ON THE ECONOMIC IMPACT OF INFORMATION SECURITY ANNOUNCEMENTS: AN EVENT STUDY ANALYSIS

Adrian Ford

School of Architecture, Computing and Engineering
University of East London
University Way, Royal Docks, London E16 2RD
{a.ford1701}@uel.ac.uk

April 2023

University of East London

A thesis submitted in partial fulfilment of the requirements of the University of East London for the degree of Professional Doctorate in Information Security

(Word count: 49,051 excluding Appendix)

Abstract: This research is concerned with the economic impact of information security events both unfavourable (data breaches and GDPR infringement fines) and favourable (CISO appointment announcements). Literature in this area was found to be sparse and with a strong US bias, therefore this study focusses on UK and European markets. Using event study methodology, the impact on share price of a hand-gathered (due to lack of a comprehensive breach database for Europe) dataset of 45 data breach announcements concerning UK/European publicly listed companies was analysed and only weak evidence was found of a negative impact overall, although the Spanish market showed a greater reaction. Regarding GDPR infringement fine announcements (25 examples), statistically significant CARs of -1% on average were observed over a three-day period. Spanish and Romanian markets were shown to be particularly reactive. Such a loss in market capitalisation was, in almost all cases, much greater than the monetary value of the fine itself, actually ca. 29,000 times greater on average. Announcements of CISO type role appointments (37 examples) showed an uplift in share price of around 0.8% on average over a three-day period before, during and after the announcement. The financial services sector was found to respond more positively (+1.8%) with statistical significance at the 1% level. As well as highlighting the benefits of transparency by publicly listed firms and disclosure regulations in early-adopter nations such as the US, the results of these studies should encourage firms to improve their cyber security posture overall to emulate highly regulated sectors such as financial services. A review of security investment strategies is also included for convenience, as well as pointers for future research. This research would be of benefit to business management, practitioners of cybersecurity, investors and shareholders, policy makers as well as researchers in cyber security or related fields.

Contents

Abstract i

Contents ii

List of figures v

List of tables vi

List of abbreviations vii

Preface and acknowledgements x

Chapter 1. Introduction 1

1.1. Research background 1

1.2. Research aims and objectives 2

1.3. Research questions 2

1.4. Research methodology 3

1.5. Research importance and contribution 3

1.6. Publications 6

1.7. Thesis layout 6

Chapter 2. Literature Review 8

2.1. Definitions 8

2.1.1. Information Security 8

2.1.2. Cyber security 9

2.1.3. Information security versus cyber security 10

2.1.4. Cyber security versus cybersecurity 12

2.1.5. Data breach 14

2.1.6. Cybercrime (and cybercriminals) 14

2.1.7. Econometric and financial terminology 16

2.1.8. Conclusion 17

2.2. Literature searches 17

2.3. Analysis and discussion 36

2.4. Addendum 41

2.5. Conclusion 43
Chapter 3. Methodology (General Approach) 45
3.1. Introduction 45
3.2. Event Study methodology (ESM) 45
3.3. EventStudyTools package 49
3.4. Data collection 50
3.5. Data analysis 50
3.6. Hypothesis development 52
3.7. Comparison of the Market Model with the Fama-French 3-Factor Model ..53
3.8. Validation of the method against literature 55
3.9. Conclusion 60
Chapter 4. The Impact of Data Breach Announcements on Company Value in European Markets 62
4.1. Introduction 62
4.2. Methodology 64
4.3. Data collection 64
4.4. Data analysis 66
4.5. Hypothesis development 67
4.6. Results and discussion 67
4.7. Conclusion 73
Chapter 5. The Impact of GDPR Infringement Fines on the Market Value of Firms 78
5.1. Introduction 78
5.2. Methodology 81
5.3. Data collection 82
5.4. Data analysis 83
5.5. Hypothesis development 83
5.6. Results and discussion 83
5.7. Conclusion 89

Chapter 6. (The CISO Effect:) The Impact of CISO Appointment Announcements on the Market Value of Firms 93
6.1. Introduction 93
6.2. Methodology 94
6.3. Data collection 95
6.4. Data analysis 96
6.5. Hypothesis development 96
6.6. Results and discussion 96
6.7. Conclusion 103
Chapter 7. The Impact of Repeated Information Security Events on Market Value 109
7.1. Introduction 109
7.2. Results and discussion 111
7.3. Conclusion 118
Chapter 8. Investment in Information Security 120
8.1. Introduction 120
8.2. Related work and discussion 120
8.3. Conclusion 126
Chapter 9. Overall Conclusion and Contribution to Knowledge 128
9.1. Introduction 128
9.2. Summary 128
9.3. Reflection on research questions 129
9.4. Contribution to knowledge 131
9.5. Research limitations 132
9.6. Pointers to future research 133
References 134
Appendix (R Code) 145

List of figures

Figure 1: Thesis chapter flow with associated publications....5
Figure 2: Comparison of information security and cyber security....9
Figure 3: Google search trends....12
Figure 4: Literature usage of security terms....13
Figure 5: Initial literature search in Scopus....18
Figure 6: Breach Level Index....41
Figure 7: Efficient market hypothesis....46
Figure 8: Event study timeline....47
Figure 9: CAARs for different types of corporate events....49
Figure 10: estudy function outline....51
Figure 11: Comparison of EST AR values with Castillo and Falzon (2018)....59
Figure 12: Comparison of EST t-test values with Castillo and Falzon (2018)....60
Figure 13: Boxplots of CAR values per event window....68
Figure 14: CAAR by industry sector....69
Figure 15: CAR versus records breached....70
Figure 16: Comparison of event windows....84
Figure 17: Comparison of event windows....97
Figure 18: Breakdown by year....99
Figure 19: Repeated events for IAG....117
Figure 20: Gordon-Loeb investment model....122

List of tables

Table 1: Literature review 20
Table 2: Comparison of MM and FF3FM 55
Table 3: Comparison of CAR calculation methods 56
Table 4: Analysis of event window (0, 2) by sector 69
Table 5: Analysis of event window (0, 2) by sector (personal data) 70
Table 6: Market effect of GDPR enactment 71
Table 7: Analysis of event window (0, 2) by sector (SPEUR350) 72
Table 8:Analysis by market index for event window (0, 1) 73
Table 9: List of Data Breaches 75
Table 10: CAAR by event window 85
Table 11: Analysis by ultimate parent company 86
Table 12: CAAR by industry sector 87
Table 13: Analysis by country 88
Table 14: Summary of GDPR fine appeals 88
Table 15: CAR by event window of fines appealed 89
Table 16: List of GDPR Infringement Fine Announcements 91
Table 17: CAAR by event window 98
Table 18: CAAR by industry sector 100
Table 19: CAAR by job title 101
Table 20: CAAR by CISO reporting line 102
Table 21: Analysis of new or established CISO roles 102
Table 22: Analysis by market currency 103
Table 23: List of CISO Appointment Announcements 105
Table 24: Summary of repeated events by company (stock symbol) 110
Table 25: Data breaches repeated events (-2, 2) 111
Table 26: Data breaches repeated events (0, 4) 112
Table 27: Comparison of CAAR for repeated breach events 113
Table 28: Repeated events (GDPR infringement fines) event window (0,3) 113
Table 29: Repeated events (GDPR infringement fines) event window (-2, 2) 114
Table 30: Repeated events (CISO appointments) event window (-1, 1) 114
Table 31: Repeated events (CISO appointments) event window (-,2 2) 114
Table 32: Repeated events (all) event window (-2, 2) 115

