

UNIVERSITY OF TARTU

Faculty of Science and Technology

Institute of Mathematics and Statistics

Gegham Fahradyan

**Statistical Testing of Claims Related to High-Frequency Stock Market Data**

Actuarial and Financial Engineering

Master's Thesis (30 ECTS)

Supervisor: Assoc. Prof. Raul Kangro

Tartu 2020

## **Päevasiseste aktsiahinna muutuste kohta käivate väidete statistiline testimine**

**Lühikokkuvõte.** Internetist on võimalik leida mitmeid väiteid päevasiseste aktsiahinna muutuste mustrite kohta. Magistritöö eesmärgiks formuleerida mõned sellised väited matemaatiliselt ning seejärel kontrollida näiteandmestike põhjal nende paikapidavust. Kontrollimine koosneb kahest etapist, millest esimeses rakendatakse matemaatilise formuleeringust tulenevaid reegleid selleks, et teha iga kauplemispäeva jaoks kindlaks, kas vaadeldaval päeval väide kehtis uuritava finantsinstrumendi korral. Seejärel rakendatakse statistilisi teste esimese etapi tulemustele, et uuritav väide lõplikult kinnitada või ümber lükata. Kinnitust leitud väidete korral viiakse läbi kauplemise simulatsioone, et uurida võimalust leitud seaduspära kasutamise abil aktsiaturul kasumit teenida. Magistritöö peamiseks väärtuseks võib lugeda aktsiaturu kohta käivate väidete teadusliku testimise protsessi etappide ja valikukohtade demonstratsiooni konkreetsete väidete analüüsi näitel.

**Märksõnad:** Aktsiaturu hindade käitumismustrid, kõrgsageduslikud päevasised aktsiaandmed, statistiline testimine, kauplemine, simulatsioonid.

**CERCS teaduseriala:** P160 Statistika, operatsioonianalüüs, programmeerimine, finantsmatemaatika

## **Statistical Testing of Claims Related to High-Frequency Stock Market Data**

**Abstract.** Nowadays one can find a number of sources, presenting their views on possible stock price behavior patterns. The objective of this thesis is to interpret some claims regarding stock price fluctuations into mathematical formulations. The latter are further tested through relevant statistical tests in two stages. The first stage includes running the test on the dataset of one trading day per financial instrument. The second stage considers running a statistical test on the outcomes of the first stage in order to finally approve or reject the claim. Trading simulations will be run through the datasets of confirmed claims. The main value of the thesis is the demonstration of various possibilities to interpret the imprecise claims into mathematical formulations and then verify them by scientific approach.

**Keywords:** patterns of stock price behavior, high-frequency intra-day stock data, statistical testing, trading, simulations.

**CERS research specialization:** P160 Statistics, operations research, programming, actuarial mathematics.

# CONTENTS

CHAPTER 1. INTRODUCTION .....	4
1.1 LITERATURE REVIEW .....	6
1.2 DATA DESCRIPTION .....	6
CHAPTER 2. CLAIM 1 (08:00-08:15) .....	7
2.1 CLAIM STATEMENT .....	7
2.2 MATHEMATICAL FORMULATION .....	8
2.3 EMPIRICAL RESULTS AND TRADE SIMULATIONS .....	15
CHAPTER 3. CLAIM 2 (08:15-08:30) .....	22
3.1 CLAIM STATEMENT .....	22
3.2 MATHEMATICAL FORMULATION .....	22
3.3 EMPIRICAL RESULTS AND TRADE SIMULATIONS .....	27
CHAPTER 4. CLAIM 3 (11:45-13:15) .....	29
2.1 CLAIM STATEMENT .....	29
4.2 MATHEMATICAL FORMULATION .....	29
4.3 EMPIRICAL RESULTS .....	32
CONCLUSION .....	36
APPENDIX .....	38
LIST OF REFERENCES .....	39

## CHAPTER 1. INTRODUCTION

It is widely known that besides serving in the role of intermediary between economic agents, financial intermediation is also viewed as a separate profitable sector of economy. Alongside their main goal of transforming the raw savings into investment capital, Financial Markets carry a huge interest in terms of speculative approach.

Traditionally, the engagement into financial intermediation by economic agents is accepted to divide it into two main categories. The first category considers acquisition of stocks with the aim of possessing the latters over a relatively long period of time, such as, for months or years (at least a couple of weeks) with the main purpose of obtaining a capital gain due to the difference in purchasing and selling prices. Secondary purposes include receiving dividend payments throughout the time of owning the stocks, a means of savings and etc. Let us call the above mentioned category as “Investing”.

The second method carries a rather speculative character (trading). In this paper we are going to mainly observe a special case of tradable instruments CDFs. The latter considers not actually purchasing and possessing the stocks, but comprising a tradable contract between the trader and the financial institution (a broker) with the aim of exchanging the difference in the current value of the financial instrument and the value of the latter at the end of the contract period. The above mentioned and described tradable contract is accepted to call “Contract for Differences” (CDF)<sup>1</sup>. The object of the CDF can be a stock, index, commodity or a currency tradable in the stock market. In this paper the CDFs are in the center of our interest.

In contemporary world, due to a large variety of online trading platforms, one can easily get access to financial trading. Some of the worldwide popular platforms are “Fidelity Investments”, “E\*TRADE”, “PLUS500” and etc<sup>2 3</sup>. The main tradable instruments on the

---

<sup>1</sup> Investopedia. (2020). “*Contract For Difference*”. Jan 12

[Online] Available at: <https://www.investopedia.com/terms/c/contractfordifferences.asp>

<sup>2</sup> Investopedia. (2020). “*Find and compare the best online trading platforms for every kind of investor*”. Mar 31

[Online] Available at: <https://www.investopedia.com/best-web-trading-platforms-4587883>

<sup>3</sup> Plus500

platforms are above described CDFs. They not only enable one to trade in real time, but also provide opportunity to open short positions, in this way, earning also on instrument devaluation. Given all these opportunities to trade without obstacles (sometimes even without commission fee), valid predictions regarding even short time future price behavior can bring the trader considerable gain.

**The hypothesis:** Throughout the historical data of financial instruments' prices, it is possible to discover patterns (at least taking place for the time being), which tend to occur at some significant probability.

**The purpose** of the paper is to verify some claims, found on open internet sources, about repeating patterns of stock price movements or characteristics of price changes, such as volatility, daily high or low price etc. The verification will be supported by mathematical formulations, statistical tests and historical intra-day stock market data. If a pattern is confirmed, possibilities of using the pattern to earn money by trading according to a proposed trading strategy are investigated by trade simulations

In order to achieve the above mentioned purpose, the below mentioned problems with the following succession are addressed:

- a) Defining the main concepts and the framework of the research,
- b) Separately presenting the initial pattern suggestions (claims) found on the open internet sources, which are yet to be verified,
- c) Verifying the pattern suggestions on the available data, based on mathematical formulation of the claims and suitable statistical tests, particularly, 1) firstly, clear decision rules are developed for deciding for each trading day if the claim was valid for that particular day 2) after collecting data about days when the claim was valid and when it was not valid, statistical tests are used for deciding if the probability of claim being valid for a trading day is higher than  $\frac{1}{2}$ .
- d) Segregating the confirmed claims. Defining trade strategies based on those claims and simulating trade deals according to verified patterns on the historical price data in order to prove feasible perspective earnings,
- e) Presenting the results and drawing conclusions.

---

## 1.1 LITERATURE REVIEW

Although a number of researchers claim that historical stock price data contains no additional information that can help to predict some future values with scientific approach<sup>4</sup>, some other sources suggest, that there are certain tendencies depending on the hour of the day, which tend to occur more often than not<sup>5</sup>. The patterns include tendencies, such as, distinct price change direction upwards or downwards over a certain time of the day, increased or decreased volatilities, activeness or passiveness by the traders, daily lowest or highest price record falling in a given hour frame, price direction depending on the earlier price behavior and etc.

---

## 1.2 DATA DESCRIPTION

A high-frequency intra-day stock data is used with real-time quotes. Each file of data represents the trade / quote details of one certain stock for one trading day. A data file is a table of 77 columns, which describe the changes in the order book. Given the problems studied in the thesis, only those quotes are used where the trade took place. The total datasets for each stock correspond to 14 trading days, and the time range is between 8:00 and 16:30. The extract of the main columns, used in the paper can be found in the table below (see *Table 1*):

Date	Symbol	Time	Price	Bid1	Ask1
20.03.2007	AZN.L	08:00:55.921	2849	2849	2859
20.03.2007	AZN.L	08:00:56.181	2849	2849	2859
20.03.2007	AZN.L	08:02:03.237	2849	2849	2854

**Table 1. Data Extract**

In the next chapters, the pattern suggestions are presented, broken down by the time of the day. The dataset is in GMT time, and the time frames of the tendencies are approximate. Since the data contains timestamps from both summer and winter times, 1 hour is added during the summer time to have the same trading period for all the days

---

<sup>4</sup> Sharpe William. (1998). “*Investment*”.(6<sup>th</sup> Edition)

<sup>5</sup> The Balance. (2019). “*Common Intra-Day Stock Market Patterns*”. Nov 8

