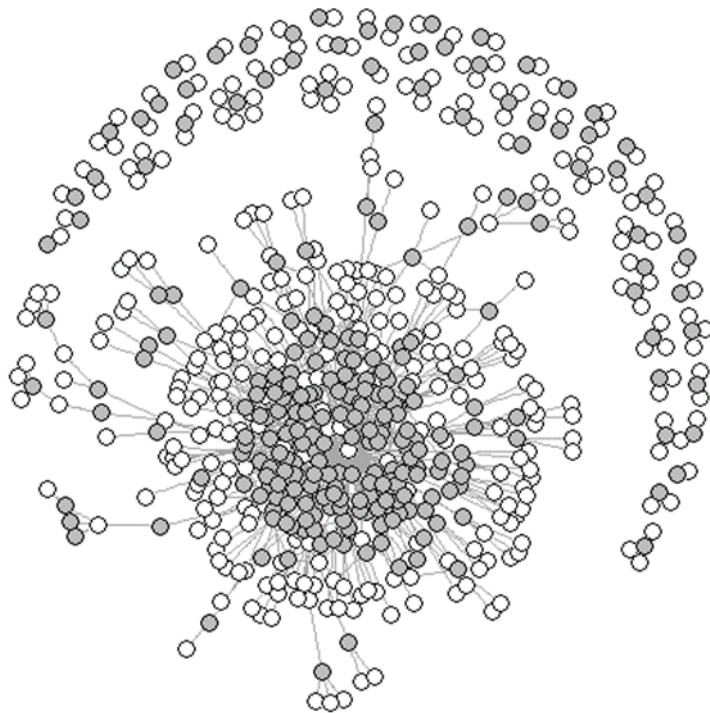


Christian Koch

**The Social Network:  
A Graph-Theoretical Perspective on  
Money, Banks and Stock Markets**



Discussion Paper

Technische Hochschule Nürnberg

13 February 2018

## **Abstract**

This discussion paper examines money, banks and stock markets from a network perspective. It aims at providing a new view on the economy and its evolution. The document starts with an analysis of Adams Smith's theory of prehistorical barter. It demonstrates the high complexity that arises in an economy that is based on a direct exchange of goods. The next section shows how this complexity could be reduced radically by introducing money as a central measure for the price of goods in ancient Egypt. Based on the work of Laszló Barabási and Réka Albert a network topology for a modern financial market is designed. This topology is in line with the historical development of European banks from medieval to modern times. Finally, a network of stock market investors is extracted from recent empirical data. The graph is used to simulate stock prices. The simulation demonstrates how fat tails in stock returns are generated by a social network, without violating the efficient market theory of Eugene Fama. The paper thus bridges the gap between classical and behavioural models.

*Keywords:* Money, Banks, Stock Market, Network, Graph, Economy, Monetary System, Fat Tails, Efficient Market Hypothesis

## Table of Contents

1.	Adam Smith and Barter .....	4
2.	The Ancient Egyptian Monetary System.....	6
3.	Barabási, Albert and Financial Markets .....	9
4.	A Network of Stock Investors .....	12
5.	Markowitz, Fama and Market Efficiency .....	18
6.	Time and Information .....	24
7.	Social Networks and Fat Tails.....	25
8.	Keynes, Herding and the Beauty Contest.....	34
9.	Conclusion .....	36
A.	References.....	37

## List of Figures

Figure 1:	Barter economy with 3 producers.....	4
Figure 2:	Barter economy with 30 producers.....	5
Figure 3:	Economy with a centralised monetary system .....	6
Figure 4:	Degree distribution of a star network with 31 nodes.....	7
Figure 5:	Scale-free network constructed with the Barabási-Albert model.....	9
Figure 6:	Degree distribution of a scale-free network with 31 nodes .....	10
Figure 7:	Network of international companies and their investors .....	12
Figure 8:	Network of international investors .....	13
Figure 9:	Degree distribution of the investor's network .....	15
Figure 10:	CDF of investor degrees .....	16
Figure 11:	Historical and simulated prices of the IBM share .....	19
Figure 12:	Historical and simulated daily returns of the IBM share.....	20
Figure 13:	Distribution function of historical and simulated returns.....	22
Figure 14:	Historical and simulated returns on a time scale .....	23
Figure 15:	Human Time System and External Events .....	24
Figure 16:	Historical and simulated prices of the IBM share .....	26
Figure 17:	Historical and simulated daily returns of the IBM share.....	28

Figure 18: Q-Q-Plot of historical and simulated returns..... 29  
 Figure 19: Historical and simulated returns on a time scale ..... 29  
 Figure 20: Distribution of signals ..... 30

**List of Tables**

Table 1: Top ten investors by degree ..... 14  
 Table 2: Results of a bootstrapping test for different distributions..... 16  
 Table 3: Measures for historical and simulated returns ..... 21  
 Table 4: Measures for historical and simulated returns ..... 28  
 Table 5: Results of 50 simulations of the IBM stock..... 31  
 Table 6: Results of 50 simulations of stocks of Dow Jones Index..... 33

**List of Equations**

Equation 1: Stock price following a Normal Random Walk ..... 18  
 Equation 2: Stock price following a Fat Tails Random Walk ..... 25  
 Equation 3: Expected return per signal of a Fat Tails Random Walk..... 26  
 Equation 4: Standard deviation of returns of a Fat Tails Random Walk ..... 27

## 1. Adam Smith and Barter

In his famous work ‘An Inquiry into the Nature and Causes of the Wealth of Nations’ Adam Smith proposes that prior to monetary systems human economies were based on barter.<sup>1</sup> In such an economy, money does not exist. Producers exchange goods directly. A butcher exchanges meet with beer produced by a brewer. From a network perspective the system of barter can be modelled as shown in the diagram below.

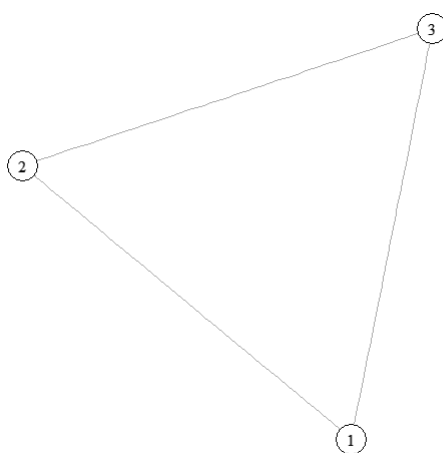


Figure 1: Barter economy with 3 producers

The network (also called graph) visualises an economy consisting of three producers of different goods (1-3). The producers/goods are represented by circles. In network science these circles are called nodes or vertices. Since no money exists in the economy, the value of a good can only be expressed in terms of other goods. The value of good 1 is expressed by the amount of good 2 and good 3. The relation between the goods is visualised by a line between the circles in the above diagram. In network science a relation between two nodes is also called an edge.

The complexity that arises from a system of barter gets obvious as soon as more producers/goods are modelled within a network. The diagram below visualises an economy with 30 goods.

---

<sup>1</sup> (Smith, 1776, p. 22)

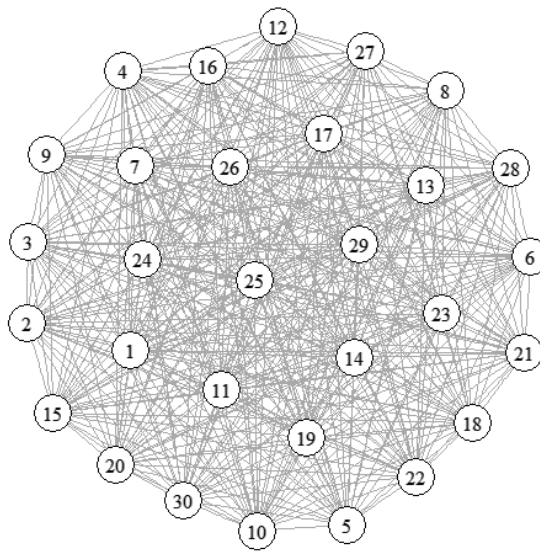


Figure 2: Barter economy with 30 producers

To exchange goods, the value of any product has to be defined by the amount of any other good in the economy. Such a graph is called a fully connected network. Every node has a relation to all other nodes.

The number of connections of a node, is called its degree. The degree is an important measure to describe the characteristics of a network. Amongst other things, it can be used to calculate the centralisation of a graph.<sup>2</sup> The above diagram has a degree centralisation of 0.<sup>3</sup> Every node has the same number of connections. The network is totally decentral. There is not a single node that has a higher degree than its associates.

With every new good that is produced in a barter market, a rising number of edges has to be added to the network. A barter economy thus gets increasingly complex and inefficient. An economy with 30 goods has 435 edges. An economy with 300 goods has 44,850 edges.<sup>4</sup> This might explain why there is no empirical evidence for the existence of a barter economy from an anthropological point of view.<sup>5</sup>

---

<sup>2</sup> (Luke , 2015, p. 93)

<sup>3</sup> Calculated with R & iGraph using the function `centr_degree`

<sup>4</sup> Calculated with R & iGraph

<sup>5</sup> (Graeber, 2011, p. 28)

## 2. The Ancient Egyptian Monetary System<sup>6</sup>

A simple solution for reducing the complexity of a barter economy is to introduce money as a measure for valuing goods. This is documented for the ancient Egyptian economy (2750-2150 BC).<sup>7</sup> The Egyptians measured the value of goods with a central unit called ‘shat’. As many other currencies, the shat was linked to gold.<sup>8</sup> The value of a shat in terms of goods was defined by government administrations. Other cultures in Asia Minor later materialised their currencies in the form of gold and silver coins.<sup>9</sup>

Providing a central currency as measure for the value of goods reduces the complexity of an economy radically. This is visualised by the diagram below.

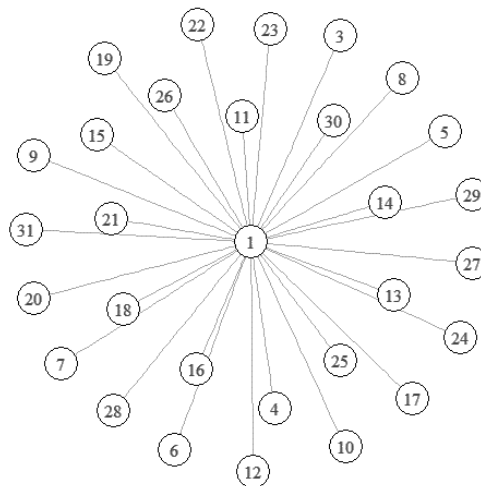


Figure 3: Economy with a centralised monetary system

In the middle of the diagram (1) the currency is modelled. Every good can be measured with this central unit. The value of one good in terms of another good can be calculated easily, without explicitly defining it. Introducing a new product to the market only requires one new edge in the network. A network with such a topology

---

<sup>6</sup> Parts of chapter 2 and 3 have been pre-published in the Wikipedia article on central banks ([https://en.wikipedia.org/wiki/Central\\_bank](https://en.wikipedia.org/wiki/Central_bank)). The corresponding paragraphs were copied into Wikipedia from an early draft of this essay.

