ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis Bachelor of Economics and Business Economics

The effects of data breaches on the stock price in the period 2016-2018

Name Student: Ralph Schuurman
Student ID Number: 434988

Supervisor: Dr. R. Cox
Second Assessor: Dr. J. Kil

Date Final Version: July 23, 2019

The views stated in this thesis are those of the author and not necessarily those of Erasmus School of Economics or Erasmus University Rotterdam.

Abstract

Event study methodology was used to study the effect of data breaches on the stock price in the period 2016-2018 for 123 firms noted on the United States stock market. In contrast to most previous works in earlier periods, no significant results were found indicating an effect in this period. This is in line with a declining trend found before when researching data breach effects. The effect of the size of a data breach on the reaction on the stock price was also studied. The size of the data breach did not have a significant effect on the stock price reaction.

Table of Contents

Introduction

The amount of data breaches has seen a rise in the past years. In 2005, there were 136 data breaches in the United States, as reported by the Privacy Rights Clearinghouse. In 2017 there were 765 breaches. Not only the amount of data breaches has increased, also the amount of records in these breaches have seen a rise. In 2005 the average amount of records breaches per data breach was 405 thousand records. In 2017, this was 2.6 million. The industry type has also changed through the years, since 2009 the medical sector is the leading sector in the amount of data breaches, even though not in records breached (Privacy Rights Clearinghouse, 2019).

Governments and supranational entities have also recognized this trend in data breaches. The European Union has introduced the General Data Protection Regulation (GDPR) to decrease the amount of data breaches and protect the privacy of its citizens. One of the main requirements in the GDPR is the data breach notification obligation. In case of a personal data breach, the controller (including companies) should report the data breach within 72 hours to the competent supervisory authority (GDPR, 2019).

Compared to the European Union, the United States has been lacking in terms of data regulation on the federal level. Most regulation on data breaches has been made in state laws. In contrast to the GDPR, no data breach notification obligation exists on the federal level and regulation therefore differs per state. All 50 states at the time of writing have a data breach notification law, with South Dakota and Alabama only passing this law in 2018.

One of the biggest data breaches in the United States was the breach of Equifax. Personal information of 143 million Americans was accessed during this breach, including names, Social Security numbers and sometimes credit cards (Federal Trade Commission, 2017). Recently it was announced that Equifax has to pay around $700 million dollars as a consequence of this breach (Wall Street Journal, 2019).

When a company is noted on a stock exchange, a breach could have a possible effect on the stock price, as a data breach is an unannounced unexpected cost. After Facebook profiles were harvested in 2018 by Cambridge Analytica, Facebook shares fell more than 24 percent in the week after the announcement the data breach occurred. The corresponding market value loss was roughly 134 billion dollars (CBS News, 2018). Even still, within two months after this decline Facebook recorded a record high share price, more than 12% above the share price before the drop.

Equifax recorded a similar drop after the announcement of their data breach. The share price of Equifax dropped 33% after the announcement (Forbes, 2017). Now, almost two years later, the stock price of Equifax has not been higher than just before the announcement. Other reasons should be considered when comparing stock prices over a longer period, still it cannot be disregarded that the data breach announcement could have had some impact on the share price.

This thesis will examine the effect of a data breach on the stock price in the years 2016 until 2018. These breaches include personal information such as social security numbers, addresses, names, passwords and sometimes credit card information and/or medical information. Data breaches and their effect on stock prices have been studied previously. Most of these studies have found that a data breach announcement has a negative effect on the stock price. These previous studies have not included events from 2014 onwards. In this paper, the more recent years will be studied to find out if recent privacy laws and a bigger amount of total breaches may have had an influence on the effect.

Previously it was found that even though data breaches had a negative reaction on the stock price, there was a decline in reaction in the later periods. It was also found that stocks from companies that had experienced a data breach before do not experience a negative reaction when a new data breach is announced. This thesis will study a later period than the research before and most companies studied have mostly already experienced data breaches before.

This is why this thesis can add value to the existing literature; following this declining trend, the reaction on the stock price to data breaches may have further declined or is possibly nonexistent in the period 2016 – 2018.

Research question

The research question is stated as:

What is the effect of the announcement of a data breach on the stock price of the concerning company in the years 2016 until 2018?

This research will focus on companies listed on the United States stock market exclusively. Previous works have also been almost exclusively focused on companies located in the United States. In contrast to existing literature, the effect of the size of the data breach in terms of records breached was also be studied.

In this study, no significant reaction was found on the stock price after the announcement of a data breach in the studied period for multiple event windows. The significance was tested using both a parametric and a non-parametric test. This is not in line with most previous research but does conform a trend that the reaction on data breaches has experienced a downward shift. Also in contrast to what was expected, the amount of records breached also had no significant effect on the data breach announcement reaction.

Theoretical framework

Data breach research

Data breaches and their effect on stock prices have been a topic of research for over more than a decade. Garg, Curtis & Halper (2003) found a significant negative reaction in the period 1999-2002 and estimated the average cost of a breach at $17-$28 million. The studies before found average costs ranging from 50 thousand dollars to 2 million dollars. Another study during that time also showed that breaches in digital security have a negative effect on stock prices (Cavusoglu, Mishra, & Raghunathan, 2004). This paper did consider all types of security events, including events where no personal data was affected as for example DOS attacks. The period 1995 till 2001 was used, where it was concluded that compromised firms had a negative abnormal return of approximately 2.1%. This findings that data breach announcements had a negative impact on the stock price was also found in the years after. Acquisti, Friendman and Telang (2006) studied the period 2000-2005 and found a negative reaction of -.58%. One study concluded that a data breach has a negative and statistically significant effect of -.84% in the period 2004 till 2006 on shareholder wealth, and more interesting the negative market reaction to a data breach was more significant in the more recent time periods (Gatzlaff & McCollough, 2010).

