Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Connectedness of high-frequency data

Author: **Bc. Petr Petras**

Supervisor: **Mgr. Tomáš Křehlík, M.A.**

Academic Year: **2015/2016**

Declaration of Authorship

The author hereby declares that he wrote this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 27, 2016

Signature

Acknowledgments

I am very grateful to my family for supporting me during whole study. And also to my supervisor who made me interested in this topic and helped me to solve problems which arose during writing this master's thesis.

Abstract

This work combines discrete and continuous methods while modeling connectedness of financial tick data. As discrete method we are using vector autoregression. For continuous domain Hawkes process is used, which is special case of point process. We found out that financial assets are connected in non-symmetrical fashion. By using two methodologies we were able to model better how are the series connected. We confirmed existence of price leader in our three stock portfolio and modeled connectedness of jumps between stocks. As conclusion we state that both methods yields important results about price nature on the market and should be used together or at least with awareness of second approach.

JEL Classification	C32, G11, G14
Keywords	Vector Autoregression, Hawkes process, High-frequency analysis, Connectedness

Author's e-mail	`petr.petras@email.cz`
Supervisor's e-mail	`krehlik@utia.cas.cz`

Abstrakt

Tato práce kombinuje diskrétní a spojité metody pro modelování propojenosti finančních tikových dat. Jako diskrétní metodologii používáme vektorovou autoregresi. Na kontinuální ose Hawkesův proces, což je speciální případ bodového procesu. Zjistili jsme, že finanční statky jsou propojeny nesymetricky. Díky použití dvou metodologií jsme byli schopni lépe modelovat propojenost těchto statků. Potvrdili jsme existenci cenového vůdce v našem portfoliu tří statků. Také jsme namodelovali propojení skoků cen mezi statky. Jako závěr práce uvádíme, že obě metody přináší důležité výsledky ohledně vývoje cen na trhu a měly by být používány dohromady nebo alespoň s vědomím existence druhé metody.

Klasifikace JEL	C32, G11, G14
Klíčová slova	Vektorová autoregrese, Hawkesovy procesy, Vysokofrekvenční analýza, Propojenost

E-mail autora	`petr.petras@email.cz`
E-mail vedoucího práce	`krehlik@utia.cas.cz`

Contents

List of Figures

Acronyms

GARCH	Generalized autoregressive conditional heteroscedasticity
MLE	Maximum likelihood estimation
VAR	Vector Autoregression

Master's Thesis Proposal

Author	Bc. Petr Petras
Supervisor	Mgr. Tomáš Křehlík, M.A.
Proposed topic	Connectedness of high-frequency data

Motivation At the beginning, modeling of market behavior was done by the simplest methods possible, by linear modeling. Every prediction was based only on some weighted previous values. After that economists started using more complex, non-linear modeling as it provided better prediction values. They started to measure connectedness of time series, which can explain how observed series are connected and how shock transfers through them. We will continue with this evolution and we will try to provide even better measure using methods designed for high-frequency data and trading.

In recent studies (Al-Deehani & Moosa 2006; Booth *et al.* 1997) researchers started to estimating spill-over effects, which describes how are multinomial series influenced by their factors. However, when using lower-frequency data, estimated values can be influenced by feedback between series. This can occur, as multiple signals can travel between series within one sampling window. This can possibly happen even when 1-day data are used and when this happens, results can suggest that series are connected together and influence each other, when in fact only one series cause changes in the second one. We are therefore using high-frequency data, in which we should be able to precisely detect direction in which the influence goes. Intra-day data feedback, if exist, should occur later then initial change, and we should be able to capture that and estimate the effects appropriately. If we find that our estimates are different from those obtained from daily or slower data, we can claim, that there exist some intra-day influences and we are able to detect them and for example our prediction for future values should account for them. This need for models that capture within day price changes are driven by real world development of the markets. As operations on the markets are faster than ever before, it is not enough to base our model on daily values, we have to capture all changes made by high-frequency trading to model them, since they can contain important information about market dynamics. These high-frequency models are getting their popularity in later years, since both academic researchers and portfolio managers are aware of fact, that models based on daily basis are not sufficient any more. We will try to measure spillovers on stock market, however our methodology can be based on already existing papers (Barunik *et al.* 2014) even when they are usually modeling spilling over at oil markets, or between oil markets and stock indices.

Since we are using intra-day data we can also, as an addition to standard time-series methods, use mathematical method based on continuous time and point processes. The Hawkes process. This process is suitable for changes that occur rather randomly in time and can be observed in exact time, when they occur and not only in specific sampling periods. (Liniger 2009) For purpose of continuous modeling it is no longer possible to use standard time-series methods. And vice versa, for normal time-series data it is not possible to use Hawkes methodology. So we are interested what are the results of continuous modeling and if they differ from discrete models. And if we can find difference in estimated coefficients there is probably also some information hidden in time which are lost, when data-frame with fixed observation periods is created.

Our main concern will be influence between cash indices. If we can find some cross-correlation, spillover effect between them and if this spillovers are similar with branching coefficients which are obtained from continuous modeling.

Hypotheses

Hypothesis #1: The existence of spillovers in high-frequency data between stocks

Hypothesis #2: The branching coefficients are non-zero.

Hypothesis #3: The indifference between spillovers and branching coefficients.

Methodology The first step will be synchronizing data into data frame with fixed sampling. Then we will use Vector auto-regression (VAR) and multi-variable models to obtain correlation between indices. Output from this part should be cross-correlation matrix in which we can see, how each of the series influence itself and also others. This between influence will be of our main interest. As addition we can use asymmetric models to take into account positive and negative changes.

In the second part we will be fitting Hawkes process on the same dataset. Hawkes process have as parameter of our interest the branching coefficient. This, similarly as in the first part, determines how much of the influence is reflected in subsequent observation.

Expected Contribution The published studies already covered the spillover effect of time series, however sampling of those series are usually daily or less frequent. In recent years we are able to obtain faster data and with those we can model spill-overs more precisely. Recent papers that operates with intra-day sampling are usually focused on oil prices and its derivates. We would like to introduce spill-overs between stock markets, which are not yet modelled on this fast frequency, at least as far as we know. Our contribution therefore is providing evaluation of this effect on high-frequency sample. We can also compare how this effects differ in comparison to daily or less frequent analysis. If we can find statistical difference between spillover estimates, we can then state, that low-frequency analysis is omitting some crucial part of the information and should not be used for deciding about connectedness of stocks.

Secondly we would like to introduce Hawkes process into economic research, which is build for continuous time, to address possible differences or problems in standard spillover computation. Branching coefficient in Hawkes process is direct analogy to cross-correlation, and therefore we can compare them. Both possible outcomes of this analysis will have direct impact on how we measure connectedness. If we find evidence, that Hawkes process have different branching coefficient, we can be sure, that even the fastest sampling is not enough to approximate for continuous time and this would violate most of economical reasoning (for example for creating diversified portfolios). On the other hand if we confirm, that discrete sampling is good approximation for continuous time, we can freely use non-continuous methods, which are both more popular and easier to calculate and approach.

Outline

1. Motivation: Why study high-frequency data, continuous or discrete approach.
2. Theory: Introduction of methods used. Connectedness of price indices.
3. Data: Description about data we are using. Data manipulation to obtain regular time-series.
4. Methods: Description of GARCH and VAR analysis, Hawkes process.
5. Results: Findings of our analysis. High-frequency spillovers. Branching coefficients.
6. Concluding remarks: Summary of the results. Final notes

Core bibliography

Al-Deehani, Talla, and Imad A. Moosa. "Volatility spillover in regional emerging stock markets: a structural time-series approach." Emerging Markets Finance and Trade 42, no. 4 (2006): 78-89.

Bauwens, Luc, and Nikolaus Hautsch. Modelling financial high frequency data using point processes. Springer Berlin Heidelberg, 2009.

Bacry, Emmanuel, Iacopo Mastromatteo, and Jean-François Muzy. "Hawkes processes in finance." arXiv preprint arXiv:1502.04592 (2015).

Booth, G. Geoffrey, Teppo Martikainen, and Yiuman Tse. "Price and volatility spillovers in Scandinavian stock markets." Journal of Banking & Finance 21, no. 6 (1997): 811-823.

Diebold, Francis X., and Kamil Yilmaz. "Measuring financial asset return and volatility spillovers, with application to global equity markets*." The Economic Journal 119, no. 534 (2009): 158-171.

Liniger, Thomas Josef. "Multivariate hawkes processes." PhD diss., Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 18403, 2009.

Barunik, Jozef and Kocenda, Evzen and Vacha, Lukas, How Does Bad and Good Volatility Spill Over Across Petroleum Markets? (May 10, 2014). Available at SSRN: http://ssrn.com/abstract=2435511

Author Supervisor

Chapter 1

Introduction

In this work we will take a look on three different stocks from the American market. We will be interested in stocks of Apple, Microsoft and AT&T to show how these stocks are connected. We will use two different methodologies to show, that these stocks influence each other. First methodology is based on discrete time with fixed sampling window. This methodology uses vector autoregression (VAR) model and error decomposition to show that connectedness of stocks is not symmetrical and it is important to model connectedness as two-way channel. Simple correlation is no longer sufficient for market analysis. Second method is based on continuous time. This method is based on Hawkes processes and models intensities. Intensity for each stock determines the probability for the price change. Because we are interested in connectedness of series it is obvious that also this methodology allows for modeling of connectedness between stocks. Both methodologies produce, in our case, 3×3 matrix of coefficients. This matrix specifies how the series are connected and how the change in price spills over the portfolio. VAR approach is being used in finance modeling for long time, Hawkes processes, on the other hand, are being used only recently on real data sets, even though the methodology exists for almost 50 years now. Problem with modeling on continuous time is the availability of data sets and computational requirements. In this work we will be using tick data set, which contains every trade made on the market in the observed time period, thanks to this we can determine actual price on market in every moment, as we assume, that every change in price would produce another new trade.

The main addition to current literature is the fact, that we will be using both methodologies on same data set. As far as we know this is the first attempt for such analysis. We will be using Hawkes processes as addition to standard spillover measure calculated via VAR. Thanks to this we will be able to model two levels of price changes. First level will be based on common price changes, which will be modeled using VAR and second level will be modeling extreme changes using continuous methodology. Thanks to this two level approach we will be able to tell how market prices react on change in one stock, and also what happens if this change is of the large magnitude.

As research questions we ask if coefficients based on the continuous time are different from those based on the discrete time, and if we are able to find any connections at all. It may be true, that true model for price changes is based on only one of those methodologies. Both provide good information, however we think that using VAR only cannot model the market fully, as it loses large portion of information, because of the sparse sampling. More frequent sampling contains more information however this model is not able to use it efficiently and can be flooded with buy-sell spread. This is where Hawkes process methodology is very useful since it is perfectly capable of using all data available, but this comes with cost of the large complexity of the model. We hope that merging these two models as suggested in this work will result in higher understanding of market behavior without too large costs caused by complexity of Hawkes process.

The thesis is structured as follows. In chapter 2 we explain more detailed reasons why we find this topic important. Chapter 3 provides introduction to core literature, which served as cornerstone for this work. Chapter 4 is core section of this thesis as it provides description of models used in this work. Chapter 5 introduces reader to data set we use in this work. In chapter 6 we present our econometric results for both methodologies. Chapter 7 aims to provide economic rationale about companies used for this work. And finally chapter 8 concludes.

Chapter 2

Motivation

Since beginning of the market trading traders tried to make predictions about future development of the market to outsmart others and make profit. Economic theory on the other hand came with idea, that markets behave randomly and therefore it is not possible to make any reliable predictions and without predictions there is no possibility for structural earnings. We call this theory Efficient market hypothesis. Traders however was not satisfied with this result, so they came up with idea not to predict prices but rather to predict volatility. As time showed even when predictions on prices revealed possibility for extra profit, this profit is usually so small that costs connected with buying or selling out-top possible profit from predictions. Predictions on volatility however showed some reasonable behavior, especially volatility clustering, where we can observe time periods with high volatility followed by times with lower volatility. This volatility modeling became important as traders are creating share baskets and portfolios in a way which lowers total volatility. However we are still very interested in price predictions as price is crucial in making market profit. Volatility predictions allow for modeling risks connected with investing into market. However this only models volatility of market and mean profit is given by overall market behavior. Models on prices could however predict mean profit rather than variance of this profit. This should be very useful in connection with volatility models.

The modern technology now allows to analyze price and volatility structures more intensively, thanks to option to sample these series with high-frequency. When we were working on one-day prices or with sampling on lower frequency, there was no possibility how to observe, if stocks or any traded goods are influencing each other within microtime structure, we were able to observe only slow adaptations across goods. We were usually able to observe changes occurring repeatedly in certain baskets of stocks or goods. If this phenomenon was persistent, we claimed that these stocks are highly correlated. Obviously highly correlated goods do not serve well as risk-reductors. When we put them in one portfolio they do not lower the risk as much as we want, since they behave almost as one type of good. With high frequency data at hand we are now able to tell how these goods are connected deep within. Not only by simple correlation measure, but by spillover measure as well. Spillovers allow us to estimate how one good influences other in the one directional fashion. This is very different to correlation measure. Correlation measure is always symmetric by definition. Thanks to spillover models, we are able to observe, that one stock is price leader, meaning that when price change occurs to price leader, others quickly follow. If that is the case, the first stock will have high spillover effect to others, while others may not influence the leader. This is natural extension of correlation measure. We say that correlation is intratemporal measure of connectedness of observed series, while spillovers serves as intertemporal one. However for high-frequency modeling time windows are so small, that this difference is almost negligible. Then spillover measure provides richer information, since it can reveal even connections which would be hidden while using correlations only. As theoretical example imagine that two stocks are connected in such fashion, that if price of stock A changes, then stock B also change its price in same manner, however when stock B change its price, stock A reacts by changing its price in opposite direction. These stocks would have correlation of zero (with very high variance), but spillovers can give us insight about the real process hidden behind this suspicious behavior and how are these stocks really connected, whereas simple correlation measure would give us very misleading information about possible independence.