<sup>7</sup> (Vantieghem, 2012)

<sup>8</sup> Early Mesopotamian economies used silver for this purpose (Goetzmann, 2017, p. 59).

<sup>9</sup> (Metcalf, 2016, pp. 43-44)

is called star graph.<sup>10</sup> The above shown network has 30 edges. A star network with 300 nodes has 300 edges accordingly.<sup>11</sup>

Compared to a fully connected network, a star network has a different degree distribution. The central currency has a relation to all other nodes. All goods have one relation to the central currency. The degree distribution of a star network with 30 goods and one currency is visualised by the histogram below.

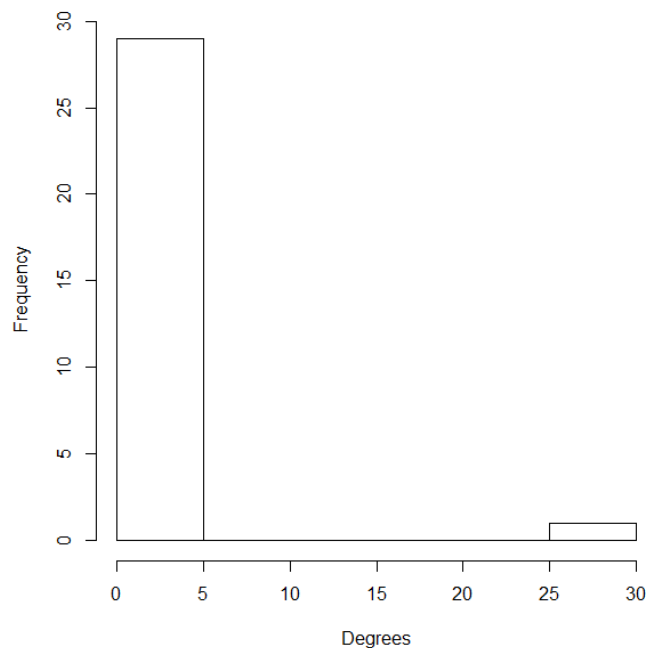


Figure 4: Degree distribution of a star network with 31 nodes

The central currency has a degree of 30 (since it is connected to all goods). All goods provided in the economy have a degree of 1. The degree centralisation is 1.<sup>12</sup> The network is totally centralised. All nodes are connected to one central government administration, managing the currency and defining the value of goods.

Besides the advantages of such a one-tier monetary system, there are drawbacks as well. The more an economy grows, the more power and responsibility is assigned to the central provider of money. The same is true for administration effort. When an economy with a centralised monetary system grows, the central administration will get more work and more power. It becomes a single point of failure. If the

<sup>10</sup> (Luke , 2015, p. 97)

<sup>11</sup> Calculated with R

<sup>12</sup> Calculated with R & iGraph using the function centr\_degree

central administration fails due to incompetency, inefficiency or fraud, the whole financial system is at risk.

In the modern world, highly centralised monetary systems can be found in communist economies. The financial system of the German Democratic Republic (GDR) was entirely controlled and operated by government institutions.<sup>13</sup>

---

<sup>13</sup> (Lipschitz & McDonald, 1990, p. 151)

### 3. Barabási, Albert and Financial Markets

In his book ‘Network Science’ Albert-László Barabási describes another type of graph that can be found in social networks as the World Wide Web. It is called scale-free network. Its degree distribution follows a power law.<sup>14</sup>

Together with Réka Albert, Barabási developed a model to explain the topology of a scale-free network by a mechanism that is called preferential attachment.<sup>15</sup> If a new node is added to a scale-free network, it will connect to a node with many connections more likely than to a node with fewer connections. The following diagram visualises a network with 31 nodes that was constructed with the Barabási-Albert model.<sup>16</sup>

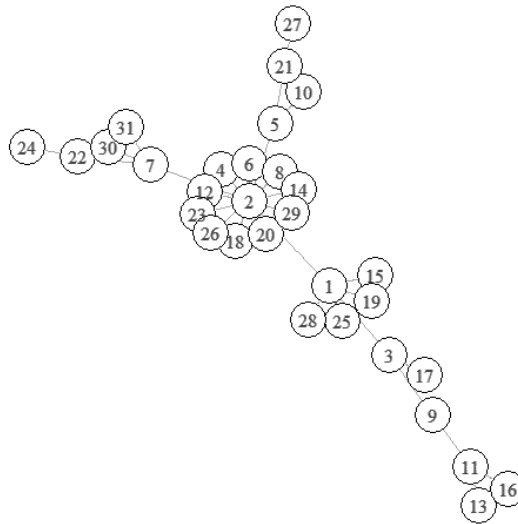


Figure 5: Scale-free network constructed with the Barabási-Albert model

The network shown above, has a central node (1) that is *indirectly* connected to all other nodes. Yet, the network is not fully centralised. The first node created in the graph (1), is connected to other nodes that again have connections to third nodes (e.g. 2, 5 and 7).

The degree distribution of the above shown scale-free network is visualised in the diagram below.

<sup>14</sup> See (Barabási, 2016, pp. 115-122). The economist Vilfredo Pareto documented such a distribution in 1897 when he analysed the structure of income in Italy (Pareto, 1897, pp. 299-345).

<sup>15</sup> (Barabási, 2016, pp. 167-174)

<sup>16</sup> Simulation with R & iGraph using the function `barabasi.game` with `power = 1.5`

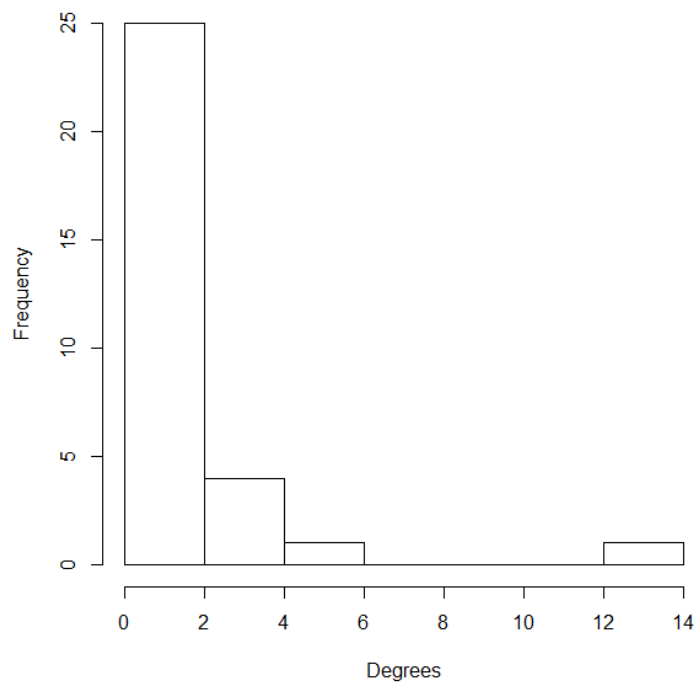


Figure 6: Degree distribution of a scale-free network with 31 nodes

The network consists of some nodes with many edges and other nodes with fewer connections. The degree centralisation of the network equals 0.39.<sup>17</sup> An economy that has a monetary system, based on this network topology, distributes the management of its currency over several institutions with varying importance. This way, a balance between centralisation and decentralisation is realised.

The topology of a scale-free network is in line with the historical development of the European financial market from the medieval to modern times. In the medieval and the early modern period, a network of professional public and private banks was established in Southern and Central Europe.<sup>18</sup> The institutes built a new tier in the financial economy. The monetary system was still controlled by government institutions, mainly through the coinage prerogative. Banks, however, could use book money to create deposits for their customers. Thus, they had the possibility to issue, lend and transfer money autonomously without direct governmental control.

A good example of those autonomous institutes are the exchange banks. In 1609 the Amsterdam Wisselbank was established as a first institute. Further exchange banks were located in Hamburg, Venice and Nuremberg. The institutes offered a

<sup>17</sup> Calculated with R & iGraph using the function `centr_degree`

<sup>18</sup> (Collins, 2012, pp. 221-225)

public infrastructure for cashless international payments.<sup>19</sup> They aimed to increase the efficiency of international trade and to safeguard monetary stability. The exchange banks thus fulfilled comparable functions to modern central banks.<sup>20</sup> They even issued their own (book) currency, called Mark Banco. A similar decentral financial system can be found in the US-American economy of the 19<sup>th</sup> century.<sup>21</sup>

Whether the topology of the banking system in the medieval and the early modern period in Europe matched a scale-free network, cannot be proven in this essay. Certainly, can be said that the responsibility for the financial system was distributed over multiple tiers from then on. The modern banking network is neither totally decentral nor fully centralised. This topology is typical for social networks. The next chapter elaborates the structure of the modern global stock market.

---

<sup>19</sup> (Collins, 2012, p. 223)

<sup>20</sup> (Kurgan-van Hentenryk, 1997, p. 39)

<sup>21</sup> (Lowenstein, 2015, pp. 199-200)

## 4. A Network of Stock Investors

Getting data about the network of public and private banks in the western economies is difficult. Extracting information on stock market investors from the Internet is not. The following diagram visualises a graph of shareholders of 259 international companies. The data was retrieved on 01 July 2017.<sup>22</sup> At that time, the enterprises were listed in the indexes Dow Jones Industrial Average, DAX 30, SMI, ATX, CAC 40, FTSE 100 and FTSE MIB.