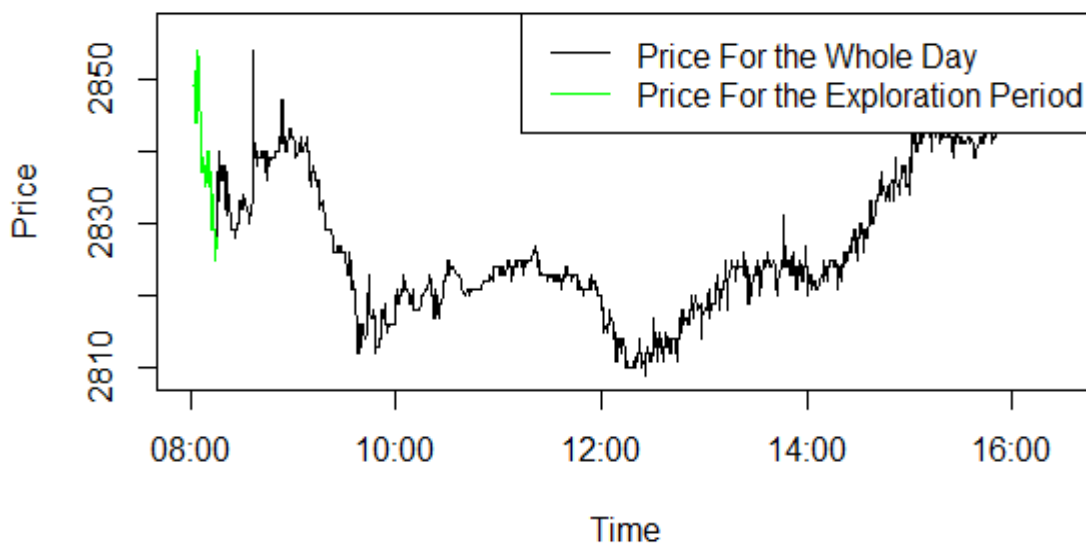
[Online] Available at: <https://www.thebalance.com/common-intra-day-stock-market-patterns-1031456>

2.1 CLAIM STATEMENT

The logic behind the claim suggests that perhaps at the beginning of a trading day some new information has arrived to market participants and therefore the trades may force the price to move in one direction for some short period of time, approximately 15 minutes, until the current price starts to reflect the actual equilibrium.

*Following the opening of the stock market, prices of financial instruments tend to move mainly in one direction for the first 15 minutes. Perhaps, it may take a couple of minutes, until the trend starts to form<sup>6</sup>.*

**Example of an "Initial Push in One Direction"**

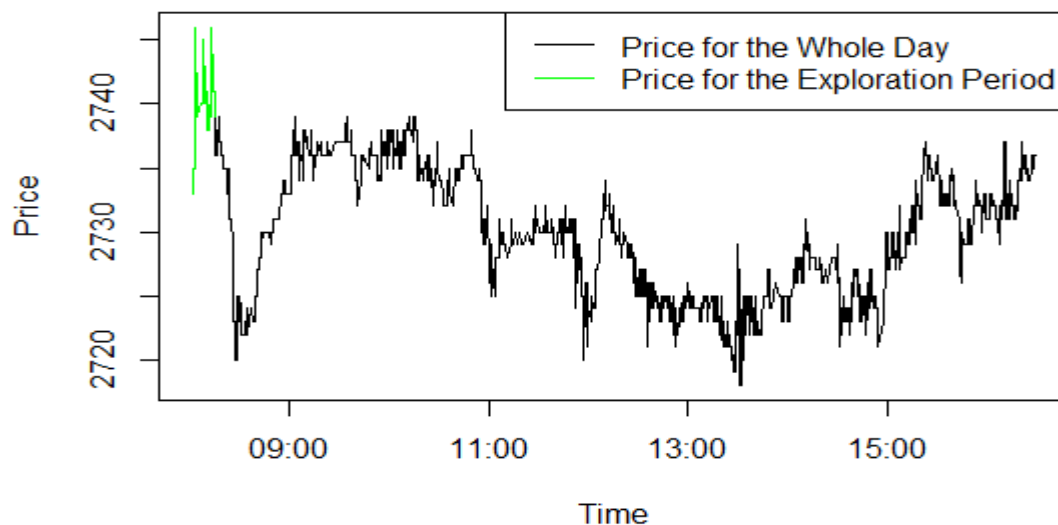


**Graph 2.1 (a). Visualization on Claim 1. The Complete Picture.** *Following opening the stock market, the prices move mainly in one direction. The claim is valid*

<sup>6</sup>The Balance. (2019). "Common Intra-Day Stock Market Patterns". Nov 8

[Online] Available at: <https://www.thebalance.com/common-intra-day-stock-market-patterns-1031456>

### Example of "Claim 1" - Invalid



**Graph 2.1 (b). Visualization on Claim 1. The Complete Picture.** *There is no clear unique moving direction during the exploration period. The claim is invalid.*

As mentioned before, the current claim refers to the first 15 minutes at the beginning of the trading day. The graph shows a situation where the claim is valid (see *Graph 2.1 (a)*). One can notice that the prices for the duration of the claim, colored in green, are pushed downwards. On the other hand the *Graph 2.1 (b)* illustrates a situation, where during the first 15 minutes prices change the movement direction several times and there is no clear direction. In the next chapters, where the mathematical approaches are described defining whether the claim is valid or not, more cases are explored, where the outcome is not very clear.

---

## 2.2 MATHEMATICAL FORMULATION

Before we present the actual formation of the problem, it is worth stating the framework of the problem.

The “rule” considers a distinct push in one direction of the price values, which correspond to the time moments of the given period. To test the above-mentioned suggestion, firstly, we separate the 15-minutes time span into three parts. We ignore the first part, which corresponds to the first 3 minutes, since as it is said in the claim formation it may take a couple of minutes until the trend starts to form. The next three minutes (3-6 min) help us

understand in which direction the price push is. We take  $l$  as the total number of trades during this period.  $l_1$  is the number of trades corresponding the first 10% of trades in the second interval of the given period, specifically:

$$l_1 = \max(l \cdot 0.1, 1)$$

Next, we compute  $P_1$  and  $P_2$  which are mean price values corresponding to trade numbers of  $(1: l_1)$  and  $(l-l_1+1:l)$ :

$$P_1 = \frac{1}{l_1} \sum_{i=1}^{l_1} Price_i$$

$$P_2 = \frac{1}{l_1} \sum_{i=l-l_1+1}^l Price_i$$

We compute the average prices of groups, instead of single values, to exclude the effects of recorded single extreme prices. Therefore, taking into account that the initial claim considered main push in one direction, we can assume if the mean of the second group is a bigger value, then our trend was supposed to be increasing and vice versa. However, there is one delicacy we should consider: buying takes place at the best ask price and selling takes place at the best bid which is lower than the best ask price, so even when the best buying and selling prices remain the same for the whole period, the average trade prices for the groups may be slightly different depending on the number of buys and sells in the corresponding trade groups. So to rule out situations when price has not actually moved in any definite direction, the claim is considered only at days when the price difference of the groups is larger than one tick .

So, we can proceed to the actual assessment of the claim, exploring the 3rd part of the given period, which is 06-15 min. During this period we consider prices' differences of trades separated by  $k-1$  observations, where  $k$  is chosen such that:

*$m =$  the number of trades in the period*

$$k = \max([0.05 \cdot m], 6),$$

where  $[x]$  is denoted as the closest integer to  $x$ .

The choice of  $k$  is justified with the idea, that the intervals should be small enough to reflect the actual tendency of the movements of the prices, yet not too small, so as it is not unnecessarily affected by the small instant fluctuations, which do not actually form the main trend. Therefore, the interval size may be later adjusted to reflect the above mentioned phenomenon depending on the data size. If we do not give a minimum value of  $k$ , in case of small data the price changes might be computed with too small intervals, or even with successive values. Next, we define  $Y$  as the vector of price changes of trades separated by  $k-1$  observations. In mathematical terms, the values of  $Y$  are, for:  $i = 1, 2, \dots, m/k$ , computed as follows:

$$Y_i = price_{j_1} - price_{j_2},$$

where

$$j_1 = i \cdot k + 1,$$

$$j_2 = (i - 1) \cdot k + 1.$$

As a result, for example, given,  $m = 82$  and  $k = \max([0.05 \cdot m], 6)$ , our array  $Y$  consists of the following 13 values:

$$Y_1 = price_7 - price_1,$$

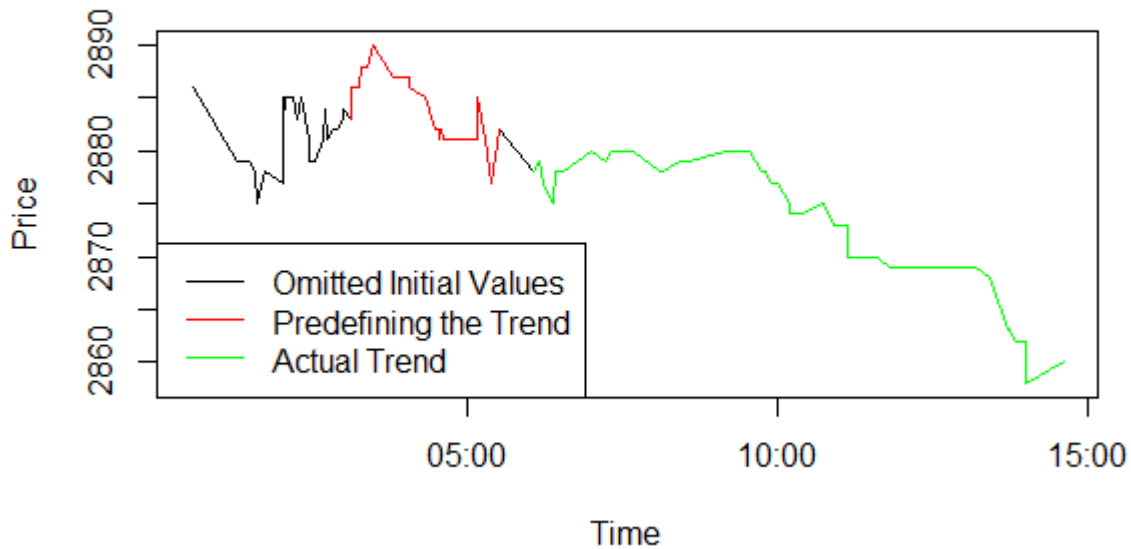
$$Y_2 = price_{13} - price_7,$$

...

$$Y_{13} = price_{79} - price_{73}.$$

If the sign of price change corresponds to the direction of price change determined by the data in the second period, we can say that the value of  $Y$  corresponds to the claim, otherwise it contradicts the claim. Let us define a new binomial variable  $X_i$  which takes values 1 and 0 when the  $Y_i$  has respectively correct or opposite direction. We ignore the cases when there is no change in the price.

