approach based on machine learning techniques to identify unnecessary volatility interruptions that are triggered within an ongoing price trend before and after the interruption, thus leading to an unnecessary delay in price discovery. Based on this identification, we further develop a deep learning model that is able to predict unnecessary volatility interruptions using ex-ante order book information, news, and other contextual features. This approach enables a more nuanced application of circuit breakers, facilitating the development of advanced market safeguards that are only triggered when necessary. Therefore, our study adds to the discussion of improved circuit breaker mechanisms, such as the forward-looking circuit breaker proposed by Bongaerts et al. (2024). Our study also contributes to the broader literature on market microstructure by applying advanced machine learning techniques to a critical market design aspect.

By analyzing the circumstances in which unnecessary volatility interruptions are triggered, we find that they are more likely to occur when liquidity is high, when there is increased activity in form of submissions and cancellations in the order book, and when relevant news reports are present. Large price fluctuations observed under these conditions rather point to well-functioning price discovery instead of erroneous price jumps. Moreover, volatility interruptions are also more likely to be unnecessary when the last price prior to the interruption is near the triggering threshold. This indicates that the existing simplistic rule-based mechanism is not capable of differentiating between plausible price changes and unsubstantiated price jumps.

The practical implications of our findings are substantial. Exchange operators can leverage our models to refine their safeguard mechanisms, avoiding or shortening unnecessary interruptions and improving overall market efficiency. Additionally, our results can inform regulatory discussions, potentially leading to more nuanced and effective rules for circuit breakers.

Future research could build on our work by applying our methodology to different markets and types of circuit breakers, further validating the robustness of our findings. Additionally, exploring other machine learning techniques and incorporating non-public information from market operators will likely enhance the predictive accuracy of our models.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT 4o in the writing process in order to check grammar and improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Abad, D., Pascual, R., 2010. Switching to a temporary call auction in times of high uncertainty. Journal of Financial Research 33, 45–75.

Abad, D., Pascual, R., 2013. Holding back volatility: Circuit breakers, price limits, and trading halts, in: Market Microstructure in Emerging and Developed Markets: Price Discovery, Information Flows, and Transaction Costs. John Wiley & Sons. chapter 17, pp. 303–324.

Bongaerts, D., Draus-De Luca, S., Van Achter, M., 2024. Circuit breakers and market runs. Review of Finance (forthcoming) .

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research 16, 321–357.

Clapham, B., Gomber, P., Haferkorn, M., Panz, S., 2017. Managing excess volatility: Design and effectiveness of circuit breakers. Working Paper .

Deutsche Boerse Group, 2024. T7 Release 12.1: Market model for the trading venue Xetra. URL: https://www.xetra.com/resource/blob/3873542/fd4c485e86257fd9d7320ded6a6c4df3/data/T7_Release_12.1_-_Market_Model%20_Xetra.pdf.

Easley, D., López de Prado, M., O'Hara, M., Zhang, Z., 2021. Microstructure in the machine age. The Review of Financial Studies 34, 3316–3363.

European Commission, 2014. Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on Markets in Financial Instruments and amending Directive 2002/92/EC and Directive 2011/61/EU Text with EEA relevance: MiFID II. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014L0065.

European Securities and Markets Authority, 2022. ESMA report on trends, risks and vulnerabilities, No. 2, 2022. URL: https://www.esma.europa.eu/sites/default/files/library/esma50-165-2229_trv_2-22.pdf.

Fama, E.F., 1988. Perspectives on October 1987, or, what did we learn from the crash? Center for Research in Security Prices, Graduate School of Business, University of Chicago.

Friederich, S., Payne, R., 2015. Order-to-trade ratios and market liquidity. Journal of Banking & Finance 50, 214–223.

Gomber, P., Clapham, B., Haferkorn, M., Panz, S., Jentsch, P., 2016. Ensuring market integrity and stability: Circuit breakers on international trading venues. The Journal of Trading 12, 42–54.

Hautsch, N., Horvath, A., 2019. How effective are trading pauses? Journal of Financial Economics 131, 378–403.

Kim, K.A., Rhee, S.G., 1997. Price limit performance: Evidence from the tokyo stock exchange. Journal of Finance 52, 885–901.

