Przegląd Statystyczny. Statistical Review, 2025, vol. 72, 3, 28–42 https://doi.org/10.59139/ps.2025.03.2
Received: 4.11.2025, revised: 31.12.2025, accepted: 4.01.2026

Empirical analysis of trade duration distributions: The WIG20 case

Agnieszka Lach[a]

Abstract. The aim of the study presented in this paper is to analyse the distributions of trade durations for WIG20 stocks using data from May 2025, with a particular focus on modelling doubly truncated data. Left-truncated distributions for trade durations have already been described in the literature, which is justified, as the values of an excessive proportion of observations were equal to zero. In this study, it is assumed that the data are also right-truncated due to time limitations between the trading sessions. Three doubly truncated continuous distributions were analysed in the study, namely the lognormal, the Pareto and the Weibull distribution. To satisfy the assumptions of stationarity and independence, the data were divided into smaller subsamples. Goodness-of-fit tests were then performed to determine which theoretical distribution best describes the empirical data. The results indicate that the quality of the fit depends on the lower truncation level – the higher the truncation threshold, the better the lognormal distribution fits the empirical trade durations.

Keywords: probability distributions, doubly truncated data, high-frequency data, trade durations
JEL: C12, C24, C41, G19

1. Introduction

The analysis of durations, defined as the waiting times between consecutive financial events, is an important aspect of market microstructure research. This analysis depends on the event type, measurement precision and the characteristics of the used data. This paper focuses on the time intervals between the successive transactions, referred to as trade durations. The timestamps of trades are recorded in milliseconds, and the research is based on tick-by-tick (intraday transaction) data.

Ni et al. (2010) distinguish between two main approaches to modelling durations: the mainstream finance approach and the econophysics approach. In the former, the most commonly used framework is the autoregressive conditional duration (ACD) model and its various extensions. In the econophysics approach, modelling is typically based on the continuous-time random walk (CTRW) framework.

A ground-breaking contribution to the mainstream finance approach was made by Engle and Russell (1998), who introduced the autoregressive conditional duration (ACD) model. Numerous modifications have since appeared in the literature,

[a] Poznań University of Economics and Business, Institute of Informatics and Quantitative Economics, Department of Operations Research and Mathematical Economics, Al. Niepodległości 10, 61–875 Poznań, Poland, e-mail: agnieszka.lach@ue.poznan.pl, ORCID: https://orcid.org/0000-0002-2831-6336.

including the logarithmic ACD model, the Markov-switching ACD model and the threshold model. However, most of them were applied to low-frequency data – that is data measured with a precision of one second or longer – and to relatively old datasets (Li et al., 2023).

An overview of the second approach might be found in Ni et al. (2010). Early studies in this field suggest that inter-trade durations may be described by the power-law, modified power-law or stretched-exponential distributions. However, as asserted by Ni et al. (2010), the Weibull distribution often provides the best fit for inter-trade durations.

Recent research in this area has evolved toward modelling random variables characterised by data inconsistencies. The ongoing digitalisation of economic processes and the growing use of algorithmic trading on stock exchanges have led to order submissions and trade executions occurring within fractions of a second. Ultra-high-frequency trading algorithms now generate transactions at microsecond or even nanosecond intervals. This introduces rounding effects, resulting in a substantial proportion of observations with zero values. At the same time, not all market activity is driven by high-frequency trading. O'Hara (2015) distinguishes two broad categories of market participants: high-frequency traders and non-high-frequency traders, emphasising that both groups have to optimise their trading strategies with respect to market design and the behaviour of other traders. Moreover, O'Hara (2015) highlights the need for new analytical tools capable of capturing the evolving market microstructure, including changes in the size of trade, the trading volume, inter-trade durations, and the interdependencies among these characteristics.

In terms of inter-trade durations, the division of market activity into high-frequency and non-high-frequency trading naturally implies a distinction between durations close to zero and those of larger magnitudes. This perspective has already begun to be reflected in the literature. A high proportion of zero-valued data necessitates partitioning trade duration distributions. Two main approaches to this issue can be distinguished: truncating data close to zero or modelling the entire distribution. Empirical findings by Kızılersü et al. (2016) for the London Stock Exchange demonstrated that the durations recorded in the order book can be described by a left-truncated Weibull distribution, with the lower truncation point set at ten milliseconds. Kreer et al. (2022) analysed entire duration distributions, also for the London Stock Exchange, and found that a mixture of one exponential and one Weibull distribution models inter-trade waiting times remarkably well across all time horizons. The Weibull component captures the behaviour in the transition and tail regions, whereas the exponential component explains the excess mass at zero. Li et al. (2023) also analysed complete duration distributions and found that inter-trade durations of stocks follow bimodal distributions, driven by switching between market-making and speculative trading strategies.