Calculating the correct estimates of spillovers brings new problems connected to high-frequency data analysis. We are nowadays able to obtain tick data. The tick data means, that we are able to observe price and time of each trade performed on the market, therefore we observe price on market as continuous variable, since in any given time we consider the price of the stock as the price of the last trade. However this fast sampling brings new phenomena of buy-sell spread and micro-structure noise. These problems were not present in slow sampling and we need to find the way how to deal with it. The buy-sell spread naturally occurs in the market, since there is always someone who is willing to buy certain amount at certain price, and someone who is willing to sell certain amount at certain price. This can be considered as some type of base price, and it is impossible to tell where true price is within the buy-sell spread. When trade occurs it is usually because buyers accept sellers price or vice versa. These trades are the most frequent for casual trading and trades with actual different prices are in a minority. Our main interest lays in this minority. These trades are ones which really influences the price on the market. And we are trying to capture exactly those. However we are flooded with trades, which do not change actual price and are only realized within the buy-sell spread. This large volume of price changes between buy and sell prices can heavily influence our estimation, as it creates non-existing variation on the market. Imagine that true price is exactly in the middle of the buy-sell spread, but investors are unable or not willing to trade for this price. There will be large amount of trades occurring with prices closely above and below true price, as investors set these price as bids. Even though true price never changes, we can observe large number of trades within the buy-sell spread. And if we then perform classical analysis on this observation, these trades will artificially increase volatility of stock even though the true price of the stock has never changed. We are therefore using methods to lower influence of this trades happening within the buy-sell spread. The most common one is to set fixed-time sampling, where longer time windows decrease problems with buy-sell spreads more significantly, but with long sampling windows we are losing some observations and with that also some information. The short sampling windows do not lose as much information, but problem with the buy-sell spread can persist. Because of that it should be obvious that setting optimal sampling window is non-trivial problem. The widely used time-frame is 5 minute sampling window and according to Andersen *et al.* (2005) it is common sampling time frame. Obviously for different time-series this can differ, but finding optimal sampling window is not a common practice, and economists tend to use the 5-minute sampling, which is usually close enough to optimal sampling window and it is easier to interpret than optimal sampling window with fractional length.

Another approach is to observe only trades below or above certain threshold and then model this occurrences. With this approach we may run into problems, when this buy-sell spread changes due to change in the true price. Also we need to use different modeling methods, when we do not have fixed sampling. Luckily for us this can be done using Hawkes process modeling, which we will introduce later and it will be the key addition to widely used VAR model.

In this work we will try to answer if it is sufficient to model only on this 5 minute sampling. We already know that spillover measures provide additional information to standard correlation measures. However question remains if this quite complex spillover measure is sufficient as model for price development. Since usual methods are interested more in one stock volatility or correlations between stocks, where it might be possible, that less information is required to fully model volatility. And as well if this sparse sampling provides unbiased results. For our last question we will use modeling on continuous time domain. We will be using point process, namely Hawkes process. This process is designed to calculate changes in continuous time dimension, and handily provides coefficients which can be directly compared with spillover measure from discrete methods. This process allows us to model not only magnitude of changes but with that also occurrences in time dimension. This solves our problem with asynchronous time series and if we set coefficients correctly also with buy-sell spread. We will use both models on same data set and we will compare coefficients obtained from these models and if we find any discrepancies we will try to explain why this occurred and if it is a common trend or not. We will conclude, whether discrete methods approximate continuous time well enough or if we should be using models design specially for continuous time series, rather then try to approximate them by discrete sampling, only to be able to use discrete methodology.

In the beginning economists were mainly building discrete models. This comes very naturally, as agents' decisions were usually also discrete, meaning there were little or no difference between having something today morning or later in evening, to the future processes this change was factored tomorrow. Also trading was occurring more discretely. Stock exchanges were closed for nights and holidays and analysis were done on opening and closing prices. These breaks offer natural points for discrete sampling. However with modern technologies, more connected world and markets, and also with 24/7 trading, we are nowadays losing these points for natural discretization. However our models, namely microeconomic ones are still stuck to this discrete times. In macroeconomics we see trends for continuous modeling, however for macroeconomics this change is not so important, since reactions there are much slower and also observables cannot be calculated so often. The crucial sector which should focus on continuous modeling is finance, since in market trading we have to take into account transitions which are faster than one second, and for that we might need continuous modeling. As well as we have to know if our discrete modeling is still useful, or we should leave it to the past.

Chapter 3

Literature review

We dedicate this section to summarize literature on topics covering spillover analysis and Hawkes processes. The main expected contribution of this work should be connecting this two methods and their direct comparison, since we have not been able to find any relevant literature comparing these methods directly, we provide to reader at least some literature covering these methods separately.

On the topic of high frequency spillovers the basic literature is Diebold & Yilmaz (2009), where authors propose return spillover measure as well as volatility spillover measure. They are using VAR models as it is very common nowadays and substitutes older calculations based on generalized autoregression conditional heteroskedasticity (GARCH) models. They created basic normalized Spillover Index using one-step ahead future forecast. Their findings are based on daily data, and from that we can see that even for larger sampling windows it is possible to obtain significant outcomes, however more frequent sampling should allow better and more precise estimates and insights how channels between series work. They also introduce indices for contributions to other variables and from other variables, which are plain summaries of spillovers across all other variables. Their analysis is based on global stock market returns, which are of their interest, and can be compared with stock prices, which will be of our interest.

Further extension is provided by Booth *et al.* (1997) where they present estimation separately for good and bad news, however in this article estimation is not yet based on VAR, but on GARCH modeling. Their markets of interest are Scandinavian countries, namely Denmark, Sweden, Finland, and Norway. They did confirmed their hypothesis, that there exists asymmetry for good and bad new, however most spillovers are not significant. Out of all 12 possible pairwise effects, they found only three of them being significant. This may be caused by many factors, one of them might be that in daily sampling they were not able to find enough information to find connections. If markets reacts in faster terms it is very hard to estimate spillovers based on daily data. On the other hand it may be possible that markets of these countries are either not very interconnected, which would result into low correlation as well as in low spillovers, or are directly influenced by some external events, which we find more probable. This would lead to high correlations between markets, but as they would react at same moment, it would be impossible to find which market spills over to which, and resulting spillovers would be close to zero or very volatile. Or finally, it may be notable case, where VAR can outperform older GARCH.

As next step we look at Barunik *et al.* (2014), where authors are estimating spillover across petroleum markets. They are again using asymmetry measure as in Booth *et al.* (1997) or Bartram *et al.* (2012). But now they are dealing with tick-data analysis, therefore this paper have more similar data set and corresponding methodology then previously discussed ones. In this paper we can find more mathematical formulation of concepts used also in this paper. Namely realized variance and semivariance and how they are calculated. They are also building on top of Diebold & Yilmaz (2012) where they define directional spillovers, which are simply spillovers corresponding to contributions from Diebold & Yilmaz (2009). New in this paper is spillover asymmetry measure, which tells us how much is asymmetry present in our spillovers. Even when this paper is researching different markets, its methods still can be modified and used for our purposes.

To note some papers which describe Hawkes process, we, first of all, mention Liniger (2009), where author in his PhD. dissertation describes thoroughly important aspects about this methodology. This work is focused on theoretical aspects connected to this topic and applications to real world issues are mentioned only briefly.

As second paper focused on Hawkes processes we mention not so theoretical paper by Embrechts *et al.* (2011), which is more oriented on practical usage of Hawkes processes. They are modeling price peaks on daily data for Dow Jones Industrial Average for period of 16 years. They are trying to model occurrence of peaks which are below or above 10% and 90% quantiles, respectively. They, in their work, provide some interesting insights how Hawkes modeling can be done, and also their results. They also provide some hints how to construct confidence intervals for Hawkes processes, and also some tests to check for correct specification and goodness of fit, which will be very useful later in this text. They finally conclude, that “Hawkes process offer a versatile class of point processes capable of modeling extremal behavior of financial time series.” (Embrechts *et al.* 2011) This shows us that this strictly mathematical concept can be used even for practical purposes.

As last but not least in this section we mention Bacry *et al.* (2015), where authors once again show flexibility of Hawkes processes and how they can be modified. For each presented model authors provide to readers another paper, where they can find more details about specific models. As models of interest we would like to pin-point models of co-jumps which can be related to problematic of Scandinavian countries mentioned earlier. When co-jumps occur, we are not able to find transition mechanics between observed series, and it is not well fitted for modeling via Hawkes processes. This movement is usually caused by some external factor which influences all series, but is either unobserved or unobservable. This co-jumping makes modeling N-dimensional Hawkes process very difficult or almost impossible. To solve this problem one can model co-jumps as one dimensional Hawkes process as shown in Bormetti *et al.* (2015). But they also provide quite extensive discussion about price and volatility models. As final note we would like to mention table of used applications in appendix of Bacry *et al.* (2015) where they provide nice summary about frequency and type of modeling on which Hawkes processes were estimated.

Another relevant topic is about synchronization of observations and connected problems. For this section we present to reader papers Aït-Sahalia *et al.* (2010) and Zhang (2011). First one describes consistent and efficient estimator of covariance of two high-frequency asynchronous assets build on generalized synchronization scheme, which should overcome problems known as Epps effect. More about Epps effect can be found in latter mentioned Zhang (2011), where author propose direct measure of this microstructure noise, when calculating integrated covariance over fixed time horizon. He states, that commonly used previous-tick covariance estimator is biased, and this bias tends to be greater for more liquid assets. Later in his work he proposes optimal sampling frequency which balances Epps effect with information loss and random effects.

Chapter 4

Methodology

In this chapter we would like to introduce basic concepts which will be used later in the text. The section 4.1 describes methods of treating the data, which we have to apply to obtain standard data set, which can be used for our discrete analysis. The section 4.2 describes methods used for our analysis. We also devote section 4.3 to address problems with stability of models, as it is important concept while modeling financial or any other real series.

4.1 Data synchronization

For our analysis we are using a high frequency data sampling from stock markets. These data are based on the tick observations, which means that every transaction which occurred on the market is recorded. In these records we have exact time of the trade as well as the price. We take this price as a true price of the observed stock at given time. This time frequency allows for very precise analysis of price movement, however it also brings some problems. For those interested in development of single price variable, this tick sampling is perfect, since if information is hidden in price development, the tick data will contain it. However our main interest is in the interaction between the different stocks and their prices. For spillover analysis (described in section 4.2) we need to synchronize our data.

Assume that for I stocks we have prices $x_{i,t}$ where t is observed time and $i \in \{1, \dots, I\}$. But we cannot directly observe price at any time t we can only observe realized trades. These trades occur in times $\tau_{i,n}$ such that $0 = \tau_{i,0} < \tau_{i,1} < \dots < \tau_{i,T-1} < \tau_{i,T} = T \quad \forall i \in I$. Note however that even when $\tau_{i,0} = 0 \wedge \tau_{i,T} = T, \quad \forall i \in \{1, \dots, I\}$. It does not have to be true, that $\tau_{i,t} = \tau_{j,t}$

for $i, j \in \{1, \ldots, I\}, i \neq j$. In fact it is only rare when this occurs. In simple words, tick observations are happening in different times for different stocks. We call this asynchronous series. The asynchronous series are not problematic for continuous modeling, since it is specially designed to treat such series, but are very problematic for discrete models, where, as input, the models assume fixed $I \times T$ matrix of observations.

For spillover analysis we have to use matrix of observations $\mathbf{X}$, which is $I \times T$ matrix, where $x_{i,t}$ is price of stock i in time t. The question here is, how to decide on price $x_{i,t}$ when we observe only $\tau_{i,n} < t < \tau_{i,n+1}$. Through the years economists came up with different procedures how to synchronize data. The most intuitive and simple one is previous-tick estimator (Zhang 2011). In this estimator we predefine sampling period to certain length Δ, which does seem to be appropriate for our analysis (usually set to 5 minutes). Then we create equally spaced sequence $\tau_{\Delta,m}$ such that $\tau_{\Delta,0} = 0, \tau_{\Delta,T} = T, \tau_{\Delta,m} - \tau_{\Delta,m-1} = \Delta$. It remains to decide how to specify the price at given time. The usual approach is to select last price that occurred at or before given $\tau_{\Delta,m}$. Then $x_{i,\tau_{\Delta,m}} := x_{\tau_{i,n}}$ such that $\tau_{i,n} < \tau_{\Delta,m} \wedge \forall \tau'_{i,n} > \tau_{i,n} : \tau'_{i,n} > \tau_{\Delta,m}$.

Note that for high number of trade realizations this method is ignoring the number of observations, while for low number it generates observation points. According to Zhang (2011) and Aït-Sahalia *et al.* (2010) one should use those synchronization techniques, that omits some observations rather than creating new ones. The former one may yield inefficient estimates, but the later can produce biased ones, which is considered worse.