Kwan, A., Philip, R., Shkilko, A., 2021. The conduits of price discovery: A machine learning approach. Working Paper .

Lauterbach, B., Ben-Zion, U., 1993. Stock market crashes and the performance of circuit breakers: Empirical evidence. The Journal of Finance 48, 1909–1925.

Lee, C.M., Ready, M.J., Seguin, P.J., 1994. Volume, volatility, and New York Stock Exchange trading halts. The Journal of Finance 49, 183–214.

Lehmann, B.N., 2019. Commentary: Volatility, price resolution, and the effectiveness of price limits. Journal of Financial Services Research 3, 205–209.

Ma, C.K., Rao, R.P., Sears, S., 1989. Limit moves and price resolution: The case of the treasury bond futures market. The Journal of Futures Markets 9, 321–335.

O'Hara, M., 2015. High frequency market microstructure. Journal of Financial Economics 116, 257–270.

Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36, 1627–1639.

Sifat, I.M., Mohamad, A., 2019. Circuit breakers as market stability levers: A survey of research, praxis, and challenges. International Journal of Finance & Economics 24, 1130–1169.

Sirignano, J., Cont, R., 2019. Universal features of price formation in financial markets: Perspectives from deep learning. Quantitative Finance 19, 1449–1459.

Subrahmanyam, A., 1994. Circuit breakers and market volatility: A theoretical perspective. The Journal of Finance 49, 237–254.

Subrahmanyam, A., 2013. Algorithmic trading, the flash crash, and coordinated circuit breakers. Borsa Istanbul Review 13, 4–9.

Syakur, M.A., Khotimah, B.K., Rochman, E., Satoto, B.D., 2018. Integration k-means clustering method and elbow method for identification of the best customer profile cluster, in: IOP Conference Series: Materials Science and Engineering. Article No. 012017.

Thorndike, R.L., 1953. Who belongs in the family? Psychometrika 18, 267–276.

U.S. Securities and Exchange Commission, 2012. Investor bulletin: Measures to address market volatility. URL: https://www.sec.gov/resources-for-investors/investor-alerts-bulletins/investoralertscircuitbreakersbulletinhtm.

U.S. Securities and Exchange Commission, 2016. Opening remarks at the equity market structure advisory committee meeting. URL: https://www.sec.gov/newsroom/speeches-statements/white-remarks-emsac-meeting-080216.

World Federation of Exchanges, 2008. Questionnaire on circuit-breakers for Q4 2008. URL: https://www.world-exchanges.org/storage/app/media/research/Studies_Reports/2008%20WFE%20Circuit%20Breakers%20Report.pdf.

Zhang, Z., Zohren, S., Roberts, S., 2019. DeepLob: Deep convolutional neural networks for limit order books. IEEE Transactions on Signal Processing 67, 3001–3012.

Appendix A: Variable description and additional descriptives

Table 5: Detailed explanation of all variables used in the models in Section 4 and Section 5

Feature	Model	Explanation
rel_spread	P1, P3	Measures the difference between the best bid and ask prices relative to the midpoint.
level_5_depth	P1, P3	Sum of the quoted Euro volume at the first five order book levels.
trade_volume	P1, P3, DL	Euro volume of all executed trades.
midpoint_return_vola	P1, P3	Measures overall volatility by calculating the standard deviation of midpoint returns.
message_count	P1, P3	Number of all messages (submissions, deletions, modifications).
news_count	P2, P3, DL	Number of news articles mentioning the corresponding company in the last 10 minutes.
news_relevance	P2, P3, DL	Average relevance score of news articles mentioning the corresponding company in the last 10 minutes.
vola_interruptions _market	P2, P3, DL	Measures the number of triggered volatility interruptions across all DAX40 constituents in the last hour. In the deep learning model, we additionally calculate this measure based on the last 24 hours.
fast_market_dummy	P2, P3, DL	Dummy variable whether the fast market indicator is set by the market operator or not.
distance_to_static _barrier	P2, P3, DL	Distance of the current price to the approximated static price range. The value is bounded between 0 and 1 with 1 being close to the barrier and 0 being far away.
vola_interruptions _instrument	DL	Measures the number of triggered volatility interruptions in the same instrument in the last hour. In the deep learning model, we additionally calculate this measure based on the last 24 hours.
count_trades	DL	Number of executed trades.
count_buys	DL	Number of buyer initiated trades.
count_sells	DL	Number of seller initiated trades.
avg_trade_volume	DL	Average trade size in euro.

continued . . .