List of abbreviations

AMEX American Stock Exchange

API Application programming interface

AR Abnormal return

BHAR Buy-and-hold Average Return

BLI Breach level index

CAAR Cumulative average abnormal return

CAPM Capital asset pricing model

CAR Cumulative abnormal return

CEO Chief Executive Officer

CERT Computer Emergency Response Team

CFO Chief Financial Officer

CIA Confidentiality, Integrity and Availability

CIO Chief Information Officer

CISO Chief Information Security Officer

COGS/S Cost of Goods Sold to Sales ratio

COO Chief Operating Officer

COVID-19 Coronavirus disease (2019)

CRAN Comprehensive R Archive Network

CRO Chief Risk Officer

CRSP Centre for Research in Security Prices (US)

CSO Chief Security Officer

CSV Comma separated variable (file format)

CTI Cyber Threat Intelligence

CTSO Chief Technology Security Officer

CUSIP Committee on Uniform Security Identification Procedures

CyBOK Cyber Body of Knowledge

DBNO Data Breach Notification Obligation

DCMS Department for Digital, Culture, Media and Sport (UK)

DoS Denial of Service

DPA Data Protection Act/Authority

ECCWS European Conference on Cyber Warfare and Security

ECIME European Conference on Information Management and Evaluation

EMH Efficient Market Hypothesis

ENISA European Union Agency for Cybersecurity

ESM Event Study Methodology

EST EventStudyTools

ETF Exchange traded fund

EU European Union

FF(3,4,5)FM Fama-French (3, 4 or 5) Factor Model

GARCH Generalized AutoRegressive Conditional Heteroskedasticity

GDPR The General Data Protection Regulation (EU)

GUI Graphical User Interface

IC3 Internet Crime Complaint Center (US)

ICCWS International Conference on Cyber Warfare and Security

ICO Information Commissioner's Office (UK)

IEC International Electrotechnical Commission

IRR Internal Rate of Return

IoT Internet of Things

ISMS Information Security Management System

ISO International Organisation for Standardisation

M&A Mergers and Acquisitions

MM Market Model

NAICS North American Industry Classification System

NASDAQ National Association of Securities Dealers Automated Quotations (US)

NCSC National Cyber Security Centre (UK)

NGO Non-Governmental Organisation

NIST National Institute of Standards and Technology (US Department of Commerce)

NPV Net Present Value

NYSE New York Stock Exchange (US)

OLS Ordinary least squares (linear regression model)

PRC Privacy Rights Clearinghouse

ROA Return on Assets

ROI Return on Investment

ROS Return on Sales

SEC Securities Exchange Commission (US)

SGA/S Selling, General and Administrative to Sales ratio

SLR Systematic Literature Review

SME Small and Medium-sized Enterprises

SOX Sarbanes-Oxley Act

US(A) United States (of America)

VCDB Veris Community Database

WEIS Workshop on the Economics of Information Security

Preface and acknowledgements

I began this journey back in 2017 and am finally reaching the end as we approach autumn 2022. Thanks are due to Sumitomo Corporation Europe Ltd (SCEU) for their support over these years. A number of SCEU colleagues have helped along the way, but special thanks go to Roger Garside for his valuable proof-reading efforts, often under major time pressure. Thanks are also due, of course, to my supervisory team at the University of East London (UEL), Dr. Ameer Al Nemrat and Dr. Seyed Ali Ghorashi of the School of Architecture, Computing and Engineering and Professor Julia Davidson OBE, Pro Vice-Chancellor, Impact and Innovation. I must also thank my wife and family for their patience and input, as well as acknowledge additional help and advice from my fellow UEL classmates and the Security Panel of the Worshipful Company of Information Technologists.

Five years is a long time. During this process, my bedtime reading has consisted only of books, trade journals or research papers related to my area of study. My wife will be pleased to hear I can now finally move onto her recommended reading list. When I began this process in 2017, I had one daughter, now I have two and we have moved home twice since. As I am writing this, my eldest daughter is preparing to move into Year 1, closely followed by her little sister who will join her in pre-school under the same roof. Just as they are beginning their formal academic education, I am writing the final chapter on mine, so to speak. However, as Isaac Asimov once said, education is not something you can finish.

Whilst we are on the subject of children and education, I recently read a version of Little Red Riding Hood to my daughters at bedtime. There were some clear information security themes in this well-known fairy tale, the social engineering skills of the wolf lead to disclosure of personal information (the description and location of her grandmother's house), despite advance warnings of possible malicious lupine activity in the woods (security awareness training), which later results in identity theft putting Little Red Riding Hood in danger as well as her poor grandmother. Even though this tale is hundreds of years old, the messages are still as relevant as ever and can easily be transposed to cyber space, the internet being the deep dark wood, the wolf a black hat hacker and Little Red Riding Hood being the weakest link in security herself (the human factor) despite the awareness training. Yes, after all this studying it is hard for me not to analyse everything through the cyber security lens. I also came across another key cyber security message in Hansel and Gretel where the eponymous children encounter a house made of gingerbread,

cake with windows of clear sugar – the fairy tale equivalent of a phishing email – as we frequently advise our user community: 'if it looks too good to be true, it usually is'.

The importance of cyber security was brought to the fore by something no one could have foreseen back in 2017, COVID-19 and the challenges of remote working. Luckily this research was always planned to be mostly internet based, nevertheless, a fundamental expectation was 'normal' behaviour of the financial markets and so data had to be capped to avoid COVID effects thereby reducing the quantity available to work with. Certainly, COVID-19 introduced some challenges in completing this thesis, not to mention being unable to attend UEL campus since the final taught module (at least I had the benefit of attending all five block modules on site and meeting classmates and networking). In fact, virtually the whole of the final two-year research component was carried out remotely (all supervisory meetings so far) reflecting the 'new normal'. I was hoping to attend WEIS 2020 as an observer (with a view to presenting research at WEIS 2021) and was looking forward to visiting Brussels after a long absence. Alas, that ended up being a virtual conference despite being postponed until December that year. My only previous experience of an academic conference (if we exclude internal UEL conferences) was ECIME 2010 where I presented a paper based on my master's dissertation concerning the adoption of IT in small and medium sized companies and I greatly enjoyed my first ever trip to Lisbon, Portugal and meeting and socialising face-to-face with other like-minded people. By the time WEIS 2021 came around in June, unfortunately the situation was no better, and I ended up attending and presenting remotely. It would, sadly, be exactly the same scenario for ECCWS 2021 (also in June). By this time, I was clearly beginning to get the hang of remote presentations as our ECCWS 2021 paper won the joint (PhD) runner-up prize. This achievement was eventually surpassed by ICCWS 2022 in March which, although a hybrid event (at the State University of New York at Albany), I was not able to attend in person due to work commitments. Thus "The CISO Effect" was presented remotely yet again but this time an outright winner in the PhD category – showing that practice makes perfect. Hopefully I will be attending a research conference in person in the near future, now we are in the twilight of the pandemic.