Gordon, Loeb & Zhou (2011) found in the period 1995-2007 a negative reaction of -.0091% but also concluded that the period 2002-2007 was insignificant. Pirounias, Mermigas & Patsakis (2014) had trouble finding statistical significant results in the period 2008-2011 but still concluded that a data breach announcement had a negative impact on the stock price. Rosati et al. (2018) is the most recent study till date. A statistical significant negative reaction was found for the period 2011-2014. This research was different than other research because it only used companies that actively use Twitter, which could maybe impact the validity of the results for companies without an (active) Twitter account.

Of 45 studies found on the impact of information security events on the stock market the majority (75.4%) reported a statistical significant impact of the event on the stock price (Spanos & Angelis, 2016). The type of breach could also influence the market reaction, as it affects the cost of the breach. In one study only when the data breach involved unauthorized access to confidential data a significant negative market reaction was found (Campbell et al. 2003).

One could argue that if an information security breach has a negative effect on the stock price, the opposite should also be true. An investment to improve information security should have a positive effect on the stock price. Substantial support for this was found. The announcement of security investments lead to positive abnormal returns in the United States stock market between 1997 and 2006 (Chai, Kim & Rao, 2011).

Downward shift in the later years

As stated in the introduction, the amount of data breaches has risen and the amount of records breached has also seen an impressive growth, this could mean that investors could react differently than earlier. When data breaches in the period 1995 till 2007 were examined, it was found that even though the data breaches had a significant negative effect on the stock price, a downward shift had occurred in the later period (Gordon, Loeb & Zhou, 2011). This means that the effect of a data breach was, in contrast to Gatzlaff & McCollough (2010), less prominent in the later years.

To explain this phenomenon, two reasons were given. The first is more effective remediation and disaster recovery used by firms because of the increase in breaches. The second is that the tendency from customers to refrain from doing business with firms experiencing data breaches has decreased. This results that investors would see security breaches as “nuisance” instead of a potentially economic threat, meaning that they have become more insensitive to breaches. A real life consequence of this could be that corporate executives are less inclined to increase digital security. This situation could be described as a principal-agent problem. The company that needs to protect the data from others has no incentive to do so. The ones whose data were breached are the only ones experiencing negative effects. The implication would be that privacy in itself will be less valued. The theory that investors are less sensitive to data breaches in later periods was also found by Pirounias, Mermigas & Patsakis (2014). They found for the period 1995-2004 an abnormal return of -2.1% and for the period 2002-2007 an abnormal return of -0.34%, but this value was not significant .

Repeated data breaches at the same company were also studied before, using events transpired between 2005 and 2013. The findings were that firstly there was a negative reaction on the stock price corresponding the first breach, but there was weak statistical evidence that the market reacts differently to a second breach of the same organization (Schatz & Bashroush , 2016), implying that second breaches have less impact than first breaches.

This also is in line with the theory that investors consider data breaches as less of a problem in more recent times periods, also described by Gordon, Loeb & Zhou (2011) and Pirounias, Mermigas & Patsakis (2014). In other words, because breaches have happened before at the same company and more breaches in general are taking place every year, the shock effect of a data breach has decreased.

The question does rise if the effect of data breaches on stock prices does matter in the long term for firms. Some research has indicated that, in junction with previous research, indeed security breaches do have a negative and significant impact on data breaches on the market value over day but then decreases and loses statistical significance (Acquisti, Friendman and Telang, 2006). This should also be taken into account when interpreting the results.

As most of the previous research indicates that a data breach will have a (negative) effect on the stock price, the first hypothesis is formulated as follows:

The announcement of a data breach has a significant effect on the stock price in the period 2016 – 2018.

Consequences of a data breach

To understand why data breaches can possibly influence the stock price, the consequences of a breach should also be looked at. The most widely source used for the costs of data breaches is the yearly report by the Ponemon Institute. This report specifies the cost of data breaches by conducting interviews with breached companies. In the United States the average cost per breach in 2018 amounts to 7.91 million dollars, with an average of 233 dollars per breached record. The biggest cost when experiencing a data breach is the lost business cost, for a big part customer turnover. The churn rate differs greatly per industry, the Health and Financial industry are at the top with both above 6% churn rate, whilst retail is at the bottom with 2.1%. Organizations in the United States pay most for losing customers, almost two times as much as the number two, Mexico. Using these figures and if available the amount of records lost, in this paper it will examined if the loss in stock price when experiencing a data breach is in proportion with the expected loss.

This is also a test for the efficient market hypothesis. The efficient market hypotheses implies that the loss in market cap value is equal to the expected financial loss. If the cost of a breach does not equal the decline in market value, the efficient market hypothesis does not hold. There has not been research done that compares these two values. By comparing the figures from the Ponemon institute with the breaches, it is possible to estimate if the possible losses on the market capitalization due to data breaches are justified.

Size of the data breach

Data breaches happen in different sizes. Some breaches will only affect records of a handful of parties. Other breaches can affect records of hundreds of millions or even billions of different parties. The amount of record breaches will have an effect on the companies' future financials, as more records breached also means bigger lawsuits and less confidence from consumers and other businesses that their records are safe with the breached company. The size of the company does have an effect on the market reaction (Gatzlaff & McCullough, 2010**)** but there has not been a paper before that tests the effect of the breach size itself. A previous study has indicated that the knowledge of the number of records breached does not have a significant effect (Rosati et al. 2018).

To test if the size of the data breach has an effect on the market reaction, a second hypothesis has been formulated to test if bigger data breaches equal a higher cost. The second hypothesis is:

The size of the breach measured in records breached has a negative significant relation with the effect on the share price.