. . . continued

Feature	Model	Explanation
news_sentiment	DL	Average sentiment score of news articles mentioning the corresponding company in the last 10 minutes.
news_similarity	DL	Average similarity score of news articles mentioning the corresponding company in the last 10 minutes.
count_add	DL	Number of messages representing the adding of a new order.
count_delete	DL	Number of messages representing the deletion of a persisting order.
count_modify	DL	Number of messages representing the modification of a persisting order.
ob_changes	DL	Number of order book updates.
minute_of_hour	DL	Number representing the current minute in the time series. Bounded between 0 and 59.
last_auction _price_return	DL	Return of the current price and the last auction price.

P1 represents the *market quality* model, P2 the *news and contextual factors* model and P3 the *full* model in the probit regressions discussed in Section 4. DL refers to the deep learning model described in Section 5.

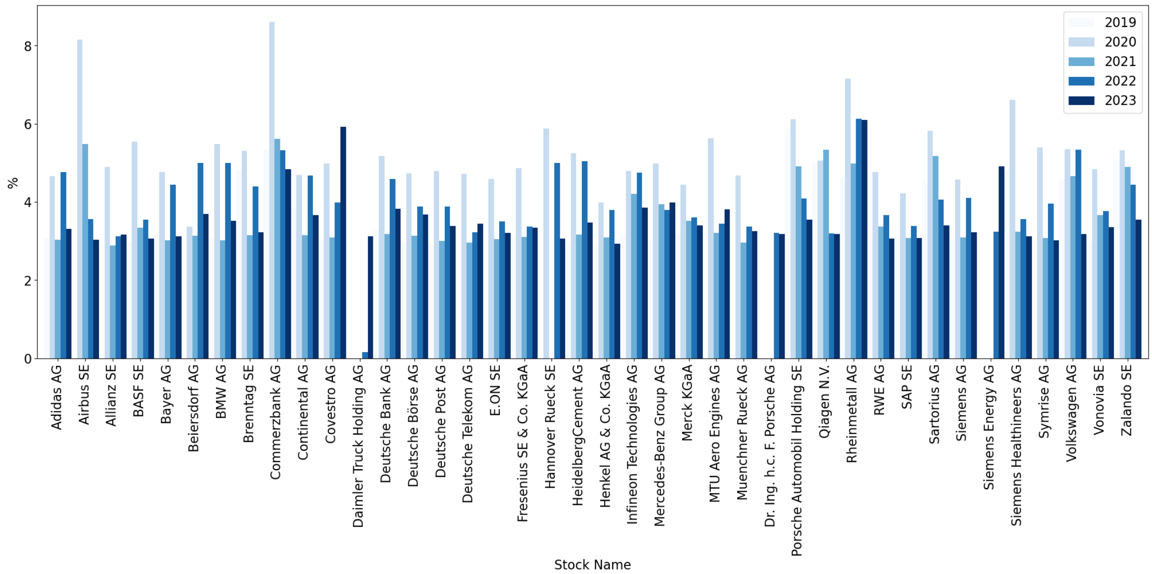

Figure 9: Approximated static price ranges for all DAX40 constituents
This figure shows the approximated static price ranges based on the largest price deviation seen from the last auction price, which serves as reference price for the static threshold triggering the interruption. For brevity and better visualization, only the yearly approximated thresholds are shown although monthly approximations are used in the models. Missing bars are due to the stock's later listing (e.g., Daimler Truck Holding AG emerged from a spin-off and was first traded in December 2021).