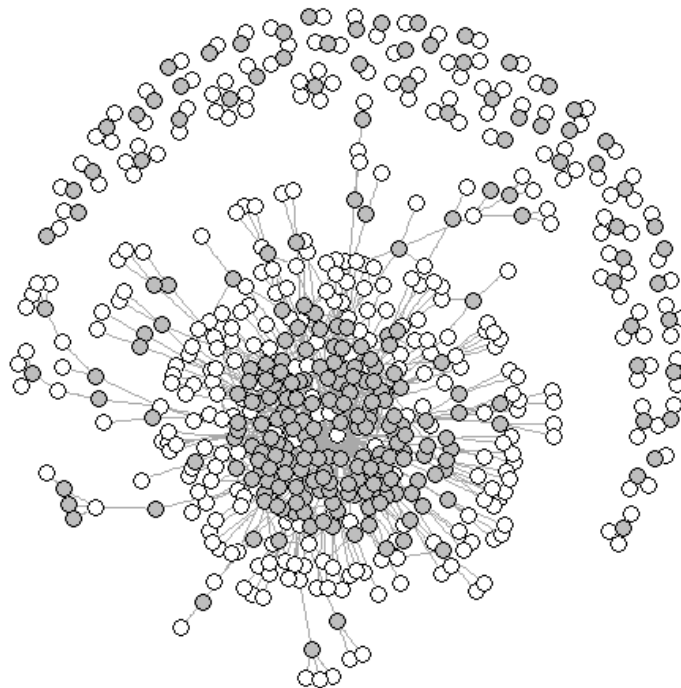


Figure 7: Network of international companies and their investors

The type of graph shown above is called a bipartite network. It contains two different kinds of nodes: the grey vertices represent companies, the white nodes investors.<sup>23</sup> The graph reveals which shareholders invested in which companies. It comprises 348 different shareholders that owned a significant part of the examined firms. Free-float shareholders are not included in the dataset. If present, they were filtered out. The graph connects investors through 13,037 edges to companies.

---

<sup>22</sup> The information was extracted with data crawled from yahoo.com, google.com and ariva.de using the tools R & Spark. All graphs were created in R with the library iGraph.

<sup>23</sup> An investor can be a person or a company. If a company that is included in the network, owns a significant share of another firm in the dataset, it is listed as investor and as a company. Treasury shares are not included in the dataset.

To produce a unipartite network that shows only investors and their relations, the graph was transformed using a specialised function in R.<sup>24</sup> The resulting investor's network is shown below. Every shareholder is represented by a node. If two investors are owners of the same company, they are connected by an edge.

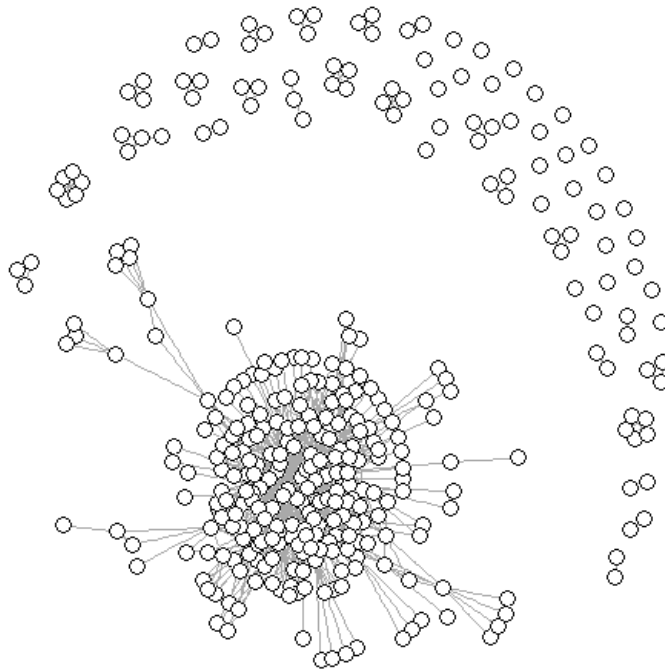


Figure 8: Network of international investors

The network models 1,069 relations. It consists of several subnetworks, called components. In the centre there is a big component connecting 244 investors. Along the margins, shareholders are shown that are not invested in many companies and only have a few relations or even none.

Table 1 lists the top ten investors of the network by their degree. The bottom of the table shows the investors with the lowest degrees. The company BlackRock is on top of the list, having connections to 183 other shareholders. BlackRock is followed by Capital Group Companies with a degree of 71. At the bottom of the ranking, several shareholders with a degree of zero are documented. These investors owned one or several companies entirely without other significant shareholders. An example is Walton Enterprises. At the time of the empirical study, the investor possessed 46.7% of the retail corporation Walmart. The remaining 53.3% were free floating

---

<sup>24</sup> The function `bipartite.projection` of the library `iGraph`

shares. Since minor investors are not included in the dataset, Walton Enterprises has a zero degree.

Rank	Investor	Degree
1	BLACKROCK	183
2	CAPITAL GROUP COMPANIES	71
3	INVESCO	54
4	NORGES BANK	53
5	FMR	37
6	HARRIS ASSOCIATES	34
7	LEGAL & GENERAL	30
8	CREDIT SUISSE	29
9	AMERIPRISE FINANCIAL	27
10	FIRST EAGLE INVESTMENT MGMT	26
...	...	...
344	ARTÉMIS	0
345	WALTON ENTERPRISES	0
346	CLUBTRE SRL	0
347	SOFIL SAS	0
348	MUTUELLES AXA	0

Table 1: Top ten investors by degree

The subsequent diagram shows the degree distribution of the investor's network.

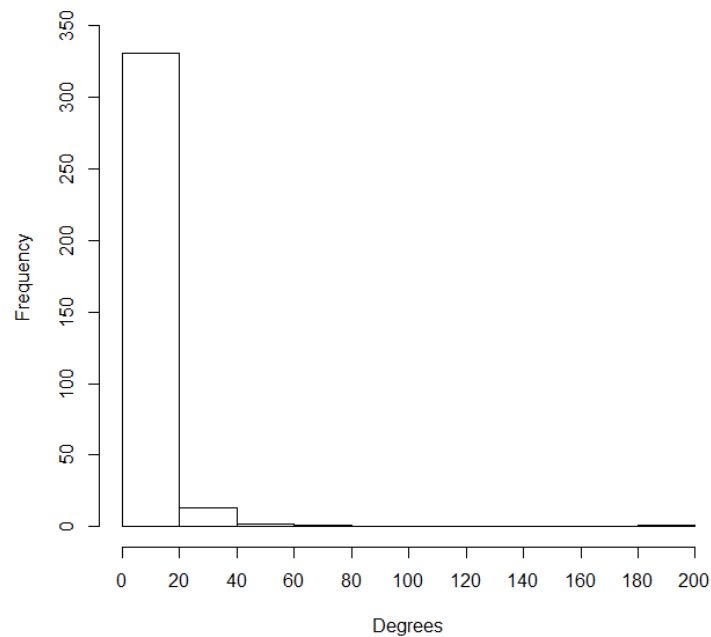


Figure 9: Degree distribution of the investor's network

BlackRock appears on the right side of the diagram with a degree of 183. It is followed by Capital Group Companies and other shareholders. The majority of the investors have a relatively low degree. The degree centralisation equals 0.51.<sup>25</sup>

The question that now arises is whether the investor's network is a scale-free graph as described in chapter 3. In this case its degree distribution should fit a power law. Other network types can be modelled more appropriately with different distribution functions as the Poisson and the log-normal distribution.<sup>26</sup>

Fitting a power-law distribution to the degrees of the investor's network requires the removal of all shareholders without connections. This is necessary since a function based on a power law cannot result in zero.

Figure 10 shows the cumulative distribution function (CDF) of the degrees of the investor's network.<sup>27</sup> The values are visualised by circles. The lines in the diagram represent the CDF of different standard distributions, fitted to the empirical values. The black solid line shows a power-law distribution, the red dashed line a log-normal distribution and the blue dotted line a Poisson distribution.<sup>28</sup>

<sup>25</sup> Calculated with R & iGraph using the function `centr_degree`

<sup>26</sup> (Barabási, 2016, pp. 147-151)

<sup>27</sup> To be strictly accurate, it is an Empirical Cumulative Distribution Function (ECDF).

<sup>28</sup> Graph created with R using different functions of the library `powerLaw`.

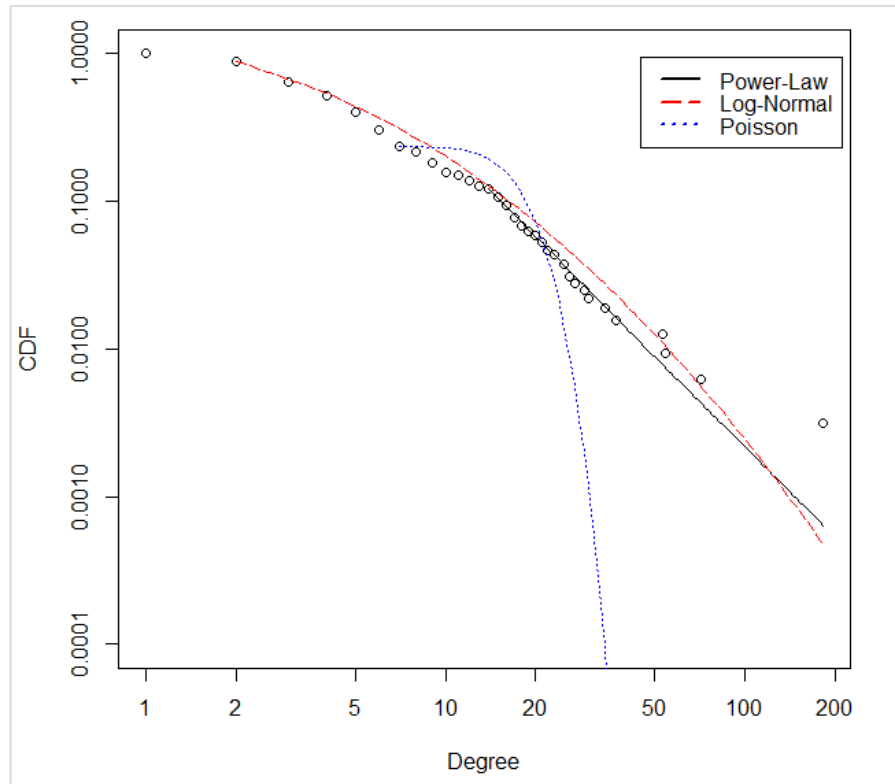


Figure 10: CDF of investor degrees

The diagram makes clear that a Poisson distribution does not fit the empirical degrees. The power-law and the log-normal distribution seem more appropriate. This impression is supported by the values listed in the following table.

Distribution	Simulations	p-value
Power-law	100	0.98
Log-normal	100	0.01
Poisson	100	0.00

Table 2: Results of a bootstrapping test for different distributions

Table 2 contains the results of a bootstrap hypothesis test.<sup>29</sup> For every standard distribution in the above diagram 100 bootstrap simulations were executed. The p-value in Column 3 indicates whether a distribution can be considered plausible to model the empirical degrees. Analogous to subsequent experiments, the hypothesis will be accepted if the p-value exceeds 0.05.<sup>30</sup>

<sup>29</sup> Performed with the R library `powerLaw`. For details see (Gillespie, 2015, pp. 3-10).