In this article, the distribution of trade durations is modelled using doubly truncated distributions. Observations below a selected lower truncation point are excluded from the analysis, since durations close to zero represent a substantial proportion of the data and require separate modelling. Additionally, an upper truncation threshold is introduced to account for the daily nature of the data. Since the analysis is conducted on a day-by-day basis, it is reasonable to assume that trade durations should also be right-truncated.

The contribution of this paper to the existing knowledge is twofold. Firstly, doubly truncated distributions are applied to model trade durations. Secondly, the research is conducted using data from the Polish stock exchange.

2. Methodology

2.1. Trade durations

To formally define trade durations, a point-process framework is adopted. Trade durations represent time intervals between successive transactions. Let $\tau \in (0, \infty)$ be a variable representing physical time and let $\tau_0 = 0$. Furthermore, let $\tau_1, \tau_2, \ldots$ be non-negative random variables satisfying condition $\tau_t \leq \tau_{t+1}$ for $t = 1, 2, \ldots$. These variables indicate successive points in time at which transactions are executed on the stock exchange. A point process is defined as sequence (τ_t), $t = 1, 2, \ldots$. A trade duration is defined as a time interval between the consecutive events:

$$d_t = \begin{cases} \tau_t - \tau_{t-1}, & \text{for } t \geq 2, \\ \tau_1, & \text{for } t = 1. \end{cases} \tag{1}$$

Process (d_t), $t = 1, 2, \ldots$, is called the duration process and is referred to as the process associated with (τ_t).

2.2. Continuous distributions for trade durations

This subsection presents three distributions that are suitable for modelling doubly truncated trade durations. The same distributions – lognormal, Pareto and Weibull – were employed by Kızılersü et al. (2016) to model trade durations after left truncation. Special attention is paid to the tails of these distributions, as they determine the behaviour of extreme durations.

The probability density function of the lognormal distribution is given by:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \tag{2}$$

where $x > 0$, $\mu \in \mathbb{R}$ is the location parameter, and $\sigma > 0$ is the scale parameter. The cumulative distribution function of the lognormal distribution has no closed-form expression and is determined using numerical methods. The tails of this distribution may range from light to medium-heavy (Čížek et al., 2005), where light tails denote those that decay at an exponential rate, and medium-heavy tails refer to those that decay more slowly than exponential but faster than power-law tails.

The Pareto distribution has the following probability density and cumulative distribution functions:

$$f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \tag{3}$$

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha, \tag{4}$$

where $x > 0$, $\alpha > 0$ is the shape parameter, and $\theta > 0$ is the scale parameter. The Pareto distribution is classified as a heavy-tailed distribution (Klugman et al., 2008), meaning that its tail decays at a power-law rate.

Finally, the probability density function and the cumulative distribution function of the Weibull distribution are given by:

$$f(x) = \alpha\beta^{-\alpha}x^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right], \tag{5}$$

$$F(x) = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right], \tag{6}$$

where $x > 0$, $\alpha > 0$ is the shape parameter, and $\beta > 0$ is the scale parameter. The Weibull distribution has a heavy tail (Haas & Pigorsch, 2009).

The parameters of all distributions examined in this study were estimated using the maximum likelihood method.

2.3. Doubly truncated distributions

According to the classification of sample types proposed by Cohen (1991), this study deals with doubly truncated samples with known truncation points. If we denote the observations by x_i, $i = 1, \dots, n$, then for each observation in a doubly truncated sample the following holds: $T_1 \leq x_i \leq T_2$, where T_1 and T_2 are known truncation points.

Now the cumulative distribution function and the probability density function of a doubly truncated random variable are defined. Let X be a random variable with cumulative distribution function F and probability density function f, and let $Y = X | T_1 \leq X \leq T_2$ denote the doubly truncated random variable. Then the cumulative distribution function F^* and the probability density function f^* of the doubly truncated random variable are given by (Krysicki et al., 2004):

$$F^*(y) = \frac{F(y) - F(T_1)}{F(T_2 + 0) - F(T_1)} \text{ for } T_1 \leq y \leq T_2, \tag{7}$$

$$f^*(y) = \frac{f(y)}{F(T_2) - F(T_1)} \text{ for } T_1 \leq y \leq T_2. \tag{8}$$

The expressions above outline the general formulation of doubly truncated random variables. In empirical applications, closed-form expressions for the probability density or cumulative distribution functions in the doubly truncated case are often highly complex and therefore typically evaluated numerically.