There is also an alternative approach to synchronizing. We can use autorefresh time (Barndorff-Nielsen *et al.* 2008). In this scheme observations are synchronized on periods for which every observed series has changed it's value. We are selecting

$$\tau_{\Delta,n} = \max\{\min\{\tau_{i,n} \in \tau_i : \tau_{i,n} > \tau_{\Delta,n-1}\}, i \in \{1, \ldots, I\}\}$$

For sequence $\{\tau_{\Delta,0}, \ldots, \tau_{\Delta,T}\}$ is now granted, that for each point in this sequence, all observed series has changed their value. This approach according to Barndorff-Nielsen *et al.* (2008) suffers from fact that most illiquid asset is the one that dictates sampling points. And that can cause bias in estimations. Barndorff-Nielsen *et al.* (2008) therefore propose to create Generalized Sampling Time synchronization scheme with changing of leading asset. In this scheme we go periodically through all observed series and require for it to be leading one. Formally: Let's assume, that series i_1 should be first one to have observation. We find $\tau_{\delta_1,n} = \min\{\tau_{i_1,n} \in \tau_i : \tau_{i_n} > \tau_{\delta,n-1}\}$ then for all other variable we find $\tau_{\delta,n} = \max\{\min\{\tau_{i,n} \in \tau_i : \tau_{i,n} > \tau_{\delta_1,n}\}, i \in \{1, \ldots, I\} \setminus i_1\}$. We repeat this procedure for $n + 1$ while selecting different series to be the first one. This will lead to sequence $\{0 = \tau_{\delta,0}, \tau_{\delta,1}, \ldots, \tau_{\delta,T-1}, \tau_{\delta,T} = T\}$. Notably this series will have lower number of terms than $\{\tau_\Delta\}$ however models on it should be unbiased.

In current economic papers it is common to use previous-tick estimator for equally spaced time series (usually 5 minutes apart). However we are introducing these methods, since for our future analysis it will be useful to have equally spaced and basic non-synchronized tick data. However unequally spaced synchronization can be used in future research for extending this methodology and testing if equally spaced synchronization on high-frequency data indeed produces biased estimations.

4.2 Model description

In this section we will describe methods used for our analysis of connectedness between stocks. First one is based on spillover index as vector decomposition from VAR as introduced in Diebold & Yilmaz (2009) and used in Barunik *et al.* (2014). Second method is based on Hawkes point process as described in Liniger (2009) and used in Bauwens & Hautsch (2009).

4.2.1 Vector Autoregression

VAR and variance decomposition is an simple and intuitive way how to approach spillover effects. First of all let's consider one lag VAR model as introduced in Diebold & Yilmaz (2009)

$$\mathbf{x_t} = \theta \mathbf{x_{t-1}} + \varepsilon_\mathbf{t} \tag{4.1}$$

where $\mathbf{x_t}$ is vector of synchronous prices, returns or volatilities at time t for I stocks, θ is $I \times I$ matrix of main interest, since from it we can observe how past values influence today's values, and $\varepsilon_\mathbf{t}$ is a vector of random errors. If we assume covariance stationary of Equation 4.1, it is possible to rewrite it into moving average form:

$$\mathbf{x_t} = \mathbf{\Theta(L)}\varepsilon_\mathbf{t} \tag{4.2}$$

where $\Theta(L)$ is infinite horizon lag operator. And furthermore use Cholesky factor decomposition and rewrite this equation as follows:

$$\mathbf{x_t} = \mathbf{\Theta(L)Q_t^{-1}Q_t}\varepsilon_\mathbf{t} \tag{4.3}$$

where, Q_t is unique lower triangular matrix.

Let's now consider one-step ahead forecast. Since we assume that $\mathbb{E}\varepsilon_t = 0$ It is easy to find out that:

$$\mathbf{x_{t+1}} = \theta\mathbf{x_t} \tag{4.4}$$

and corresponding error vector

$$\mathbf{e_{t+1}}, \mathbf{t} = \begin{bmatrix} a_{0,11} & \cdots & a_{0,1I} \\ \vdots & \ddots & \vdots \\ a_{0,I1} & \cdots & a_{0,II} \end{bmatrix} \begin{bmatrix} u_{1,t+1} \\ \vdots \\ u_{I,t+1} \end{bmatrix} \tag{4.5}$$

Let's remind, that $\mathbf{u}_t = \mathbf{Q}\varepsilon_\mathbf{t}$ and elements $a_{0,ij}$ are part of $\mathbf{A_0}$ matrix which is constructed as $\mathbf{A}(L) = \Theta(L)\mathbf{Q}^{-1}$.

Such decomposition allows us to directly estimate the influence of changes in one variable to other variables as well as to own variable. We define own variance shares as variance due to own shocks ($a_{0,ii}^2$). And also spillovers from and to variable j as $\sum_{i,i\neq j} a_{0,ij}^2$ and $\sum_{i,i\neq j} a_{0,ji}^2$ respectively. We can also define total spillover as $\sum_{i,j} a_{0,ij}^2 - \sum_i a_{0,ii}^2$. And as spillover index we define total spillover normalized by sum of variances.

$$S = \frac{\sum_{i,j} a_{0,ij}^2 - \sum_i a_{0,ii}^2}{tr(\mathbf{A_0A_0'})} \times 100 \tag{4.6}$$

Furthermore this concept can be extended to allow for high order of VAR and p-step ahead forecast by simply extending (4.1) with further lags.

$$\mathbf{x_t} = \theta_\mathbf{1}\mathbf{x_{t-1}} + \theta_\mathbf{2}\mathbf{x_{t-2}} + \cdots + \theta_\mathbf{p}\mathbf{x_{t-p}} + \varepsilon_\mathbf{t} \tag{4.7}$$

Then we arrive to following measure of spillover

$$S = \frac{\sum_{h=0}^{H-1} \left(\sum_{i,j} a_{h,ij}^2 - \sum_i a_{h,ii}^2\right)}{\sum_{h=0}^{H-1} tr(\mathbf{A_hA_h'})} \times 100 \tag{4.8}$$

We also define directional spillovers as measures how one series spills over to other, and how other contributes to first one.

$$S_{i\to\bullet} = \frac{\sum_{j,j\neq i} a^2_{0,ij}}{tr(\mathbf{A_h}\mathbf{A'_h}) \times 100} \tag{4.9}$$

will be referred as spillover from i, and

$$S_{\bullet\to i} = \frac{\sum_{j,j\neq i} a^2_{0,ji}}{tr(\mathbf{A_h}\mathbf{A'_h}) \times 100} \tag{4.10}$$

will be referred as spillover to i, we again normalized this measure by total variance, for better comparison across different models or datasets, if needed. However we need to keep this in mind when we will be comparing this measures with branching coefficients from Hawkes modeling, since normalization of coefficients in Hawkes models are not so straight forward. Therefore we should compare only magnitudes and relative sizes rather than absolute values.

Since relevant literature suggests, that prices tend to react differently when facing positive and negative news, and with that positive and negative price changes, we might also need asymmetric measures. For this we introduce indicator function $\mathbb{I}_{(u<v)}$, where term in lower index is some arbitrary expression, for values or functions u, v. Whenever this expression is true indicator function is equal to 1, and 0 otherwise. With help of this notation we can easily define spillovers driven by positive changes and negative changes separately, where our indicator function is: $\mathbb{I}_{(\Delta p_{t,i}>0)}$ for positive changes across time t and variables i and $\mathbb{I}_{(\Delta p_{t,i}<0)}$ for negative changes. Asymmetric measures can be easily implemented using modified methodology even for Hawkes process. This will be discussed in next section.

4.2.2 Generalized Vector Autoregression

As extension to basic VAR we will, in this section, describe enriched methodology as described in Diebold & Yilmaz (2012). This article presents generalized version of VAR and we will use methodology instead of simple VAR, since basic VAR can suffer from dependence on ordering. This comes from process of diagonalization while doing Cholesky decomposition, which is necessary while calculating spillover matrix. This is rather large drawback, since there is no given order in which we should input our time series. Therefore methodology which solves this problem is highly preferred.

Diebold & Yilmaz (2012) also states, that spillovers based on generalized VAR can be used while comparing and calculating spillovers across different asset classes. This benefit is not crucial for our analysis, however it is useful to know about this fact. So generalized VAR can be used for calculating spillovers between e.g. stock and bond markets or different countries.

To solve the problem of the ordering dependence we use generalized impulse response function as introduced in Koop *et al.* (1996).

$$GI_Y(n, v_t, \omega_{t-1}) = \mathbb{E}\left[Y_{t+n}|v_t\omega_{t-1}\right] - \mathbb{E}\left[Y_{t+n}|\omega_{t-1}\right] \tag{4.11}$$

where ω_{t-1} is information set available at time $t-1$. If we now consider MA(∞) representation as stated in (4.2) we obtain

$$GI(n, V_t, \Omega_{t-1}) = \Theta_n \varepsilon_t$$

We can now obtain distribution of GI based on assumed distribution of ε_t. The shocks to the system is now reflected via diagonals of matrix $\Theta_n \Sigma \Theta_n'$, where $\varepsilon_t \sim N(0, \Sigma)$. By "(...)averaging the squares of the GI component by component against the joint distribution of these system-wide shocks. A pseudo-traditional impulse response function for linear model can be constructed by taking the square root of the diagonal of $A_n \Sigma A_n'$ ($\Theta_n \Sigma \Theta_n'$ in our case) for each horizon. This solution deals with the composition problem (ordering dependence) by ignoring it through focusing on the effect of system-wide shocks." (Koop *et al.* 1996)

This is due to construction of GI function. If we consider shock to only j^{th} variable, it is possible to integrate out shocks to other variables. (Pesaran & Shin 1998)

Drawback of this methodology is, that shocks that reflects to other variables or not normalized to one, as it happens with Cholesky decomposition due to non-orthogonalized shocks. (Pesaran & Shin 1998) However this can be solved by normalizing with respect to each row. Formally as stated in Diebold & Yilmaz (2012) elements of A matrix are defined as follows

$$a_{ij}(P) = \frac{\sigma_j j^{-1} \sum_{p=0}^{P-1} (I_i' \Theta_p \Sigma I_j)^2}{I_i' \Theta_p \Sigma \Theta_p' I_j)} \tag{4.12}$$

where P is horizon of VAR(P), elements a_{ij} are directly compared to matrix A from Equation 4.5. Σ is error variance matrix, σ is standard error for the j^{th}

equation, and I_i is vector of zeros, except on i^{th} place, where it is one.

This is now normalized to one as follows

$$\tilde{a_{ij}} = \frac{a_{ij}(P)}{\sum_{j=1}^{N} a_{ij}(P)} \tag{4.13}$$

This redefined version of matrix A is now order independent and normalized in a way that shocks to each variable sum up to one. Now we can proceed as before to construct all measures as stated in previous section.

4.2.3 Hawkes process

"The first type of point processes proposed in the context of market microstructure is the ACD (autoregressive conditional duration) model introduced by Engle and Russel (Engle & Russell 1998). This model and its variants remains, by far, the most used model in high frequency econometrics (Bauwens & Hautsch 2009). In this class of models, the process is defined by the means of its 'hazard function' that specifies the conditional law of inter-event (or duration) intervals. However, point processes (or counting processes) can alternatively be represented by their 'intensity function' that represents the conditional probability density of the occurrence of an event in the immediate future"(Bacry *et al.* 2015).

This autoregressive conditional duration models were introduced as a way to model irregularly spaced trades, which are however not independent. This is exactly what we need for our estimation. Every observation creates event called point, we treat each of these points as observation and we assume that distribution of these points are not absolutely random, but there exists underlying process which can describe behavior of these points. Because of that, we call this underlying process the point process. Hawkes process, which is of main interested of this work, is a special type of point process as described above. This process allows for modeling of discrete events in continuous time. Occurrences of observed events, points, are distributed across continuous time frame according to some underlying process, which takes into account relations between observed events. We aim to describe this underlying process, since it can be used to provide predictions about future. Usual assumption from efficient market hypothesis is, that this events occur according to normal distribution, however data suggest that tail events occur more often than they should, and also that they tend to cluster together. This is known as volatility clustering.

We have chosen the Hawkes process as our underlying process because Hawkes process allow to model this more precisely, since Hawkes self-intensity allows for this type of clustering. Hawkes intensity is not only derived by assumption about distribution, but evolves as points occur. Therefore, when point event occurs, Hawkes intensity increase and with that also probability of occurrence of another point. As opposing force we introduce to reader decay function, which controls stability of Hawkes process and rate by which intensity decreases to base level. Current size of Hawkes intensity can be used in calculation of probability that next point occurs. As we can foresee, this self-excitation can lead to explosive processes, therefore it is important to have balance between increase of Hawkes intensity and decrease due to decay function. More about this problematic can be found in section 4.3, where we in detail describe, what exactly this relation should fulfill.

Hawkes process as a point process, can be used for modeling uni-variate events as in seismology and earthquake modeling, however we will use its variation for multivariate case. Note that we will describe two different Hawkes processes, one which is usually called *vector* point process and is more similar in its nature to uni-variate type and second one which is called *multivariate* point process. Vector point process is used for already synchronous time series. This type of process consists of tuplets of the form $(t_i, x_{i,1}, \ldots, x_{i,d})$, where $x_{i,j}$ is observed value for component j and countable value i, t_i stands for time of a event numbered i. It is obvious, that for equally spaced sampling as mentioned in section 4.1 is $\{t_i\}_{i=1}^{T}, T \in \mathbb{R}^{*+}$ an arithmetic progression, and we are losing part of the power which is offered to us by Hawkes process. For unequally spaced sampling is $\{t_i\}_{i=1}^{T}, T \in \mathbb{R}^{*+}$ a random sequence and therefore using Hawkes process has deeper meaning. Second type of Hawkes process will be used for non-synchronized time series. This multivariate process is perfectly designed for having multiple time series with different observation times. This multivariate Hawkes process will be using triples of following form (t_i, x_i, d_i), where again t_i stands for time of event i, x_i is a value at time t_i and d_i specifies which component we are referring to.