<sup>30</sup> The interpretation of p-values is discussed in detail by (Wasserstein & Lazar, 2016).

The high p-value of 0.98 suggests that a power-law distribution fits the investor degrees very well, even though it covers only a part of the empirical values. This is in line with comparable empirical studies, e.g. conducted by Garlaschelli et al.<sup>31</sup> The low p-value of 0.01 indicates that a log-normal distribution is not suitable to model the investor degrees. It should, however, be considered as alternative, since it covers a large part of the empirical values.<sup>32</sup>

Modelling the stock market as a social network could shed light on a phenomenon that is intensely discussed in the financial literature: fat tails in stock returns. This topic is examined in the next chapters.

---

<sup>31</sup> (Garlaschelli, Battiston, Castri, Servedio, & Caldarelli, 2005)

<sup>32</sup> This is visualised by Figure 10. The power-law distribution fits best for  $x_{\min} = 14$  with a coverage of the empirical degrees of 12%. The log-normal distribution fits best for  $x_{\min} = 2$  with a coverage of 88%.

## 5. Markowitz, Fama and Market Efficiency

In his paper ‘Portfolio Selection’ Harry Markowitz proposes to interpret the return of a stock as a random variable.<sup>33</sup> A return is the change of the stock price from one point in time (e.g. a day) to the next point in time in percent. Based on this hypothesis, Markowitz explains why rational investors construct portfolios in favour of picking a single stock that promises the highest return.

The theory of Markowitz raises an essential question: why should the trades of rational investors result in stock returns that are randomly distributed? Should rational behaviour not lead to deterministic outcomes? This seeming paradox is resolved by the efficient market hypothesis of Eugene Fama.

In his essay ‘Random Walks in Stock Market Prices’ Fama suggests that on a fully efficient market all available information is reflected in the fair market value of a stock.<sup>34</sup> Only new information that could not be foreseen by investors (that is in fact random) can lead to price changes. Assuming that big international stock markets are efficient, their returns should follow a probability distribution. Their prices should follow a so-called Random Walk.

To simulate daily stock prices with a random walk, the following equation can be used.

$$P_{t+1} = P_t * (1 + r_t)$$

Equation 1: Stock price following a Normal Random Walk

The stock price of the next day ( $P_{t+1}$ ) is determined by the price of the current day ( $P_t$ ), changed by a random return ( $r_t$ ). The price of a stock tomorrow can consequently be calculated by the price of today adding a price change.

According to Fama’s theory, the return/change of a stock price is caused by new information that arrives on the market. Investors interpret the information and start trading the stock. This leads to a price adjustment that is reflected by the return. Newly arriving information on a market will be called *signal* throughout the rest of this paper. *Investors receive a signal that leads to a price change.*

---

<sup>33</sup> (Markowitz, 1952)

<sup>34</sup> (Fama, 1965)

To simulate stock prices with Equation 1, pseudo-random numbers produced by a computer can be used. To do that, the probability distribution of the random numbers has to be chosen preliminary. This topic is deeply discussed in the financial literature. Several models suggest a normal distribution.<sup>35</sup>

The following diagram shows the history of the stock price of the IBM share, traded in New York from 02 July 1992 to 01 July 2017.<sup>36</sup> In addition to the historical price, the results of a simulation using Equation 1 are visualised.

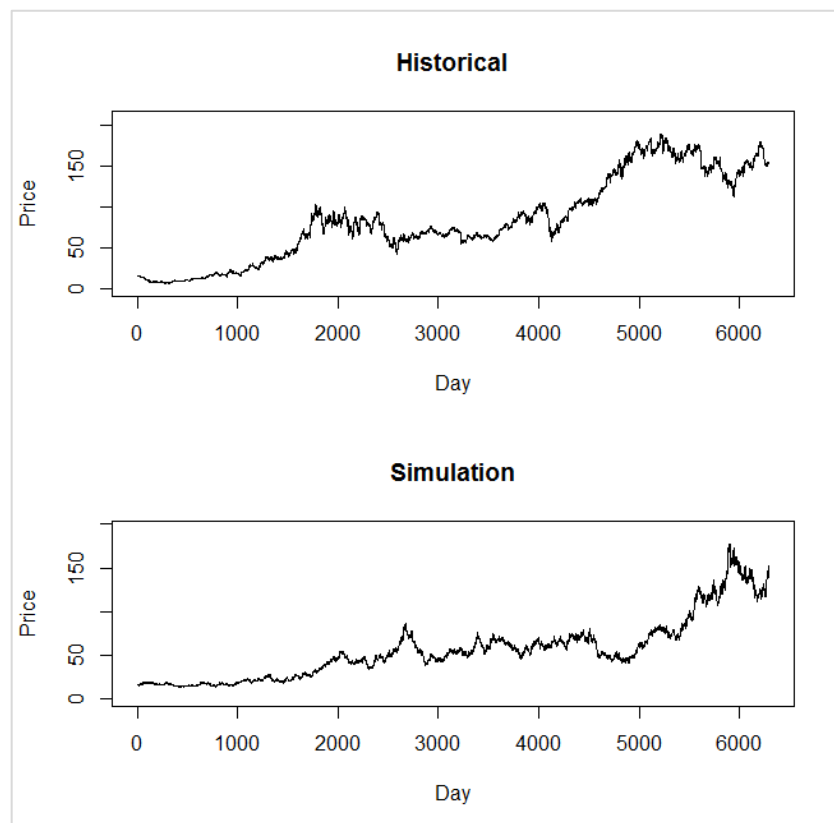


Figure 11: Historical and simulated prices of the IBM share

The simulation started with the historical price of 02 July 1992. For every subsequent trading day, it has generated normally distributed returns ( $r_t$ ) and added them to the price of the previous day. The pseudo-random numbers were generated with R.<sup>37</sup> Mean and standard deviation of the underlying normal distribution were calculated from historical returns.

<sup>35</sup> In fact, many models suggest a normal distribution for logarithmic returns. With regard to small price changes, this distinction can be ignored.

<sup>36</sup> Stock prices retrieved from Yahoo Finance.

<sup>37</sup> Generated with R using the function `rnorm(n = n, mean = changes.mean, sd = changes.sd)`; the parameters `n`, `mean` and `sd` were calculated from historical returns.

As the diagram above shows, a Normal Random Walk can be used to produce stock prices that appear similar to real quotes. Since a Normal Random Walk is based on random numbers, the development of the price can of course differ from the real history. The structure, however, seems equivalent. The following chart shows a comparison of the distribution of the historical and the simulated daily returns of the stock. It reveals significant differences.

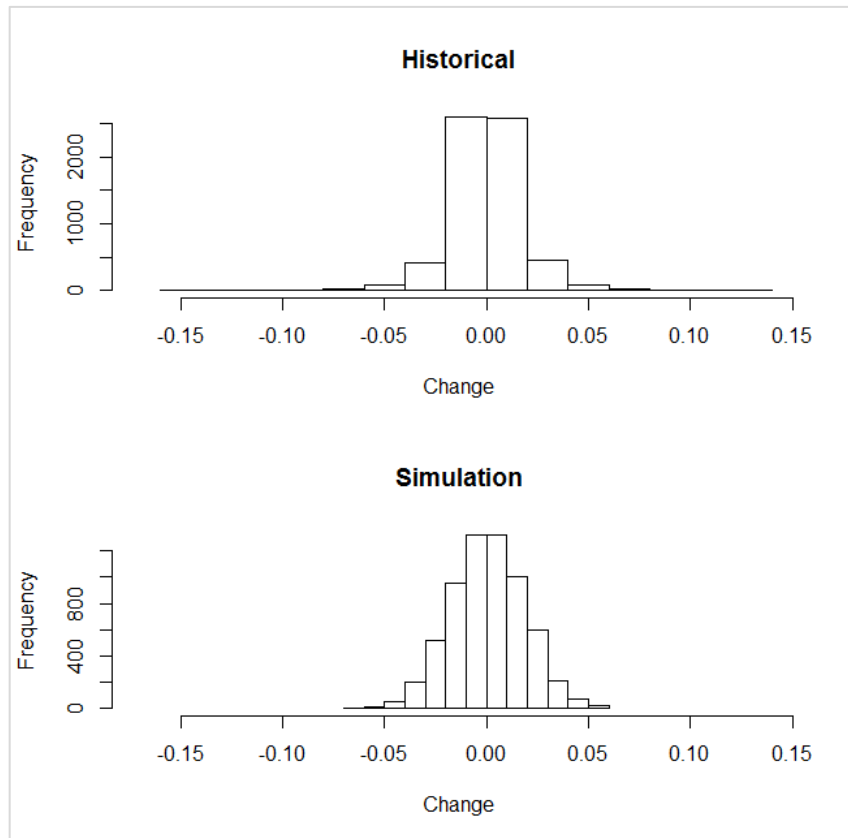


Figure 12: Historical and simulated daily returns of the IBM share

The histogram of the simulated daily returns (bottom of the diagram) follows a normal distribution. This is verified by the measures listed in Table 3. The average simulated return equals 0.05% (0.0005). The standard deviation (also called volatility) equals 1.79% (0.0179). The measures can be interpreted in the following way: the shareholders of the simulated stock would on average have earned 0.05% per day. This profit, however, would have deviated on average by 1.79% points from this value. On a good day, the return could have been higher, on a bad day it could have been lower. This is shown by the tails on the left and the right side of the histogram. In the best case the price increased by 6% (0.0569). In the worst case it fell by 7% (-0,0682) on one day.

The shape of the historical distribution is different. The mean of the historical returns is 0.05% (0.0005). The standard deviation is 1.78% (0.0178). This is compliant with the simulation. The risk of high deviations from the average return is, however, much higher. This is visualised by the fat tails on the left and right side of the histogram. In the best case the price rose by 13% (0.1316). In the worst case it fell by 16% (-0.1554) on a day. The real IBM stock was much riskier than the normal distribution suggests. This is caused by the fat tails.