2.4. Stationarity and independence tests

The goodness-of-fit between empirical distributions and selected continuous distributions can be assessed only when the stochastic processes under consideration are stationary and the observations are independent and identically distributed. Trade durations are typically stationary; however, they frequently attest to significant autocorrelation (Doman, 2011). Stationarity here means weak (covariance) stationarity, where the process has constant mean, variance and covariance over time.

In order to verify the stationarity of the time series, the Augmented Dickey–Fuller test was conducted. The null hypothesis of this test assumes that the analysed time series is non-stationary, while the alternative hypothesis presumes that the series is stationary.

The independence of the examined time series was tested with the Ljung-Box test. Under the null hypothesis, the observations are independent, i.e. all autocorrelation coefficients equal zero, while the alternative hypothesis assumes that at least one autocorrelation coefficient is nonzero.

2.5. Goodness-of-fit tests

This subsection presents the statistical tests used to evaluate how well the theoretical distributions fit the empirical trade duration data. Two classical goodness-of-fit tests were applied: the Kolmogorov-Smirnov and the Cramér-von Mises tests. The Anderson-Darling test, which is also popular, was intentionally omitted. The power of this test stems from its weighting function, which assigns greater weight to the tails of the distribution. When the distribution is doubly truncated, as is the case in this study, the weighting function would need to be modified accordingly, which is beyond the scope of this analysis.

Let us assume that we have a sample $\boldsymbol{X} = (X_1, \dots, X_n)'$ of i.i.d. random variables with an unknown distribution function F. To construct a goodness-of-fit test, both the empirical cumulative distribution function and the theoretical cumulative distribution function are required.

The empirical cumulative distribution function is defined as (Krzyśko, 2004):

$$F_n(x; \boldsymbol{X}) = \frac{\#\{1 \leq j \leq n : X_j \leq x\}}{n}, \tag{9}$$

where $x \in \mathbb{R}$, $\boldsymbol{X} \in \mathbb{R}^n$ and # denote the number of elements satisfying the condition. The theoretical cumulative distribution functions are presented in Subsection 2.2, and their doubly truncated forms are defined in Subsection 2.3.

The null and alternative hypotheses to be tested are as follows:

$$\begin{aligned} &H_0 : F = F_0, \\ &H_1 : F \neq F_0, \end{aligned} \tag{10}$$

where F_0 denotes the assumed theoretical distribution function.

The two goodness-of-fit statistics used in this study are defined below. The Kolmogorov-Smirnov statistic is defined as:

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x; \boldsymbol{X}) - F_0(x)|, \tag{11}$$

and the Cramér-von Mises statistic as:

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x; \boldsymbol{X}) - F_0(x)]^2 dF_0(x). \tag{12}$$

In the empirical part of this study, the parameters of the theoretical distributions are estimated from the data, therefore composite hypotheses are tested. As noted by Pewsey (2018), in such cases the sampling distributions of the test statistics depend on several factors – including the form of the theoretical distribution, the estimated parameters, the estimation method and the sample size. Consequently, the exact distributions of the Kolmogorov-Smirnov and Cramér-von Mises statistics are unknown, and their sampling distributions were approximated using bootstrap methods.

The decision regarding the null hypothesis is based on the value of the test statistic. If the calculated value of the statistic exceeds the corresponding critical value, null hypothesis H_0 is rejected. Otherwise, there is no sufficient evidence to reject H_0.

2.6. The choice of truncation points

The selection of the left truncation point corresponds to identifying a boundary between high-frequency and non-high-frequency trading activity. This boundary is not fixed and has evolved over time with advances in computational speed and other operational and technological factors (O'Hara, 2015). Empirical evidence (Li et al., 2023) suggests that trade durations are characterised by bimodal distributions, with modes occurring at the millisecond and second levels; the former commonly associated with high-frequency traders and the latter with non-high-frequency traders. Accordingly, the boundary is assumed to lie between these regimes, and four left truncation points were considered (in seconds): 0.001, 0.01, 0.1, 1.

In addition to left truncation, right truncation was applied for practical reasons. When analysing daily data, trade durations cannot exceed the length of the trading session itself. Therefore, a right truncation point of 28,800 seconds was imposed, corresponding to eight hours. This choice does not substantially alter the modelling assumptions but yields favourable computational properties by mitigating some numerical issues.

Overall, the adopted truncation pattern reflects both the empirical characteristics of trade durations and the practical considerations related to data resolution and numerical estimation.