Anyway, now the scene is set and without further ado, let us move on to the core content of this thesis, which I very much hope you will enjoy reading.

Adrian Ford BSc(Hons) MBA MBCS CChem MRSC
London, September 2022 (Revised January 2023)

Chapter 1. Introduction

1.1. Research background

Information security has come very much to the fore in recent years. Not only have there been some very high-profile examples of data breaches reported in the media such as that of British Airways (Bloomberg, 2018), Marriott (Forbes, 2018) and LinkedIn (Fortune, 2021) but, more recently the COVID-19 pandemic has further increased the profile of cyber security with Interpol (2022) reporting that *"cybercriminals are taking advantage of the widespread global communications on the coronavirus to mask their activities"*, an observation echoed by the UK Government (Cabinet Office, 2022) in its National Cyber Strategy: *"the past year has seen cyber attacks on hospitals and oil pipelines, schools and businesses, some brought to a standstill by ransomware, and commercial spyware used to target activists, journalists and politicians".* According to CyBOK, these cyberattacks cost global economies an estimated $400bn and remain *"an increasing political, societal and economic concern"* (CyBOK, 2022).

This increase in cyberattacks over the years has clearly not gone unnoticed by governments, who have taken measures to protect critical national infrastructure, their citizens and both public and private enterprises from cyber hostility through awareness campaigns, industry standards and introduction of new legislation (such as the GDPR) aimed at increasing cyber resilience. For example, the DCMS (2020) reports that *"75% of companies from the UK reported investment in cyber security due to the GDPR requirements"*.

With all this heightened interest and activity in the area of information security, it is only natural to ask what lessons might be learnt from security incidents such as data breaches and how might any negative effects be accurately measured? Similarly, how might one quantify any benefits of positive security events such as investment in protection measures or the introduction of new legislation? Various sources report difficulties in measuring the cost of security events accurately (i.a. Makridis and Dean, 2018). Visibility of private enterprises is naturally limited to what they are willing or legally obliged to declare. It makes sense, therefore, to focus on publicly listed companies as there is greater transparency through market reactions to security related events which is the basis for this work.

An initial literature review of the economic impact of information security events (see Chapter 2 for more detail) uncovered some gaps – specifically a very strong US bias in studies of this type (Spanos & Angelis, 2016) and a lesser proportion of papers reporting on the impact of favourable security events (Ali et al., 2021), such as the introduction of new legislation or investment in security measures. This thesis begins to plug these gaps by focussing on UK and European markets (Chapter 4), the introduction of the GDPR (Chapter 5) and investment in human capital related to information security (Chapter 6).

1.2. Research aims and objectives

The primary aim of this research is to investigate and analyse the economic impact (company market value) of such information security events (specifically data breaches, GDPR infringement fines and CISO recruitment) to provide supporting evidence for business cases concerning investment in security measures.

A secondary aim is expansion of the existing knowledge base in this area which is addressed by the UK and EU focus of this thesis to help offset the US bias of previous work as well as contributing to a lack of studies on the GDPR and favourable information security events in general.

The overall aims are be achieved through the following objectives:

RO1. To investigate the impact (if any) of information security events on the market value (share price) of companies.
RO2. To identify any patterns/correlations between market value and other factors such as cyber breach categories or industry sectors.
RO3. To investigate the economic impact of the introduction of legislation such as the GDPR
RO4. To investigate whether current frameworks for business investment decision making are taking into consideration the importance of cyber security and, if not, to highlight such gaps.

1.3. Research questions

Based on the research objectives above, the following research questions were proposed and have been revised and expanded following the literature review in Chapter 2. Specifically, RQ1 was extended to consider both favourable and unfavourable security events (in light of the paucity of studies regarding the former) and RQ3 became more focussed on infringement fines due to the availability of such data and, again, lack of

existing such studies. Subsequently, it was also extended to incorporate the results of GDPR fine appeals both successful and unsuccessful.

RQ1. What is the impact (if any) on share price of a security event, be it favourable or unfavourable and how do these findings compare with the literature?

RQ2. Are there any patterns in the data, such as correlations between drop in market value and category of cyber-attack, data breach, industry sector etc.?

RQ3. Regarding the introduction of the GDPR, what is the economic impact of infringement fines on the market value of firms, including those appealed and overturned?

RQ4. How can these findings be incorporated into the security investment strategies of organisations?

1.4. Research methodology

In order to answer the above research questions, based on the extensive literature review in Chapter 2, Event Study Methodology (ESM) was chosen as the most appropriate approach and is described in detail in Chapter 3. This was the most prevalent method used in previous studies regarding the impact of information security events on the share price of publicly listed firms. As RQ4 is a question requiring an answer of a more qualitative nature, however, that is addressed separately in Chapter 8.

1.5. Research importance and contribution

Through a deeper understanding of the economic impact of security events, businesses are better positioned in decision making concerning information security investment, which is explored in Chapter 8. This research makes a contribution to the knowledge base (as indicated above) by focussing on UK/EU markets thereby offsetting the US bias seen in existing studies (Chapters 4 and 5). Due to relatively recent introduction of the GDPR (2018) there is also a paucity of studies regarding its introduction to which this thesis makes a contribution. The dearth of literature concerning CISOs in general (Karanja & Rosso, 2017) is also addressed in Chapter 6, along with the recognised lack of studies focussing on favourable information security events (Ali et al., 2021). Studies regarding the impact of repeated data breaches were also found to be lacking (Schatz & Bashroush, 2016a) and Chapter 7 contributes to this area.

In addition to the above contributions, there is also useful input here into event study methodology in general (choice of event window and market reference, approach to handling confounding events, comparison of MM and FF3FM), including usage of the

EST package (and comparison of thesis results with literature). This package was not reported as having been used before in studies of this type.

Pointers to future research are also given in Chapter 9. Areas of possible interest noted include the lack of a comprehensive breach database for Europe (Chapter 4), the apparent increased sensitivity of the Spanish markets (Chapters 4 and 5), the large magnitude of GDPR fine appeal abnormal returns (Chapter 5) as well as the positive impact of CISO appointment announcements and need for more transparency in this area (Chapter 6). These studies were all hampered by market effects of COVID-19 and would all benefit from being revisited in future once the markets re-stabilise after the pandemic.

The research contributions are also summarised along with justifications for each in Chapter 9. This research would be of benefit to business management, practitioners of cyber security, investors and shareholders, policy makers as well as researchers in cyber security or related fields.