Let us now define both Hawkes processes, vector and multivariate:

$$\lambda(t) := \eta + \vartheta \int_{(-\infty,t)\times\mathbb{R}} \omega(t-s)g(x)N(\mathrm{d}s \times \mathrm{d}x), \quad t \in \mathbb{R} \tag{4.14}$$

where $\omega : \mathbb{R}_+ \to \mathbb{R}_+, g : \mathbb{R}^d \to \mathbb{R}_+$, and $\eta, \vartheta \geq 0$

$$\lambda_j(t) := \eta_j + \sum_{k=1}^{d} \vartheta_{jk} \int_{(-\infty,t)\times\mathbb{R}} \omega_j(t-s) g_k(x) N_k(\mathrm{d}s \times \mathrm{d}x), \quad t \in \mathbb{R} \tag{4.15}$$

where $\omega_j : \mathbb{R}_+ \to \mathbb{R}_+, g_k : \mathbb{R}^d \to \mathbb{R}_+$, and $\eta_j, \vartheta_{jk} \geq 0$. Then we call this vector and multivariate Hawkes process respectively, with $\lambda(t)$ being appropriate Hawkes intensity.

Important note is, that in well established Hawkes processes each point is either an immigrant or a descendant. Meaning that each observation is either because of influence from other components (immigration) or because of influence from own previous points (descendancy). We also assume natural occurrence of observations. This is to establish stable relation, however this natural points are assumed to be before time period of interest and do not further influence our estimation. However it is important to know that natural intensity exists. Occurrences of immigration and descendancy events are defined by functions occurring in Equation 4.14 and Equation 4.15 and we will provide detailed description of them.

Immigration intensities. Values η, η_j are connected to frequencies at which new immigrants arrive.

Decay functions. ω_j governs the intensity for components other then j, which is directly influenced by time difference. This function lowers Hawkes intensity and it is the only term which naturally does so. Modeling of correct values of decay function is crucial for stability of the Hawkes process.

Impact function. g_k and g are functions which determines how large the influence will be related to value of mark value x_i. Impact function is a part which takes into account importance of points rather than quantity of them. In threshold Hawkes models this reduces to degenerate function, as we will specify later.

Branching coefficients. ϑ_{jk} tells us how the intensity of component j is increased when point event occurs in component k. In vector case this reduces to one-dimensional value increasing overall intensity. For multivariate case this is the coefficient which directly shows how series influence each other. Branching coefficients will be of our main interest, since they are directly comparable with spillovers in discrete case. For better manipulation with them we define *branching matrix*.

Branching matrix. Let $\mathbf{Q}$ be a $(d \times d)$ matrix, where $\{\mathbf{Q}\}_{j,k} := \vartheta_{j,k}$.

See Embrechts *et al.* (2011) for more details.

Now, when we have defined all functions in Equation 4.14 and Equation 4.15, let us describe in more detail how the Hawkes intensity is calculated. The term η is base intensity, which do not change across time and is connected to base immigration. Values of ϑ scale how influential will be value of integral. Thanks to this term immigration can differ across components. Integral is evaluated across real axis as observed value can be any real number. This number enters to function g, which captures importance of different values. Second axis for integration is time line for the time until now. Decay function is integrated across this time spectrum and thanks to this, it can have different weight for past and recent times.

It should be now clear, that Hawkes process allows for complete modeling of trades on the market. Since it can capture both prices in terms of x_i from triplet (t_i, x_i, d_i) as well as time t_i from same triplet. However we should again proceed with caution, since large portion of trades will be within the buy-sell spread. Hawkes process is capable of modeling even this large volume of trades within small price spectrum, however it may lead to insignificant conclusions about larger changes and also immerse computational requirements. As a solution to problem with flooding of unimportant observations we can use counterpart to sparse sampling in case of discrete time series. Embrechts *et al.* (2011) suggest to use point process structure which corresponds with peaks above 90% quantile and below 10% quantile. Other authors suggests that changes above 1/2 of standard deviation should be used. We will use basic descriptive statistics to try to estimate appropriate quantiles for our purpose. Note also, that we do not have to use same quantiles across stocks or even for upper and lower boundary. If one has reasonable suspicion that stocks of interest behave differently or rather react with different magnitude, it may be appropriate to use different quantiles for each observed component.

Another alternative is to view this quantile selection as opportunity to model market prices as summation of threshold Hawkes processes on different frequencies, instead of modeling full process with prices, we can model simple point processes with different thresholds and then sum them up to create model of prices. As suggested for volatility in Wyart *et al.* (2008).

In the literature few different decay functions are already used. Their selection is also rather arbitrary, however usually we can choose from two function families. Either we model using exponential decay function or we use power law decay function. The decay functions also allow for precise specification of each component, as decay functions can be different for each component, however functional family should remain the same for all series, as selection of decay function is driven mainly by assumption about underlying process, which should be same for all elements, which are traded on same market. For exponentional specification we have:

$$\omega_j(t) = \delta_j^1 \exp\{-\delta_j^2 t\} \tag{4.16}$$

If we would like to simplify our estimation or we do not have any prior information about possible differences in stocks, we can always set $\delta_j^{\cdot} = \delta_k^{\cdot}, \forall j, k \in \{1, \ldots, d\}$. Arbitrary parameters and functions will be set to best fit the data we will be using for our analysis. Exponential function has been used in many early applications, for example Filimonov & Sornette (2012), however nature of trades on financial markets could be better described by long-term memory processes, therefore using power-law decay functions. As specified in Hardiman *et al.* (2013), where using of power-kernel is recommended. These authors are showing on simulations, as well as on real data, that branching coefficients while using exponential kernels may have significant downward bias. To provide reasoning for power-law kernel, they have simulated series with branching ratio close to one and exponential kernel estimation suggested that true value is below 0.7. This can have significant effect on estimated model. While power-law kernel provided estimation much closer to true value of one.

The power law kernel in Hardiman *et al.* (2013) is having following form

$$\omega_j(t) = \Theta(t - t_0)\phi_0 \frac{t_0^{\varepsilon}}{t^{1+\varepsilon}} \tag{4.17}$$

where t_0 is short-time cutoff and $\Theta(\cdot)$ is Heavside function, which is same as indicator function $\mathbb{I}$ defined earlier.

In this section we have specified many options how is possible to estimate Hawkes processes on financial data. For our future analysis we however need to specify one baseline model, which will be used for computation. This decision significantly affects results from our estimation and it would be interesting to redo this analysis with different types of specifications. Specification which we will be using in future analysis is based on simple threshold model on prices with symmetric influences with exponentional kernel. This is in fact the most simple one, but it should be sufficient for our analysis. However we have introduced broad enough methodology, so interested reader should be able to continue with different base line model.

To fit Hawkes process on jump part of price, we will modify Equation 4.14. To remind this equation to the reader we repeat definition here and will follow with specification for our case.

$$\lambda(t) := \eta + \vartheta \int_{(-\infty,t)\times\mathbb{R}} \omega(t-s)g(x)N(\mathrm{d}s \times \mathrm{d}x), \quad t \in \mathbb{R} \tag{4.18}$$

and we specify $g : \mathbb{I}_{p \notin q(0.95, n=50000)}$ where, p stands for price, and q is quantile function with two variables, where first specifies bandwidth for quantiles and second number of considered observations. To enlighten this approach even more this indicator function works as filter function, which leaves only those observations which value is above 97.5% quantile or below 2.5% quantile considering all observations in 50000 long rolling window.

Both values for quantile function are arbitrary selected. Threshold for extreme values was selected as usual 95% interval, which can be found in any econometrical analysis and also reflects traditional models on VAR. Size of rolling window was selected based on number of observations in one 5-minute window. If prices were more stable it would be preferred to compute extreme values based on comparing whole data sample or larger rolling window, but due to nature of price changes in our data, this would yield large periods of time being exceptionally good and others exceptionally bad. However Hawkes process is not suitable for such behavior, if we want to model high-frequency changes. Therefore we had to create methodology to capture changes within these good and bad times and rolling quantiles method fits our purpose, since price jumps are compared to prices in close proximity. As another arbitrary decision we decided for symmetry of jumps, therefore prices in lower quantiles are treated in same manner as prices in high quantiles. This is again to match our VAR methodology. If we would decide to expand our analysis to treat asymmetry, we would have to compare asymmetric VAR with Hawkes process where g function would have also two parts g^+ and g^- specified as $g^+ : \mathbb{I}_{p>q(0.975, n=50000)}$ and $g^- : \mathbb{I}_{p<q(0.025, n=50000)}$.

4.2.4 Maximization procedures

As procedure for estimating parameters in our models we will be using Maximum Likelihood estimation (MLE). This is very popular choice, while maximizing non-linear models. Objective function is directly connected with model estimation. For Hawkes processes we are using exponential kernel maximization. Here we note that Hardiman *et al.* (2013) re-estimate parameters of power law kernel with non-parametric methods, to verify if MLE provides unbiased results. They found out that MLE and non-parametric method provide same results, and MLE is also robust to data anomalies, which they are facing around year 1998. Therefore we can assume, that MLE will provide valid results even for our case.

We will again follow procedure and notation which can be found in Embrechts *et al.* (2011). Let us assume, that we do have observations from time T_* to time T^*, which we call period D. This is natural restriction since in Equation 4.15 and Equation 4.14 we have integrals from time $-\infty$, but we do not have any data for period $(-\infty, T_*)$. Therefore we want to modify our intensity functions such that they will account for this fact. Let us reformulate this by introducing Compensator.

$$\Lambda_j(t) := \int_{T_*}^{t} \lambda_j(s)\mathrm{d}s \quad \textit{and} \quad \Lambda(t) := \int_{T_*}^{t} \lambda(s)\mathrm{d}s \tag{4.19}$$

for multivariate and vector-valued cases respectively.

We can then define Hawkes likelihood function on D as follows for multivariate case:

$$\begin{aligned}
\log L \;&=\; \sum_{j=1}^{d} \int_{D\times\mathbb{R}} \log \lambda_j(t) N_j(\mathrm{d}t \times \mathrm{d}x) + && (4.20)\\
&+\; \sum_{j=1}^{d} \int_{D\times\mathbb{R}} \log f_j(t) N_j(\mathrm{d}t \times \mathrm{d}x) - \sum_{j=1}^{d} \Lambda_j(T^*) && (4.21)
\end{aligned}$$

and for vector case:

$$\begin{aligned}
\log L \;&=\; \sum_{j=1}^{d} \int_{D\times\mathbb{R}^d} \log \lambda(t) N(\mathrm{d}t \times \mathrm{d}x) + && (4.22)\\
&+\; \int_{D\times\mathbb{R}^d} \log \prod_{i=1}^{d} f_i(x_i) N(\mathrm{d}t \times \mathrm{d}x) - \Lambda(T^*) - && (4.23)\\
&-\; \int_{D\times\mathbb{R}^d} \log c(F_1(x_1), \ldots, F_d(x_d)) N(\mathrm{d}t \times \mathrm{d}x). && (4.24)
\end{aligned}$$

This log-likelihood can then be numerically optimized by standard numerical algorithms since all integrals are finite for finite values and finite time period. Also logarithmic version of likelihood is strictly preferred, as reader can notice, that multiplication within integrals from (4.14) and (4.15) are now transformed to summation, which lowers complexity for numeric algorithms.

4.3 Stability of models

In this short section we would like to address problem of stability of our models. Stability of systems are so important for viability of models, that parameters of decay functions are selected in a manner which yields desired behavior. However we should be aware which parameters are to be set and how to ensure, that our models are stable. The selected parameters also should not go against economic intuition. When these parameters have to be set in such manner, that there is no economic rationale behind them, we should re-think our underlying model.

When calculating VAR model we set only number of lags considered. VAR do not use any deep parameters, which would require our attention, therefore we check the stability of the VAR estimate only ex post. We call the system stable if resulting matrix from Vector Error Correction Models (VECM) are stable, which is if and only if eigenvalues of this system is lower than one.

For Hawkes process situation is more complicated, since we can set parameters of decay function such as following system will be stable, critical or unstable. Following parameters can and will be estimated be MLE, however we should keep in mind meaning of our results. If we assume exponential kernel as specified in Equation 4.16, then parameters δ_j^1 and δ_j^2 decides on stability of our system. To prove that let us consider one dimensional case with intensity function Equation 4.14. We simplify this intensity even more by assuming that price distribution is unimportant for deciding about stability. This holds because process driving prices do not depend on occurrences of events, and we are interested in frequency of events when we are talking about stability. Therefore let's assume that $g(x) \equiv 1 \quad \forall x \in \mathbb{R}$. Then we can calculate average intensity as in Hardiman *et al.* (2013):

$$\Lambda = \eta + \Lambda \int_{0,\infty} \omega(\tau) \mathrm{d}\tau \tag{4.25}$$

Then we state, that stability of the system does not depend on value η but only on value of the integral. If the value of integral is lower then one, then each of parents will have less then one descendant and the system will be stable and decaying and Hawkes intensity of such process will converge towards zero.

If the value of integral is higher than one, we have explosive system, since each parent will have more than one descendant and each new descendant will again have more than one descendant and Hawkes intensity will grow over time and it will be unbounded. Λ is then infinite with finite probability. The edge case, the critical case, is exactly when our integral is equal to one. Then the process is self-sustainable. In Hardiman *et al.* (2013) authors showed that for real financial markets branching coefficient is close to one or equal to one. This has strong implications for estimation of Hawkes processes. First of all, this suggests that original parents are or should be out of the sampling window, since process is (almost) self-sustainable, and with increasing number of parents, this would result intoincreasing number of price events. This implies that we should observe η close to zero in Equation 4.14 and Equation 4.15. They also state that number of events did not increase between years 1999 and 2009 substantially. Please note that this refer to mid-point price changes not to absolute number of trades, which increased significantly with introduction of electronic trading.