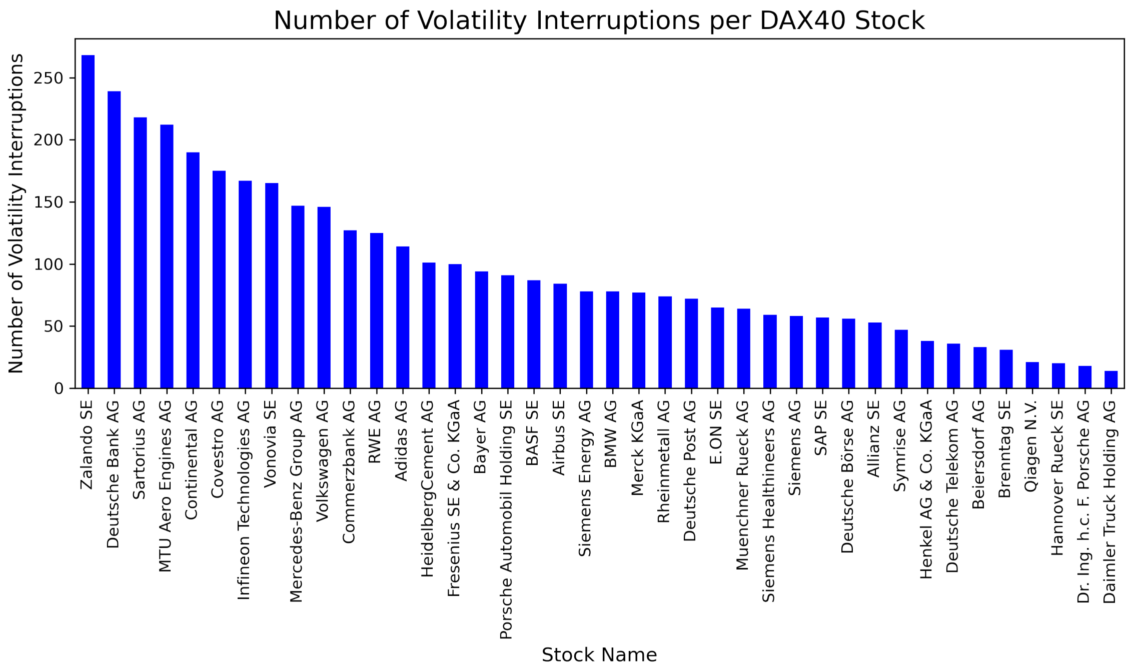

Figure 10: Total number of volatility interruptions for each DAX40 stock between 2019 and 2023