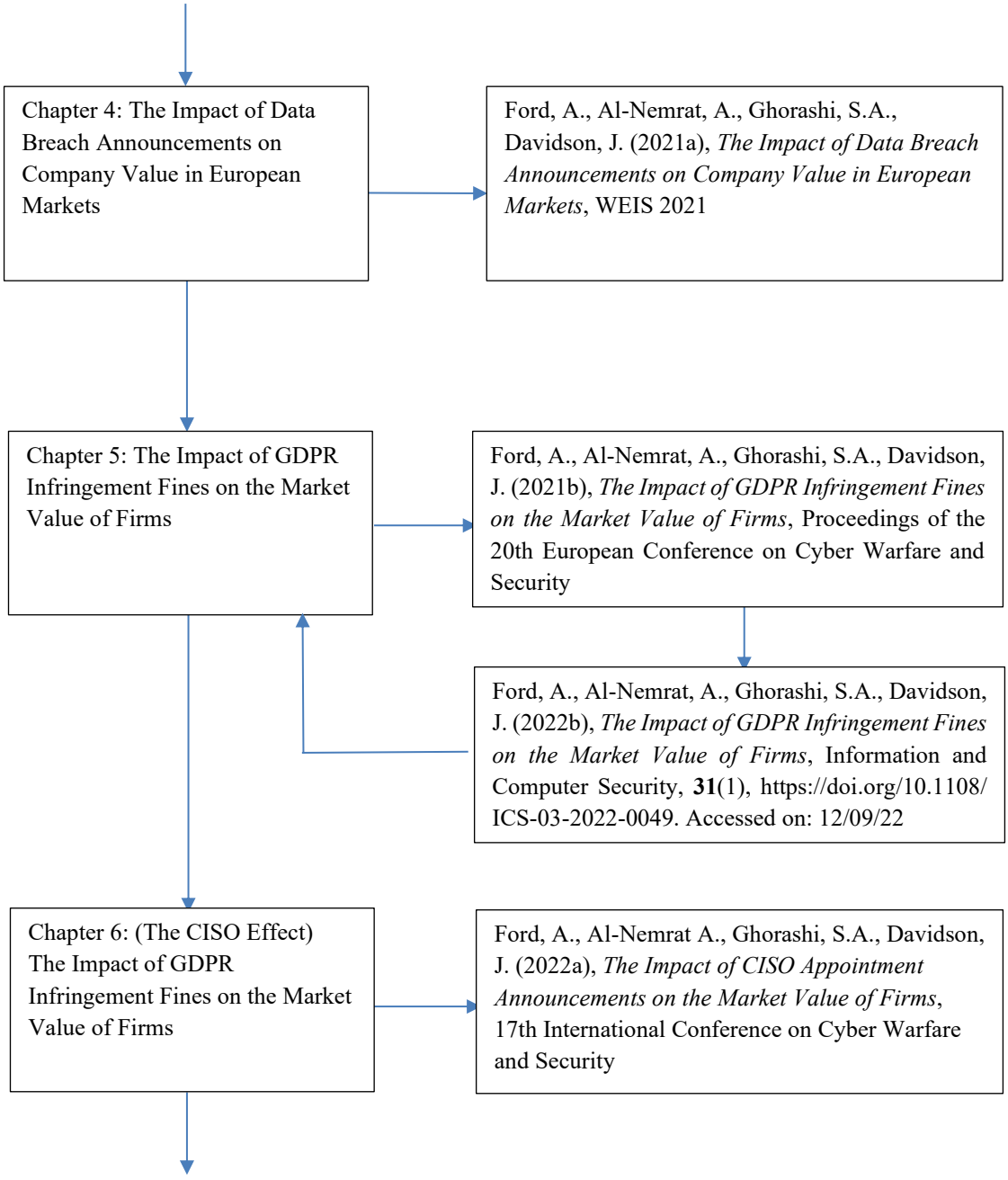


Figure 1: Thesis chapter flow with associated publications

1.6. Publications

The list of publications associated with this thesis is shown (along with the chapter flow) in **Figure 1**. In summary, Ford et al. (2021a) was based on Chapter 4, Ford et al. (2021b) on Chapter 5, and Ford et al. (2022a) on Chapter 6. Subsequently, Ford et al. (2021b) was revised and expanded into a journal paper (Ford et al., 2022b).

1.7. Thesis layout

It is useful here to expand on the thesis chapter flow shown in **Figure 1** to give an overview of the structure of this manuscript in its entirety.

This chapter (Chapter 1) has given some background to the research along with the aims and objectives, why it is relevant and the intended audience. An explanation of how the research questions have been developed from the aims and objectives is included, along with an overview of the methodology used to answer them. Contributions to knowledge and associated publications are also listed.

The next chapter (Chapter 2) gives a detailed overview of the literature searches carried out, along with a review and definitions of terminologies in use and how gaps in the knowledge base shaped this work.

The event study methodology and common approaches to data collection and analysis as well as hypothesis development applied in Chapters 4, 5 and 6 are described in detail in Chapter 3. Also contained therein is a validation of the software package used by comparing both with literature and other methods.

The core research chapters (Chapters 4, 5 and 6) each examine the economic impact of different types of information security events through the calculation of abnormal returns on the share prices of publicly listed companies. Chapter 4 is concerned with data breach announcements in EU markets to offset the strong US bias seen in the literature whereas Chapter 5 concentrates on the GDPR, beginning with infringement fine announcements and then researching the impact of fine appeals and the impact of the introduction of data protection legislation overall. The lack of literature on favourable information security events is the focus of Chapter 6 which uses the same approach (as described in Chapter 3) to measure positive returns from CISO appointment announcements.

These three core chapters are all linked together in Chapter 7 which examines repeated events on companies appearing more than once in the three core chapter datasets.

Chapter 8 is a brief review of investment in information security with a view to advising organisations how much to spend on information security and what to spend it on.

Chapter 9 begins with a brief summary of this thesis as a whole, followed by a reflection on each of the research questions. The contributions to knowledge are then described and justified and, subsequently, challenges during the process are highlighted (research limitations) before pointers to future research are listed. After listing references, the thesis concludes with an appendix showing the R code used to generate the relevant figures and tables.

Chapter 2. Literature Review

2.1. Definitions

In advance of completing literature searches, to ensure better results, it was necessary to review terminologies in some detail as it was found very early on in the process that there were multiple variations in use.