## Example of "Claim 1"



**Graph 2.2. Visualization on Claim 1. Period Breakdown.** *The price behavior corresponding to the period marked in red helps us predefine the expectations for the future trend. The trend itself is marked in green.*

Currently our problem is to statistically prove that the probability of price movement in the direction of the initial push is more than 0.5. In case it is proven, we can say that the claim works for the specific day being tested. For the solution of the above-mentioned problem a one-tailed test of population proportion is going to be used. The null hypothesis of the test we express as:  $H_0: p \leq p_0$ , and as an alternative hypothesis  $H_a: p > p_0$ . In our problem,  $p_0 = 0.5$ , that is, hypothetically we assume as the null hypothesis that the probability of the price movements in the direction of the initially assumed trend is not large than the probability of movements in the opposite direction. Consequently, the initial claim will be considered verified in the case when the hypothesis is rejected with confidence level of 0.8<sup>7</sup>. The confidence level is not chosen very high, as the aim of the test is to assess the possibility of earning money on stock price fluctuations, so, for a trader it is not necessary to

---

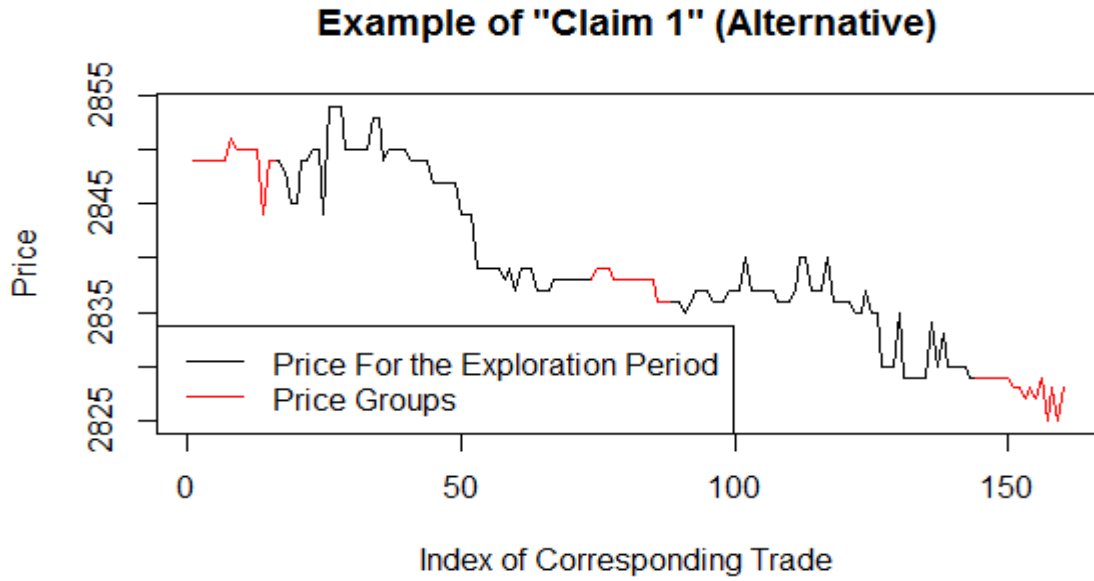
<sup>7</sup> The Pennsylvania State University. (2018). "Introduction to Applied Statistics". Department of Statistics Online Programs

[Online] Available at: <https://newonlinecourses.science.psu.edu/stat800/node/45/>

wait until exactly the end of the observed period, but one can end the position in desired moment, in this way, in practice omitting some of the theoretically observed values.

However, this method may sometimes be problematic. For example, firstly, it is up to our judgment which  $k$  value we choose, whether it is static, depends on the size of price values or the variance. Therefore, the outcome of the exploration is dependent on our choice of interval. Second, let us assume that we have such a situation when in the given period the price has increased insignificantly several times, but mostly it has been static and no decreases in stock price. If the up movements are enough to make the statistical test significant, we will have a situation that according to our test there has been a distinct push in one direction, however the price has mostly been near one value in reality. And last but not least, in order for this method to be efficient, we need a time interval with sufficiently densely filled data (enough number of trades within first 15 minutes) in order to make the statistical test significant.

In this terms, we are going to introduce an alternative way of mathematical formation of the *claim 1*. The method is not that thorough as the previous one, however it is much easier to implement. We take 3 different groups of price values from the given period at the beginning (it considers taking not from 8:00 but from 8:03 as it is said it may take a couple of minutes for the main trend to form), at the middle and at the end. The quantities are equal to rounded 10% of the number of price values. Number of values is taken instead of an interval of time, as it is possible to have too few values at the beginning of the period due to insufficient activeness of the stock market at the very beginning of the day. Further on, we calculate the means of the price groups and compare them. The groups are explored instead of single values as elements of comparison to exclude the effects of recorded single extreme prices (see *Graph 2.3*). The figure shows the stock price for the first 15 minutes on Y axis, and the number of trades for the period indexed by order on X axis. The parts colored in red represent price groups at the beginning, at the middle and at the end of the exploration period.



**Graph 2.3. Visualization on Claim 1. Alternative Approach.** Means of price groups, colored in red, are used to determine whether there is a trend or not.

Denote  $P_1, P_2$  and  $P_3$  are respectively the means of price groups at the beginning, at the middle and at the end of the given period. Also, denote  $[x]$  as the closest integer to  $x$ . If we take  $m$  as the total number of price values in the given interval, then  $P_1, P_2$  and  $P_3$  respectively correspond to the following indexes of numbers of trades:  $(1: l_1), (l_21: l_22)$  and  $(l_3:m)$ , where

$$l_1 = [m \cdot 0.1]$$

$$l_{21} = [m \cdot 0.46]$$

$$l_{22} = [m \cdot 0.55]$$

$$l_3 = [m \cdot 0.9]$$

Additionally, we again introduce the variable  $min\_tick$  which is the minimum possible change in the price according to the features of a certain stock. So, our test, which has 3 possible outcomes, is defined as follows:

- Positive, if all of the below mentioned take place
  - a.  $(P_2 - P_1) \cdot (P_3 - P_2) > 0$
  - b.  $|(P_2 - P_1)| > min\_tick$

- c.  $|(P_3 - P_2)| > \text{min\_tick}$
- Insignificant, if  $|(P_2 - P_1)| < \text{min\_tick}$
- Negative, otherwise

By using the rule 2 – *claim 1 alternative approach*, we determine one of the three possible options for each stock.

### **Defining the Claim Validity**

After we have defined the framework of validating the claim for each separate day, we can proceed to introduce the formulations of determining whether the claim on the whole is valid or not. Our available data consists of 14 trading days for each stock. Here we introduce a new hypothesis, where the inputs are the actual outcomes of the tests for each separate day of one stock. The possible values can be positive, negative and insignificant depending on the price behavior and trade frequency on the given day. Let's define a vector  $\chi_i$ , which can have values 1, -1 and 0 respectively corresponding the values of test outcomes for each day ( $i = 1, 2, \dots, 14$ ). To determine the probability of claim being valid we ignore the days where the direction of movement cannot be determined.

Denote:

$$n = \sum_{i=1}^{14} |\chi_i|, \text{the size of the sample}$$

$n_s$

= number of valid days (theoretically with distribution  $\text{Bin}(n, p)$  with unknown  $p$ )

$H_0: p \leq 0.5$ , The numbers of valid and invalid claims do differ

$H_a: p > 0.5$  The probability of claim being valid is higher,

than the probability of it to be invalid,

$$\alpha = 0.2.$$

As we expect claims for more days to be valid than invalid, we can afford to use the one-sided proportion test. The claim is considered verified on the whole, in case the p-value of the above mentioned statistical test is less or equal than 0.2.

---

## 2.3 EMPIRICAL RESULTS AND TRADE SIMULATIONS

In the given section, the empirical explorations will be carried out and the results presented. Further on, a statistical test will be carried out on the outcome of the claims, with the aim of finding out whether the number of valid claims for each day significantly exceeds the number of invalid claims.

In order for each stock dataset to be suitable for the analysis for claim 1, 3 criteria are checked. The datasets are required to satisfy all the below mentioned points during the timestamp 8:03-8:15:

- the timestamps of transactions are spread across at least 8 different minutes,
- at least 30 transactions are made during the period,
- the timestamps of transactions during 8:03-8:06 period are spread across at least 2 different minutes.