Data

The sample of data breaches was collected using the list of data breaches maintained by the Privacy Rights Clearinghouse in the period 2016 till 2018. This list is not an exhaustive list of all data breaches ever occurred in the United States. The disclosure of data breach is regulated on the state level, most state laws follow the state law of California as this was the first state to pass these regulations. The state law of California requires businesses to disclose any breach of security to any Californian residents whose unencrypted personal data was, or could have been, acquired by an unauthorized person. Because this is a state law, only corporations with a physical location in California are subject to this law. A New York corporation with no location in California is not subject to this law, even though the subjects of the data breach could very well be Californian residents. This means that until the other states implemented a comparable law, not all data breaches were reported and therefore not all breaches could have been added to the Privacy Rights Clearinghouse database.

The list also almost exclusively includes companies in the United States. Previous research has also used the database from the Privacy Rights Clearinghouse, although none have used the breaches later than 2014 (Gatzlaff & McCollough, 2010, Schatz & Bashroush, 2016; Rosati et al., 2018).

In 2016, there were 768 breaches reported, in 2017 681 and in 2018 414 breaches for a total of 1863 breaches. As this research focuses on the effect of a data breach on the stock price, only breaches of companies that were public at the time of the breach were included in the sample. This leads to 137 data breach events in the period. To better measure the effect of the data breach on the stock price, events taken place on days where the stock market was closed were excluded. One event was excluded because no trades concerning the stock were done on the date of the event and the days after. This was also the only stock noted on an illiquid exchange. The total amount of events is therefore 123 events**.** As recommended by others the firm names and event dates are in the appendix in Table 1A (McWilliams & Siegel, 1997; De Jong & Naumovska (2015).

Under the classification 'data breach' the same definition will be used as the Privacy Right Clearinghouse uses, also used in Rosati et al. (2018). "A data breach is when a company inadvertently leaks your personal information as a result of a hack attack, lost or stolen computers, fraud, insider theft, and more" (Privacy Rights Clearinghouse, 2011).

Of the 123 data breach events, 48 had an unknown number of records breached. Of the breaches where the amount of records breached were known, the average amount of record breaches was 56,095,195 records, the median was 3674 records. This big difference is explainable because of outliers in the amount of records breached. Interesting to note is that the two biggest data breaches were from the same company, Yahoo. On 22 September 2016 Yahoo had a breach of 500 million records and on 14 December 2016 Yahoo had a data breach of 3 billion records. According to the used database, 83 of the 123 data breach events were data breaches at companies that have experienced a breach before the event period.

The Privacy Rights Clearinghouse sorts the type of organization using eight classifications: Business-Financial and Insurance Services (BSF), Businesses – Other (BSO), Business-Retail/Merchant Including Online Retail (BSR), Educational Institutions (EDU), Government & Military (GOV), Healthcare, Medical Providers & Medical Insurance Services (MED), Nonprofits (NGO) and Unknown (UNKN). As Governments and Nonprofits are not publicly traded, these do no not appear in the sample.

The distribution for the different classifications is described by Table 1.

Table 1. Overview of breaches.
The table reports summary statistics for 132 data breaches between 2016 and 2018. Mean, number of records breached, minimum and maximum are included. The data and classification is retrieved from the Privacy Rights Clearinghouse.

	N	# records breached known	Average records breached	Median records breached
Business-Financial and Insurance Services	21	11	7,048,565	1253
Businesses – Other	44	13	83,036,009	431,000
Business-Retail/Merchant Including Online Retial	15	12	26,980,203	1292
Healthcare, Medical Providers & Medical Insurance Services	43	39	19,357	2426
Total	123	75	56,095,195	3674

Both the average amount and the median amount of records breached is the highest in the Business – Other category. This is explainable because social networks and websites such as Facebook and Twitter experience the biggest data breaches and are included in this category. One thing that should be considered about the amount of record breached is that in the sample this is often the possible amount of record breached, not per se the true amount of records breached. When it is unknown how many records were breached but known how many possible records could be breached, the Privacy Rights Clearinghouse counts the amount of possible records breached. A bug in a website that could expose data of 100 million customers is counted as 100 million records breached. This does not mean that these 100 million records were actually accessed by unauthorized parties.

Methodology

Efficient market hypothesis

To examine the effect of a data breach, event study methodology based on the efficient market hypothesis (EMH) will be adopted. Coined by Fama in 1970, the efficient market hypothesis is described as a hypothesis that states that in an efficient market all available information is reflected in the prices (Fama & Malkiel, 1970). Three forms of market efficiency were discussed in this paper: the weak form, the semi-strong form and the strong form. In a later paper, Fama revised the definitions (in the same sequence) to predictability, event studies and test for private information (Fama, 1991).

In the weak form (predictability) past price information is priced in the stock and thus can't be used to predict the future of the stock. This means that future securities' prices are a random walk. In the semi-strong form (event studies) prices reflect all public information, including public information about the future. This implies that it is only possible to earn abnormal returns by using non-public information. In the strong form (test for private information) all information, public and private, is incorporated in the market prices. This implies that insider trading cannot be used to earn abnormal returns.

If the strong form of the EMH prevails, making information such as data breaches public will have no effect on the stock price. If the semi-strong form of EMH prevails, the new information of the data breach will be immediately reflected in the stock price of the corresponding company. If the weak form of the EMH prevails, the announcement of a data breach will be incorporated in the stock price although not immediately. Event studies are often used to test the semi-strong form and generally do support the semi-strong form. As this thesis will perform an event study, the semi-strong form of the EMH will be tested. If the semi-strong EMH hypothesis holds, the reaction on the event from investors on the stock price should reflect the actual costs incurred because of the data breach.

Event study methodology

Event studies have been used a long time to measure the impact of a specific event on a firms value. Applications of the event study have been among others announcements of mergers and acquisitions, change in regulatory measure and legal liability cases. Since its introduction, event study methodology has become the standard method of measuring the reaction of a security price to an event (Binder, 1998). The reasoning behind an event study is that because of rationality in the market, the effects of an event will be immediately reflected in security prices (MacKinlay, 1997). In research where the effect of data breaches on firms and their stock prices are studied, event studies are used almost exclusively. This is also partly done because firms often do not disclose the actual cost of a data breach.