2.1.1. Information Security

To begin with, the term *"information security"* (sometimes shortened to *"infosec"*) is defined by (ISO/IEC, 2009) as the *"preservation of confidentiality, integrity and availability of information"*. This concept, commonly referred to as the *"CIA triad",* is generally attributed to NIST[1] (Neumann, Statland & Webb, 1977: 11-3,4). It has become well established and is frequently quoted in security literature (i.a. Edgar & Manz, 2017; Karanja & Rosso, 2017; Ali et al., 2021). This model has been updated more recently to incorporate later developments in the field, such as that of non-repudiation in the context of blockchain. The 'Parkerian hexad' (Parker, 1998), for example, expands the CIA triad to include the additional elements of possession, authenticity and utility. Indeed, the ICO reflect the additional characteristic of authenticity in their definition: *"the ability of network and information systems to resist, at a given level of confidence, any action that compromises the availability, authenticity, integrity or confidentiality of stored or transmitted or processed data or the related services offered by, or accessible via, those network and information systems"* (ICO, 2022). The ISO/IEC, NIST and ICO definition are clearly very much aligned. A more practical definition is given by CSOOnline which defines information security as "*a set of practices intended to keep data secure from unauthorized access or alterations, both when it's being stored and when it's being transmitted from one machine or physical location to another*" (CSO, 2020). This definition also introduces the concept of an alternative term 'data security', both 'at rest' and 'in-transit' – important considerations for any practitioner. The kinds of data involved are indicated more explicitly in the SANS definition: *"Information Security refers to the processes and methodologies which are designed and implemented to protect print, electronic, or any other form of confidential, private and sensitive information or data from unauthorized access, use, misuse, disclosure, destruction, modification, or*

[1] Known as the National Bureau of Standards at that time.

disruption" (SANS, 2022). Here, a reference is made to 'print' – a physical form of information as well as electronic[2].

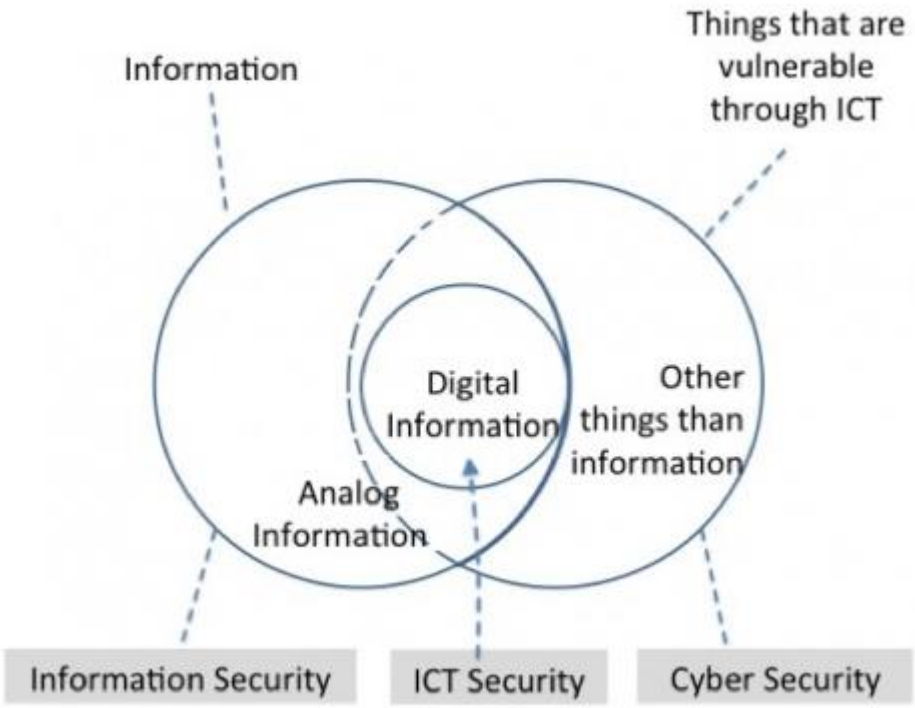


Figure 2: Comparison of information security and cyber security (Source: https://en.wikipedia.org/wiki/File:Cybersecurity_vs_information_security.png. Accessed on: 04/04/23)

2.1.2. Cyber security

Now that the scope of the term "*information security*" has been clarified, it makes sense to move on to that of *"cyber security"*. The UK Government in their Cyber Strategy (Cabinet Office, 2022) define cyber security as the protection of information assets restricted within the domain of cyberspace, therefore it is also necessary to clarify the meaning of the expression "*cyberspace*" beginning with the use of the prefix "*cyber*". The use of this term in English[3] was first reported by the American mathematician Norbert Wiener (Wiener, 1948) where he coined the term "*cybernetics*" as meaning *"the science of communications and automatic control systems in both machines and living things"* (OED, 2022). Coupled with 'space' the resulting term *"cyberspace"* is defined by NIST as *"the complex environment resulting from the interaction of people, software*

[2] The author recalls being reminded "The Data Protection Act (1998) never mentions the word *computer*". Sensitive data on paper is just as important.
[3] The term "cybernétique" had previously been reported in French literature (1834) by André-Marie Ampère but was in a political context (the science of government).

and services on the Internet by means of technology devices and networks connected to it, which does not exist in any physical form" (NIST, 2022a).

It is interesting to note that the definitions of both terms cybernetics and cyberspace include a human factor, *"living things"* or *"people"*, so cyberspace is more than just the concept of interconnected digital technology - Merriam-Webster (2022) define it simply as *"the online world of computer networks and especially the Internet"*, for example – a point which is revisited later in consideration of how security breaches occur and the motivation (of the bad actors) behind them. One could argue the same would apply to information security and the concept of 'unauthorised access' yet Von Solms and Van Niekerk (2013:97) opine that *"in cyber security this [human] factor has an additional dimension, namely, the humans as potential targets of cyber attacks or even unknowingly participating in a cyber attack"*. It could further be argued that a victim of identity theft through analogue (information security) means is subject to harm, but Von Solms and Van Niekerk (2013) observe that in the case of a cyberattack the impact on the individual is more direct, whereas in the case of identity theft (information security) it is purely the information that is compromised. Consider the case of cyberbullying, for example, here the psychological effect on the target can have a profound and lasting effect. How such interactions in cyberspace (and technology in general) affect humans has given rise to the growing science of cyberpsychology[4] and what has become known as *"The Cyber Effect"* (Aiken, 2017). As both individuals and organisations increase their adoption of digital technology and become part of this interconnected community known as cyberspace, the more they become at risk from this (more direct) "*cyber effect*".

2.1.3. Information security versus cyber security

There is clearly a difference, therefore, between the two terms information security and cyber security. Whilst Von Solms and Van Niekerk (2013:97) note that *"the term cyber security is often used interchangeably with the term information security",* they go on to argue that *"although there is a substantial overlap between cyber security and information security, these two concepts are not totally analogous"*. For comparison purposes, a useful visualisation is shown in **Figure 2**. Starting with the example of an analogue information only breach such as the loss of UK Ministry of Defence papers in Kent (The Independent, 2021), this is clearly classed as an information security event (as opposed to a cyber security event) as no digital data was involved. However, in the case

[4] Perhaps this could be shortened to 'cychology' although, as 'cyber' is a morpheme, it may not be linguistically appropriate.

of devices connected to the internet (IoT), an attacker could take control of, for example, a driverless car, a passenger aircraft or a nuclear power station. Such a breach would appear on the right-hand side of the diagram (**Figure 2**) and not necessarily be in scope for information security, rather a cyberattack. The cyberbullying referred to above would also, again, be out of scope for information security. It becomes clear at this point that the terms are, indeed, not analogous and nor is information security a superset of cyber security and analogue information. IT (or ICT) security refers specifically to digital information assets within an organisation which would again exclude any analogue information but may well be connected to the internet and thus be vulnerable to cyber-attackers as well as internal threats.