2.7. Description of the empirical study

The methodological framework was adapted from Kızılersü et al. (2016). The analysis was carried out separately for each company according to the following procedure:

1. Data collection and preliminary processing, which involved removing records with no trading volume and observations recorded outside the regular trading hours of the stock exchange (9:00 a.m. to 4:50 p.m.). Additionally, it was assumed that the time between the stock exchange's closing and reopening was equal to zero;
2. Calculation of trade durations according to formula (1);
3. Truncation of data below a specified lower threshold T_1, where $T_1 \in \{0.001, 0.01, 0.1, 1\}$, and above the upper threshold $T_2 = 28{,}800$ (equivalent to 8 hours in seconds);
4. Division of the doubly truncated sample obtained in step 3 into smaller samples of n = 125 observations each. Subsamples containing fewer than 125 observations (typically the last one) were excluded;
5. Testing for stationarity and independence in the subsamples obtained in step 4;
6. Conducting goodness-of-fit tests for doubly truncated lognormal, Pareto and Weibull distributions (applied only to stationary and independent subsamples).

This systematic approach ensured consistency across all the analysed companies and allowed a reliable comparison of goodness-of-fit results between different truncation levels.

3. Empirical study

3.1. Data

The study was carried out for 20 companies listed on the Warsaw Stock Exchange (WSE), included in the WIG20 index as of 31st May 2025. These were: Alior Bank (ALRR), Allegro (ALEP), Bank Pekao (PEO), Budimex (BDXP), CCC (CCCP), CD Projekt (CDR), Dino Polska (DNP), KGHM (KGH), Kęty (KTY), Kruk (KRU), LPP (LPPP), mBank (MBK), Orange Polska (OPL), Pepco (PCOP), PGE (PGE), PKN Orlen (PKN), PKO Bank Polski (PKO), PZU (PZU), Santander Bank Polska (SPL1) and Żabka (ZAB). The time and sales data for these companies over the period of May 2025 were obtained from the Eikon Refinitiv database. The time data were measured with a millisecond accuracy. All the results presented in this section are sorted in an ascending order by company market capitalisation as of 30th December 2025.

Descriptive statistics of trade durations are presented in Table 1. Although all the stocks included in the WIG20 index are a part of the same benchmark, they differ in terms of liquidity. For the most liquid company in the index, buy-sell transactions occur on average every 1.9 seconds, whereas for the least liquid one, approximately every 14.6 seconds. In general, higher market capitalisation is typically associated with higher liquidity. All the distributions are right-skewed and leptokurtic. The share of zero values in the total number of observations ranges from 38.4% to 50.7%. The high proportion of such observations suggests that zero-inflated models might be appropriate, meaning that observations close to zero can be treated differently from the remaining data. The following analysis focuses exclusively on the part of the dataset that remains after excluding the observations with values close to zero.

Table 1. Descriptive statistics for trade durations

Ticker[a]	N	Zero values (in %)	Min (s)	Max (s)	Mean (s)	St. Dev. (s)	Skewness	Excess kurtosis
KTY	36,559	42.7	0.0	1,106.0	14.6	42.2	6.2	68.1
CCCP	79,770	46.7	0.0	566.4	6.7	20.6	6.7	76.3
KRU	81,668	38.4	0.0	374.5	6.5	15.4	5.0	43.8
OPL	68,045	40.4	0.0	843.4	7.8	25.4	8.1	117.8
ALRR	122,392	47.8	0.0	362.8	4.4	14.7	6.8	71.6
BDXP	77,738	48.6	0.0	694.4	6.9	19.9	5.8	62.2
PCOP	133,632	43.1	0.0	281.8	4.0	11.4	5.5	48.0
PGE	108,336	41.6	0.0	363.0	4.9	14.4	5.6	50.1
ZAB	102,632	42.7	0.0	454.7	5.2	15.8	6.9	78.1
CDR	127,531	48.4	0.0	302.9	4.2	12.3	5.5	46.5
ALEP	216,991	43.7	0.0	267.1	2.5	7.5	6.4	67.3
LPPP	38,159	40.1	0.0	1,033.7	14.0	43.6	7.0	84.6
DNP	135,978	43.4	0.0	378.1	3.8	12.2	6.7	73.7
MBK	59,185	39.3	0.0	898.3	9.0	28.6	6.9	79.3
PEO	169,257	45.8	0.0	308.0	3.2	10.6	6.8	73.9
KGH	167,696	50.7	0.0	234.2	3.2	9.4	5.7	51.5
SPL1	74,748	47.9	0.0	610.6	7.1	22.2	6.1	60.1
PZU	123,814	47.8	0.0	309.8	4.3	11.9	5.1	40.4
PKO	247,259	47.4	0.0	278.0	2.2	7.1	7.1	85.6
PKN	279,210	43.6	0.0	168.9	1.9	5.1	5.7	57.5

a Tickers are ordered in an ascending order according to the company market capitalisation as of 30th December 2025.

Source: author's calculations.