Therefore to ensure criticality of the system we set parameters as follows:

For case of exponential kernel we have:

$$\int_{0,\infty} \omega(\tau)\mathrm{d}\tau = \int_{0,\infty} \delta_j^1 \exp -\delta_j^2 \tau \mathrm{d}\tau = \frac{\delta_j^1}{\delta_j^2} \tag{4.26}$$

This shows that if we set $\delta_j^1 = \delta_j^2 = \delta_j$ we force our system to be exactly critical.

For power law kernel situation is similar:

$$\int_{0,\infty} \omega(\tau)\mathrm{d}\tau = \int_{0,\infty} \Theta(t - t_0)\phi_0 \frac{t_0^\varepsilon}{t^{1+\varepsilon}} = \frac{\phi_0}{\varepsilon} \tag{4.27}$$

This is of course case for single dimension Hawkes process, to demonstrate how parameters in kernels interact for simple cases. We are working with multidimensional processes where situation will be more complicated, however with intuition from single dimension we can easily translate this to higher dimensions. If we consider number of offspring in any dimension to be less than one, then we have stable system. As this number approaches one we are getting to critical case, and when this number exceeds one we have explosive system. Let us note here that Embrechts *et al.* (2011) normalize Equation 4.15 and Equation 4.14 in Condition 1 such that

$$\int_0^\infty \omega_j(t)\mathrm{d}t = 1 \quad and \quad \int_{-\infty}^\infty g_k(x) f_k(x)\mathrm{d}x = 1$$

and

$$\int_0^\infty \omega(t)\mathrm{d}t = 1 \quad and \quad \int_{\mathbb{R}^d} g(\mathbf{x})f(\mathbf{x})\mathrm{d}x = 1$$

This secures that numerical calculations are finite, in sense that optimalization do not diverge, as it could happen without this limitation. Without stability condition it is possible that MLE finds as optimum degenerate case for intensity rapidly converging to zero or exploding.

This restriction is very strong and ensures stability. Modeling without this condition requires more complicated function g.

4.4 Analysis on rolling window

This section describes reasons for and benefits of rolling window analysis as well as approach, that we will use in this work.

The rolling window analysis is based on repetitive usage of base model on subset of available observations. Size of the subset is usually selected while having in mind econometric and economic reasoning. Number of observations within selected subset should be large enough, so we can perform analysis with large number of degrees of freedom, so we can rely on asymptotic behavior. This is especially true for sparse sampling and large models. VAR models require high number of degrees of freedom, which rapidly grows with number of series and lags included in the model. Even when data set is rich enough we should not run analysis on too short window, since then we are losing connection to economic interpretation and possibility for predictions. Short time samples can be heavily influenced by disturbances on market. This can be very useful for finding these disturbances and we in fact will use rolling methodology exactly for this purpose in later section. However predictive power of such models is usually very low, since results from these times can be used only when market is in very similar state. On the other hand too wide rolling window can be too similar to full sample analysis and will provide no new information.

For our work we have selected 300 ahead rolling window, which should be wide enough to provide sufficient observations so we can rely on asymptotics, and it is also small enough so we will obtain large amount of estimates on which we can perform descriptive statistics such as mean and quantile values.

For synchronous methodologies it is easy to specify window based on number of observations, since in each sampling time we have observation for every time series. However actual time difference between first and last observation in the rolling window may widely differ, therefore we suggest that for analysis on rolling window it is better to specify size of the window on the time line and let number of observations vary, rather than vice versa. This problem is obviously solved naturally when we also restrict our sampling to be fixed to time period. In our case to 5-minute sampling.

In second part we are using asynchronous methodology, where specifying size of the rolling window as number of observations has no sense since number of observations differ not only as time progress but also varies in each series. Because of before mentioned reasons we suggest, that rolling window based on time interval is more appropriate for asynchronous sampling and continuous methodologies. So part about Hawkes process will use this approach and we will set two different windows one to contain one day and second one to contain one week.

Rolling window analysis serves also as a good robustness check. Since we are selecting subsamples of our data set, outliers can distort the estimation only when they are included in subsample. This way we can compare full sample estimates with estimates from rolling window and if subsample analysis is similar to full sample one, then we say, that our results are robust.

Chapter 5

Data

In this work we are working with three stock data series. They are ranging from January 2 from the beginning of the trading day until September 29 and end of the trading day. The trading days starts at 4 a.m. or 5 a.m. and ends at 8 p.m. We have at hand tick data with trade prices and volumes. Overall this sampling window of nine months provides to us total of 52,525,432 observations for Apple , 28,013,120 for Microsoft and 18,110,152 for AT&T . For illustration we provide sparse (5-minute) sampling graph for each of series in Appendix. As Figure A.1 for Apple, Figure A.2 for Microsoft and Figure A.3 for AT&T.

Even though we do not have continuous trading days, we will treat our data as continuous. We follow intuitive approach how to remove gaps since for our VAR analysis we need to have data in one time frame, therefore we constructed interval D as follows:

$$D := \bigcup_{i=1}^{N} [d_i^0, d_i^1) \tag{5.1}$$

where N is number of available trading days, d_i^0, d_i^1 are start and end of i^{th} trading day. Therefore we glue our days together without gaps. Same approach will be used if we lack observation in specific time window or trading day. We treat this as if this window is non-existent. There are methods how to treat gaps in sampling, however application of these methods are beyond scope of this text. General idea is either to model non-existing trades with some approximation. The easiest way is to fill these gaps with linear approximation between closing and opening price. Other approach is to introduce special term which changes its value during the day. In this case we assume that opening trades may be more volatile than trades in the evening. For Hawkes process estimation we assume, that trades would have occurred, if it was possible, so data for Hawkes methodology remains untreated, and enter the estimation with gaps.

Chapter 6

Results

Following chapter presents results we have obtained from our analysis on data set. It is again divided by methodology and overall results are presented in chapter 8.

6.1 Vector Autoregression

We have modeled VAR system on 5-minute sparse sampling with synchronous data. For this process we used two time lags for our VAR matrix. Results are presented as Table 6.1.

Readers should be aware of the fact that numbers in Table 6.1 are entries as percentage explained in direction from columns to rows normalized to 100%. E.g. 25.29% in cell (3,1) is spillover index from Apple to Microsoft. As expected result we present, that explanatory powers are strongest for diagonal values. In other words price of the stock mostly depends on own past values, and other stocks are secondary. What is also important to note, is that spillover matrix is no close to being symmetrical, which is expected result and serves as confirmation that simple correlation measure would not be sufficient for our case. From this we conclude, that spillover models should be used when modeling connectedness of prices with discrete times. From Table 6.1 we can see

To \ From	Apple	Microsoft	AT&T
Apple	97.09	1.90	1.01
Microsoft	18.26	79.54	2.20
AT&T	25.29	7.57	67.14

Table 6.1: Overall spillovers

that stocks of Apple are leading in the market, stocks for Microsoft and AT&T do not influence Apple as strongly as Apple influences them back. This just confirms, that idea of directional spillovers is important for financial analysis using VAR.

As additional result we provide graph with rolling estimation of before mentioned spillovers. It is important to see, how spillovers change in time, and rolling window estimation is perfect tool for this. Basically for each estimation we restrict our sample to 300 consecutive observations and recalculate spillover table. Results are presented as graphs which evolve with time. And can be found in Appendix. Spillover to Apple as Figure A.5, to Microsoft as Figure A.6, and to AT&T as Figure A.7.

Lower line represents spillovers from Apple in case of red line, and from Microsoft in case of green line. Upper line represents summation of spillovers captured in lower line and auto-correlation. Different way to look at it what is missing from upper line is spillover from AT&T if line is blue, and spillover from Microsoft if line is green.

This moreover splits graph into segments. Lower segment represents dependence of stock in title on stock based on color. Segment in middle is always self-dependence. And upper segment is how much is stock influenced by stock again color coded.

From graphs we can see, that Apple is very stable with self dependence. Apart from few spikes term (1,1) from Table 6.1 is above 70%. In fact 1st quartile is 72.78, median 82.56. This only confirms that Apple is the market leader, at least in our settings with only three stocks. Statistics for other two stocks are widely different. For Microsoft, we have median at 51.53 for on-diagonal element, this means that only little above half of the price is determined by past prices of Microsoft. Median of spillover from Apple to Microsof is at 38.24, which is surprisingly high. This means that in construction of price of stocks of Microsoft, stocks of Apple have significant role. Even 1st quartile is very high at 22.16. AT&T has very low influence on Microsoft, only around 5%. If we analyze spillovers to AT&T we will find out, that median for (3,3) is below 50%, very close to (3,2). Values for this two cells at median are 39.48 and 39.02 respectively. Median for spillover from Microsoft to AT&T stands at almost 17%. As an interesting observation we present the fact, that median and mean for elements on diagonal is lower than estimated value on whole sample, but it is exactly the opposite for off-diagonal elements. This may suggest that longer samples overestimates on-diagonal elements and underestimates off-diagonal elements, but we will not investigate this observation any further.

From rolling window analysis we can now draw conclusions. It should be obvious, that these stocks are connected to each other in more complex way than with simple correlation. Apple is price leader in this three stock portfolio. If we observe change in price of Apple it is more then likely, that Microsoft and AT&T will move in same direction. Moreover this do not propagate in same magnitude in opposite direction. Therefore if we observe changes in price of either Microsoft or AT&T, probability that price of Apple will change is rather low. Moreover connectedness between Microsoft and AT&T is also unsymmetrical. Microsoft is price leader to AT&T. Price of AT&T is in more than half of the times dependent more on prices of Apple and Microsoft than on auto-correlation.

Here we mention that, when we calculate simple correlation between Apple and Microsoft we arrive to value -0.013[1]. This would suggest, that these stocks are more or less independent on each other. However we now already know that this is not true. Simple correlation between Apple and AT&T is computed to 0.282. This estimate is close to our computed spillover from Apple to AT&T, however it does not capture the fact, that AT&T almost do not influence Apple.

We finish this section with summary of previous results. We have estimated classical correlation and spillover measure and we have obtained following results. Spillover measure provides asymmetric information about connectedness of series, which we should take into account when predicting future development. Also simple correlation fails to show dependence between stocks of Apple and Microsoft, which is clearly visible thanks to spillover effect.

6.2 Hawkes process

Results in this section are based on specification of Equation 4.14 with function g as specified in subsection 4.2.3. To ease computational requirements we have specified basic intensity close to zero based on Hardiman *et al.* (2013), where authors state, that in financial series Hawkes approximation is based on long memory process with unknown origin, therefore initial intensity should be close to zero, since occurrences of events should be based only on descendancy and immigration rather than on base intensity.

We have done numerical optimization based on MLE method and obtained

[1]Note that simple correlation is measure which returns values in [-1,1].

very interesting results which are vastly different from our findings from VAR model.

$$\vartheta = \begin{pmatrix} 0.496 & 0.707 & -0.501 \\ -2.376 & 0.097 & 2.512 \\ -3.392 & 3.813 & 2.841 \end{pmatrix} \tag{6.1}$$

This suggests that dynamics for extreme values are more connected across series and moreover that connections can also be negative. Since we have obtained very volatile results we present them in non-normalized form. From this it can be seen that Apple (the price leader from VAR analysis) behave the most according to expected behavior. It has stable non-explosive excitations. Positive value for self-excitation is must, because negative values would suggests for behavior not observable on market, where extreme values would be rare and probability of next extreme event would be smaller after different extreme event. And because we observe volatility clustering, we assume that extreme values produce more extreme values. Overall intensity for Apple sums to 0.7, which perfectly corresponds to findings in Hardiman *et al.* (2013) for exponential kernels.

Branching coefficients on non-diagonals can have different signs and magnitudes, this is not against economic intuition. As stocks can be interconnected in various ways. Negative terms are bit surprising, since it means, that extremes in one stock can effectively suppress probability of shock in different series.

We would like to pinpoint terms $\vartheta_{2,1}$ and $\vartheta_{3,1}$. We interpret this high negative values as suppressing power for intensity process of Microsoft and AT&T. Probability , that extreme value will occur is smaller if we have witnessed extreme of Apple in recent past. This would suggest that shocks in Apple's stocks do not propagate to others and even serves as stability measure. We however already know that spillover effect is still present, but it will not lead to extreme values.

Important to notice are also terms $\vartheta_{2,3}$ and $\vartheta_{3,2}$. This suggests that extreme values of Stock 2 and Stock 3 are highly connected. Intensity function of either of these stocks will increase if the other one experiences extreme value.

In following paragraphs we will discuss results from rolling window analysis. It is important to note, that we will draw conclusions based on estimations that covers one whole day and one whole week. Therefore this is only subsample of full rolling analysis. We decided for this step for two major reasons. First of all, in day analysis we have no systematic gaps, this makes estimated underlying process more connected to observed values. And secondly, since our rolling window is time based and not observation based, it may happen that some windows would contain exactly same observations, due to before mentioned systematic gaps. This would invalidate descriptive statistics, since this estimates would be more influential.

First of all let us state that spillovers to own asset stick to positive values. This comes very naturally and it would require further investigation if we would have found otherwise. Moreover low estimate from full sample about $\vartheta_{2,2}$ is probably influenced by some cluster of observations or maybe by systematic gaps, since daily estimates are very similar to $\vartheta_{1,1}$ and it is also expected for diagonal elements to have more positive value, then the one we estimated on full sample. From this analysis we can also see, why $\vartheta_{3,3}$ is estimated as such high number. We can see cluster of very high estimates between May and beginning of July, with one more high peak in early August. Also estimates of $\vartheta_{3,1}$ and $\vartheta_{3,2}$ are much higher for full sample than for daily values. This can be explained be high number of large peaks, which are also present in estimates of $\vartheta_{3,1}$ and $\vartheta_{3,2}$ but also by higher estimate for decay function for full sample. This, as we already know serves as counterpart to matrix ϑ, and reduces intensity of the process. This can be also the reason for low value of $\vartheta_{2,2}$, since decay function for Microsoft is rather small in comparison to Apple and AT&T for full sample. However mean values in rolling analysis for decay functions ω are very comparable, and for medians ω_2 is again lower than ω_1 and ω_3 but not as much as for full sample estimate.