To perform the event study, multiple steps will be followed, as described by Campbell et al. (1997) and McWilliams & Siegel (1997). The event and event window will be defined and the relevant firms will be selected (as described in the Data section). Then a model to estimate the normal and abnormal returns will be chosen. Using this model, the cumulative abnormal returns will be estimated and tested.

The model

The model used to estimate normal and abnormal returns will be the Market Model. When comparing the Mean-Adjusted Returns Model, the Market-Adjusted Returns Model and the Market Model, there is a slight preference for the Market Model (Dyckman et al. 1984). Most other literature regarding event studies on data breaches uses the market model (including Gatzlaff & McCollough, 2009 and Rosati et al. 2018).

This model is specified as:

$$R_{i,t} = \alpha + \beta_i R_{m,t} + \varepsilon_{i,t}$$

Where

$R_{i,t}$ is the return on security i in period t

α_i is a measure of risk-adjusted performance

β_i is a measure of systematic risk

$R_{m,t}$ is the return on market portfolio in period t

$\varepsilon_{i,t}$ is the disturbance term for security I in period t

The abnormal return (AR) is represented by $\varepsilon_{i,t}$ as this is the difference between the actual return and the expected return based on the estimation. This can also be written as the following formula:

$$AR_{i,t} = R_{i,t} - (\alpha + \beta_i R_{m,t})$$

These abnormal returns will then be summed for the event period to obtain the cumulative abnormal returns (CAR).

$$CAR(t1, t2) = \sum_{i=t1}^{t2} AR_{i,t}$$

Event window

To determine the event window, past literature in data breach event studies has been used. Gatzlaff & McCullough (2010), Cavusoglu, Mishra, & Raghunathan (2004) and Rosati et al. (2018) mostly used a two day window (0,+1). A data breach announcement could come at the end of the day and therefore a two day window would be expected to best capture the response of the stock market. One of these previous papers also reported that the difference between a one day window or a two day window was only slight (Gatzlaff & McCullough, 2010). Other research has used only a one day window: only the day of the announcement (Acquisti, Friedman & Telang, 2006) citing Hendricks & Singhal (1996). According to this paper, there are two reasons why a one-day event period should be used. Firstly, one-day event period will provide a better estimation because it reduces the possibility of other factors not related to the incident having an effect on the price. Secondly, it would increase the power of statistical tests. Even though Acquisti, Friedman and Telang (2006) argue to use a one-day window, they also compensated for a delay in the news cycle and examine a short period following the event day, thus admitting that a one-day window is probably too small.

Even though security breaches are generally unanticipated (Cavusoglu, Mishra, & Raghunathan, 2004) some literature suggests that leakage of the news can occur. Some research found negative AR values in the days before the event (Goel & Shawky, 2009). This is debatable, as other research has concluded that no statistically significant results are found in the days before the event (Gatzlaff & McCullough, 2010). Therefore, no dates before the event are included in the event window. To check if maybe it takes some time between the announcement of the data breach and the reaction on the stock price, bigger event windows will also be tested. This can also be useful to see if data breach announcements have a longer term effect, also done by Rosati et al. (2018). A problem with a bigger event window is that it increases the chances of taking confounding events into calculation, reducing the reliability and validity of this study (Brown & Warner, 1985). It is shown that confounding events can have a big impact on the reliability of the study and therefore measures should be taken to limit this influence (De Jong & Naumovska, 2015). Taken all this into account, an event window of two days (0,+1) will be used primarily and an event window of (0,+9) will be used to check for a possible delay in the news of the breach.

Estimation window

The standard estimation window of 92 trading days will be used (-100,-8) (Acquisti, Friedman & Telang, 2006). Most other relevant research has also used an estimation window of around a 100 days.The software Eventus was used to perform the analysis using the database from the Center of Research in Security Prices (CRSP). The exact date of the announcement of the breach was retrieved from the Privacy Rights Clearinghouse database.

The retrieved cumulative abnormal returns will first be aggregated and then be tested for significance using a standard cross-sectional t-test and a non-parametric Wilcoxon Signed Rank test. To aggregate the cumulative abnormal returns, the following formula will be used to calculate the cumulative average abnormal return:

$$CAAR = \frac{1}{N}\sum_{i=1}^{N} CAR$$

To test for significance, both a standard cross-sectional t-test and a Wilcoxon Signed Rank test will be performed.

The standard cross-sectional t-test will use the following formula. The null hypothesis for the t-test test is that the cumulative abnormal result is equal to zero, thus implying no effect of the event on the stock return.:

$$t_{CAAR} = \sqrt{N}\frac{AAR_t}{S_{CAAR_t}}$$

Where

$$S_{CAAR}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(CAR_i - CAAR)^2$$

A significance level of 5% is used.

The non-parametric Wilcoxon Signed Rank test will be performed to check for robustness of the results. As it is a non-parametric test, it does not rely on the assumption that the cumulative abnormal returns are normally distributed. The Wilcoxon Signed Rank Test will also negate the effect of outliers on the data. The firms in the sample are not thinly traded as they are all on major stock exchanges. If the sample is thinly traded, a preference would be the Generalized Sign test above the Rank test as it is more powerful under conditions of thin trading (De Jong & Naumovska, 2015).

The Wilcoxon Signed Rank test will be performed using the following formula:

$$W_t = \sum_{i=1}^{N} rank(A_{i,t})^{+}$$

rank($A_{i,t}$) denotes the positive rank of the absolute value of abnormal returns $A_{i,t}$ for firm i at time t.
The test statistic for testing H_0 : CAAR = 0 is defined as:

$$Z_{wilcoxon,t} = \frac{W - N(N-1)/4}{\sqrt{\frac{(N(N_{!} + 1)(2N+1))}{12}}}$$

To study if the efficient market hypothesis holds, the latest report on data breach costs done by the Ponemon Institute will be used to examine if the loss on the stock price is in proportion with the cost of a data breach. In previous research, it was found that the market reaction after the crash of the Challenger on the stock of the company at fault reasonably corresponded with the subsequent losses (Maloney & Mulherin, 2003).