The following table lists different measures that describe the characteristics of the historical and the simulated stock returns of IBM.<sup>38</sup>

Measure	Historical returns	Simulated returns
Mean	0,0005	0,0005
Standard deviation	0,0178	0,0179
Min	-0,1554	-0,0682
Max	0,1316	0,0569
Kurtosis	7,2426	-0,0847
Shapiro-Wilk normality test: p-value	0,0000	0,3269
Two-sample Kolmogorov-Smirnov test (historical vs. simulated): p-value	0,0000	

Table 3: Measures for historical and simulated returns

The means and the standard deviations are nearly identical. The highest profits and losses (Min & Max) differ enormously. Real stock returns are more volatile than normally distributed price changes. The very low p-value ( $\sim 0.0$ ) of a Two-sample Kolmogorov-Smirnov test, comparing the two distributions, confirms that they are significantly different.

The intensity of fat tails can be measured with the so-called Kurtosis of a distribution. The higher the Kurtosis of a return distribution is, the higher is the risk of radical price changes. With a Kurtosis of 7.2 (7.2426) the return distribution of IBM

<sup>38</sup> Calculated with R and the library e1071.

clearly has fat tails. This is also verified by a Shapiro-Wilk normality test. The very low p-value ( $\sim 0.0$ ) indicates that the returns do not follow a normal distribution.

The subsequent diagram shows a quantile-quantile plot (Q-Q plot) of the historical and simulated returns. It reveals as well that the distributions have different characteristics.

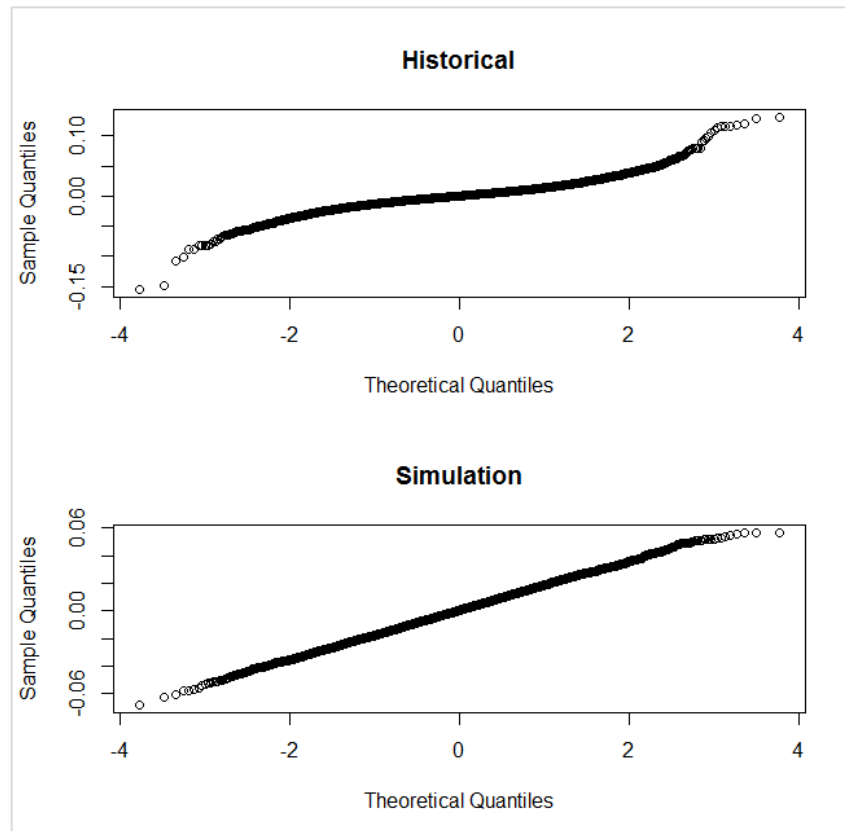


Figure 13: Distribution function of historical and simulated returns

Q-Q plots compare the quantiles of a sample to the theoretical quantiles of a normal distribution. The simulated returns (bottom of diagram) follow a straight line. This indicates that they are normally distributed. The historical returns are not linear and have fat tails.

Figure 14 visualises the daily percental changes of the historical and simulated prices on a time scale. It reveals that the simulated returns are more stable than the historical ones. The real IBM share had periods of relatively low volatility, mixed with periods of high volatility. In economics, this phenomenon is known as volatility clustering. A possible reason could be an autocorrelation of returns.<sup>39</sup>

<sup>39</sup> See e.g. (Bollerslev, 1986), (Cont, 2007).

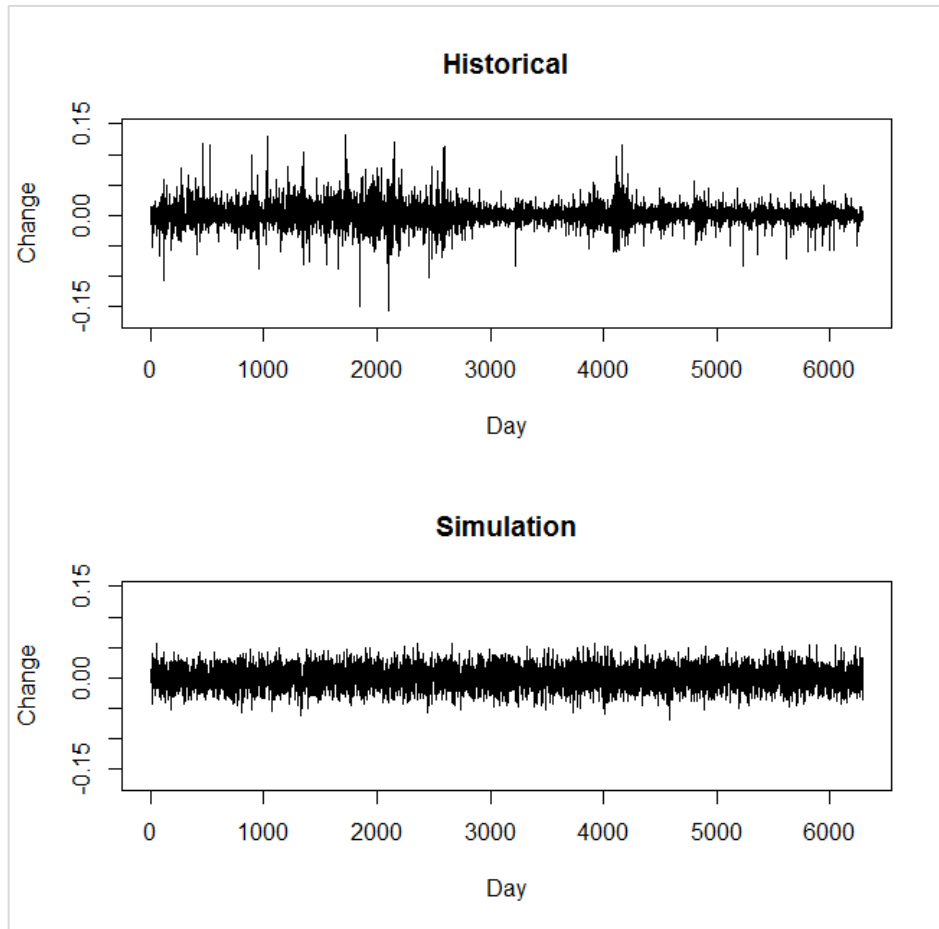


Figure 14: Historical and simulated returns on a time scale

Summarising, it can clearly be said that financial models expecting stock returns to be normally distributed underestimate risk. Several papers come to this conclusion.<sup>40</sup> Chapter 7 provides a network view on fat tails in stock returns. Preliminary, a concept of time and information is introduced in the next section.

<sup>40</sup> See (Mandelbrot, 1963), (Mandelbrot, 2006), (Taleb, 2008)

## 6. Time and Information

In his book ‘Zeit’ (English: Time) Rüdiger Safranski proposes a concept of time. He defines time as the distance between two events. No events, no time. The human time system, consisting of seconds, minutes and hours, is based on a clock that produces a sequence of equidistant events.<sup>41</sup>

A clock tries to measure time by generating ticks that have precisely the same distance. Thus, the human time system is normed and made independent from external influences. Compared to external events, the human time is, however, relative. If a human being recognises many external events in one time period (e.g. an hour), the time is running. If there are only a few events, time passes by slowly.

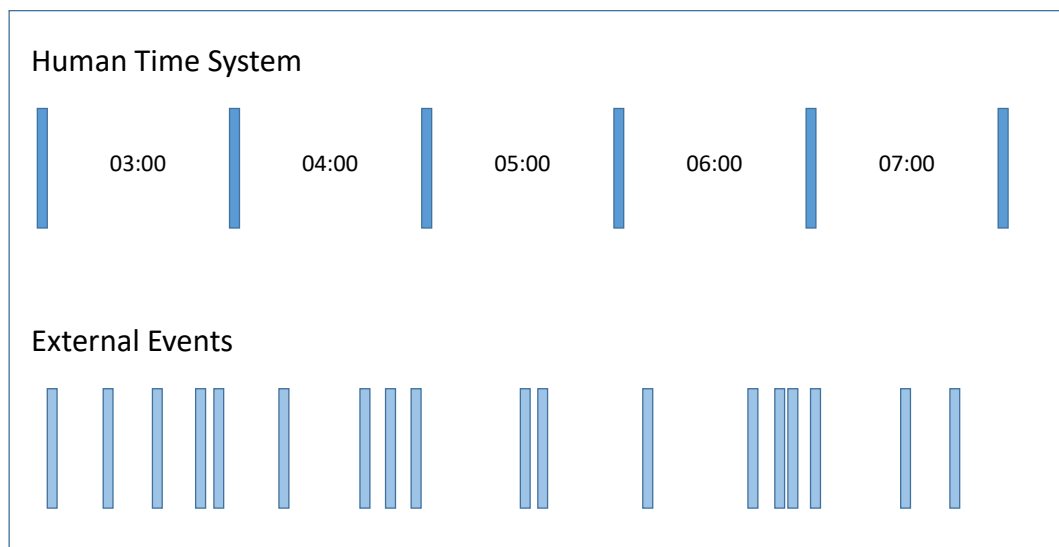


Figure 15: Human time system and external events

The figure above visualises the human time system in comparison to external events. Within a time period (hour, day etc.) there can be many events or only a few. The same is true for signals arriving at the stock market.

Stock returns are measured for a specific period of time. In the previous chapter, daily returns were analysed. Why should a new piece of information arrive only once a day on a stock market? It is very likely that on some days there are many signals, on other days there are only a few.<sup>42</sup> This idea will be applied to Random Walks in the next chapter.