To comment on off-diagonal elements, it seems that estimates from long sample have bigger magnitude than estimates based on daily prices. This can be caused by systematic gaps, however this requires further research. Otherwise estimates have usually correct signs and fluctuates around long-term estimate. As exception $\vartheta_{1,2}$ seems to fluctuate more around zero than around estimate from previous part, as almost half estimates are below zero. And $\vartheta_{1,3}$ stay most of the times in positive numbers while long-term estimate was negative. Reason for this however remains unknown.

Plotted results can be found in Appendix. Figure A.8 presents matrix for all branching coefficients for daily values. Median for every branching coefficient is plotted as dot dashed line in appropriate region. We also included estimate from full sample and plotted it as dotted line. Figure A.9 shows all branching coefficients in one plot. Thanks to this plot we can visually check how estimates for branching coefficients behave together. Colors used in this plot are same as in Figure A.8. We can see that when one extreme happens usually estimate for different coefficient is also extreme. This can be caused either by some global event on the market, or by very influential stock, since branching coefficients may offset each other.

In the analysis which uses one week as a window we observe large spikes in weeks 29, 36, and 39. We suspect that in these weeks some major events occurred. If and what events caused this will be analyzed in next section.

To summarize this part we conclude, that results from Hawkes methodology are more surprising than for VAR methodology and also different from our findings in VAR. This suggests, that there is different underlying process under extreme distribution and common prices. We also recommend using rolling window analysis as robustness check for estimation, since we have found out, that these estimates for different time periods may vary.

Chapter 7

Economic background

It may seem unusual that we present economic background of our analysis after econometric results, but we do this because this work should serve mainly as introduction to different methodologies, and results from econometric part are therefore more important. Another reason is that we will in this section use results from chapter 6. Namely those from the rolling sample analysis, as we will try to match shocks, that can be found in time series, to announcements and events of our companies.

7.1 Company introduction

In this section we will briefly introduce companies which are used in this work.

Apple Inc. in previous sections referred to as Apple, is an American technological company founded on April 1, 1970 with initial public offering on December 12, 1980. It started as company for selling personal computers and nowadays is manufacturing and selling various types of electronics. It is the largest IT company based on revenue and total assets Chen (2015).

Microsoft Corporation in previous sections referred to as Microsoft, is again American technological company, founded on April 4, 1975 with initial public offering in 1986. Main focus is on software technologies namely operating systems for personal computers and mobile devices. And also develops office suites. Company also produces hardware technologies, such as gaming devices and after acquisition of Nokia it is also interested in mobile phones market.

AT&T Inc. is an American telecommunication company founded on October 5, 1983. It is the one of the largest providers of telecommunication services in the United States.

We can see that all companies used in our analysis are active in technological industry, have long history and can be considered as representatives in hardware, software and services. As an addition to those three stocks we will use index S&P 500 as a representative of whole economical sector. Thanks to this we should be able to distinguish whether events influenced whole market, for case of changes in all assets, or only technological segment, if changes are more visible in our three stocks than in S&P 500, or only single sector in technological segment, in case of major changes in estimates of only one stock.

To finish introduction of used time series we present:

S&P 500 is an American composite index of 500 large companies traded on New Your Stock Exchange or on NASDAQ Stock Market. Stocks on the list need to fulfill given requirements. It is important to note, that all three stocks of our interest are parts of S&P 500, but S&P 500 represents whole spectrum of the market, where our stocks can be viewed as representatives of IT industry. Price development as plot is presented as Figure A.4 in Appendix.

7.2 Periods of interest

In this section we will try to find outlying estimates for Hawkes methodology. We focus more on Hawkes process part instead of VAR part, because we are interested in shocks and events on market, which should be more visible this way, since Hawkes processes, as we are using them, are modeling exactly these extreme values. As a tool for finding extremes we have constructed threshold based on normal distribution as mean plus 1.96 times standard deviation, this should contain 95% of observation, and noted all days which have estimates above this threshold. We will be mainly focused in high values of estimates, since they cause extreme price changes on market, which is more economically important, than extremes in opposite direction, as extremely low values for branching coefficients only lowers probability of extreme occurrences, rather than causing negative price changes as one could have thought.

We will be finding extremes for each branching coefficient separately, but then we will use only one set of suspicious days for each stock, which will consist of all suspicious days across rows of branching matrix. We will look for extremes in $\vartheta_{(i,\cdot)}$ for i$^{\text{th}}$ stock.

We present results of this analysis in Table 7.1.

Additionally we provide Figure A.9 in appendix with all estimates plotted together, where peaks are visible, and it also can be seen, that extremes are

Apple	Jan-16	Feb-16	Feb-26	Mar-02	Mar-10
	Mar-27	Apr-10	Apr-14	May-01	May-06
	May-21	May-22	May-29	Jun-04	Jun-19
	Jul-13	Sep-07			
Microsoft	Jan-13	Feb-03	Feb-16	Apr-08	May-01
	May-21	Jul-16	Aug-28		
AT&T	Jan-13	Feb-16	Apr-14	May-01	Jun-04
	Jun-12	Jul-02	Jul-16	Sep-10	
Multi-stock dates	Jan-13	Feb-16	Apr-14	May-01	May-21
	Jun-04	Jul-16			

Table 7.1: Days with extreme estimates

usually occurring across multiple coefficients. And also that negative extremes are usually accompanied with positive ones for different coefficient, therefore they will appear in our dates. Figure A.8 provides color codes for Figure A.9, as each branching coefficient is plotted on its own graph. Additionally dotted line represents estimate from full sample, and dot dashed line median of plotted values.

7.3 Events

We assume, that market reacts to events almost instantly, so we will be looking for events in times specified above and we hope that we will find some correlation. For major events as quarter reports and big releases we will also allow for days close to our estimate, since there may be leaking of information to the market, which can cause early reactions based on expectations. Delayed reaction can be explained by mixed signals from announcements or the fact that news can happen after trading day and trades will be realized next day. As source for list of events we will be using official websites of corporations since they provide list of all press releases made by corporations.

7.3.1 Apple Inc.

Apple's first quarter results were presented on January 27, which do not correspond with any date with extreme behavior. This report do not present any shocking information. Apple reports highest-ever revenue, which was driven by sales of iPhone, and it was expected in longer horizon.

Main and the only main announcement in February was quite large investment of €1.7 billion into new European data center. This new data center is powered by renewable energy, which can have influence on investors. This information was released on February 23, which is an day in which Apple hit the highest price per stock in year 2015, and February 26 is first day after announcement where price was higher than day before. Connection of these events is probably small, but maybe still relevant.

March 10 is a day right after chunk of announcements about new 12-inch MacBook, updates to 13-inch MacBook, and also about date when Apple Watch will be available. All these events have positive impact on expectation about Apple's earnings so we assume, that occurrence of extreme prices were highly probable, and because of it, it appeared on our list.

On April 10 Apple opened pre-order for Apple Watch as well as preview availability in stores of all big markets. This did not visibly influenced stock prices, however our analysis found this day precisely. The announcement was made day earlier, but we assume that trying it in stores produced volatile reviews on this product and with it also some extreme trades on the market, which are however not visible on simple graph.

The second quarter report was announced on April 27, with again positive news as Apple reported strong revenue growth, again because f sales of iPhone and Macs but also by all-time best performance of the AppStore, the software part of the company. This was accompanied with announced increase of dividends per share by 11 %. This also resulted in peak price of stocks, but it did not showed up in our analysis, so we assume, that change in price was very smooth.

In May Apple introduced new MacBook, however date do not line up well with our estimation. We have quite large chunk of extreme intensity coefficients in May, but Apple's press release do not present any reason why should be so. Only relevant announcement is on May 19, introducing new computer set-ups, but timing does not line up very nicely. On the market, stocks of Apple seem to change directions very often across whole May. This may be reason for this suspiciously high branching coefficients.

In June Apple introduces new operating system on 8^{th}, but it seems that it does not influence market or branching coefficients for this period. On 4^{th} Apple Watch is promised to enter also not so large markets, and it seems that, Apple Watch news are connected to high estimates of branching coefficients. They also state, in their article, that sales are exceeding estimates of Apple, which can also influence market processes.

On July 20 Apple stock reached last peak before significant drop, which happened across following month. This market downturn surprisingly occurred after positive fiscal report for third quarter.

Last suspicious date is September 7, which is two days before introduction of new iPhone 6s and new iPad. It may be possible that these events are connected, as introduction of new iPhone was anticipated, and changes on market could be influenced by discussion about new iPhone. Apple in the new model also incorporated new technologies for touch screens. Every new possibility about interface brings high discussions with many pros and cons. Therefore price changes on stock could have been influenced by this and resulted in producing high estimates for branching coefficient.

It seems that except for Apple Watch announcements, connection between high branching coefficient and official press releases is rather small. This serves as confirmation about no possibility for predictions on market. It would be interesting to see on larger sample, if it is possible that introducing the new line of products, not only new versions of products in already known line, can really influence market in such a way, that extremes are more probable and this reflects to branching coefficients

7.3.2 Microsoft

The hunt for events of Microsoft begins on January 13, right before this day Microsoft partners with two large restaurant companies to bring technologies to dinning tables. In one case servers are using Microsoft tablets to speed up ordering process, in second case Microsoft introduces kiosks for self-ordering. This may seem as minor events, but it further sets position of Microsoft as partner to many.

Major event was dated to January 21, when Microsoft introduces new operating system Windows 10. We would expect that this event will be heavily influential, but it is not the case for this release. Neither in our estimates nor on market can be found any unusual development.

February and March are quite uninteresting in news as well as in market price development. Microsoft stocks are traded for price around $42, after major fall at the end of January from $47 mark. And this low price window persists until April 22. As the major events of this time period we mention dividend announcement on March 10, which however did not influence market price in visible way, neither it appeared in our analysis.

On April 7 Microsoft announced date for third quarter report, which is April 23, even though April 7 is just one day before one of dates with extreme values, this announcement is probably of low importance for market, since it has no real value in sense of possible economic expectation. Reaction to actual announcement is very different story. After announcement price for Microsoft stock rose to values similar to those before January 27. This announcement possibly ended period of low value of Microsoft stock. Still, it do not appear as suspicious day for changes in underlying process.

In may we have not found any significant announcements and minor ones do not fall into our estimated dates.

June was more interesting for Microsoft, as it introduced new phones (June 3) and also declares quarter dividend (June 9). Still the stocks of Microsoft had declining tendency through whole May and June. We did not found any dates with outlying estimates in June and really there is no reason to think that there should be one. Here we note, that quarter dividends, which can be view as major event, are well-expected and also dividend per share is not an surprising number as it did not changed since last dividend payments.

Launch of Windows 10 on July 28 is surely large event in Microsoft history however it do not influenced market or our estimation in any visible way.

Starting on August 18 Microsoft faces big drop in price on market, however we did not find any company specific reason why this should happen, neither in official press release nor by our analysis. We will bring this date back in section about S&P500.

To summarize this section, Microsoft stock prices and press announcements points to different days than our analysis about probability of extreme prices. This further supports theory about unpredictability of the market. However our analysis still remains relevant as our computation remains valid, only we were not able to match these days with higher occurrences of extremes with real life events.

7.3.3 AT&T

AT&T news starts with report for fourth quarter of previous year on January 27. In the report we can find growth in wireless revenue and in number of subscribers. This gives idea about future growth of AT&T.

In March company announced on March 23, that first quarter report will be available on April 22. On March 27 they declared quarterly dividends with standard value of $0.47 per share and quarter. This value per share increased in December 2014 from $0.46. Market do not seem to react to these news in any way and our analysis do not consider these days to be important.

On April 22 AT&T reported results from first quarter. More or less positive. They reported that transformation of company towards future needs goes well and also demand for services grows. Two days after company's board of directors is reelected. This two events probably lead to increase in stock price after period of rather low price.

On June 26 another round of dividends in standard value was announced and also introduced new member to board of directors. In this time period stocks of AT&T had highest market price from observed period. After July 7 we observe systematic drop in price of stocks, with biggest drop again on August 25.

At the end of July AT&T finished acquisition of DIRECTV and with this step AT&T became the largest pay TV provider. This also allowed AT&T to offer better internet services to wider spectrum of customers. Therefore it should be seen as very positive action. However we are unable to join this with any day from our analysis.

We end this analysis with quarter dividend on September 25 again with same value as in previous quarters and no suprising market developments.

7.3.4 S&P 500

This part is devoted to peaks and pitfalls of S&P500 which corresponds to dates from Table 7.1. We are interested if major changes on the market can influence estimation of branching coefficient. We inspect this hypothesis by looking at the dates that occurred in more than one stock in our previous analysis. They can also be found in Table 7.1 in the last row.

On January 13 market reports more than 75% of stocks ending in red numbers, dragged by falling price of crude oil in energy sector, as of July 25, 2016 Zacks equity research states in report for yahoo.finance.com (Zacks) .

April 14 again had most of the sectors in red numbers, so S&P 500 was again declining, but not so much as in January. (Zacks)

No significant events happened on May 1, but May 21 was quite interesting, as S&P 500 reached its maximum for observed period, but this increase happened slowly across longer period of time and it is not a sharp peak.