This is important to make sure that the period has enough number of trades, at the same time to assure that the trades are enough spread across the period to check for the trend. The last bullet point reasserts that the initial period of 8:03-8:06 is well covered across time, since that period's trade data is responsible for future trend direction assumption. As a result, from the available datasets 17 suitable stocks' data are chosen.

For the claims, which are proven valid according to the statistical testing, trade simulations will be carried out in order to test whether it would have been possible to earn money through trading via the strategies suggested by our claims

The simulation considers:

1. Following the stock movement direction from the first price average to the second price average, which are described in the mathematical formation of the claim.
2. In case the difference of the above mentioned values is greater than the minimum price tick for a given stock, one should short sell with the aim of further buying, and vice versa (profit might be adjusted to reflect possible commissions). The starting point is considered to be the first ask/bid value after the last value based on which the second price average is computed.
3. In case the difference is less than the minimum tick, no trade is suggested to make for the day concerning the given claim.

With the aim of illustrative result presentation, the details of the statistical tests will be listed on one stock data: *AstraZeneca plc* (AZN). The outcome is placed on *table 2*:

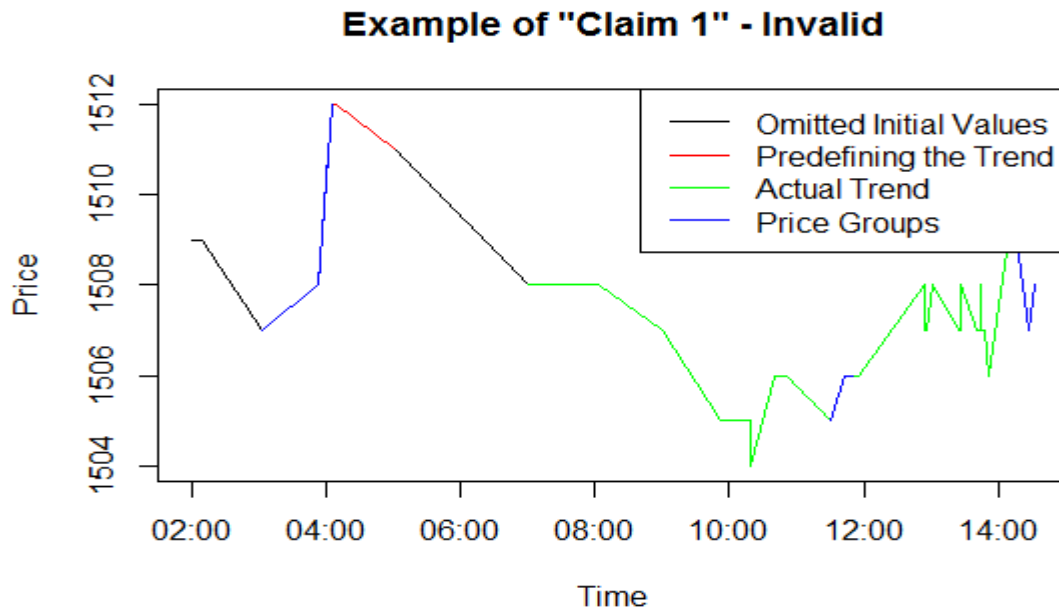
Date	Satisfying	Profit(\$), commission = 0\$	Profit(\$), commission = 1\$
March 19	Yes	13	12
March 20	Yes	8	7
March 21	N/A	0 (no investment)	0 (no investment)
March 22	Yes	3	2
March 23	No	-3	-4
March 26	Yes	3	2
March 27	Yes	1	0
March 28	No	-11	-12
March 29	No	-3	-4
March 30	Yes	1	0
April 02	Yes	4	3
April 03	No	-1	-2
April 04	Yes	-1	-2
April 05	Yes	0	-1
<b>P-Value</b>	<b>0.1336</b>	<b>Profit = 14\$</b>	<b>Profit = 1\$</b>

**Table 2. Empirical Results for “Claim 1 – Alternative Approach”. AZN**

As a result, we have 9 days when the suggested claim proves to be valid and 4 days when the claim is not valid and one day, when the difference of the starting and ending values for the first interval did not exceed the minimum tick, consequently, the data for the day is omitted. As the probability of having such outcome when actual probability of claim being valid is not larger than 0.5 is quite small, we conclude that the claim is valid more often than not and we could try to use it for trading decisions. As we do not have more data to carry out a trade simulation on an independent set of days, we use the same data for simulations. The outcome is shown on *table 2*.

Conclusion on illustrative results: one can notice that by following the suggestions of the given claim, it would have been possible to earn on stock price fluctuations.

On the contrary, we have a number of stocks, where the claim proved to be invalid. Let us carry out some exploratory analysis on some of those. Amongst the invalid claims, it is worth segregating them into two groups: first group would include those trading days, where the price behavior obviously contradicts the suggested claim, after being tested with all presented techniques (see *graph 2.4*), while the second group would be those trading days, which have outcome depending on the selected approach (see *graph 2.5* and *graph 2.6*).

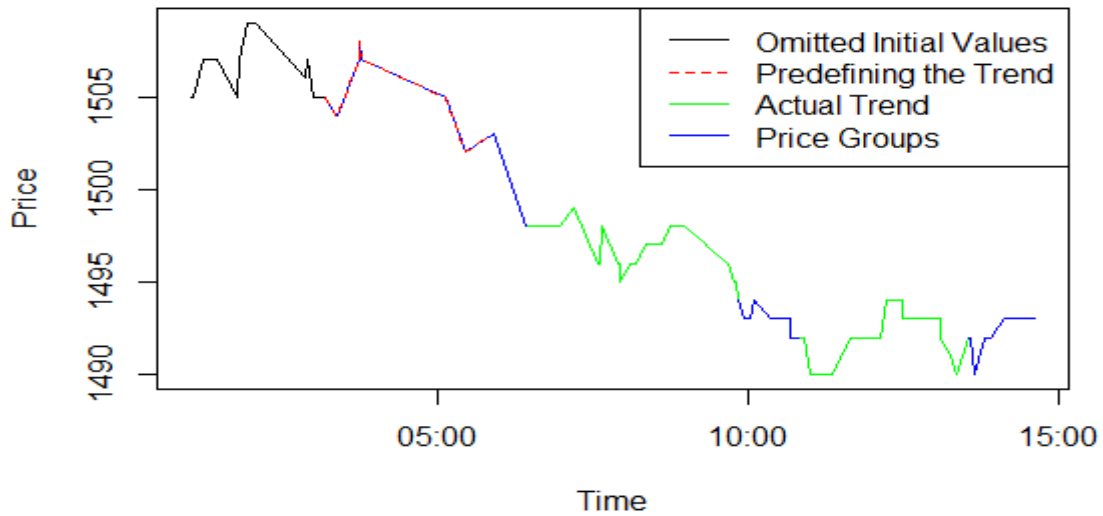


**Graph 2.4. Visualization on Claim 1.** *The claim is invalid regardless of the approach*

Let us consider the following stock datasets: *Unilever plc* (ULVR). According to claim 1, the dataset contains 12 trading days suitable for our test, out of which 2 have positive results and 10 negative, resulting in overall p-value being equal 0.98. Yet, with the alternative approach we have equally 6 positive and negative daily results, resulting in p-value being equal 0.5. Eventually, the overall p-value is not satisfactory and we will reject the claim, however, let us visually explore to see the differences.

As earlier discussed and an example shown on *Graph 2.4*, for some trading days, the test outcome is negative independent of the approach, including by visual inspection, since there is no distinct price push in one direction. However, there are trading days, where depending on what approach we choose, we can have different results. For example, let us look at the price fluctuations on March 20 (see *graph 2.5*).

### Example of "Claim 1" - Indefinite

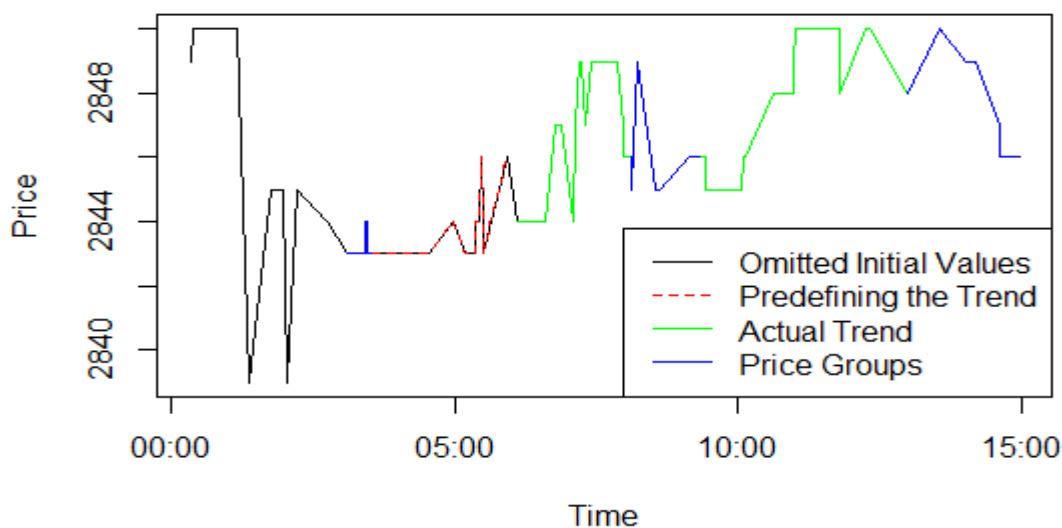


**Graph 2.5. Visualization on Claim 1.** *Claim validity depends on the employed approach. It is valid according to the first approach, but invalid according to the alternative one.*

From first sight, it seems feasible that the claim may be valid. While applying the first approach on claim 1, we get 7 stock price downwards movements out of total 10, which results in a valid claim with p-value being equal 0.17. On the contrary, while applying the alternative approach, we get the following price group means: 1504.75, 1493 and 1492.154 respectively for the beginning of the period (starting from 8:03) for the middle and for the end. The minimum price tick for current stock is 1\$, hence, one can notice that the 3<sup>rd</sup> price group average is not significantly less than the second group, as opposed to expected outcome, since the second price group average is considerably less than the first one.