The difference between information security and cyber security definitions was also researched by Schatz (2018) who reported that "*the scope of the term 'cyber security' is closer to that of systemic or macroeconomic concerns, whereas 'information security' is more focused at the organisational level*". It seems that, perhaps, the scope of these terms is changing over time with cyber security becoming more relevant for individuals as well as organisations as they become more connected with cyberspace (and thereby cybercrime which is defined later).

It would also be interesting to look if there is any change over time in usage of these terms. Indeed, Schatz (2018)[5] presented a Google search trend chart which has been updated in **Figure 3**.

[5] https://trends.google.com. Accessed on: 04/04/23

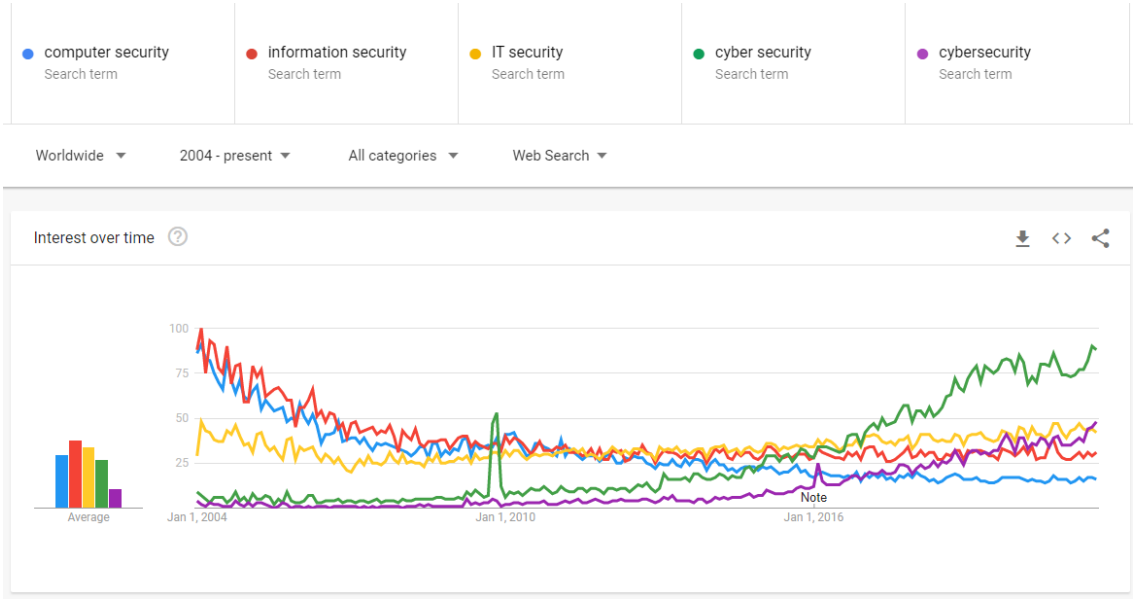


Figure 3: Google search trends

2.1.4. Cyber security versus cybersecurity

Although "*information security*" has been searched for more overall, the trending terms are "*cyber security*" or "*cybersecurity*". There seems to be some variation on the exact syntax here, which is worth noting. Although use of the hyphenated term "*cyber-security*" was negligible, it appears that "*cyber security*" outnumbers the undisjointed term "*cybersecurity*", so this spelling is used wherever possible throughout this thesis (except direct quotations, of course). This approach is consistent with i.a. Edgar and Manz (2017) who state: "*there are varying perspectives on how to write out cyber space. The etymology of the word comes from joining the words cybernetics and space. As you see throughout this book we chose to use the two-word version. Cyber has become a commonly used adjective that relates things to metaphysical, virtual, or digital representations. It has become a modifier similar to physical. More awkwardly, sometimes the word cyber is used as shorthand for cyber security, which while semantically untrue, has been gaining traction. "Oh you work in cyber" means cyber security and not cyber space or cyber space-related fields. Finally, real-world examples of security can be used to help explain one word or two. For example, National Security, Social Security, physical security, home security, network security, computer security, are all two words. The lack of understanding of the word cyber seems to force some to merge it into one, misleading concept, cyber security. For all of these reasons, we prefer, and will continue to use cyber security, as two words.*" It appears, therefore that cyber is not just a modifier or prefix but a word (actually a morpheme) in its own right as evidenced by its use as both an adjective

and a noun. Cyber has certainly overtaken other examples of prefixes (and morphemes) relating to computers and internet such as "e-", "i-" and "virtual".

To compare with Google search trends, literature searches for these terms were also carried out using Scopus and the results shown (**Figure 4**).

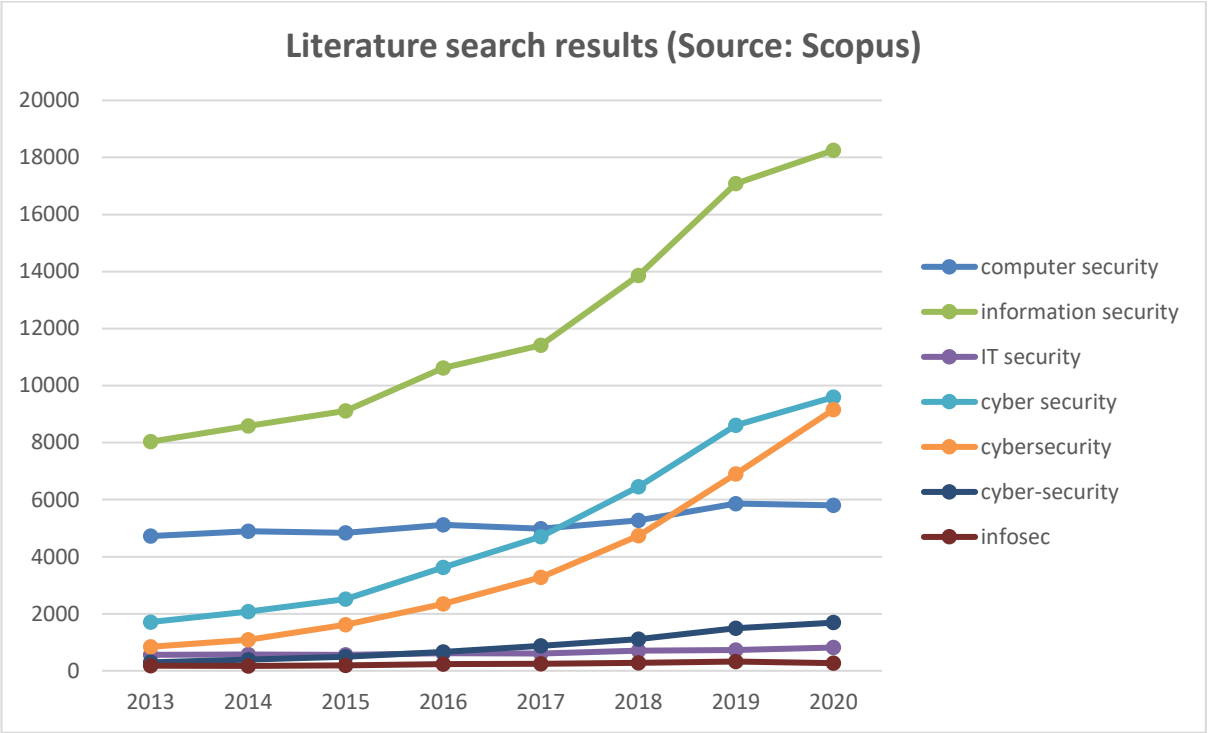


Figure 4: Literature usage of security terms (Source: Scopus)