3.2. Results

First, the stationarity and independence of the entire samples were examined separately for each company. The augmented Dickey-Fuller and Ljung-Box test results revealed that all the analysed samples are non-stationary and show significant autocorrelation. This justified dividing the samples into smaller subsamples.

To determine the appropriate subsample size, the data for the PGE company, after being doubly truncated, were sequentially divided into subsamples of the sizes of 30, 50, 75, 100, 125 and 150 observations. Stationarity and independence tests were then performed on each subsample, and the sample size boasting the highest passing rates for both tests simultaneously was selected. Kızılersü et al. (2016) adopted a sample size of 30 observations in their study; however, for the PGE data, the simultaneous pass rates for both tests did not exceed 16% for subsamples of this size. Subsequent sample sizes were chosen arbitrarily, starting from 50 and increasing by 25 each time. At the size of 150, the passing rates began to decrease again. Ultimately, the full samples were split into subsamples of 125 observations each, as this division ensured a reasonable proportion of stationary and independent subsamples. The number of subsamples of the size of 125 corresponding to various lower truncation thresholds for each company is presented in Table 2.

Table 2. Number of samples of the size of 125 for various lower truncation thresholds

Ticker[a]	Lower truncation threshold			
	0.001 s	0.01 s	0.1 s	1 s
KTY	145	131	111	89
CCCP	307	285	231	186
KRU	370	340	298	253
OPL	287	243	200	154
ALRR	460	398	320	226
BDXP	284	258	210	171
PCOP	556	497	403	308
PGE	454	385	317	242
ZAB	439	405	333	246
CDR	481	440	355	277
ALEP	887	778	615	434
LPPP	158	144	120	90
DNP	552	481	371	270
MBK	253	221	175	132
PEO	670	590	466	302
KGH	603	541	438	342
SPL1	277	250	197	145
PZU	468	420	350	282
PKO	958	844	633	436
PKN	1,170	1,049	887	599

a Tickers are ordered in an ascending order according to the company market capitalisation as of 30th December 2025.

Source: author's calculations.

The passing rates of the stationarity and independence tests for the samples of the size of 125 are presented in Table 3. Depending on the lower truncation threshold, the proportion of subsamples that passed the stationarity test ranges from 89.5% to 98.6%, while for the independence test from 73.4% to 93.7%. The proportion of the subsamples that passed both the stationarity and independence tests at the same time varies between 69.6% and 89.8%.

Table 3. Passing rates of stationarity and independence tests for the samples of the size of 125 (in %)

Ticker[a]	Lower truncation threshold											
	0.001 s			0.01 s			0.1 s			1 s		
	(1)[b]	(2)[c]	(3)[d]	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
KTY	98.6	82.8	82.1	97.7	78.6	77.1	97.3	88.3	86.5	96.6	86.5	83.2
CCCP	94.5	81.4	77.9	89.5	81.4	74.0	93.5	83.6	78.4	95.7	83.9	81.2
KRU	97.3	86.8	84.3	97.9	88.5	87.4	97.7	86.9	84.9	98.0	88.9	87.8
OPL	94.4	85.0	82.2	93.0	84.4	80.3	90.0	80.0	77.0	95.5	81.8	80.5
ALRR	94.4	86.3	82.8	96.2	87.4	84.2	94.7	84.1	80.6	96.9	92.0	89.8
BDXP	94.0	84.5	80.3	94.6	81.0	78.3	96.2	81.4	80.5	97.7	85.4	84.2
PCOP	96.0	83.3	81.3	96.6	81.5	79.1	95.8	80.9	78.2	96.1	82.1	79.9
PGE	95.4	81.7	79.3	94.6	81.6	79.0	95.3	83.3	82.0	97.5	87.6	86.8
ZAB	93.9	78.6	74.9	92.4	79.3	75.6	94.0	76.9	74.5	94.7	82.9	80.1
CDR	94.8	86.1	83.0	93.9	85.2	81.6	92.4	87.0	83.7	96.0	91.0	88.8
ALEP	94.1	83.7	80.1	94.3	83.6	80.3	93.3	78.4	75.1	95.6	84.1	81.3
LPPP	94.9	85.4	82.9	91.0	81.9	76.4	91.7	76.7	73.3	93.3	84.4	78.9
DNP	96.4	84.1	82.3	92.5	82.3	78.8	94.1	83.0	79.5	96.3	91.5	88.2
MBK	94.9	93.7	89.7	97.3	88.2	87.3	96.0	85.7	82.9	96.2	86.4	85.6
PEO	96.1	84.2	82.2	94.9	84.1	80.5	94.0	83.5	81.1	96.4	89.4	86.8
KGH	95.7	82.3	80.1	95.9	82.8	80.2	95.4	79.9	76.9	95.9	84.2	81.0
SPL1	96.8	86.6	85.2	95.2	86.8	84.4	94.9	83.3	80.2	98.6	89.7	89.0
PZU	94.0	84.6	81.4	94.8	83.1	80.5	95.1	82.9	81.4	95.4	84.0	81.2
PKO	97.2	83.8	81.7	96.8	83.8	81.5	96.4	83.3	81.4	96.1	89.5	87.2
PKN	93.6	81.1	78.0	93.0	78.1	74.8	90.6	73.4	69.6	96.8	88.2	86.6