It is also possible to have the opposite picture in regards with the approach of mathematical formulation (see *graph 2.6*).

### Example of "Claim 1" - Indefinite

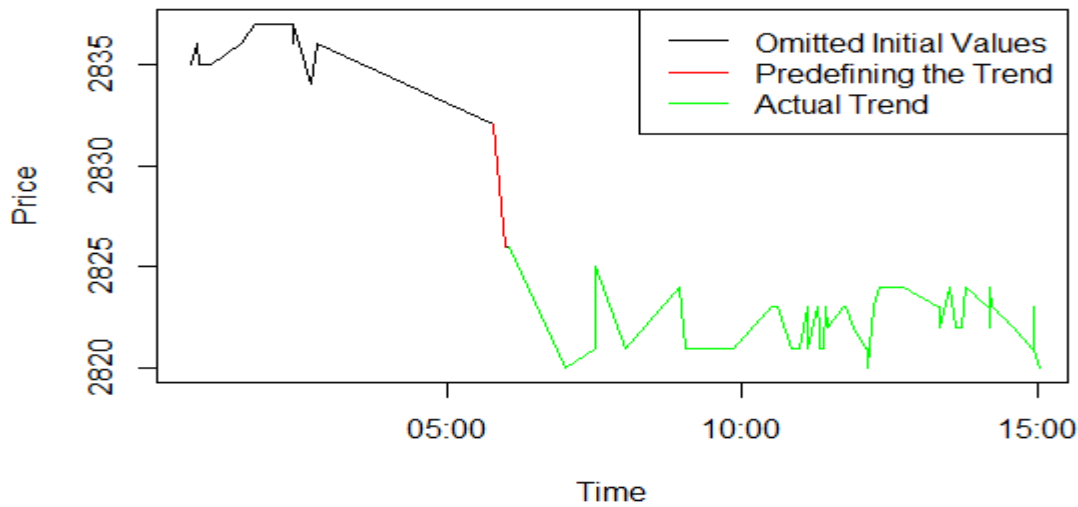


**Graph 2.6. Visualization on Claim 1.** *Claim validity depends on the employed approach. It is invalid according to the first approach, but valid according to the alternative one.*

Let us explore the datasets of the following stock: *Rio Tinto plc* (RIO). The claim in general is rejected with 11 invalid results out of 12 suitable days with p-value being equal 0.995 for claim 1 approach, and equally 6 valid and invalid results for claim 1 alternative approach. Now let us proceed to the example: for the trading day March 28, the claim 1 is rejected with p-value being equal 0.25. However, the claim 1 alternative is proven valid, since the means of the price groups are as follows: 2843.214, 2845.615 and 2848.067.

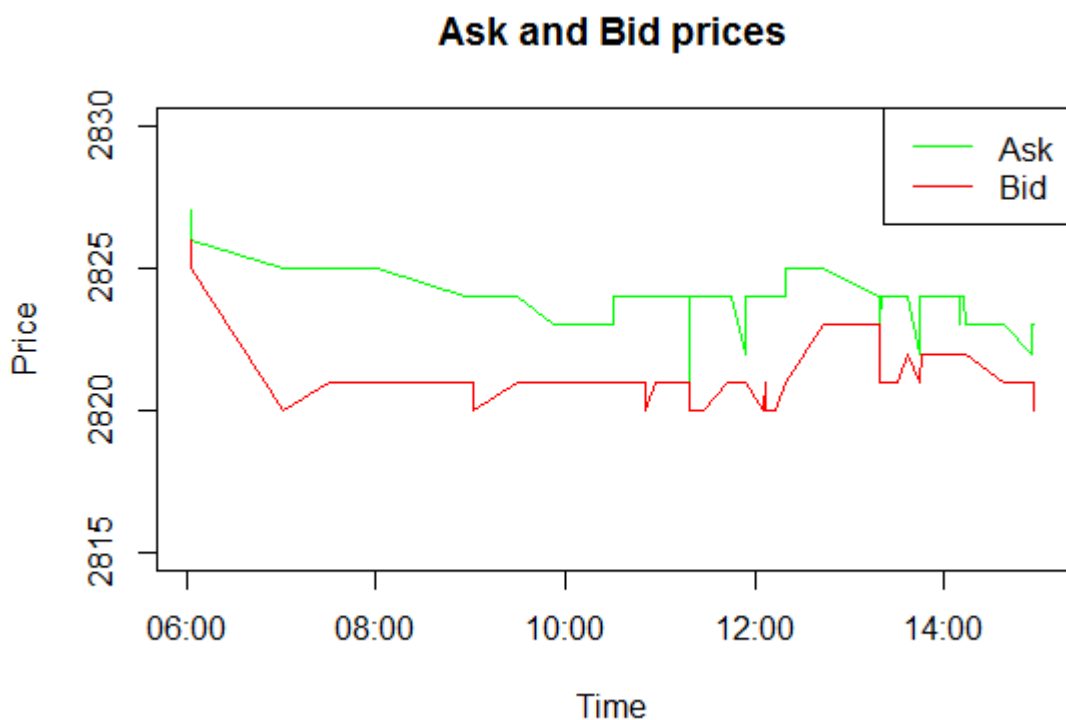
Last, but not least, let us introduce an example, where the statistical test outcome on a specific trading day depends on the  $k$  value, earlier introduced in the mathematical formulations part. The results of claim 1 for AZN with common  $k$  value equals to 6, only 1 day has a positive result out of total 11 with p-value being equal 0.99, while, in case we take  $k=8$ , the number of positive results is 3 out of 11 with p-value of 0.886. Let us visually explore one of the trading day features (March 23), where the outcome depends on the value of price differences' interval (see *graph 2.7*).

### Example of "Claim 1" - Indefinite



**Graph 2.7. Visualization on Claim 1.** Claim validity depends on  $k$ . It is valid while using the intervals obtained with  $k = 8$ , whereas it is invalid when  $k = 6$ .

The first 3 minutes, marked red on the graph, suggest us that we should expect a price push downwards. In case our  $k$  is 8, we have a positive test with  $p$ -value=0.144, as opposed to  $p$ -value=0.273 in case of  $k$  being 6. As we have different outcome depending on the  $k$ , also by visual inspection, it is unclear whether there is a price push in a certain direction, we will simulate a trading for the day. Since according to the initial move, the trend is expected to be downwards, one should short sell the instrument, which corresponds to the bid price, and then buy it back at the end of the period, which corresponds to the ask price (see *graph 2.8*). Hence, the trade simulation shows, that the strategy would have brought us 3\$ per trading share.



**Graph 2.8. Visualization on Claim 1. Trade simulation on an indefinitely proven claim.**

*The simulation results in gain*

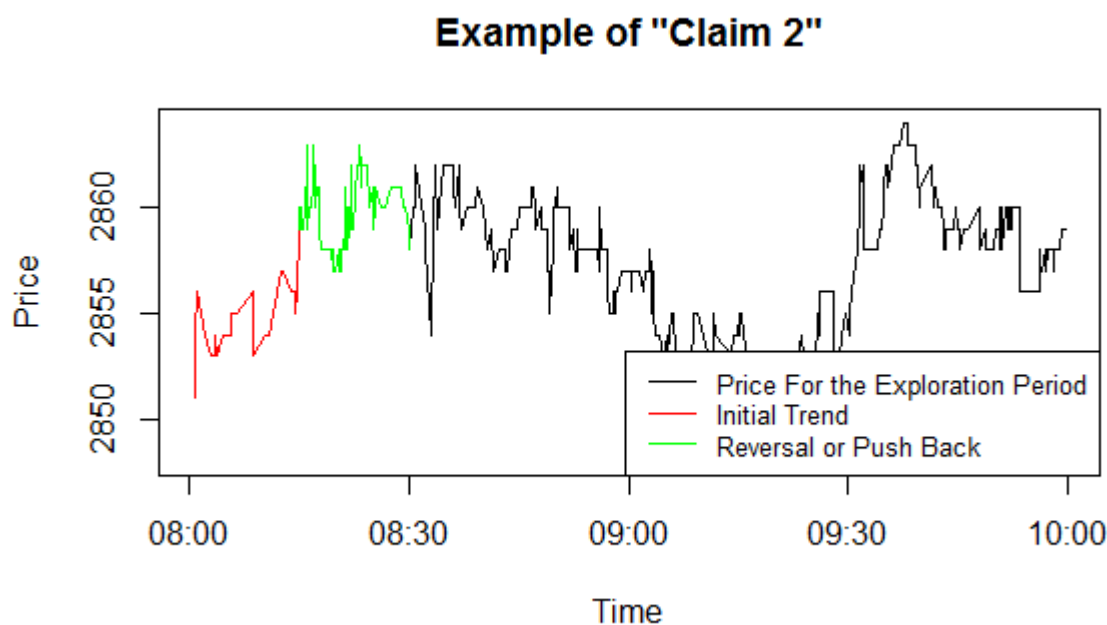
To conclude with claim 1 in general, we can firstly state that a valid claim was found: *Claim 1 – Alternative approach* is verified for AZN. In summary, we can say that one can develop a suitable technic to match the nature of a given stock price behavior, since we noticed that the outcome can vary depending on the approach to the mathematical formulation, also some specifics inside a certain formulation. Additionally, it turned out that the first rule in claim 1 is very conservative and by using it the claim was considered to be valid only for a very small number of days even when visually the claim seemed to be valid much more often

The detailed outcome of statistical tests for each stock is stored in the table *appendix 1*.