The cumulative abnormal return of the stock times the market value will indicate the loss of the value of the firm due to the breach. The figures of the Ponemon Institute will be used in combination with known facts about the data breach such as amount of records concerned. These two figures will be compared to conclude if the market reacts efficiently.

Cross-section analysis

The second hypothesis will be tested using a cross-section regression. Even when the mean CAR is zero, cross-sectional tests are relevant (Kothari & Warner, 2007) but should be carefully interpreted (Campbell, Lo, and Mackinlay, 1997). The economic effect of the event often differs per firm and therefore abnormal returns vary cross-sectionally. The following regression will be performed using Ordinary Least Squares.

$$CAR = \alpha + \beta_1 \ln(Records) + \beta_2 \ln(Firmsize) + \beta_3 MTBValue + \beta_4 Medical + \beta_5 Card + \beta_6 Year + \varepsilon$$

- The dependent variable is the cumulative abnormal returns of the firms experiencing a data breach using the event window (0,+1).
- Records will be the natural logarithmic function of the amount of (possible) records breached. The logarithmic function will be used as the variance of records breached is big, with a minimum of 1 and a maximum of 3 billion.
- The size of the firm will be measured as the logarithmic market value of the firm. This will be calculated by using the share price times the outstanding shares on the day of the breach. This control variable is included because the cost of a data breach may differ for small and big firms. The use of market capitalization to control for size was also used by papers by Gatzlaff and McCullough (2010) and Cavusoglu, Mishra, & Raghunathan, (2004) which both found a positive and statistical relationship indicating that the reaction to a breach at a smaller firm is more severe than at a bigger firm.
- The market to book value will be used to control for the fact that companies with higher growth potential might undergo a more negative reaction on the stock price. This variable was also used by Rosati et al. (2018). This value was calculated using the following formula:

$$Market\ to\ Book\ ratio = \frac{Price\ Per\ Share}{Book\ Value\ per\ Share}$$

- If medical information was breached, the dummy variable Medical will be equal to 1. Otherwise the dummy will be zero. Most of the medical companies that experienced a data breach did not publicly disclose if medical information was breached. Of the 14 events where companies did disclose this information, 11 of these events included a breach of medical information. As such, medical companies that did not disclose the exact type of information included in the data breach will be treated as if the event included a breach of medical information.

- If payment card fraud was reported the dummy variable Card will be 1. Previous research has indicated that the market reacts most severely to credit card information theft (Garg, Curtis & Halper, 2003). To control for this reaction this variable is used.
- Year is a dummy variable to control for characteristics specific for a year.

A correlation matrix of these variables can be found in Table 2. The presence of medical and credit card information leakage has the highest correlation, which means that often both types of information have been breached.

The summary statistics can be found in Table 3. The mean of the cumulative abnormal returns is almost zero, with on the minimum and maximum almost the same number with a different sign. The amount of records breached was included as a logarithmic function because of outliers. The amount of records breached now range from 0 (when only one record was breached) to 21.8, with a mean of 9.07. In 56% of the breaches, medical information was breached. Credit card information was breached in 9.3% of the breaches. The Market to Book ratio ranges from -.83 to 135.5, with a mean of 6.7.

Table 2. Correlation Matrix.

The table reports the correlation between the variables used in the regression.

	CAR	lnRecords	lnMarketCap	Medical	Creditcard	MarketToBookRatio
CAR (0,+1)	1					
lnRecords	-.1933	1				
lnMarketCap	-.1501	.1929	1			
Medical	.1311	.1989	-.0909	1		
CreditCard	-.2972	.1034	-.1585	-.3620	1	
MarketToBookRatio	.0999	0.148	0.059	.0849	-.0177	1

Table 3. Summary Statistics for the used variables for cross-section analysis.

The table reports summary statistics for 75 data breaches between 2016 and 2018. Only data breaches that had a known number of records breached were included. Mean, standard deviation, minimum and maximum for all relevant variables.

	N	Mean	St.Dev	Min	Max
CAR (0,+1)	75	-.0024	.0292	-.1247	.11404
lnRecords	75	9.07	4.78	0	21.821
lnMarket_cap	75	23.877	1.729	18.13351	26.857
Medical	75	.56	0.499	0	1
Creditcard	75	.0933	.292	0	1
Market To Book Ratio	75	6.735	16.814	-.8284	135.541

Results

According to the method described in the Methodology section, the following results found in Table 4 were found.

Table 4. Cumulative Abnormal Returns (CAR).
The table reports the cumulative abnormal returns 123 data breaches between 2016 and 2018. Mean cumulative returns, cross-sectional t-test values with probability and Wilcoxon signed rank test statistics with probability are included.

	N	Mean Cumulative Abnormal Return	Cross sectional t-statistic	Cross sectional t-statistic probability	Wilcoxon signed rank test statistic	Wilcoxon signed rank probability
Event window						
(0,0)	123	-0,00119078	-0,791494126	0,430191846	-278	0,4851207
(0,+1)	123	-0,003070071	-1,312518141	0,19180922	-548	0,1675537
(0,+2)	123	-0,002644886	-0,845179742	0,399664977	418	0,2932678
(0,+3)	123	-0,001362463	-0,434987883	0,664339725	544	0,1707011
(0,+4)	123	-0,004075807	-0,930185911	0,354111767	-387	0,3306926
(0,+5)	123	0,003181241	-0,706178049	0,481423821	42	0,9160873
(0,+6)	123	-0,003512548	-0,714342692	0,476379852	-7	0,9859897
(0,+7)	123	-0,00408256	-0,802463575	0,423845295	-113	0,7767793
(0,+8)	123	-0,004944443	-0,94167467	0,348220136	-238	0,5501833
(0,+9)	123	-0,005852842	-1,002420546	0,318124599	-10	0,9799863

In the primary event window (0,+1) the cumulative abnormal returns have no significant value when evaluating the cross sectional t-test, meaning that the cumulative abnormal returns do not significantly differ from zero. In the bigger event windows for all event windows also no significant value for the cumulative abnormal returns was found using the cross sectional t-test.