June 4 was a positive day for S&P 500 as seven sectors ended in the green. (Zacks)

July 16 was again positive in terms of market value, but without any large changes on market.

Fall between August 17 and August 25 is not pinpointed by our model, however it is large drop in terms of S&P 500 value. Because this event influenced broad spectrum of sectors and also all companies in our analysis, it is not surprising that branching coefficients did not climb over threshold. This is due the fact, that if market moves together occurrence of extreme values remain relatively same. Hawkes model reacts to such event theoretically only by change in decay coefficient, because extremes do not spillover differently, only probability of extreme event is overall higher.

7.4 Summary for economic background

To conclude over this section we state, that markets seem unpredictable even when using Hawkes methodology as we were using it in this work. Moreover even major announcements do not seem to have predictable influence on market. We conclude, that this may be due to rational expectations of investors, since we did not found any shocking news in our observed period. Only company, where we were able to match some events with extreme values for extreme coefficients, was Apple, and it happened only in the beginning of time period of our interest. Still we do not consider analysis about extreme branching coefficient worthless. It may reveal deeper structure on the market, and connections to real announcements would be beneficial but it is not required. If we assume that extreme transactions occur on the market not because of the external events, but because of the underlying process, it is important to have model for that, and Hawkes process is designed specially to serve this purpose. With this in mind, we still suggest for future methodology to keep economic background in check, to observe how models behave when facing structural changes. In our case we were facing large market drop in late August, but it seems that both models are suitable for this situation, as they not show any large changes for estimates in this period.

Chapter 8

Conclusion

To conclude about connectedness of examined financial series and reviewed literature we suggest that modeling of high-frequency data can yield important information about nature of price development and underlying processes. Both models bring important conclusion about behavior of the market. Since our findings from VAR model and Hawkes process model are so vastly different, we suggest that appropriate model for stock prices would consist of at least two parts, where one is based on discrete VAR model and serves as basic price evolution model and second one is based on continuous modeling. The first, more basic model, should then be enriched with special model of jumps using continuous Hawkes processes, which can explain majority of sudden jumps in prices. We recommend for future research to also use robustness check. Rolling window analysis proved itself as valuable addition to both models. Thanks to it we were able to confirm, that our results from VAR are robust. For Hawkes process we have discovered, that on full sample branching coefficients were slightly higher than median for daily values, however direction in which extreme values influence itself was confirmed, thanks to this additional analysis. We therefore confirmed that stocks on market are connected both by spillover measure from VAR as well as by branching coefficient from Hawkes analysis. This results strongly suggest that simple correlation measure is not sufficient when calculating connectedness of stocks on the market. Both because estimate of the correlation may underestimate connectedness of series and because connection is not symmetrical.

Conclusion about our series is following. Apple is the price leader in our three stock portfolio and AT&T is the price follower. Changes in price of Apple will propagate to both other stocks, while opposite direction will not have as significant effect. Moreover Apple serves as suppressor of extreme values for Microsoft and AT&T. If we combine this two findings together we conclude, that when extreme price change occurs to Apple both other stocks will react, however probability that these changes will be of extreme nature is rather small. As also interesting finding we present, that prices of Microsoft and AT&T are not so connected on every day basis, however are very influential in extreme values. This would suggest that these stocks floats more or less independently, but when some shock happens to either of them, the second one has also high probability of extreme value to occur. Connection to real events on the market is not visible, we suggest, that announcements are already incorporated in market process in time of announcement.

As open questions for further analysis we leave methodology of selecting appropriate kernel for Hawkes process as power-law kernel should yield better results, but have more complicated implementation. Also the possibility for modeling full price development using Hawkes process with complex function g, which can even incorporate VAR methodology. As other way how to enrich our analysis we suggest using asymmetric models for both VAR and Hawkes process. In the modern literature we can usually find, that market behaves differently when facing positive and negative price changes, we would suspect, that branching coefficient for large negative price changes would be overall higher than for positive changes, so this asymmetric methodology can provide even better estimates for the true model of the market.

Bibliography

Aït-Sahalia, Y., J. Fan, & D. Xiu (2010): "High-frequency covariance estimates with noisy and asynchronous financial data." *Journal of the American Statistical Association* **105(492)**: pp. 1504–1517.

Al-Deehani, T. & I. A. Moosa (2006): "Volatility spillover in regional emerging stock markets: a structural time-series approach." *Emerging Markets Finance and Trade* **42(4)**: pp. 78–89.

Andersen, T. G., T. Bollerslev, & N. Meddahi (2005): "Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities." *Econometrica* **73(1)**: pp. 279–296.

Bacry, E., I. Mastromatteo, & J.-F. Muzy (2015): "Hawkes processes in finance." *Market Microstructure and Liquidity* **1(01)**: p. 1550005.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, & N. Shephard (2008): "Designing realized kernels to measure the ex post variation of equity prices in the presence of noise." *Econometrica* **76(6)**: pp. 1481–1536.

Bartram, S. M., G. Brown, & R. M. Stulz (2012): "Why are us stocks more volatile?" *The Journal of Finance* **67(4)**: pp. 1329–1370.

Barunik, J., E. Kocenda, & L. Vacha (2014): "How does bad and good volatility spill over across petroleum markets?" *Available at SSRN 2435511* .

Bauwens, L. & N. Hautsch (2009): "Modelling financial high frequency data using point processes." In "Handbook of financial time series," pp. 953–979. Springer.

Booth, G. G., T. Martikainen, & Y. Tse (1997): "Price and volatility spillovers in scandinavian stock markets." *Journal of Banking & Finance* **21(6)**: pp. 811–823.

Bormetti, G., L. M. Calcagnile, M. Treccani, F. Corsi, S. Marmi, & F. Lillo (2015): "Modelling systemic price cojumps with hawkes factor models." *Quantitative Finance* **15(7)**: pp. 1137–1156.

Chen, L. (2015): "The world's largest tech companies: Apple beats samsung, microsoft, google." Accessed July 25, 2016. http://www.forbes.com/sites/liyanchen/2015/05/11/the-worlds-largest-tech-companies-apple-beats-samsung-microsoft-google/#68a1aa80415.

Diebold, F. X. & K. Yilmaz (2009): "Measuring financial asset return and volatility spillovers, with application to global equity markets." *The Economic Journal* **119(534)**: pp. 158–171.

Diebold, F. X. & K. Yilmaz (2012): "Better to give than to receive: Predictive directional measurement of volatility spillovers." *International Journal of Forecasting* **28(1)**: pp. 57–66.

Embrechts, P., T. Liniger, L. Lin *et al.* (2011): "Multivariate hawkes processes: an application to financial data." *Journal of Applied Probability* **48**: pp. 367–378.

Engle, R. F. & J. R. Russell (1998): "Autoregressive conditional duration: a new model for irregularly spaced transaction data." *Econometrica* pp. 1127–1162.

Filimonov, V. & D. Sornette (2012): "Quantifying reflexivity in financial markets: Toward a prediction of flash crashes." *Physical Review E* **85(5)**: p. 056108.

Hardiman, S. J., N. Bercot, & J.-P. Bouchaud (2013): "Critical reflexivity in financial markets: a hawkes process analysis." *The European Physical Journal B* **86(10)**: pp. 1–9.

Koop, G., M. H. Pesaran, & S. M. Potter (1996): "Impulse response analysis in nonlinear multivariate models." *Journal of econometrics* **74(1)**: pp. 119–147.

Liniger, T. J. (2009): *Multivariate hawkes processes.* Ph.D. thesis, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 18403, 2009.

Pesaran, H. H. & Y. Shin (1998): "Generalized impulse response analysis in linear multivariate models." *Economics letters* **58(1)**: pp. 17–29.

Wyart, M., J.-P. Bouchaud, J. Kockelkoren, M. Potters, & M. Vettorazzo (2008): "Relation between bid–ask spread, impact and volatility in order-driven markets." *Quantitative Finance* **8(1)**: pp. 41–57.

Zacks Equity Research (2015): "Stock market news." Accessed July 25, 2016.
https://www.zacks.com/stock/news/203636/stock-market-news-for-january-13-2016
https://www.zacks.com/stock/news/170617/stock-market-news-for-april-14-2015
https://www.zacks.com/stock/news/177159/stock-market-news-for-june-04-2015.

Zhang, L. (2011): "Estimating covariation: Epps effect, microstructure noise." *Journal of Econometrics 160* **no. 1**: pp. 33–47.