### 3.1 CLAIM STATEMENT

The claim 2 is an immediate sequel of claim 1. It is supposed that the initial push in one direction, conditioned with new information at the beginning of the trading day, does not continue after the first 15 minutes. Perhaps, a new equilibrium is achieved and traders reassess their positions.

*The initial push usually carries on circa 15 minutes and after that it is challenged: it either takes a reverse direction, or just flattens. This phenomenon perhaps is followed for another 15 minutes, followed by the original trend asserting itself<sup>8</sup>.*



**Graph 3.1. Visualization on Claim 2.** *The figure represents the initial push upwards during the first 15 minutes. As it can be seen, for the next 15 minutes (colored in green), the price does not follow the same trend.*

### 3.2 MATHEMATICAL FORMULATION

<sup>8</sup> The Balance. (2019). "Common Intra-Day Stock Market Patterns". Nov 8

[Online] Available at: <https://www.thebalance.com/common-intra-day-stock-market-patterns-1031456>

The “rule” suggests that the trend which began with the start of the trading day will not be in force after 15 minutes. The current rule will be implemented only on those stock data, which turned to prove claim 1 valid. It should be noted, since the main object of the consideration in this claim is the 2<sup>nd</sup> 15 minutes period which describes that during that period the original trend is not holding, the last part of the claim, regarding original trend re-asserting itself after the 2<sup>nd</sup> 15 min is not considered. Analogically, as in the case of claim 1, we will introduce two methods of assessing the validity of claim 2: the first method will be more thorough and carry a stronger statistical background, at the same time requiring densely filled data in order to hold statistical significance, while the second method will be easier to implement, free of heavy statistical formulas and still possible to utilize in case of scarce data.

Since the assessment of the claim 2 will be highly dependent on the outcome and procedure of the claim 1, we will use some of the earlier defined variables. As in order to prove the validity of claim 2 we need to show that the stock prices during the mentioned period have no specific moving direction or go opposite the initial trend, it is firstly necessary to recall the initial trend direction. Thus, in case we want to observe which direction the price push was during the first 15 minutes, we can find the price movement direction during the period of min 3-6, and it will be enough to say that the same trend continued till min 15, as for verifying the claim 2, we use only the verified for the 1<sup>st</sup> claim data. Further on, we proceed to using the values  $P_1$  and  $P_2$  that were already introduced and described in the Mathematical Formation part for the claim 1 (See section 2.2). At this point we have already acquired all the desired information from claim 1 and we can start specifically the assessment of the claim 2. Since again we are going to deal with assessments of trends, the procedures are going to be very much analogical.

Firstly, a  $k$  value is defined, which equals 5% of the number of the price values (the same as number of trades during the mentioned period) in the second 15 minutes (denote  $m$ ), rounded to an integer value, yet we define that  $k$  cannot be less than 6. The choice of  $k$  is justified with the idea, that the intervals should be small enough to reflect the actual tendency of the movements of the prices, yet not too small, so as it is not unnecessarily affected by the small instant fluctuations, which do not actually form the main trend. Giving a minimum value to  $k$  is vital in order to exclude the scenarios when in case of small data the price changes might be computed with too small intervals, or even with successive values. Next,

we define  $Z$  array of values, corresponding to the price changes over consecutive intervals. So, in case we take the intervals too small,  $Z$  may be affected by possible random changes, which do not form the main trend, yet, if we take too large intervals, there may not be enough values of differences to make the further statistical test significant. To reflect both above-mentioned problems, we will define the array  $Z$  which is computed according to a  $k$  value and interval indexes  $j1$  and  $j2$  as follows:

$$Z_i = price_{j1} - price_{j2},$$

where

$$j1 = i \cdot k + 1,$$

$$j2 = (i - 1) \cdot k + 1,$$

for  $i = 1, 2, \dots, m - 1$

As a result, for example, given,  $m = 138$  and  $k = \max(0.1 \cdot m, 6) = 7$ , our array  $Z$  is consisted of the following 19 values:

$$Z_1 = price_8 - price_1,$$

$$Z_2 = price_{15} - price_8,$$

...

$$Z_{19} = price_{134} - price_{127}.$$

The result is an array of values  $Z$ , which represents values of desired changes in price movements. Further on we can proceed to distinguish the price movement directions according to change in the mathematical sign with the intent of further assessing the probability in which direction prices tend to move within the second fifteen minutes of the trading day. We denote an array  $X$  of 0-s and 1-s, where the value is 1 if the change was in the opposite direction of the original push or there was no price change and 0 if it was in the direction of the original push.

Currently, we already have all the necessary parameters for formulating and carrying out the statistical test in order to verify or reject the initially suggested claim. In order to form the claim question as a statistical test, we introduce a one-tailed population proportion test

and we say that our null hypothesis  $H_0: p \leq p_0$  and the alternative hypothesis  $H_a: p > p_0$ , that is, that is, our null hypothesis states that the probability of the price changes in the same direction as the original push is larger than the probability of moving in the opposite direction, while our alternative hypothesis suggests that the probability of price movements during the minutes 15-30 in the opposite direction as it used to hold previously is greater.

1.  $n = \text{number of values contained in array } X$
2.  $p_0 = 0.5$
3.  $\alpha = 0.2$
4.  $n_s = \text{the number of successes.}$

In the test formulas and methods of population proportion are used in classical way. Examples of formulas and implementation in R software can be obtained at the enclosed source link<sup>9</sup>.

As a conclusion we say, that in case p-value is less than 0.2, we can confirm that the number of price changes in the assumed direction is significantly less than the changes in the opposite direction or no changes. The latter results in the validity of claim 2, that is, in 15 minutes after the beginning of the trading day, the stock prices tend to lose the assumed push in one direction during the first 15 minutes, or even it may go with the opposite trend.

However, analogically with the claim one method one, the above mentioned procedure of assessing the validity of the claim has several drawbacks depending on the data nature, such as,

1. It requires a minimum level of sample size in order to make the statistical test significant
2. In case of scarce data, a larger interval size in defining the price changes may result in insignificant statistical test, while a narrower interval may reflect the unimportant small changes in stock prices which do not form the main trend.

Thereafter, we introduce an alternative way of mathematical formation of the *claim 2*. The method is not that thorough as the previous one, however it is much easier to implement

---

<sup>9</sup> Kassambara A. (2016). “One-Proportion Z-Test in R”. *Statistical tools for high-throughput data analysis*

[Online] Available at: <http://www.sthda.com/english/wiki/one-proportion-z-test-in-r>

and it is supposed to cover the drawbacks of the method 1. We take 3 different groups of price values from the given period at the beginning, at the middle and at the end. The quantities are equal to rounded 10% of the number of price values (number of trades). Number of values is taken instead of an interval of time, as it is possible to have too few values at the beginning of the period due to insufficient activeness of the stock market at the very beginning of the day. Further on, we calculate the means of the price groups and compare them. The groups are explored instead of single values as elements of comparison to exclude the effects of recorded single extreme prices.

Denote  $PM_1$ ,  $PM_2$  and  $PM_3$  are respectively the means of price groups at the beginning, at the middle and at the end of the given period. Also, denote  $[x]$  as the closest integer to  $x$ . If we take  $m$  as the total number of price values in the given interval, then  $PM_1$ ,  $PM_2$  and  $PM_3$  respectively correspond to the following indexes of numbers of trades:  $(1: l_1)$ ,  $(l_1: l_2)$  and  $(l_2:m)$ , where

$$l_1 = [m \cdot 0.1],$$

$$l_{21} = [m \cdot 0.46],$$

$$l_{22} = [m \cdot 0.55],$$

$$l_3 = [m \cdot 0.9],$$

Additionally, we are again going to use the variable `min_tick` which is the minimum possible change in the price according to the features of a certain stock. So, our test, depending on the initial trend during the first 15 minutes, is defined as follows:

a) Initial trend downwards:

$$\text{Test} = \left\{ \begin{array}{l} \text{Negative,} \quad \text{if } (PM_2 - PM_1) \cdot (PM_3 - PM_2) > 0 \\ \& \\ |(PM_2 - PM_1)| > \text{min\_tick} \\ \& \\ |(PM_3 - PM_2)| > \text{min\_tick} \\ \& \\ PM_3 < PM_1 \\ \text{Positive, otherwise} \end{array} \right.$$

b) Initial trend upwards:

$$\text{Test} = \left\{ \begin{array}{l} \text{Negative, if } (PM_2 - PM_1) \cdot (PM_3 - PM_2) > 0 \\ \& \\ |(PM_2 - PM_1)| > \text{min\_tick} \\ \& \\ |(PM_3 - PM_2)| > \text{min\_tick} \\ \& \\ PM_3 > PM_1 \\ \text{Positive, otherwise} \end{array} \right.$$

### 3.3 EMPIRICAL RESULTS AND TRADE SIMULATIONS

In this section empirical analysis will be carried out and results will be tested on the basis already described in section 3.2. It was already mentioned, that the statistical tests for claim 2 will be carried out only on the datasets which were proven valid for claim 1. As a result, we have stock AZN which was proven valid for claim 1 – alternative approach for 9 days out of 14 (p value = 0.1336). The outcome of claim 2 testing is presented on *table 3* below.