To check robustness of the results a nonparametric test was also performed, the Wilcoxon signed-rank test. This test shows the same results as the cross sectional t-test, none of the event windows have significant cumulative abnormal returns. These results are in contrast to most previous works on the announcement of data breaches. Because of these results, the first hypothesis, stating that data breaches have a significant effect on the stock price, will be rejected.

This could be explainable by multiple things. One is that the trend, described by Gordon, Loeb & Zhou (2011), that the effect of a data breach has become less over the years has continued. This trend was also confirmed by Pirounias, Mermigas & Patsakis (2014), using data from 2008 until 2011. The possible explanation for this is that investors have become less sensitive to security breaches, as more are taking place every year.

This declining trend was not found in the previous most up to date work of Rosati et al. (2018), using data up to 2014 from the same database. Rosati et al. (2018) only studied companies that had an active Twitter account, thus maybe skewing the results and as a consequence these results are not applicable to all companies experiencing data breaches, only to certain types of companies (the ones with an active Twitter account).

Another explanation can be found in the fact that when looking at data breaches in the years 2016-2018, most of the companies in the sample have already experienced data breaches before. Schatz & Bashroush (2016) showed that a second data breach has no significant reaction on the stock price of said firm. As 83 of the 123 events were companies that have experienced at least one breach prior to 2016, this could also have an effect.

That the announcement of a data breach on the stock price has no effect according to this study may mean that the reaction is now in line with the efficient market hypothesis. Previous event studies found values for the cost of data breaches to be ranging from $17-28 million (Garg, Curtis & Halper, 2003), to $356-$381 million (Pirounias, Mermigas & Patsakis, 2014). The figures from the Ponemon institute, not gathered with an event study but with interviews with breached companies, estimate the average cost of a data breach in the US around $7.91 million. There is a big gap between this value and the value of the cost calculated by previous event studies. It should be taken into account that there is a chance that the figures reported by the Ponemon Institute are underreported, as it is better for companies to downplay the cost of a data breach.

One should also note that event studies only entail companies listed on the stock exchange, which are often the bigger companies. The figure from the Ponemon is an average across all companies. When looking at the so called ‘mega breaches’ do the figures come closer to the figures found using event studies. A mega breach of 1 million records has an estimated cost of $39.39 million, a breach containing 50 million records has an estimated cost of $350.44 million. When looking at the sample of 75 companies that have a known amount of records breached, their average cost of a breach using the Ponemon figures of cost per breached record is $13 million, which is not in line with the figures found by most previous event studies.

Is possible that in recent years the reaction on data breaches is more in line with the figures from the Ponemon Institute. This would mean that the efficient market hypothesis holds for the more recent periods. When data breaches were new, investors did not know what the effect of a data breach would be on a company. In recent years, it was shown that the effects of a data breach were less severe than investors expected. The expected value for a data breach has declined in recent years in such a way that it has no longer an effect on the share price.

Table 5. OLS regression on Cumulative Abnormal Returns.

The table represents the results from the regression of the amount of records as a logarithmic function and the control variables on the Cumulative Abnormal Returns. The Cumulative Abnormal Returns are those of the (0,1) event window. Market capitalization is the logarithmic function of the market value of the firm as a proxy for size. Market to book ratio is created by dividing the share price with the book value per share. Medical and Creditcard are both dummies which turn 1 if medical or credit card data has been part of the breach. Year is a dummy variable to control for yearly effects. In the parentheses are the robust t-statistics with standard errors clustered.

*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(0,+1) CAR
lnRecords	-.000638 (0.69)
lnMarketCap	-.00320 (-1.58)
Medical	-.000543 (0.08)
Creditcard	-.0307 (-1.65)
MarketToBookRatio	0.000250* ** (3.20)
2017.year	-0.0110 (-1.50)
2018.year	0.00290 (0.35)
cons	0.0819 (1.54)
R2	0.1888

Even though the cumulative abnormal returns (CAR) were not significant, a cross section analysis was done to evaluate what could have an effect on the cumulative abnormal returns. Because of the insignificance of the CAR these results should be carefully interpreted. As shown in Table 5, only the Market to Book ratio has a significant effect on the CAR. This effect is significantly positive, meaning that firms with a higher Market to Book ratio experience higher cumulative abnormal returns. This is in contrast to previous research which found that the Market to Book Ratio had a significant negative impact on the CAR results (Gatzlaff & McCullough, 2010; Rosati et al. 2018). The amount of records breached did not have a significant effect on the CAR values. This is not in line with what would be expected, as a bigger data breach will often cost more for the company. Garg, Curtis & Halper (2003) found that breaches of credit card information had a much higher negative return, this was not found in the results here. The size of the company did not have an effect on the CAR values, in contrast to Gatzlaff and McCullough (2010) and Cavusoglu, Mishra, & Raghunathan, (2004) who did found a significant relationship between these two. Therefore, the second hypothesis, stating that the size of the breach has a positive significant relation with the effect on the share price, will be rejected.

Conclusion and discussion

In this paper, the reaction on a data breach announcement on the stock price in the period 2016 until 2018 was studied using event study methodology. In most previous research, using data from earlier periods, a data breach had as a consequence a significant negative reaction on the stock price. In the more recent studies, some found that in the more recent periods this reaction was less or non-existent. Here, no significant effect of a data breach on the stock price was found in the period 2016 until 2018.