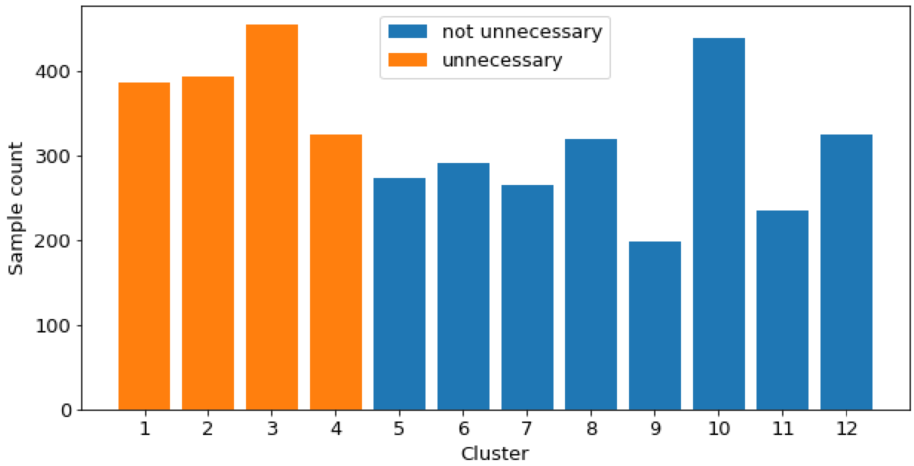

Figure 11: Histogram of the clustering results

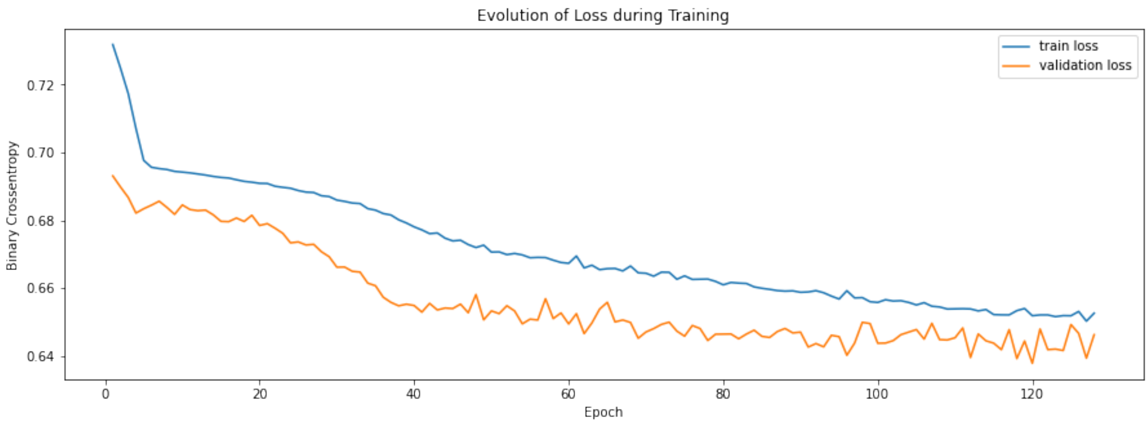

Figure 12: Loss during the training of the classification model for the training and testing subset