a Tickers are ordered in an ascending order of company market capitalisation as of 30th December 2025.
b Augmented Dickey-Fuller test for stationarity. c Ljung-Box test for independence. d Both tests (b and c).
Source: author's calculations.

The next step involved conducting goodness-of-fit tests, applied only to samples that had successfully passed the stationarity and independence checks. The tests were performed using the Kolmogorov-Smirnov and Cramér-von Mises statistics. Given the similarity of the results, only the passing rates from the Cramér-von Mises test have been reported (Table 4). The interpretation depends on the value of the lower truncation point. For the smallest truncation point of 0.001 s, the highest passing rates were observed for the Weibull distribution (for 17 out of the 20 companies analysed), although the passing rates for the lognormal distribution were also notably high. At the next truncation level, 0.01 s, both the lognormal and Weibull distributions recorded the highest passing rates for 10 firms each. For the remaining truncation levels, i.e. 0.1 s and 1 s, the passing rates for the lognormal distribution reached 100% across all firms. The Pareto distribution represented the weakest overall fit.

Table 4. Passing rates of Cramér-von Mises test for doubly truncated samples of the size of 125 (in %)

Ticker[a]	Lower truncation threshold											
	0.001 s			0.01 s			0.1 s			1 s		
	(1)[b]	(2)[c]	(3)[d]	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
KTY	73.1	0.0	96.5	90.1	3.1	90.8	100.0	13.5	86.2	100.0	56.2	79.3
CCCP	88.9	1.3	96.7	92.3	10.9	95.1	100.0	18.2	84.4	100.0	68.1	82.4
KRU	98.1	3.2	85.6	100.0	14.1	75.5	100.0	50.3	78.3	100.0	79.7	82.9
OPL	58.9	1.4	99.3	93.4	6.6	96.3	100.0	18.0	82.7	100.0	59.5	82.3
ALRR	74.6	3.3	100.0	93.2	5.3	98.7	100.0	26.6	90.9	100.0	80.1	87.3
BDXP	85.9	1.8	98.2	95.4	7.0	91.5	100.0	23.9	82.3	100.0	64.7	79.5
PCOP	91.4	0.9	98.4	96.0	8.7	94.7	100.0	36.2	81.0	100.0	87.3	84.6
PGE	69.4	2.2	99.6	97.4	4.2	95.8	100.0	16.4	80.1	100.0	79.8	82.2
ZAB	96.1	2.3	91.0	93.6	5.9	87.6	100.0	23.7	75.3	100.0	72.5	83.2
CDR	90.2	1.5	96.6	98.4	7.1	91.1	100.0	40.0	76.1	100.0	84.5	86.6
ALEP	89.3	2.5	97.9	99.1	8.9	96.2	100.0	54.3	88.5	100.0	92.4	88.0
LPPP	75.3	1.9	99.4	90.3	4.9	100.0	100.0	14.2	93.3	100.0	50.0	86.1
DNP	71.0	2.7	99.6	90.6	14.8	99.0	100.0	24.3	88.1	100.0	87.4	90.1
MBK	63.6	1.6	100.0	85.5	10.0	100.0	100.0	21.7	86.3	100.0	61.4	85.0
PEO	85.2	2.7	98.9	95.9	11.0	97.1	100.0	35.8	89.7	100.0	87.0	83.0
KGH	90.7	1.7	98.3	97.6	8.3	93.1	100.0	49.1	82.7	100.0	91.7	89.9
SPL1	70.0	1.4	100.0	85.6	9.6	98.4	100.0	15.8	75.0	100.0	61.1	83.9
PZU	94.4	1.5	97.4	98.3	5.0	91.4	100.0	37.8	82.3	100.0	85.4	81.4
PKO	90.3	2.0	98.5	95.9	14.6	96.0	100.0	42.2	90.5	100.0	95.9	84.5
PKN	99.1	1.2	94.2	99.9	6.1	91.5	100.0	75.3	91.0	100.0	99.7	93.8

a Tickers are ordered in ascending order of company market capitalisation as of 30th December 2025. b Lognormal distribution. c Pareto distribution. d Weibull distribution.