It appears that many more articles are continuing to use the term "*information security*" perhaps reflecting its broader, more historical definition (analogue as well as digital). Nevertheless, the sum of the three search terms involving "*cyber*" is higher. Also, there is a visible trend that the single word version is going to overtake that of the disjointed version in 2021, perhaps going towards consistency with e.g. *"cyberspace"*. Note that these results are rather different to the Google search terms results (above) where "*information security*" was a much less popular term – it appears that there is a difference in terminology between academic literature and Google searches which, again, may change over time as more futuristic terms come into common use such as "*metaverse*"[6].

[6] *"The Metaverse is a collective virtual open space, created by the convergence of virtually enhanced physical and digital reality. It is physically persistent and provides enhanced immersive experiences"* (Gartner Inc., 2022).

2.1.5. Data breach

Going back to the definition of "*information security*" above, this is the preservation of the CIA Triad (or, if one prefers, Parkerian Hexad). Therefore, where there has been a security failure, and one or more of these elements is/are compromised, this consitutes a security incident or event known as a "*data breach*" (also known as a data leak or data loss). A "*data breach*" as defined by NIST (2022a) is *"an incident that involves sensitive, protected, or confidential information being copied, transmitted, viewed, stolen or used by an individual unauthorized to do so".* The NCSC (solely concerned with cyber, of course) refers to such an incident as a "*cyber incident*" and defines the same as follows: *"A breach of the security rules for a system or service - most commonly; Attempts to gain unauthorised access to a system and/or to data, Unauthorised use of systems for the processing or storing of data, Changes to a systems firmware, software or hardware without the system owners consent, Malicious disruption and/or denial of service."* (NCSC, 2022a). Nevertheless, one could argue this definition is not restricted entirely to cyberspace as, for example, "*unauthorised access to a system and/or to data"* could involve analogue (e.g. confidential paper records being accessed), and how does one categorise the theft of a company laptop as a computer is, indeed, involved?

2.1.6. Cybercrime (and cybercriminals)

At this point it would be useful to understand the concept of cybercrime, recently researched by Phillips et al. (2022: 382) who report that there *"is no single clear, precise and universally accepted definition of cybercrime*[,] *a fact that is acknowledged by both academics and organizations alike"*. The authors go on to say that the two most commonly cited definitions of cybercrime are Thomas and Loader (2000: 3), "*computer-mediated activities which are either illegal or considered illicit by certain parties and which can be conducted through global electronic networks*" and Gordon and Ford (2006: 14), *"any crime that is facilitated or committed using a computer, network, or hardware device"*. Under both these definitions it would appear the case of laptop theft is out of scope for cybercrime and more an information security incident.

A natural progression from cybercrime is that of a cybercriminal and Phillips et al. (2022: 392) comment that *"Due to the breadth of behaviors that constitute 'cybercrime', there is also no obvious corresponding profile of what constitutes a 'cybercriminal'. Additionally, it is unclear whether cybercriminals ought to be conceptualized as individuals, groups, organizations/institutions, or, even, nation-states. Previous attempts to classify cybercrimes have focused on the criminal act itself; however, clarity could be*

gained by accounting for the characteristics of perpetrators (e.g., individuals, organized crime groups, and coordinated individuals) and their motivations".

Although cybercriminal profiling is not in scope for this work, the above quotation provides useful insight for practitioners in considering the origin and nature of security threats[7]. It should be borne in mind that, from the perspective of an organisation, these threats are not always entirely external and may come from e.g. a disgruntled employee or from up/down the supply chain. In fact, Verizon (2022) report that 62% of intrusion incidents in 2021 were initiated through business partners. Furthermore, as per Von Solms and Van Niekerk (2013:97), the threat actor could even be participating in the attack inadvertently.

Briefly following on from the point regarding motivation of cybercrime, a useful, concise summary is the "*Three Ps of motivation*" (Neville-Rolfe, 2020), namely Pride, Political and Profit, with Pride being the need for hackers to gain self-esteem and the respect of their peers through their achievements. Interestingly, despite hacktivism[8] being on the rise (Political) up to 3%, Verizon (2022) reports that 96% of all intrusions in 2021 were motivated by *"financial or personal gain"* (Profit). Indeed, Png, Wang and Wang (2008) remark that "*the trend is toward attacks for pecuniary gain, rather than to show off technical prowess or gain peer approval*". Although different threat actors may target different types of organisations, this statistic must send a strong message to practitioners in general as to what types of attack to expect.

Concluding on another very important statistic from the Verizon (2022) report, specifically that 82% of breaches involved a human element. As known from the definitions of cyber security highlighted above, the human factor is a very important one, stressing the need for practitioners to ensure adequate end-user security awareness training programmes are in place in their organisations. Indeed, this human factor is reflected in the NIST (2022b, emphasis added) definition of an organisation's "*security posture*", another term worthy of mention here, viz. "*the security status of an enterprise's networks, information, and systems based on information security resources (e.g., **people**, hardware, software, policies) and capabilities in place to manage the defense of the enterprise and to react as the situation changes*".

[7] Otherwise known as Cyber Threat Intelligence (CTI)
[8] Hacktivism is the act of misusing a computer system or network for a socially or politically motivated reason (Source: https://www.techtarget.com/searchsecurity/definition/hacktivism. Accessed on: 04/04/23)

2.1.7. Econometric and financial terminology

In view of the economic focus of this thesis, it was also necessary to review definitions of key econometric and financial terms to aid literature searches.

Firstly, the term 'abnormal return' (AR) is defined as the difference between the actual return and the expected return of a security (see e.g., MacKinlay, 1997). Such ARs are usually associated with a corporate 'event' such as an earnings, acquisition or divestiture announcement. For these types of 'event studies' (ESM), ARs are usually calculated on a daily basis and summed over a number of days before (possibly), during and after the event (a period known as the 'event window') resulting in a cumulative abnormal return (CAR) figure. For cross-sectional analyses involving multiple firms or events, an average CAR figure (CAAR)[9] is often quoted.

To calculate the expected return, a regression analysis is carried out on the share price over a time period usually much longer than the event window (the 'estimation window') based on a mathematical model, the most commonly used of which, in ESM studies, is the 'market model' (MM). This model is a simple single-factor, linear model based on the capital asset pricing model (CAPM) as described by i.a. Sharpe (1964). Unlike the CAPM, the MM expected return is based on a suitable 'reference market' return multiplied by the firm's individual β factor, offset by the risk free rate which is assumed to be constant (α). The CAPM and MM are described in more detail in Chapter 3. More granular expected return models have been developed, such as the Fama-French 3 Factor model (FF3FM) which are purported to predict returns more accurately than a single-factor model such as the MM (Fama & French, 1992). The FF3FM is also explained in Chapter 3.