Date	Satisfying	Profit(\$), commission = 0\$
<b>March 19</b>	Yes	3
<b>March 20</b>	Yes	3
<b>March 22</b>	Yes	-14
<b>March 26</b>	Yes	-15
<b>March 27</b>	Yes	1
<b>March 30</b>	Yes	0
<b>April 02</b>	Yes	-2
<b>April 04</b>	Yes	9
<b>April 05</b>	Yes	-16
<b>P-Value</b>	<b>0.0038</b>	<b>-31</b>

**Table 3. Empirical Results for “Claim 2 – Alternative Approach”. AZN**

As one can see, according to the p-value, *Claim 2 – Alternative approach* is verified for AZN. Since the main cornerstone for claim 2 is that, if there is a trend for the first 15

minutes, it does not continue for the next 15 minutes, we can try to employ a trading strategy where we bet against the initial trend. However, as it is shown on the table, the trade simulation ends up in loss.

In conclusion, unfortunately, we do not have many stocks to verify for the claim 2, since only 1 stock dataset satisfied the statistical test for claim 1. Yet, it is worth mentioning, that although a claim may be proven valid according to the test, the negative outcomes may prevail, resulting in loss in trading.

### 2.1 CLAIM STATEMENT

Unlike the other previous claims, claim 3 can hardly be applied to build a trading strategy to earn money. The current claim carries a rather advisory character. The logic behind the claim suggests that during the lunch time, including some buffer time before and after it, the activity level in the stock market goes down, due to which price fluctuations tend to extinguish.

*During the lunch time with a little buffer the stock market is supposed to be the quietest throughout the day according to the claim. Hence, the traders should consider avoiding transactions during this period<sup>10</sup>.*

### 4.2 MATHEMATICAL FORMULATION

The expression quietest time throughout the day, supposedly, refers to the stock market being least volatile during the period. The latter can be interpreted in mathematical terms as follows: during the mentioned period the variance of the price changes is relatively smaller than during other equivalent periods.

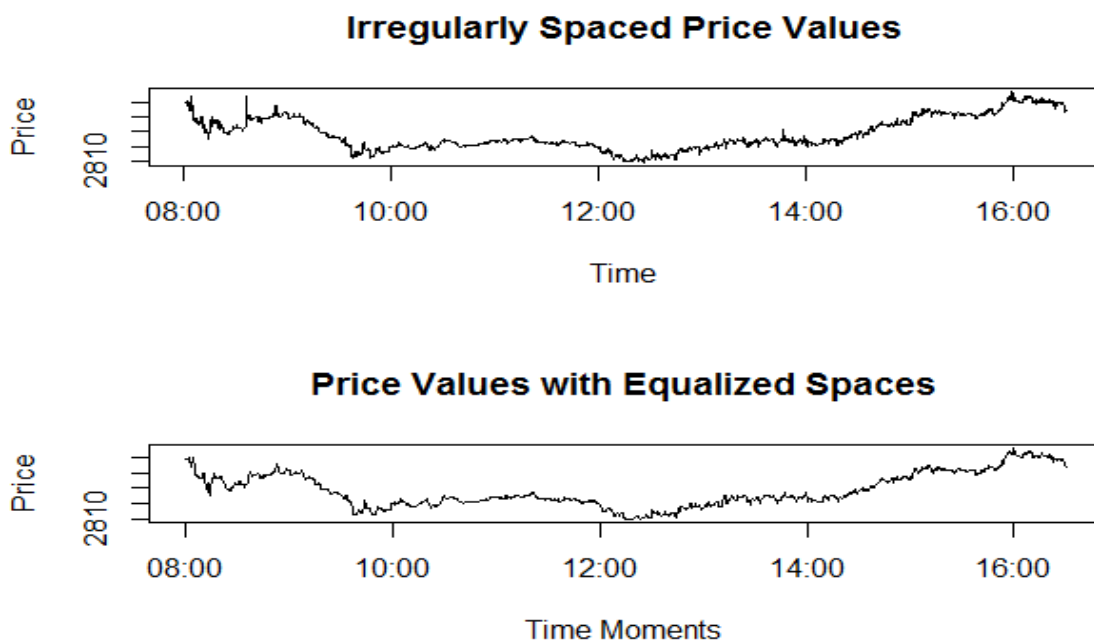
With the aim of testing the claim, we divide the time span from 8:45 till 16:15 into 90 minutes intervals. The interval length is conditioned by the length of the period we are assessing. It should be noted that we intentionally omitted the first 45 and last 15 minutes intervals in order to have equally long periods of exploration. As a result, we get 5 equal intervals to compare, where the 3<sup>rd</sup> one should have the least variance according to the suggested claim. However there is a problem in comparing the variances in current form: although the intervals are equal, the trades are not equally spaced as they are real time trades. Specifically, it is possible to have significantly varying number of values in various intervals, which will make the comparison inaccurate.

---

<sup>10</sup>The Balance. (2019). “Common Intra-Day Stock Market Patterns”. Nov 8

[Online] Available at: <https://www.thebalance.com/common-intra-day-stock-market-patterns-1031456>

To start the process of equalizing the time intervals, we firstly define what we already have, that is, let  $P_j$  be the price of  $j^{\text{th}}$  trade and  $T_j$  corresponding time of the trade, where  $j = 1, 2, \dots, m$ , the total number of real time price values in the mentioned period of time. Next we introduce a new vector  $\theta_i$ , which consists of equally spaced time moments, separated by 30 seconds, where  $i = 1, 2, \dots, n$ , where  $n = 12 \cdot 75 + 1 = 901$  (including the opening and closing time moments of the period). 30 seconds is chosen as the length of the equally spaced time interval in order to include enough number of price values to have the whole picture. At the same time, in case too narrow interval is chosen, there is possibility to have many empty intervals, where there were no trade deals in the given interval. As it can be observed on *Graph 4.1*, the general picture of the stock fluctuations are preserved after equalizing the time intervals. Yet, depending on the nature of the specific stock fluctuations and frequency of trade deals, one can adjust the time interval, taking into account the above explained conditions.



**Graph 4.1. Visualization on Claim 3. Comparison of Price Values with Irregularly and Regularly Spaced Intervals.** *As it can be concluded by visual observation, the nature of price fluctuations is not affected by regular sampling*

After we have the equally spaced time moments, now we need to adjust them corresponding price values  $\pi_i$ . The system of adjusting values works in such a way that to each of the time moments the last value of the price is attached from previous 30 seconds time interval. As a result, naturally, we will have 901 equally spaced price values for the time period from 8:45 till 16:15. It can be noted that due to this technique we omit some of the price values. Nevertheless, from mathematical point of view, it means that we do not follow the price process in continuous time but use regularly spaced sampling, which makes it easier to estimate the volatility. Moreover, our general aim is to find out whether the initially suggested claim tends to happen more often at some significant probability than not.

R Scripts for creating equally spaced “Price Vector”

```
price_out=rep(NA, length(time_moments))

for (t in time_moments){

  while(time[j] < t & j <= length(time)){

    j = j+1

  }

  price_out[i] = price[max(j-1,1)]

  i=i+1

}
```

After we have solved the issue with irregular time intervals, we can proceed to comparing the above described volatilities. As a result, each of the time periods will have  $180+1$  time moments and corresponding price values, where the last value will coincide with the beginning value of the consecutive period. Let us denote those price values according to two indexes, reflecting the sequence in periods, and the sequence in each period, such as:  $Y_{ij}$ , where,  $i = 1, 2, \dots, 5$  is the  $i^{\text{th}}$  90-minute long time period, and  $j = 1, 2, \dots, m$  is the  $j^{\text{th}}$  price value in a specific time period, and  $m$  is the number of prices in each interval (it depends on the interval lengths of the time moments).

As it is highly feasible that each consecutive stock price value is dependent on the previous price value, we will obtain the price returns with the aim of comparing the volatility, let us denote:

$$X_{ij} = (Y_{i,j+1}/Y_{ij}) - 1,$$

where

$$i = 1, 2, \dots, 5,$$

$$j = 1, 2, \dots, m - 1.$$

Therefore, the variance  $V_i$  of price returns for each of the period would be calculated as follows:

$$V_i = \frac{1}{m-1} \sum_{j=1}^{m-1} (X_{ij} - \mu_i)^2,$$

where  $\mu_i$  is the mean of X values in  $i^{\text{th}}$  period.

The claim 3 will be considered valid in case the 3<sup>rd</sup> period has the least variance, such as:

$$\min(V_i) = V_3$$

---

#### 4.3 EMPIRICAL RESULTS

As it was mentioned in the claim statement, the current claim cannot be used to earn money on stock market, yet it is rather to advise to avoid trading on the time span covered by the claim. The p-values of statistical tests for each of the stocks are stored in table *appendix 1*. The statistical tests carried out on the outcomes of each trading days are based on the principles described at the beginning of section 2.2 under the formulation part “*Defining the Claim Validity*”. Since we compare 5 different intervals and assuming the computed volatilities for all periods are random with the same distribution, the probability of the volatility of the third period to be the smallest is 1/5. Therefore, it makes sense applying  $p0 = 0.2$ . The detailed results for each stock are stored on *appendix 1*. As it can be noticed, the claim is verified for quite many stocks.