Appendix B: Architectures of the deep learning models

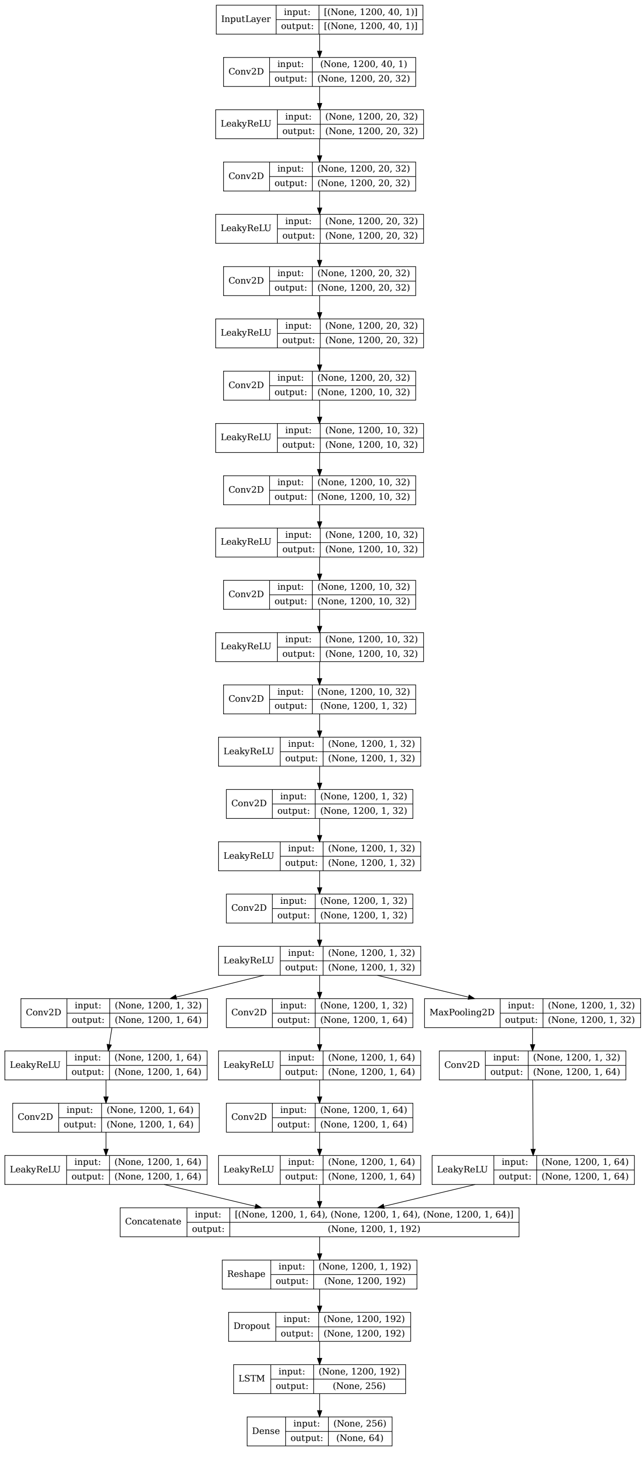

Figure 13: Architecture of the implemented encoder part of the autoencoder model

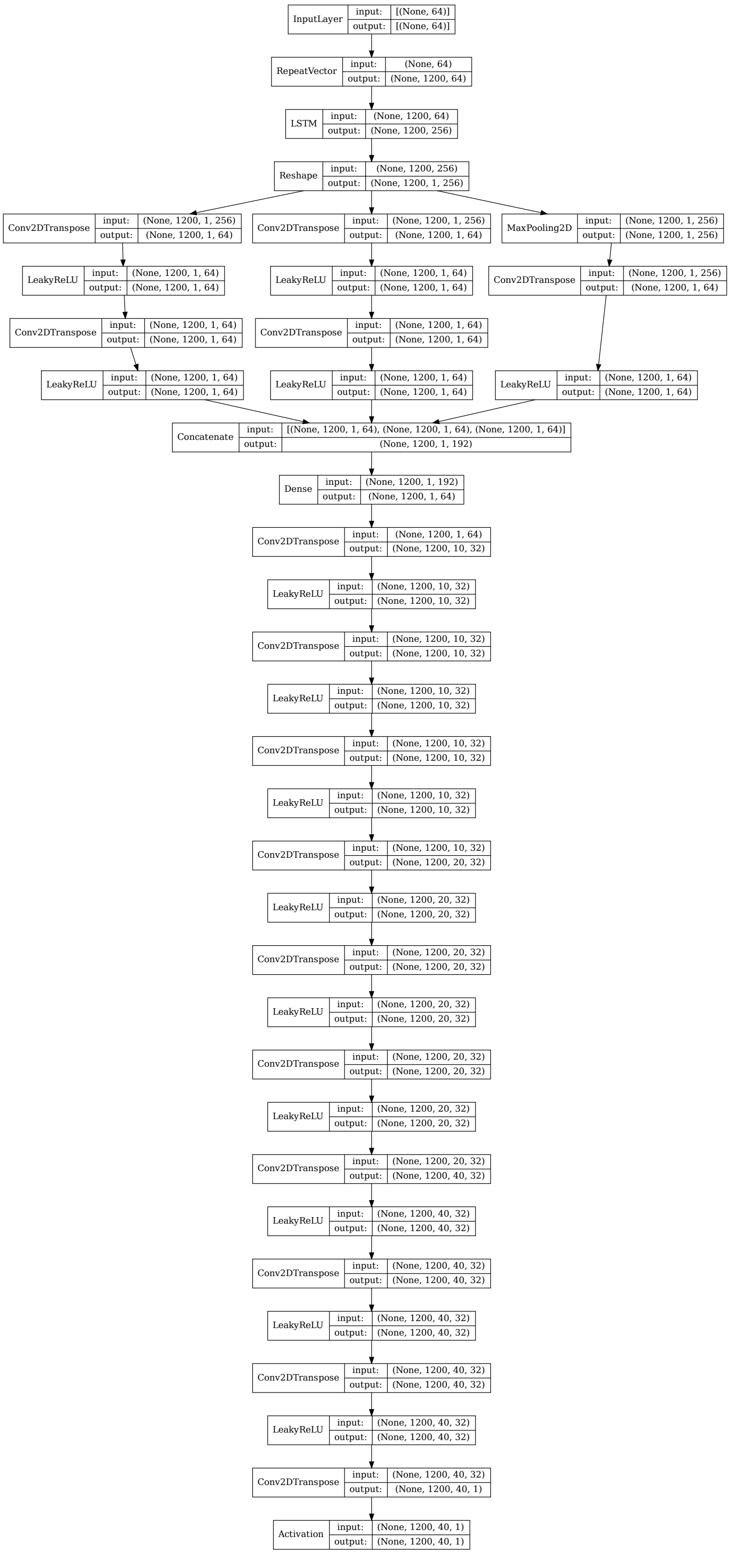

Figure 14: Architecture of the implemented decoder part of the autoencoder model

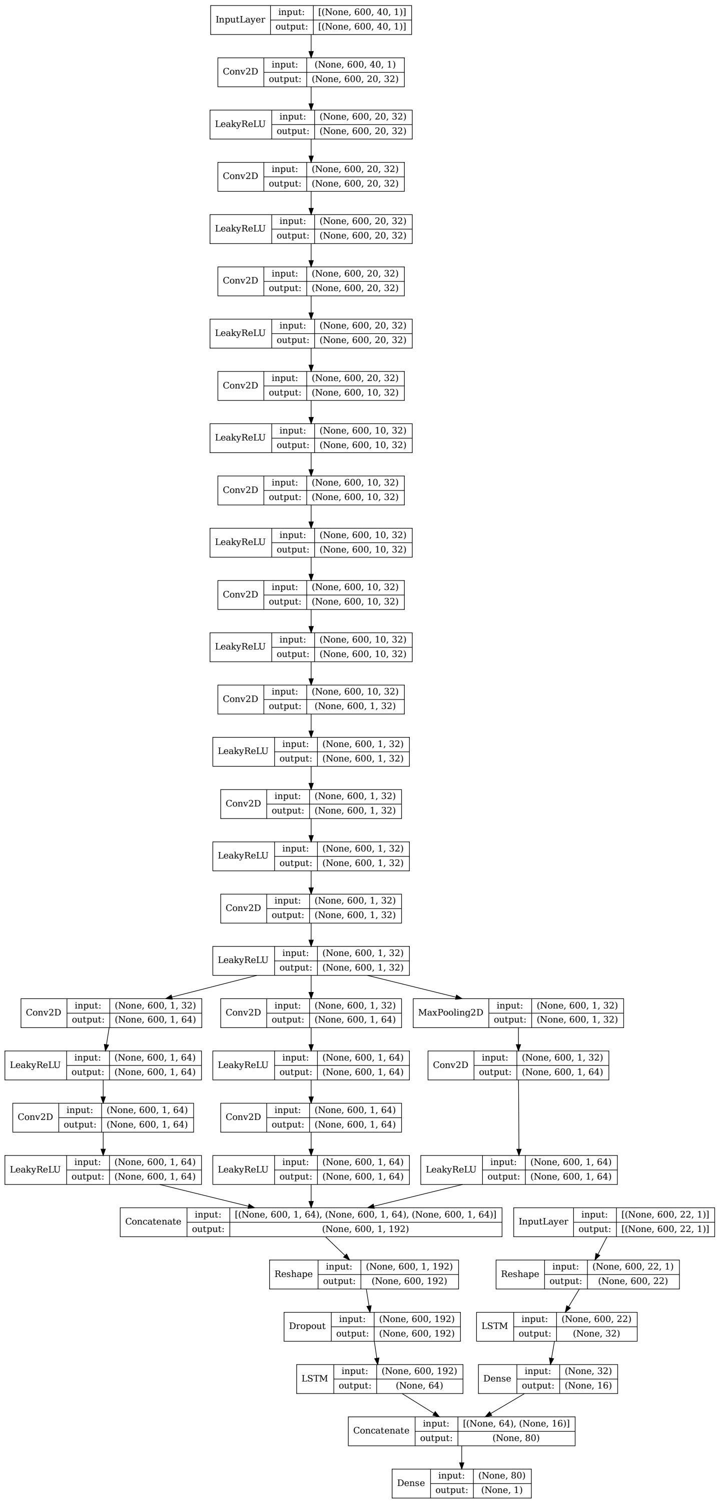

Figure 15: Architecture of the implemented classification model

Appendix C: Additional results and robustness tests

Table 6: OLS regression results with stock fixed effects

This table shows the results from OLS regressions analogous to the models described in Section 4. Here, the results are not based on probit regressions, but on OLS regressions as the inclusion of stock fixed effects introduces 39 dummy variables and probit regressions do not perform well when comprising a large amount of parameters. The relative spread and midpoint return volatility is given in basis points (bps). The level 5 depth and the trade volume is given in hundred thousand Euro. The message count is given in thousands.