<sup>41</sup> (Safranski, 2015, pp. 86-90)

<sup>42</sup> This idea is also discussed by (Mandelbrot, 2006)

## 7. Social Networks and Fat Tails

### 7.1 Concept

The stock market simulation in chapter 5 was based on an algorithm that only generated one random return per day. In line with the ideas of the previous section, a realistic algorithm should instead generate a varying number of signals for each day. On some days there should be many signals, on other days only a few. This algorithm will be called *Fat Tails Random Walk* (FT Random Walk) and is defined by the following equation.

$$P_{t+1} = P_t * \prod_{i=1}^s (1 + r_i)$$

Equation 2: Stock price following a Fat Tails Random Walk

A Fat Tails Random Walk is similar to a Normal Random Walk (Equation 1). The stock price of the next day ( $P_{t+1}$ ) is calculated by the price of the current day ( $P_t$ ) and adjusted by random returns ( $r_i$ ). The only difference is that the FT Random Walk produces not one signal per day, but several. The number of signals is defined by the variable  $s$ . Every signal leads to a percental price change, reflected by  $r_i$ .

In many cases, the different price changes within one day will balance out. They are normally distributed and therefore include positive and negative values as well. On some days, the returns will, however, go entirely in the same direction and thus lead to a major price change. This is how fat tails are produced. The phenomenon can be compared to waves in the ocean. When small waves collide, their powers normally balance out. Though, in some cases their powers sum up and create a freak wave.

This idea encourages another important question: how is the number of signals per day distributed? A normal distribution does not fit, since its values include positive and negative numbers. The number of signals instead has to be positive or zero. Experiments with the Fat Tails Random Walk have shown that the empirical degree distribution of the investor's network, discussed in chapter 4, leads to very good results.

The following diagram visualises stock prices, generated with a Fat Tails Random Walk using the degree distribution of the investor's network to produce the number of signals per day ( $s$ ).

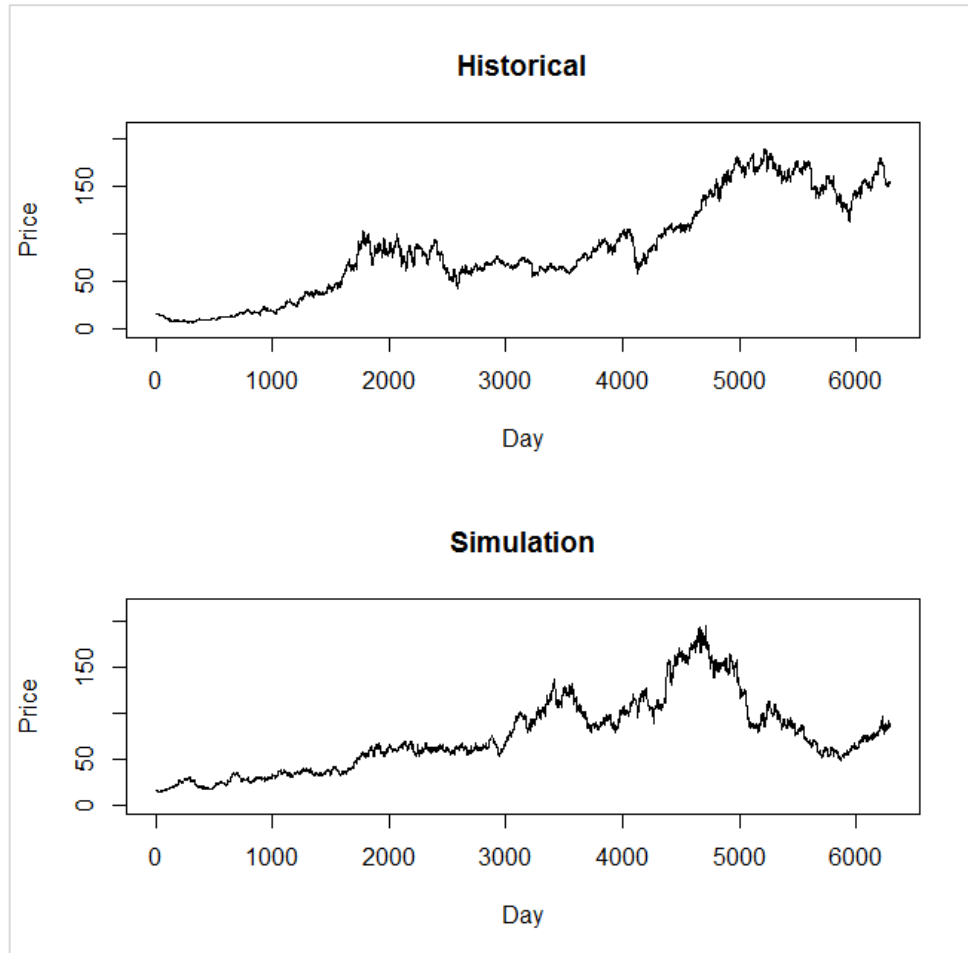


Figure 16: Historical and simulated prices of the IBM share

Analogous to chapter 5, the diagram compares historical prices of the IBM share with the results of the simulation.<sup>43</sup> The returns per signal ( $r_i$ ) were modelled with normally distributed random numbers. The expected return per signal  $E(r_i)$  was calculated with the following equation.

$$E(r_i) = \frac{\mu_{ht}}{\mu_s}$$

Equation 3: Expected return per signal of a Fat Tails Random Walk

<sup>43</sup> Stock prices retrieved from Yahoo Finance on 01 July 2017. History from 02 July 1992 to 01 July 2017.

The expected return per signal ( $E(r_i)$ ) is based on the average historical return ( $\mu_{ht}$ ), divided by the average number of signals per day ( $\mu_s$ ). A similar equation was used to calculate the standard deviation of a return per signal.

$$\sigma_{ri} = \frac{\sigma_{ht}}{\sqrt{\mu_s}}$$

Equation 4: Standard deviation of returns of a Fat Tails Random Walk

The standard deviation of a return per signal ( $\sigma_{ri}$ ) is based on the standard deviation of the historical daily returns ( $\sigma_{ht}$ ), divided by the square root of the average number of signals per day ( $\mu_s$ ).

Figuratively speaking, the simulation works in the following way: an algorithm chooses randomly one shareholder from the investor's network for every simulated day. The number of its neighbours (its degree) is interpreted as the number of signals arriving on the market. For every signal, the program generates a random return and adds it to the previous price. The algorithm produces at least one signal per day. If an investor with a degree of zero is selected, one signal is generated nevertheless.<sup>44</sup> This avoids trading days without price changes. The consequences that arise from this figurative view are discussed in chapter 8.

The above line chart reveals that historical and simulated prices have a similar structure. Figure 17 exposes that this is also true for the distribution of the returns. Both, the historical and the simulated distribution, clearly have fat tails. This is confirmed by the measures, listed in Table 4.<sup>45</sup>

The means (0.05% vs. 0.04%) and the standard deviations (1.78% vs. 1.76%) documented in Table 4 are nearly identical. The minimum and maximum values as well as the Kurtosis indicate that the simulated and historical returns have fat tails. The very low p-value of the Shapiro-Wilk normality test confirms that the returns are not normally distributed. The relatively high p-value (0.56) of a Two-sample Kolmogorov-Smirnov test, considers the two distributions as identical.

<sup>44</sup> Degrees of zero are consequently transformed to 1. An alternative approach would be the removal of all investors with a degree of zero from the network as described in chapter 3.

<sup>45</sup> Calculated with R and the library e1071.

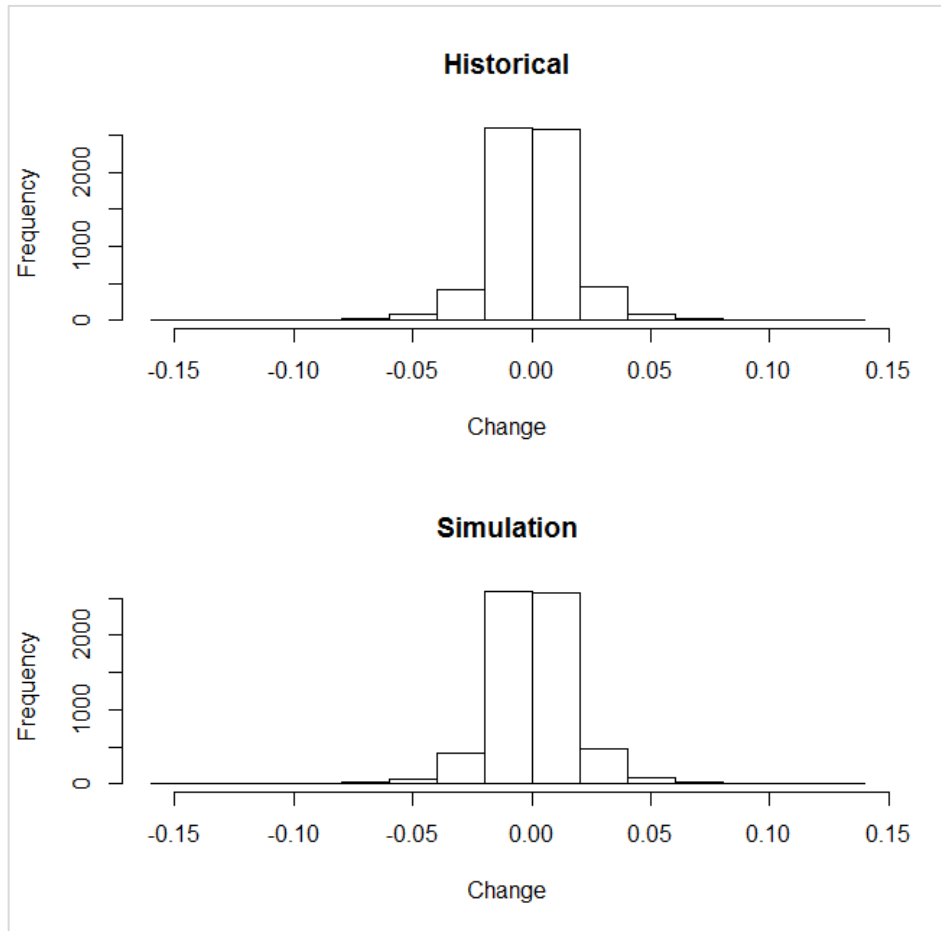


Figure 17: Historical and simulated daily returns of the IBM share

Measure	Historical returns	Simulated returns
Mean	0.0005	0.0004
Standard deviation	0.0178	0.0176
Min	-0.1554	-0.1457
Max	0.1316	0.1368
Kurtosis	7.2426	6.5233
Shapiro-Wilk normality test: p-value	0.0000	0.0000
Two-sample Kolmogorov-Smirnov test (historical vs. simulated): p-value	0.5554	

Table 4: Measures for historical and simulated returns

The similarity of the datasets is further visualised by Q-Q-Plots (Figure 18), comparing the two distributions with a normal distribution. Figure 19 shows the daily changes of the historical and the simulated prices on a time scale. It reveals that both datasets include massive price changes.