Appendix A

Appendix

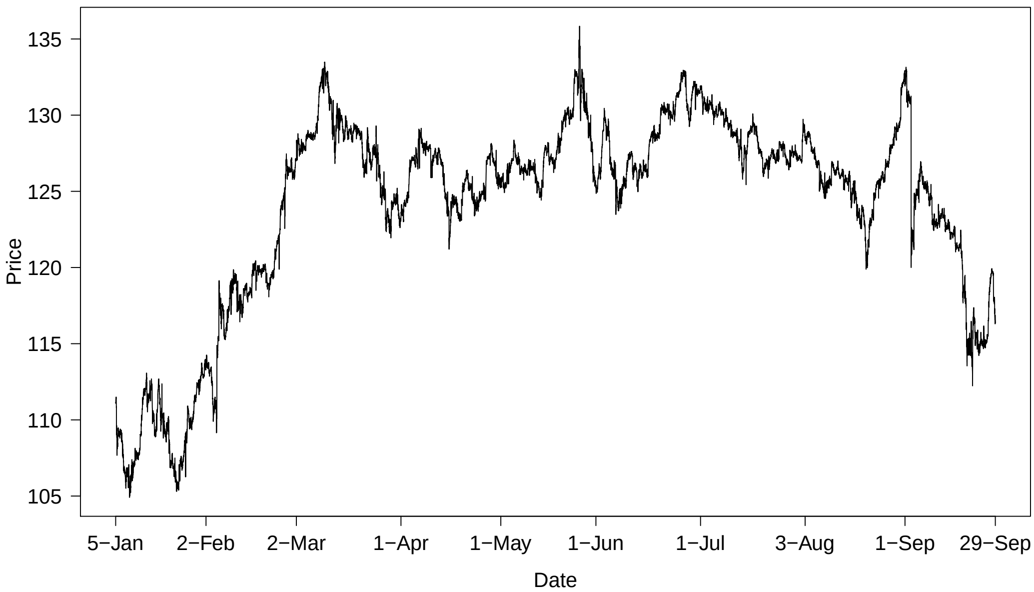

Figure A.1: Price of Apple

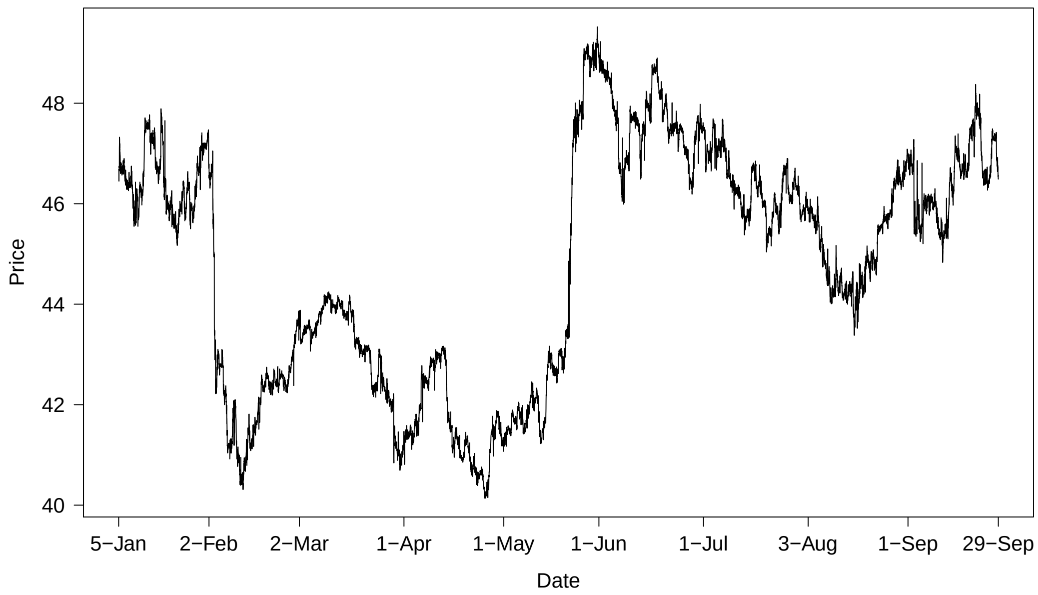

Figure A.2: Price of Microsoft

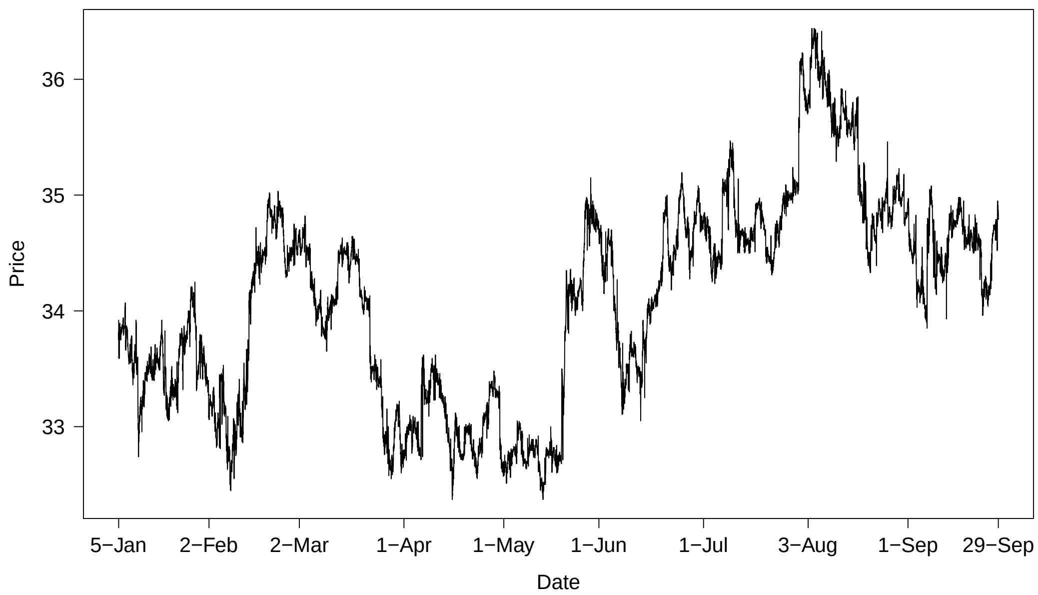

Figure A.3: Price of AT&T

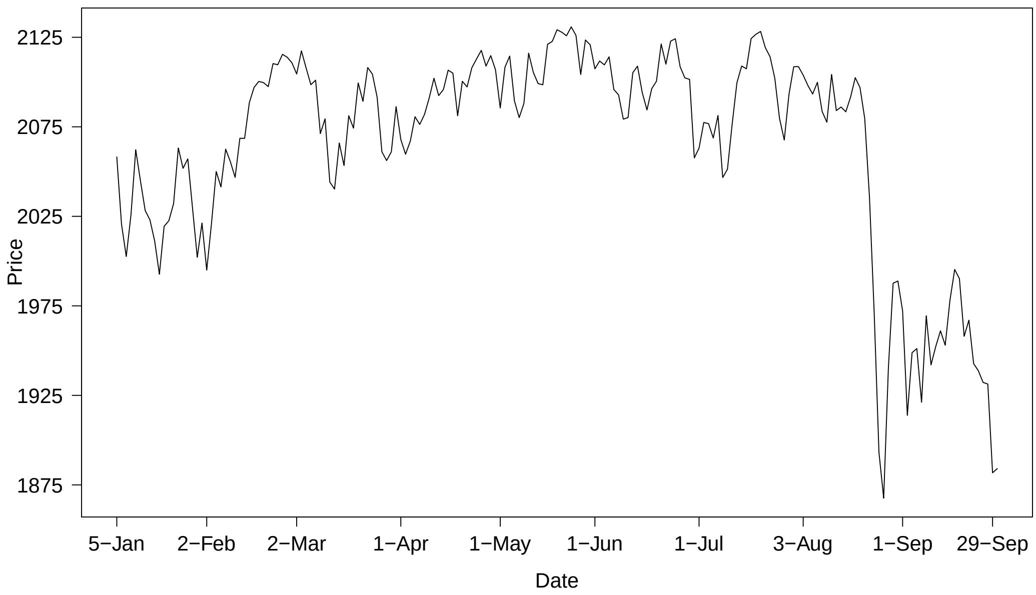

Figure A.4: Price of S&P 500

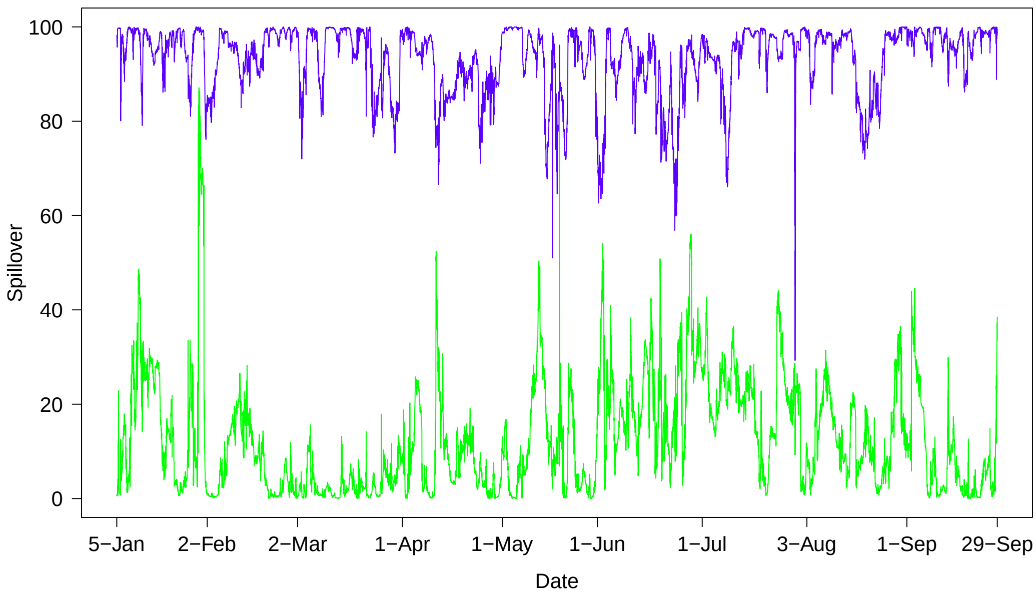

Figure A.5: Rolling spillover to Apple's stock

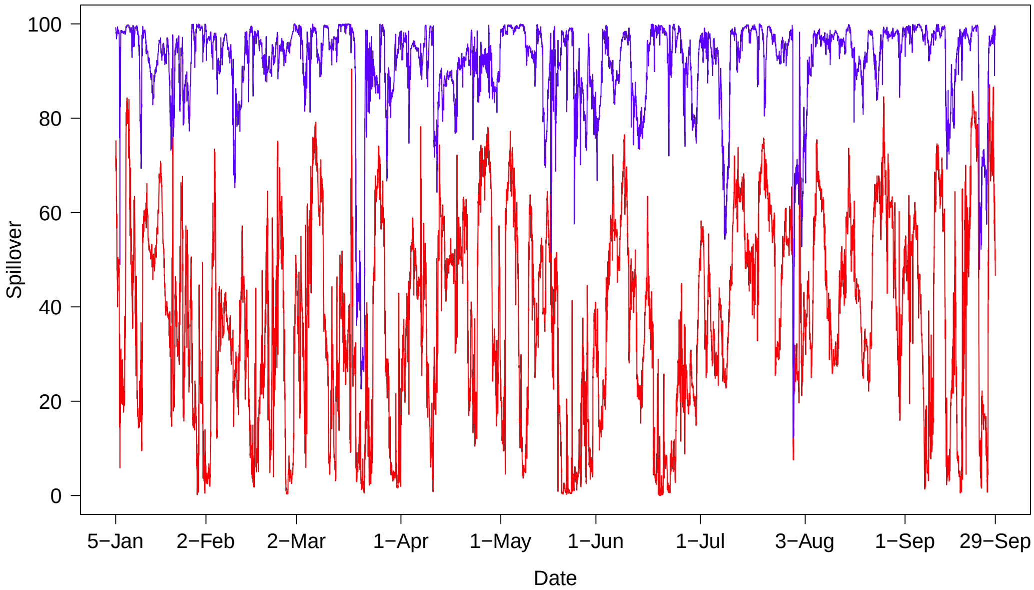

Figure A.6: Rolling spillover to Microsoft's stock

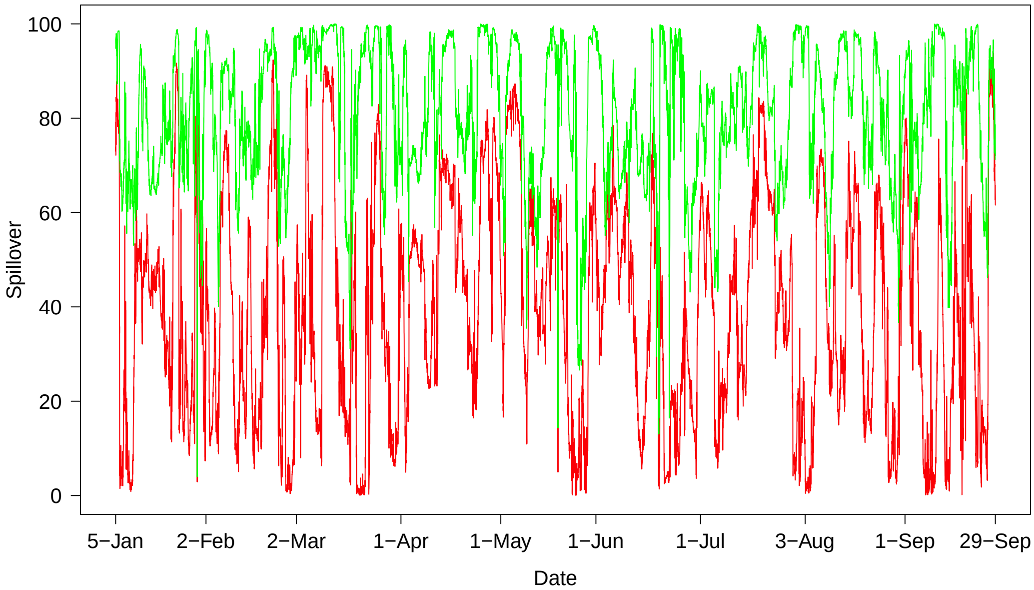

Figure A.7: Rolling spillover to AT&T's stock

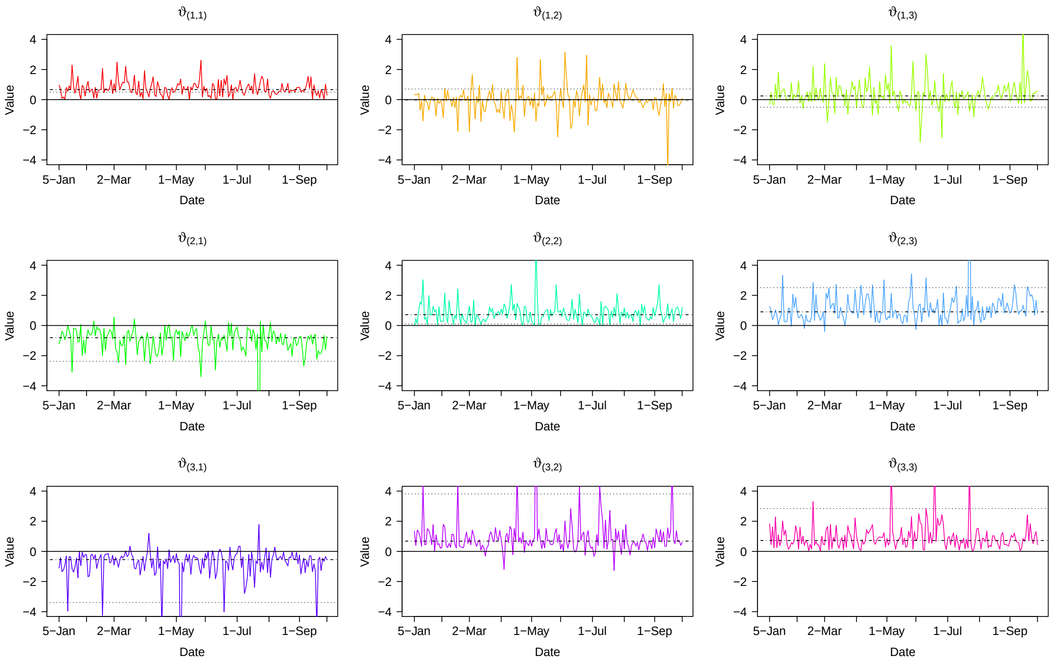

Figure A.8: Matrix of branching coefficients

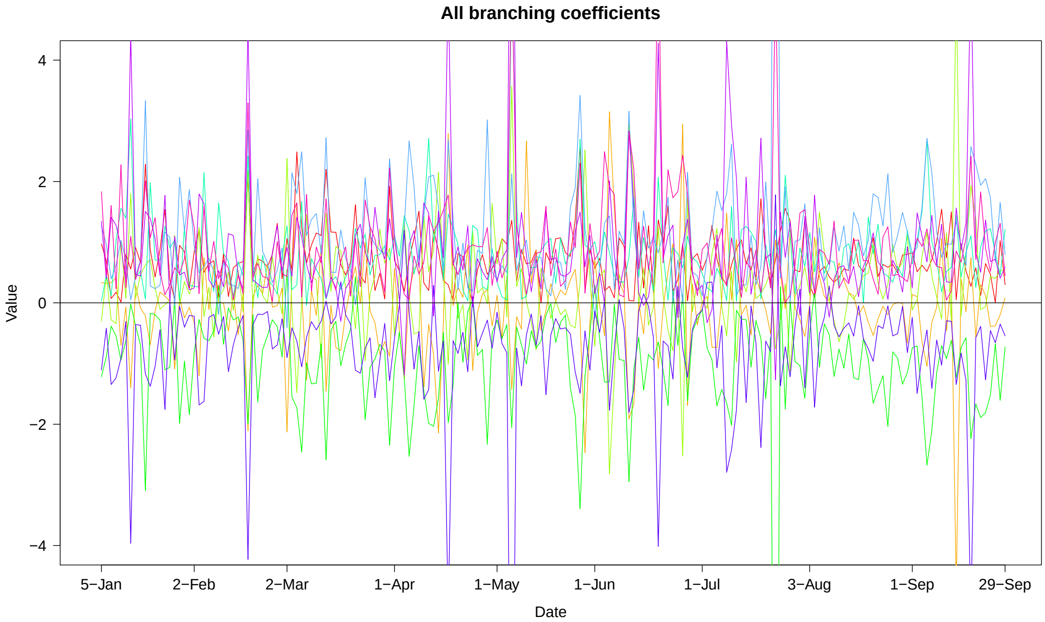

Figure A.9: Branching coefficients