	Market Quality	News and Contextual Factors	Full Model
const	0.4460***	0.3637***	0.3342***
rel_spread	-0.0046**		-0.0031
level_5_depth	0.0058**		0.0042*
trade_volume	-0.0034***		-0.0034***
midpoint_return_vola	0.0136		0.0190**
message_count	0.0061		0.0049
news_count		-0.0053	-0.0042
news_relevance		-0.1263***	-0.1134***
news_relevance_interaction		0.0113**	0.0106*
vola_interruptions_market		-0.0017	-0.0018
fast_market_dummy		-0.0012***	-0.0012***
distance_to_static_barrier		0.1811***	0.1605***
Stock FE	Yes	Yes	Yes
R-squared	0.0291	0.0296	0.0355
R-squared Adj.	0.0180	0.0183	0.0230

Standard errors in parentheses.
* $p < .1$, ** $p < .05$, ***$p < .01$

Table 7: Marginal effects of probit regression models including a COVID-19 dummy variable

This table shows the marginal effects of the proposed probit models. Marginal effects describe the absolute change in the output probability given a change in the independent variables. The COVID-19 dummy variable marks all observations between February and May of 2020. The relative spread and midpoint return volatility is given in basis points (bps). The level 5 depth and the trade volume is given in hundred thousand Euro. The message count is given in thousands.

	Market Quality	News and Contextual Factors	Full Model
rel_spread	-0.0037*		-0.0037*
	(0.0020)		(0.0020)
level_5_depth	0.0040*		0.0039*
	(0.0022)		(0.0022)
trade_volume	-0.0033***		-0.0031***
	(0.0007)		(0.0007)
midpoint_return_vola	0.0116		0.0151*
	(0.0091)		(0.0091)
message_count	0.0083**		0.0087**
	(0.0035)		(0.0037)
covid_dummy	-0.0412*	-0.0073	0.0102
	(0.0215)	(0.0244)	(0.0259)
news_count		-0.0044	-0.0040
		(0.0034)	(0.0035)
news_relevance		-0.1248***	-0.1129***
		(0.0359)	(0.0361)
news_relevance_interaction		0.0154***	0.0140***
		(0.0052)	(0.0052)
vola_interruptions_market		-0.0019	-0.0023*
		(0.0013)	(0.0014)
fast_market_dummy		-0.0010**	-0.0010**
		(0.0004)	(0.0004)
distance_to_static_barrier		0.1411**	0.1334**
		(0.0568)	(0.0572)
Pseudo R-squared	0.0071	0.0069	0.0115

Standard errors in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 8: Marginal effects of probit regression models including dummy variables for each year

This table shows the marginal effects of the proposed probit models. Marginal effects describe the absolute change in the output probability given a change in the independent variables. The relative spread and midpoint return volatility is given in basis points (bps). The level 5 depth and the trade volume is given in hundred thousand Euro. The message count is given in thousands.

	Market Quality	News and Contextual Factors	Full Model
rel_spread	-0.0037*		-0.0032
	(0.0020)		(0.0020)
level_5_depth	0.0041*		0.0037*
	(0.0022)		(0.0022)
trade_volume	-0.0030***		-0.0029***
	(0.0007)		(0.0007)
midpoint_return_vola	0.0097		0.0130
	(0.0091)		(0.0091)
message_count	0.0075**		0.0079**
	(0.0036)		(0.0037)
year_2020	-0.0328	0.0093	0.0018
	(0.0380)	(0.0390)	(0.0396)
year_2021	-0.0459	-0.0296	-0.0372
	(0.0436)	(0.0430)	(0.0435)
year_2022	0.0071	0.0408	0.0146
	(0.0371)	(0.0365)	(0.0372)
year_2023	0.0395	0.0626	0.0425
	(0.0388)	(0.0383)	(0.0390)
news_count		-0.0042	-0.0038
		(0.0034)	(0.0035)
news_relevance		-0.1138***	-0.1048***
		(0.0360)	(0.0362)
news_relevance_interaction		0.0147***	0.0136***
		(0.0052)	(0.0052)
vola_interruptions_market		-0.0017	-0.0019
		(0.0013)	(0.0013)
fast_market_dummy		-0.0009**	-0.0009**
		(0.0004)	(0.0004)
distance_to_static_barrier		0.1364**	0.1210**
		(0.0551)	(0.0556)
Pseudo R-squared	0.0088	0.0091	0.0127

Standard errors in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$