This paper is the first to also look into the relationship of the size of the breach and the reaction. There was also no significant relationship found, meaning that the size of the breach had no influence on the reaction on the stock price. This contradicts with the principles of the efficient market theory, as one would expect that a bigger data breach would mean a bigger reaction on the stock price, as a bigger breach often means higher costs.

Some limitations existed when conducting the research. One is that the date of announcement as described in the Privacy Rights Clearinghouse database may in fact not be the date most of the investors received the news. Due to inside information or news leakage it is possible that some investors already incorporated the news before it actually released.

Secondly, the values of the cumulative abnormal returns were not significant. This means that the interpretation of the cross-sectional analysis can be problematic and has a low reliability. The non-significance could exist because of a few factors. One is that indeed the data breach announcements do not have an effect on the stock price in the period. It is also possible that the sample is too small to find a significant effect, even though other papers mostly do not use more events than in the used sample. Another is that the dates of the announcement of the breaches were not exact enough or that the breaches in this period are different than the breaches before.

Thirdly, some papers have only found significant results when using different models instead of the Market Model (Gordon, Loeb & Zhou, 2011; Schatz & Bashroush, 2016). This paper only factors in the Market Model, it could be possible that the results are significant when using a different model to perform the event study.

For future research a recommendation is to create a sample of data breaches in a longer period. Then it can be evaluated where the turning point is between a significant reaction on the stock price and no significant reaction. It would also be interesting to study the announcement of the consequences (the announcement of a fine for example) of the data breach instead of the data breach itself.

References

Andriotis, A. (2019, July 19). Equifax to Pay Around $700 Million to Resolve Data-Breach Probes. *The Wall Street Journal.* Retrieved from: https://www.wsj.com/articles/equifax-to-pay-around-700-million-to-resolve-data-breach-probes-11563577702

Acquisti, A., Friedman, A., & Telang, R. (2006). Is there a cost to privacy breaches? An event study. *ICIS 2006 Proceedings*, 94.

Art. 33 GDRP: Notification of a personal data breach to the supervisory authority. Retrieved from: https://gdpr-info.eu/art-33-gdpr/

Binder, J. (1998). The event study methodology since 1969. *Review of quantitative Finance and Accounting*, *11*(2), 111-137.

Campbell, K., Gordon, L. A., Loeb, M. P., & Zhou, L. (2003). The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security*, *11*(3), 431-448.

Campbell, J. Y., Champbell, J. J., Campbell, J. W., Lo, A. W., Lo, A. W. C., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. princeton University press.

Cavusoglu, H., Mishra, B., & Raghunathan, S. (2004). The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *International Journal of Electronic Commerce*, *9*(1), 70-104.

Chai, S., Kim, M., & Rao, H. R. (2011). Firms' information security investment decisions: Stock market evidence of investors' behavior. *Decision Support Systems*, *50*(4), 651-661.

De Jong, A., & Naumovska, I. (2015). A note on event studies in finance and management research. *Review of Finance*, *20*(4), 1659-1672.

Dyckman, T., Philbrick, D., & Stephan, J. (1984). A comparison of event study methodologies using daily stock returns: A simulation approach. *Journal of Accounting Research*, 1-30.

Fama, E. F. (1991). Efficient capital markets: II. *The journal of finance*, *46*(5), 1575-1617.

Gatzlaff, K. M., & McCullough, K. A. (2010). The effect of data breaches on shareholder wealth. *Risk Management and Insurance Review*, *13*(1), 61-83.

Garg, A., Curtis, J., & Halper, H. (2003). Quantifying the financial impact of IT security breaches. *Information Management & Computer Security*, *11*(2), 74-83.

Goel, S., & Shawky, H. A. (2009). Estimating the market impact of security breach announcements on firm values. *Information & Management*, *46*(7), 404-410.

Gordon, L. A., Loeb, M. P., & Zhou, L. (2011). The impact of information security breaches: Has there been a downward shift in costs?. *Journal of Computer Security*, *19*(1), 33-56.

Gressin, S. (2017). *The Equifax Data Breach: What to Do.* Federal Trade Commission,

Hendricks, K. B., & Singhal, V. R. (1996). Quality awards and the market value of the firm: An empirical investigation. *Management science*, *42*(3), 415-436.

Kam, K. (2017, September 21). After Falling 33%, Equifax Is Still Overvalued. *Forbes*. Retreived from: https://www.forbes.com/sites/kenkam/2017/09/21/after-falling-33-equifax-is-still-overvalued/#418257b62b88

Kothari, S. P., & Warner, J. B. (2007). Econometrics of event studies. In *Handbook of empirical corporate finance* (pp. 3-36). Elsevier.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of economic literature*, *35*(1), 13-39.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383-417.

Maloney, M. T., & Mulherin, J. H. (2003). The complexity of price discovery in an efficient market: the stock market reaction to the Challenger crash. *Journal of corporate finance*, *9*(4), 453-479.

McWilliams, A., & Siegel, D. (1997). Event studies in management research: Theoretical and empirical issues. *Academy of management journal*, *40*(3), 626-657.

Mirhaydari, A. (2018, May 10). Facebook stock recovers all $134B lost after Cambridge Analytica data scandal. *CBS News.* Retrieved from: https://www.cbsnews.com/news/facebook-stock-price-recovers-all-134-billion-lost-in-after-cambridge-analytica-datascandal/

Pirounias, S., Mermigas, D., & Patsakis, C. (2014). The relation between information security events and firm market value, empirical evidence on recent disclosures: An extension of the GLZ study. *Journal of Information Security and Applications*, *19*(4-5), 257-271.