However, in order to be able to make decisions based on the claim the probability of the claim being valid should be high enough. Hence, we carry out another analogical statistical test with  $p0 = 0.5$ . This would imply that according to the hypothesis the claim is valid more often than not. In case we get stocks for which the claim is valid, we can conclude that traders had better avoid making transactions on the stock market during the period of the claim, since low volatility is less likely to create enough turbulence in the price fluctuations in order for market participants to earn on differences of ask and bid prices.

There are only 2 stocks which are very close to be valid with p-values 0.21 in case  $p0 = 0.5$ : BARC and LLOY. Below, for an illustrative purpose, the details of those stocks for each day will be presented on *table 4*.

Date	BARC / Satisfying	LLOY / Satisfying	TSCO / Satisfying
March 19	Yes	Yes	No
March 20	Yes	Yes	No
March 21	No	No	No
March 22	Yes	Yes	No
March 23	No	No	No
March 26	Yes	Yes	No
March 27	No	No	No
March 28	Yes	Yes	No
March 29	Yes	Yes	No
March 30	Yes	Yes	No
April 02	No	Yes	Yes
April 03	Yes	No	Yes
April 04	Yes	Yes	Yes
April 05	No	No	No
<b>P-Value (<math>p0 = 0.5</math>)</b>	<b>0.21</b>	<b>0.21</b>	<b>0.97</b>
<b>P-Value (<math>p0 = 0.2</math>)</b>	<b>0</b>	<b>0</b>	<b>0.5</b>

**Table 4. Empirical Results for “Claim 3”. BARC, LLOY and TSCO.** *The table shows the outcomes of the tests for each day of the mentioned stocks and p-values in accordance with the applied  $p0$  value.*

Even though in the case of  $p_0 = 0.2$  for most of the stocks the claim turns to be valid, there are some stocks with strictly invalid outcome, for example, *Tesco PLC* (TSCO). For TSCO the claim 3 results include valid outcome for 3 days out of 14 with p-value of 0.969 and 0.5 for  $p_0$  of 0.5 and 0.2 respectively. This refers to the fact that with high chance the period of our current claim is not the calmest throughout the day. The visual explorations wouldn't help us much, as we have used the price returns while calculating the variances instead of actual price values: most of the price returns are very close to 0 and visualization of those values would create a chaotic graph. Hence, we will build a table of variances per time periods, for each trading day (see *table 5*):

Date	Period 1	Period 2	Period 3	Period 4	Period 5	Min period
<b>March 19</b>	18.9553	10.3095	12.5767	29.8839	25.7132	2
<b>March 20</b>	45.7106	15.4098	25.1469	17.2359	23.6531	2
<b>March 21</b>	10.5787	13.2823	18.4923	20.0573	17.1607	1
<b>March 22</b>	20.3539	14.1694	17.0303	20.788	29.0583	2
<b>March 23</b>	38.9751	30.1067	55.2413	41.0557	30.0926	5
<b>March 26</b>	22.7578	14.4435	10.278	9.6883	53.1409	4
<b>March 27</b>	13.8244	12.4368	19.7155	18.0492	27.2844	2
<b>March 28</b>	19.4159	28.2832	25.8809	47.1878	41.7779	1
<b>March 29</b>	22.1243	9.6579	15.5515	19.4509	23.4854	2
<b>March 30</b>	29.1636	6.5914	9.5821	22.6167	39.9946	2
<b>April 02</b>	18.0616	32.2994	9.9152	14.1383	30.885	3
<b>April 03</b>	10.4802	10.8505	8.4949	13.9854	20.1648	3
<b>April 04</b>	27.3065	13.3788	9.6669	23.0054	37.8272	3
<b>April 05</b>	11.5919	9.4727	11.156	7.6800	10.1699	4

**Table 5. Claim 3. Tesco PLC (TSCO). Variances per time periods.** For the current stock, the claim is rejected, with only 3 valid outcomes out of possible 14.

In conclusion for claim 3, it is worth mentioning, that the claim statement can hardly be used as a basis for a trading strategy, yet, it can serve for possible time suggestion for traders when to avoid making transactions or when not necessarily. In the current claim we applied two approaches with  $p_0$  being equal 0.2 and 0.5. In the first case, for most of the

stocks the claim proved to be verified. It could be concluded that in those cases, the probability of the volatility of exploration period to be the smallest is more than  $1/5$ , the proportion of the length of the observed interval among the whole period considered. In the second case, only for two stocks the claim was close to be valid with p-values being equal to 0.21 while our alpha is 0.2. Should the p-value be considered acceptable, we can say that for those two stocks, the probability of the volatility of exploration period to be the smallest is more than 0.5. Hence, we can suggest traders avoiding making transactions during the claim period.

## CONCLUSION

In financial literature and internet sources, there are wide varieties of hypothesis and judgments concerning stock price behavior. Many of those assume that at some points of historical price movements, those fluctuations may carry repetitive characters at some significant probability. Perhaps, those patterns may be applied in trading strategies to earn a guaranteed return at least for the time being. At this stage, a vital issue is interpreting such claims into mathematical formulations in order to make the imprecise claims applicable on numeric datasets and form trading strategies.

In this paper, 3 of such claims are taken into consideration. As it was repeatedly mentioned and showed throughout the research, there can be various ways of implementation of one statement regarding stock price behavior. Moreover, the observation has proven that the final results of the statistical tests may differ depending on the employed techniques of interpreting the claim into numeric formulations. Nevertheless, as it may seem, this does not lead to a deadlock: a claim may not necessarily be applicable to many stocks. Moreover, some claims may reflect the actual fluctuations of some stocks, while others may be suitable for different financial instruments. Hence, by testing different approaches, one ought to find the correct mathematical formulation which matches the actual behavior of the given stock, should the trader desire to implement the strategy on real trading. To test the functionality of the claim approaches, which were proven significant due to the statistical tests, trade simulation is carried out. As a result, as shown on *table 2*, it would have been possible to earn “*risk-free*” income by applying the trade strategy described by the claim on a specific financial instrument.

Furthermore, a claim may not be necessarily applicable to build a trading strategy, yet it can be rather used to find advisable periods to trade throughout the day or avoid adverse periods. In this paper I have discovered 2 stocks’ datasets, which were compliant with such a suggestion.

In conclusion, it is worth mentioning that the cornerstone of this thesis is the interpretation of claims into mathematical formulations. In the thesis, it was proven that it is possible to find stocks, which behavior is coherent with the statements of the claims. The research may be continued in efforts to match the formulations to other periods of the trading

day, adjust mathematical formulations' parameters to better reflect the actual behavior of stock price fluctuations, or explore new claims with a new mathematical point of view.

## APPENDIX

Stocks / P values	Claim 1	Claim 1 Alternative	Claim 3 (p0=0.5)	Claim 3 (p0=0.2)
<b>AAL</b>	0.96	0.62	0.61	0.04
<b>AV</b>	1	0.62	0.79	0.13
<b>AZN</b>	0.99	0.13	0.5	0.01
<b>BARC</b>	0.96	0.5	0.21	0
<b>BG</b>	0.94	0.83	0.61	0.04
<b>BLT</b>	0.98	0.91	0.91	0.32
<b>BP</b>	0.99	0.99	0.39	0
<b>DGE</b>	1	0.81	0.61	0.04
<b>GSK</b>	0.98	0.64	0.61	0.04
<b>HBOS</b>	0.98	0.91	0.91	0.32
<b>HSBA</b>	0.99	0.99	0.5	0.01
<b>LLOY</b>	0.96	0.5	0.21	0
<b>RBS</b>	1	0.5	0.39	0
<b>RIO</b>	1	0.5	0.5	0.01
<b>TSCO</b>	0.98	0.5	0.97	0.5
<b>ULVR</b>	0.98	0.61	0.91	0.32
<b>XTA</b>	1	0.61	0.5	0.01

### Appendix 1. Empirical Results.

## LIST OF REFERENCES

- The Balance. (2019). “*Common Intra-Day Stock Market Patterns*”. Nov 8  
[Online] Available at: <https://www.thebalance.com/common-intra-day-stock-market-patterns-1031456>
- Investopedia. (2020). “*Contract For Difference*”. Jan 12  
[Online] Available at: <https://www.investopedia.com/terms/c/contractfordifferences.asp>
- Investopedia. (2020). “*Find and compare the best online trading platforms for every kind of investor*”.  
Mar 31 [Online] Available at: <https://www.investopedia.com/best-web-trading-platforms-4587883>
- Kassambara A. (2016). “*One-Proportion Z-Test in R*”. *Statistical tools for high-throughput data analysis*
- The Pennsylvania State University. (2018). “*Introduction to Applied Statistics*”. *Department of Statistics Online Programs*  
[Online] Available at: <https://newonlinecourses.science.psu.edu/stat800/node/45/>
- Plus500 [Online] Available at: <https://www.plus500.com/>
- Sharpe William. (1998). “*Investment*”. (6<sup>th</sup> Edition)

## **Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks**

Mina, Gegham Fahradyan,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose “Statistical Testing of Claims Related to High-Frequency Stock Market Data” mille juhendaja on Assoc. Prof. Raul Kangro reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commons'i litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

autori nimi Gegham Fahradyan  
**pp.kk.aaaa 20.04.2020**