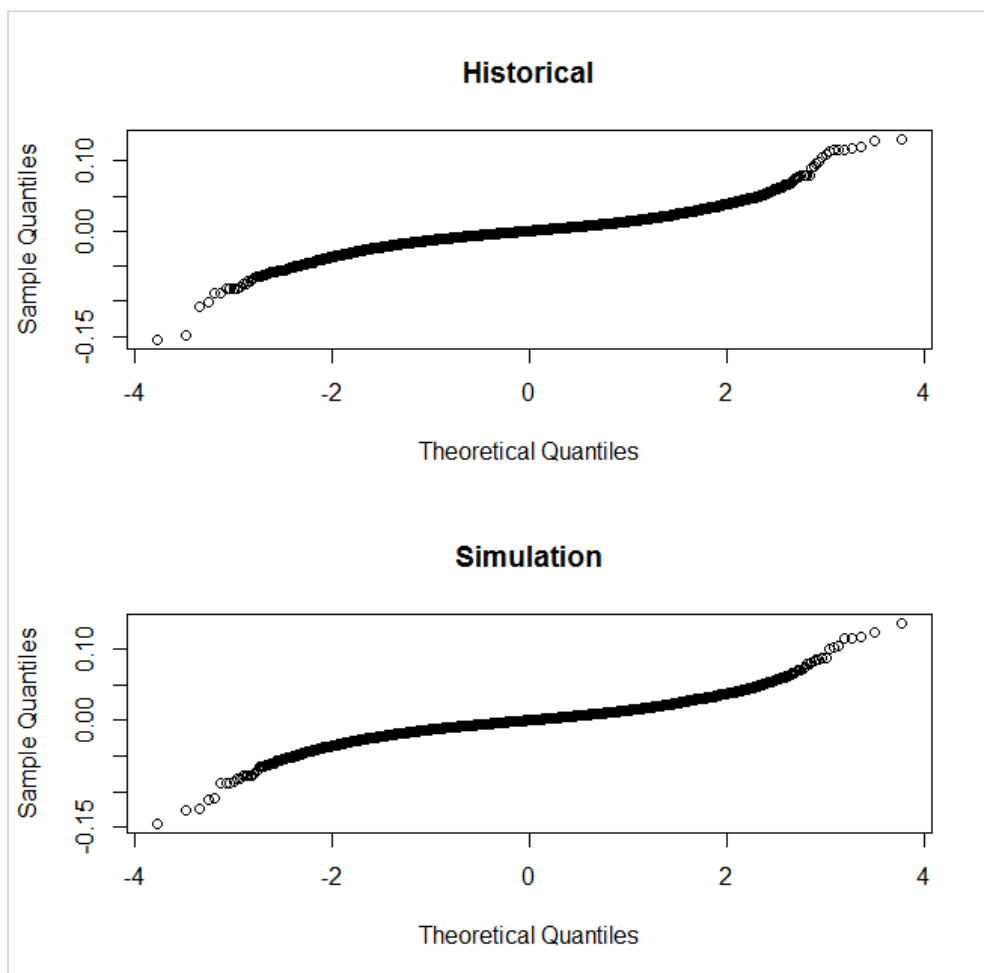


Figure 18: Q-Q-Plot of historical and simulated returns

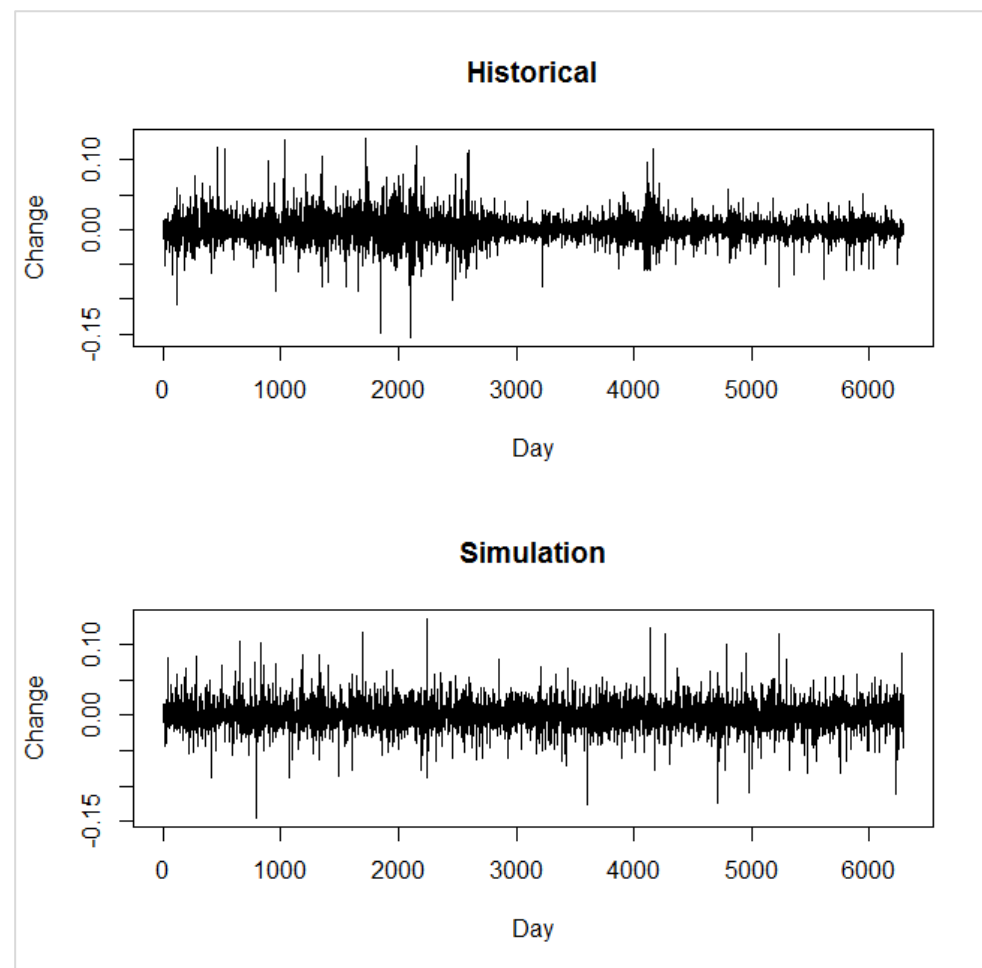


Figure 19: Historical and simulated returns on a time scale

While the diagrams show high similarities, there are differences as well. Figure 19 reveals clusters of high and low volatility that are persistent over a period of time for the historical returns. The volatility of the simulated returns is not clustered. The returns change randomly without a correlation to past values. This issue will be discussed in the next section.

The subsequent diagram shows the distribution of signals for the simulated days.

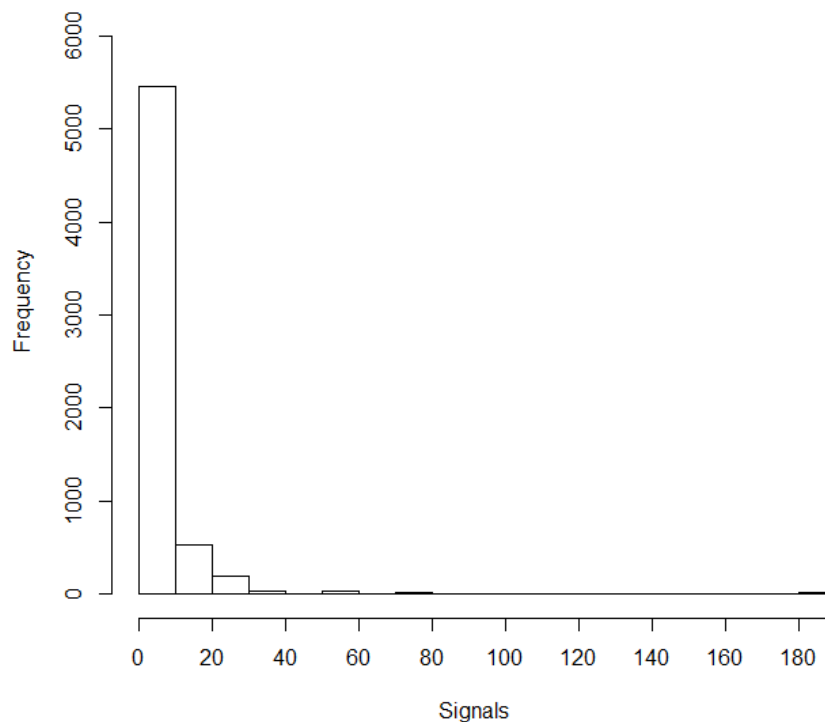


Figure 20: Distribution of signals

Since the signals are based on the degree distribution of the investor's network, it is not surprising that the histogram has a similar shape.

## 7.2 Further Evidence

The previous chapter presented the results of a single simulation. From a statistical point of view this provides weak evidence. The observations could be incidentally. To gain more confidence in the hypothesis, a Fat Tails Random Walk was simulated 50 times for the IBM stock for different time periods. For every run, a Two-sample

Kolmogorov-Smirnov test was performed, comparing the historical and the simulated return distributions. The results are summarised in the table below.<sup>46</sup>

Simulated time period	KS-Test historical vs. simulated; Mean	KS-Test historical vs. simulated; Median
5 years 02 July 2012 – 01 July 2017	0.4459	0.4245
10 years 02 July 2007 – 01 July 2017	0.3175	0.3611
15 years 02 July 2002 – 01 July 2017	0.6055	0.6500
20 years 02 July 1997 – 01 July 2017	0.3171	0.2970
25 years 02 July 1992 – 01 July 2017	0.5813	0.6147

Table 5: Results of 50 simulations of the IBM stock

The above table lists the mean and the median of the p-values of the performed KS tests. Despite minor variations, the p-values suggest consistently a very high similarity of the historical and the simulated distributions.

Table 6 documents the results of 50 FT Random Walks for all companies listed in the Dow Jones Industrial Average on 01 July 2017.<sup>47</sup> For every simulation a KS test was performed. The according p-values are shown in the columns 2-4. If no data was available for a certain company, the corresponding cell contains 'N/A'. The last row of the table documents the average.<sup>48</sup>

<sup>46</sup> IBM share simulated 50 times with a Fat Tails Random Walk and a Normal Random Walk. Measures calculated with R and the library e1071.

<sup>47</sup> Source: Yahoo Finance. The simulated returns were also tested with the Anderson-Darling method. The results were similar.