Privacy Rights Clearinghouse. Data breaches. Retreived from: https://www.privacyrights.org/data-breaches

Privacy Rights Clearinghouse (June 27 2011).. DATA BREACHES: WHY YOU SHOULD CARE AND WHAT YOU SHOULD DO Retrieved from:
https://www.privacyrights.org/blog/data-breaches-why-you-should-care-and-what-you-should-do)

Schatz, D., & Bashroush, R. (2016). The impact of repeated data breach events on organisations' market value. *Information & Computer Security*, *24*(1), 73-92.

Spanos, G., & Angelis, L. (2016). The impact of information security events to the stock market: A systematic literature review. *Computers & Security*, *58*, 216-229.

Appendix

Table A1: Firm names and event dates

This table shows the data breach event dates and companies used in this paper. This is included because of recommendation by papers from McWilliams & Siegel (1997) and De Jong & Naumovska (2015).

Event date	**Company**
01/08/2016	Time Warner Cable
01/11/2016	Blucora (TaxAct)
01/13/2016	HSBC Bank USA, National Association
01/15/2016	Hyatt Hotels
01/25/2016	HealthEquity
01/26/2016	Centene
01/27/2016	Wendy's
02/03/2016	Rite Aid
03/01/2016	Walmart Stores, Inc.
03/04/2016	Walgreen Co.
03/16/2016	Advanced Auto Parts
03/16/2016	PerkinElmer, Inc.
03/24/2016	Verizon Enterprise Solutions
03/28/2016	Sprouts Farmers Market
04/05/2016	Pacific Gas and Electric Company
04/05/2016	Target Corporation Health Plan
04/21/2016	Voya Financial Advisor's Inc.
05/03/2016	Charles Schwab
05/05/2016	ADP, LLC.
05/06/2016	Equifax Inc.
05/06/2016	Google Inc.
05/16/2016	Noodles and Company
05/20/2016	Aflac
06/08/2016	Wal-Mart Stores, Inc.
06/13/2016	Twitter
06/16/2016	Multi-Color Corporation
08/26/2016	Toyota Motor Corporation
09/22/2016	Yahoo
10/12/2016	Vera Bradley
10/25/2016	Cisco
10/26/2016	Anthem, Inc.
11/04/2016	Wal-Mart Stores, Inc
11/22/2016	The Madison Square Garden Company
11/23/2016	Hewlett Packard Enterprise Services
11/30/2016	Google Android

12/05/2016	CVS Health
12/07/2016	T-Mobile
12/12/2016	Quest Diagnostics
12/14/2016	Yahoo
12/19/2016	Humana Inc.
12/20/2016	Western Union
12/28/2016	InterContinental Hotels Group (IHG)
01/27/2017	WellCare Health Plans, Inc.
02/02/2017	Sunrun
02/03/2017	InterContinental Hotels Group (IHG)
02/03/2017	Walgreen Co.
02/06/2017	Capital One
02/08/2017	The Boeing Corporation
02/27/2017	Boeing
03/06/2017	Barclays Bank
03/08/2017	CVS Health
03/13/2017	Tyler Technologies Inc.
03/16/2017	Aflac
04/18/2017	Humana Inc
04/19/2017	Northrop Grumman Systems Corporation
04/26/2017	Chipotle Mexican Grill
05/02/2017	Sabre Corporation
05/04/2017	Gannett Co
05/04/2017	Google Docs
05/11/2017	Intuit
05/17/2017	MolinaHealthcare.com
05/17/2017	Sabre Corporation
05/19/2017	Rite Aid
06/02/2017	Game Stop
06/19/2017	Bed Bath & Beyond
06/20/2017	The Buckle Inc.
07/05/2017	DXC Technology
07/24/2017	Anthem, Inc.
07/31/2017	Anthem
07/31/2017	Wells Fargo
08/14/2017	Performant Financial Corporation
08/29/2017	CoreLogic/Credco
09/07/2017	Equifax Corporation
09/18/2017	W. W. Grainger, Inc.
09/20/2017	Viacom
09/25/2017	Adobe
09/29/2017	Briggs & Stratton Corporation
10/12/2017	T-Mobile
10/13/2017	CVS Pharmacy

10/17/2017	Insulet Corporation
11/02/2017	Kimberly-Clark
11/14/2017	ABM Industries
11/16/2017	Hyatt Hotels
11/21/2017	Humana Inc
12/21/2017	Molina Healthcare
01/08/2018	CyrusOne, Inc.
01/17/2018	Ameriprise Financial, Inc.
01/22/2018	The Coca-Cola Company
01/29/2018	Nevro
02/02/2018	Triple-S Advantage, Inc.
02/07/2018	Nevro
02/09/2018	Intuit Inc.
02/09/2018	OneMain Financial
02/12/2018	Goldman Sachs & Co. LLC
02/13/2018	Bed Bath & Beyond, Inc.
02/15/2018	Dollar General Corporation
02/16/2018	Marriott International Inc.
02/16/2018	Navistar, Inc.
02/20/2018	OneMain Financial
02/22/2018	Walmart, Inc.
03/15/2018	UnitedHealth Group Single Affiliated Covered Entity
03/26/2018	Walmart Inc.
04/06/2018	Best Buy
04/06/2018	Delta Air Lines, Inc.
04/06/2018	Walgreen Co.
04/17/2018	Inogen, Inc.
04/17/2018	MAXIMUS, Inc. / Business Ink, Co.
04/20/2018	W. W. Grainger, Inc.
04/20/2018	SunTrust Banks, Inc.
04/27/2018	Walgreen Co.
05/24/2018	T-Mobile
05/29/2018	Aflac
06/12/2018	Facebook, inc.
06/12/2018	HealthEquity, Inc.
06/12/2018	Nuance Communications
06/13/2018	WellCare Health Plans, Inc.
06/22/2018	InfuSystem, Inc.
07/17/2018	One main financial
09/28/2018	Facebook, Inc.
10/08/2018	Alphabet, Inc. - Google+
10/10/2018	Cigna
10/25/2018	CNO Financial Group, Inc.
11/30/2018	Marriott International