Source: author's calculations.

For comparison purposes, the Kolmogorov-Smirnov and Cramér-von Mises tests were also applied to the left-truncated samples, and again only to those which had successfully passed the stationarity and independence tests. As before, only the passing rates from the Cramér-von Mises test have been reported (Table 5). Several observations can be made on their basis. For the left-truncated samples with left truncation points set at 0.001 s or 0.01 s, the highest passing rates in the majority of cases were reported for the Weibull distribution. Notably, at the lowest left truncation point of 0.001 s, the passing rates for the lognormal distribution were very low. For higher left truncation levels, namely 0.1 s and 1 s, the results were similar to those obtained for doubly truncated samples, with the lognormal distribution showing the best overall fit. In general, the passing rates for doubly truncated samples were higher than those for left-truncated samples.

Table 5. Passing rates of Cramér-von Mises test for left truncated samples of the size of 125 (in %)

Ticker[a]	Lower truncation threshold											
	0.001 s			0.01 s			0.1 s			1 s		
	(1)[b]	(2)[c]	(3)[d]	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
KTY	10.3	4.8	20.3	17.6	9.2	25.2	94.6	17.3	19.6	98.9	77.1	58.2
CCCP	2.6	1.0	61.4	28.8	10.0	64.6	98.7	17.8	55.6	99.4	70.6	67.3
KRU	7.3	6.2	51.0	56.5	20.4	44.7	100.0	55.3	65.7	100.0	91.3	76.2
OPL	7.0	16.9	44.2	28.0	9.4	54.6	97.5	14.7	48.8	100.0	62.7	71.5
ALRR	9.8	7.5	58.7	37.7	6.4	86.4	87.5	11.6	63.7	99.5	82.8	77.2
BDXP	6.0	2.5	51.8	24.8	9.2	48.2	99.1	30.2	58.6	100.0	75.4	68.2
PCOP	9.0	1.1	66.5	53.9	13.7	79.9	99.8	30.5	65.7	100.0	92.1	75.1
PGE	7.3	8.3	63.6	43.4	4.7	72.3	99.4	17.8	55.7	100.0	80.0	69.4
ZAB	12.8	3.0	70.4	43.2	18.2	69.2	96.7	22.0	39.5	99.1	70.1	58.5
CDR	13.3	2.3	68.6	44.3	10.9	65.8	99.7	36.9	61.0	100.0	88.8	80.8
ALEP	13.4	3.9	72.2	71.7	14.7	89.6	100.0	39.1	75.5	100.0	90.0	80.3
LPPP	4.4	2.6	36.5	18.1	10.5	48.3	81.7	13.9	43.9	92.2	59.4	59.0
DNP	10.3	7.3	54.6	41.4	16.6	90.6	99.2	19.0	68.5	100.0	87.9	84.3
MBK	7.5	11.5	39.8	24.4	12.8	54.1	96.0	18.3	34.3	96.8	66.0	69.9
PEO	17.6	4.6	63.8	53.2	13.6	92.4	95.3	23.2	68.3	100.0	85.1	76.9
KGH	9.0	3.0	73.0	52.0	12.1	72.8	100.0	46.1	71.4	100.0	95.2	84.1
SPL1	11.2	5.5	53.3	21.6	13.8	55.2	90.8	11.4	30.9	100.0	80.0	73.1
PZU	8.6	2.1	69.8	47.9	5.1	55.8	99.7	33.9	70.4	100.0	93.8	77.3
PKO	13.4	2.3	60.7	62.3	16.9	91.6	100.0	32.9	77.1	100.0	94.2	74.1
PKN	17.5	1.5	84.9	95.5	34.4	82.5	100.0	56.5	85.3	100.0	100.0	85.7

a Tickers are ordered in ascending order according to the company market capitalisation as of 30th December 2025. b Lognormal distribution. c Pareto distribution. d Weibull distribution.
Source: author's calculations.

4. Limitations of the study

The analysis presented in this paper is exploratory in nature. Its main limitations are discussed below, along with directions for future research.

To start with, this study focuses exclusively on trade durations, which represent only one, albeit important, component of market microstructure. Other relevant characteristics, including price dynamics, trade size and trading volume, should be examined in further research.

Moreover, due to the precision of the available data, recorded at a millisecond level, the analysis is restricted to non-high-frequency trading activity. In this setting, all zero-valued durations can be clearly attributed to high-frequency traders who account for approximately 40–50% of the observations. Although attempts have been made in the literature to model such mixed data at a millisecond precision using exponential and Weibull distributions mixed together (Kreer et al., 2022), such an approach is considered to involve substantial simplification. Accurately modelling the distributional mass close to zero would require data recorded with a higher time precision, which were not available to the author at the time of conducting this study. Consequently, the analysis focuses on modelling the left-truncated part of the distribution.