<sup>48</sup> Calculated with R. Quotes retrieved from Yahoo Finance. Results rounded to 4 decimal places.

For all simulated companies the average median p-value equals 0.27 for 5 years, 0.25 for 15 years and 0.17 for 25 years. The results suggest a significant similarity between the historical and the simulated returns. For some stocks, however, the figures are not convincing. All p-values below a threshold level of 0.05 are highlighted in grey and italics in the subsequent table.

Company	5 years	15 years	25 years
	KS test	KS test	KS test
	Median	Median	Median
Apple	0.3962	0.2337	0.3685
American Express	0.1801	<i>0.0000</i>	<i>0.0123</i>
Boeing	<i>0.0228</i>	0.0567	<i>0.0462</i>
Caterpillar	0.2878	0.5352	<i>0.0265</i>
Cisco Systems	0.6206	0.1654	0.0813
Chevron	0.2521	0.0640	<i>0.0009</i>
DuPont	0.5160	0.4350	0.0971
Walt Disney	0.1801	0.2815	0.2356
General Electric	0.4693	<i>0.0009</i>	0.1705
Goldman Sachs <sup>49</sup>	0.0505	0.0501	N/A
Home Depot	0.1376	0.4288	0.3341
IBM	0.2138	0.5729	0.6446
Intel	0.3562	0.3354	0.2721
Johnson & Johnson	0.1141	0.1491	0.2721
JPMorgan Chase	0.2730	<i>0.0000</i>	<i>0.0010</i>
Coca-Cola	0.2804	0.2942	0.2107
McDonald's	0.1254	<i>0.0414</i>	<i>0.0294</i>
3M Company	0.3432	0.0964	0.1606

<sup>49</sup> Data available from 04 May 1999 to 01 July 2017

Merck & Co	0.3309	0.0860	0.1671
Microsoft	0.5263	0.2123	0.2029
Nike	0.5587	0.4806	0.3805
Pfizer	0.2730	0.3807	0.0279
Procter & Gamble	0.3562	0.5306	0.2356
Travelers	0.1700	0.0006	0.0343
UnitedHealth	0.0769	0.0198	0.0337
United Technologies	0.1376	0.4985	0.2818
Visa <sup>50</sup>	0.1141	N/A	N/A
Verizon	0.0105	0.4896	0.1106
Wal-Mart	0.3962	0.4123	0.1606
Exxon Mobil	0.2730	0.3213	0.0182
<b><u>Average</u></b>	<b><u>0.2681</u></b>	<b><u>0.2473</u></b>	<b><u>0.1703</u></b>

Table 6: Results of 50 simulations of stocks of Dow Jones Index

For a time period of 5 years 28 stocks could be modelled realistically enough to result in a median p-value greater than 0.05. For 15 years 24 stocks could be simulated successfully, for 25 years the p-value is significant for 18 companies. Apparently, the quality of the simulation declines with an increasing time period. This could have several reasons.

Firstly, the FT Random Walk and its input parameters could be inaccurate. This would lead to a distortion that gains weight as the number of simulation runs increases. Another possible explanation is a dynamic network topology that changes over time. As shares are bought and sold frequently in the stock market, the ownership structure evolves. This would be an explanation for volatility clustering as well as market cycles. Both topics are studied comprehensively by Robert Shiller in his book “Irrational Exuberance”.<sup>51</sup> The next chapter discusses the consequences of understanding the stock market as a social network.

<sup>50</sup> Data available from 19 March 2008 to 01 July 2017

<sup>51</sup> (Shiller, 2016)

## 8. Keynes, Herding and the Beauty Contest

Why is the degree distribution of the investor's network suitable to simulate stock prices? A possible explanation can be found in the 'General Theory of Employment, Interest and Money' by John Maynard Keynes.<sup>52</sup>

According to Keynes, investors do not primarily pick stocks by their fundamental (inner) value. Speculators are rather driven by the expected behaviour of other investors, trying to foresee their trades. In the financial literature this phenomenon is referred to as *herding*. An investor is likely to trade if his fellow investors also do. To buy or sell a share just before a competitor is an effective strategy. Committed as criminal act, based on insider knowledge, this pattern is called 'front running'.

Keynes compares the stock market to a beauty contest. When gambling on the results of a beauty contest, it is not the best approach to bet on the candidate that one self considers most beautiful. Instead, a speculator should bet on the nominee that is considered most attractive by the majority of the judges.

Applied to the investor's network from chapter 4, this would mean that an investor trades when his neighbours in the graph trade as well. The number of neighbours of an investor is measured by its degree. This makes the behaviour of herding transparent. It furthermore makes clear that herding is not necessarily irrational. A trade of one investor produces a signal on the market that will be recognised by others.<sup>53</sup> They will process the signal, interpret it and carry out trades as well. The dealings eventually result in a price change of the stock. If one investor has many neighbours, its signal will be recognised broadly. If a speculator has only a few or even none neighbours, the effect is limited.

Whether the trades of investors in the network will go in the same direction, depends on their view. The purchase of IBM shares by BlackRock could in principal trigger two reactions. Firstly, another investor expecting the stock price to rise could interpret the trade as buying signal. On the other hand, a shareholder with negative expectations for the future could see it as an opportunity to sell his position. As

---

<sup>52</sup> (Keynes, 1936, p. 156)

<sup>53</sup> This might not happen immediately, but with a time lag.

described in chapter 7.1, price changes often balance out. Yet, from time to time they go entirely in the same direction and cause a major price correction.

Imitating the behaviour of a neighbour in the network can be seen as heuristic. Tversky & Kahnemann describe several of these rules of thumb in their essay 'Judgment under Uncertainty: Heuristics and Biases'.<sup>54</sup> Relying on the knowledge of a competitor can save own analysis effort. This is particularly useful for complex financial products like stocks. Herding is less attractive with regard to instruments that can be valued straightforwardly.<sup>55</sup>

Seeing the stock market as a social network is compliant with behavioural finance and the efficient market hypotheses of Eugene Fama at the same time. Herding is not necessarily irrational. Thinking in terms of a network can thus help to bridge the gap between classical and behavioural theories.

---

<sup>54</sup> (Tversky & Kahneman, 1974)

<sup>55</sup> This could be interesting with regard to the so-called Equity Premium Puzzle. For details see (Mehra & Prescott, 1985).

## 9. Conclusion

Summarising, a network perspective on money, banks and the stock market grants interesting insights. As discussed in chapter 1, a graph shows the high complexity that arises in an economy based on barter. This complexity can be reduced radically by a centralised monetary system. The price to pay is a single point of failure in form of a central administration that provides the currency. The economic development of communist countries emphasises this issue. A scale-free network offers a balance between a fully centralised and a decentral economic system. This might be the reason why it is found largely in social networks.

A network of stock investors shows a high similarity to a scale-free graph. It can be used to explain and to simulate fat tails in stock returns. Using the empirical degree distribution of a network that was constructed with empirical market data, a majority of the shares of the Dow Jones Industrial Average can be simulated realistically.

Explaining stock markets from a network perspective can shed light on questions that have been discussed intensely. It explains the behaviour of herding without violating the efficient market hypothesis of Eugene Fama and thus bridges the gap between classical and behavioural models.

Besides the answers given, this paper raises many questions. Several ideas were formulated, but could not be analysed in detail. If some of the questions inspire readers to study the topic further, the goal of this paper is achieved: to provide a new view on the economy and its evolution.

## A. References

- Barabási, A.-L. (2016). *Network Science*. Cambridge: Cambridge University Press.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, pp. 307-327.
- Collins, C. (2012). *The Oxford Encyclopedia of Economic History, Volume 3. BANKING: Middle Ages and Early Modern Period*. Oxford: Oxford University Press.
- Cont, R. (2007). In G. Teyssière, & A. Kirman, *Long memory in economics* (pp. 289-309). Berlin Heidelberg: Springer.
- Fama, E. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*, pp. 55-59.
- Garlaschelli, D., Battiston, S., Castri, M., Servedio, V. D., & Caldarelli, G. (2005). The scale-free topology of market investments. *Physica A* 350 (2-4), pp. 491-499.
- Gillespie, C. S. (2015, Februar). Fitting Heavy Tailed Distributions: The poweRlaw Package. *Journal of Statistical Software*, 64 (2), pp. 1-16.
- Goetzmann, W. N. (2017). *Money Changes Everything: How Finance Made Civilization Possible*. Princeton: Princeton University Press.
- Graeber, D. (2011). *Debt. The first 5,000 Years*. New York: Melville House.
- Keynes, J. M. (1936). *The General Theory Of Employment Interest And Money*. London: Macmillan.
- Kurgan-van Hentenryk, G. (1997). *Banking, Trade and Industry: Europe, America and Asia from the Thirteenth to the Twentieth Century*. Cambridge: Cambridge University Press.
- Lipschitz, L., & McDonald, D. (1990). *German Unification: Economic Issues Occasional Paper*. Washington D.C.: International Monetary Fund.

- Lowenstein, R. (2015). *America's Bank: The Epic Struggle to Create the Federal Reserve*. London: Penguin Books.
- Luke, D. A. (2015). *A User's Guide to Network Analysis in R*. Heidelberg: Springer.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, 394-419.
- Mandelbrot, B. (2006). *The (mis)Behaviour of Markets: A Fractal View of Risk, Ruin and Award*. New York: Basic Books.
- Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, pp. 77-91.
- Mehra, R., & Prescott, E. C. (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics* 15, pp. 145-161.
- Metcalf, W. E. (2016). *The Oxford Handbook of Greek and Roman Coinage*. Oxford: Oxford University Press.
- Pareto, V. (1897). *Cours d'Économie politique - Tome Second*. Lausanne: F. Rouge.
- Safranski, R. (2015). *Zeit*. München: Carl Hanser Verlag.
- Shiller, R. J. (2016). *Irrational Exuberance*. Princeton: Princeton University Press.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan and T. Cadell.
- Taleb, N. N. (2008). *The Black Swan: The Impact of the Highly Improbable*. London: Penguin Books.
- Tversky, A., & Kahneman, D. (1974, September 27). Judgment under Uncertainty: Heuristics and Biases. *Science, New Series*, Vol. 185, No. 4157, pp. 1124-1131.
- Vantiegheem, C. (2012). *Monetary practices In ancient Egypt*. Retrieved from Money Museum National Bank of Belgium: <http://www.nbbmuseum.be/en/2012/05/nederlands-geldgebruik-in-het-oude-egypte.htm>

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician Vol. 70 , No. 2*, pp. 129-133.