Also, this study should be regarded as an initial step towards more comprehensive analyses. The data span was arbitrarily chosen and limited to a single month, serving primarily as an illustrative sample. Future research should aim to investigate the potential differences across time scales and to identify calendar-related effects such as variations across months, weeks of the month or days of the week.

It must also be remembered that the studied data referred only to 20 companies from the WIG20 index, so the scope of the sectoral analysis was limited. Most companies in the sample belonged to different sectors, according to the classification provided by the WSE. Moreover, no clear patterns were identified with respect to the market capitalisation of the companies analysed in this study. A more detailed analysis would require a broader set of companies.

Finally, the static model presented here may be embedded in dynamic models. Such attempts have already been reported in the literature, for example in Li et al. (2023), where static distributional components were combined with dynamic mechanisms. Extending the analysis in this direction in future research might also be worthwhile.

5. Conclusions

The comparative analysis indicates that the observed differences in trade duration distributions are driven by both market-specific factors and truncation choices. At higher left truncation levels, the lognormal distribution is preferred regardless of whether left-truncated or doubly truncated samples are examined. This finding distinguishes the WSE from the London Stock Exchange, where a consistent preference for the Weibull distribution was shown by previous studies.

For smaller values of the left truncation point, the results depend on whether left or double truncation is applied. In the case of the left truncation, the Weibull distribution provides the best fit, whereas in the case of the double truncation, the Weibull and lognormal distributions provide a similar fit.

Overall, the passing rates of the goodness-of-fit test are generally higher for doubly truncated samples than for left-truncated ones, suggesting that explicitly accounting for the natural upper bound imposed by the trading session length may lead to a more appropriate modelling of trade durations. These conclusions should be interpreted in the light of the data and modelling limitations discussed in the preceding section.

References

Cohen, A. C. (1991). *Truncated and Censored Samples. Theory and Applications.* CRC Press. https://doi.org/10.1201/b16946.

Čížek, P., Härdle, W. K., & Weron, R. (Eds.). (2005). *Statistical Tools for Finance and Insurance.* Springer. https://doi.org/10.1007/b139025.

Doman, M. (2011). *Mikrostruktura giełd papierów wartościowych.* Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Engle, R. F., & Russell, J. R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica*, *66*(5), 1127–1162. https://doi.org/10.2307/2999632.

Haas, M., & Pigorsch, C. (2009). Financial Economics, Fat-Tailed Distributions. In M. A. Meyers (Ed.), *Encyclopedia of Complexity and Systems Science* (pp. 3404–3435). Springer. https://doi.org/10.1007/978-0-387-30440-3_204.

Kızılersü, A., Kreer, M., Thomas, A. W., & Feindt, M. (2016). Universal behaviour in the stock market: Time dynamics of the electronic orderbook. *Physics Letters A*, *380*(33), 2501–2512. https://doi.org/10.1016/J.PHYSLETA.2016.05.035.

Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2008). *Loss Models: From Data to Decisions*. John Wiley & Sons. https://doi.org/10.1002/9780470391341.

Kreer, M., Kızılersü, A., & Thomas, A. W. (2022). Censored expectation maximization algorithm for mixtures: Application to intertrade waiting times. *Physica A: Statistical Mechanics and its Applications*, *587*, 1–11. https://doi.org/10.1016/J.PHYSA.2021.126456.

Krysicki, W., Bartos, J., Dyczka, W., Królikowska, K., & Wasilewski, M. (2004). *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach: część 1. Rachunek prawdopodobieństwa.* Wydawnictwo Naukowe PWN.

Krzyśko, M. (2004). *Statystyka matematyczna*. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.

Li, Z., Chen, X., & Xing, H. (2023). A multifactor regime-switching model for inter-trade durations in the high-frequency limit order market. *Economic Modelling*, *118*, 1–42. https://doi.org/10.1016/J.ECONMOD.2022.106082.

Ni, X.-H., Jiang, Z.-Q., Gu, G.-F., Ren, F., Chen, W., & Zhou, W.-X. (2010). Scaling and memory in the non-Poisson process of limit order cancelation. *Physica A: Statistical Mechanics and its Applications*, *389*(14), 2751–2761. https://doi.org/10.1016/J.PHYSA.2010.02.040.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, *116*(2), 257–270. https://doi.org/10.1016/J.JFINECO.2015.01.003.

Pewsey, A. (2018). Parametric bootstrap edf-based goodness-of-fit testing for sinh-arcsinh distributions. *TEST*, *27*(1), 147–172. https://doi.org/10.1007/S11749-017-0538-2.