Should the actual return be less than the expected return, the resulting AR would be negative, signifying a pessimistic market response whereas positive ARs are indicative of market optimism. One would naturally expect an 'unfavourable' information security event, such as a data breach announcement, to result in negative CAR for the firm in question and, conversely, a 'favourable' event, such as investment in security measures, to yield positive CAR as the market incorporates this information according to the efficient market hypothesis (EMH), which is elucidated in Chapter 3 (i.a. Fama, 1970).

[9] One could argue ACAR might be more appropriate, however, CAAR seems to be the generally favoured term in ESM studies.

2.1.8. Conclusion

It would be convenient here to reiterate some key points arising in this section so far before progressing on to literature searches:

1. Information security and cyber security are not one and the same – this thesis is concerned with both, and the choice of terminology reflects the context wherever possible. If it were necessary to choose one and only one here, then the closest fit would be "*information security"* as that which is pure cyber, and neither information security nor IT/ICT security (**Figure 2**), is not the primary focus of this research[10].
2. Use of the disjointed term "*cyber security*" is preferred throughout this thesis although, to be consistent with the literature, the original terminology used in direct quotations has been retained. It is necessary however, to be mindful that *"cybersecurity"* is also in frequent use and that preferences appear to be changing over time. On that basis, for the purposes of literature searches, of course, it is necessary to consider as many variations in nomenclature (in general) as is practical to ensure maximising search results.

2.2. Literature searches

Now that an understanding of terminology in the areas of information security, econometrics and finance has been gained, the next step was to carry out initial literature searches. Spanos and Angelis (2016: 219), in their systematic literature review of the impact of information security on share price began with a search string as follows: *(("Information Security" OR "Computer Security" OR "Network Security" OR "Internet Security" OR "Information System Security" OR "Web Security" OR "Software Security" OR "Application Security") AND ("Market Value" OR "Stock Value" OR "Stock Market" OR "Stock Price" OR "Market Price"))*. Although several digital sources were searched, Scopus[11] returned by far the greatest number of articles. The authors manually filtered a total of 191 studies down to a set of 27 which was subsequently expanded to 37 using the backward snowball technique (supplementing with relevant embedded references iteratively) and including only ESM related work – one could argue there would be some benefit in incorporating ESM terminology in the initial search string to avoid less relevant matches.

[10] Hence the term "information security" has been reflected in the title of this work.
[11] https://www.scopus.com. Accessed on: 04/04/23

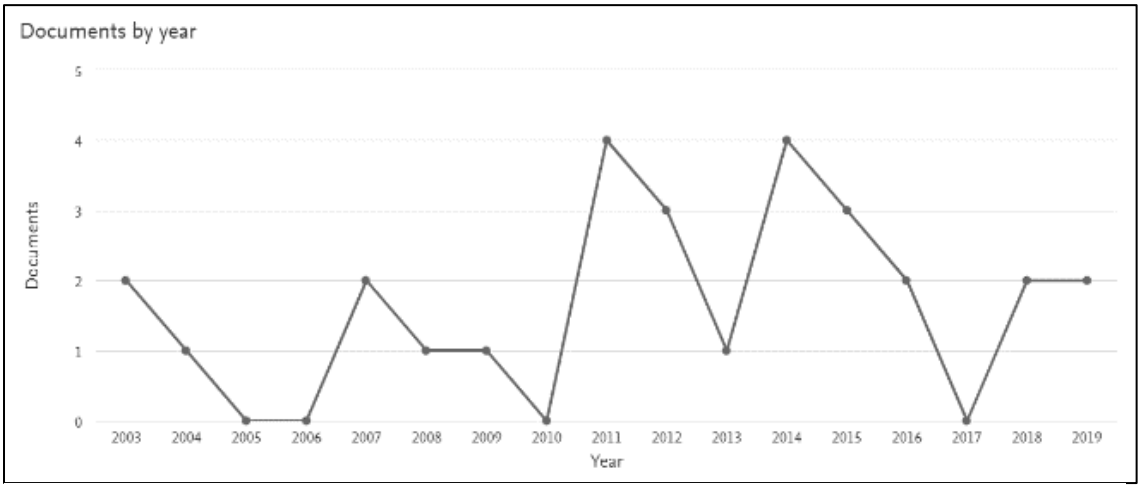


Figure 5: Initial literature search in Scopus

Adopting this approach, the first Scopus search string of *"((TITLE-ABS-KEY ("information security" OR "IT Security" OR "infosec" OR "cyber security" OR "cybersecurity") AND TITLE-ABS-KEY ("event study" OR "event studies"))) AND (LIMIT-TO (SRCTYPE , "j"))*" yielded 28 journal articles[12], a very close match with the 27 identified initially by Spanos and Angelis (2016). The number of articles returned per year is shown in **Figure 5**. It can be seen that only four such articles have been published since 2017. An analysis of the geographic breakdown also reveals a strong US bias with 19 examples (68%) and only 2 originating from the UK. A large proportion of retrieved articles here were focussing on data breaches and reporting a negative impact. Revisiting this search string in May 2020 yielded only one additional article since 2017 which did have specifically a European focus (Roškot, Wanasika & Kroupova, 2020). This study focussed on Wannacry and Petya attacks in 2017 and concluded there were, actually, positive impacts on market value. It was also noted that some embedded references which were relevant here were not included in the search results, so it was necessary to broaden the search. Therefore, the term "data breach" was incorporated into the search string as follows (and expanded from title and abstract only to all search fields): *"((TITLE-ABS-KEY ("information security" OR "it security" OR "infosec OR "cyber security" OR "cybersecurity" OR ("data" AND ("breach" OR "breaches"))) AND ALL ("event study" OR "event studies"))) AND (LIMIT-TO (SRCTYPE , "j"))"* retrieved 120 examples, 45 of which were since 2017.

[12] The search was restricted to journal articles to ensure results were peer-reviewed and is consistent with e.g. Schatz and Bashroush (2016b).

By searching within the 120 retrieved records for the keyword "GDPR" only two matching results were returned, only one of which was relevant here, Corbet and Gurdgiev (2020) who highlight "*the lack of robust regulatory mechanisms for systematic prevention, mitigation, and enforcement of data security breaches*" and advocate the use of "*white knight*" hackers to identify security weaknesses in organisations. Clearly there was a dearth of literature existing concerning GDPR in this area of study which was not entirely surprising considering GDPR was only introduced mid-2018. A sub-search for the keyword "*CISO*" also was unproductive in that only one article was retrieved (Johnson & Goetz, 2007).

Based on the search results above, supplemented by the backward snowball technique, as per Spanos and Angelis (2016), each paper was reviewed manually for suitability and relevant information summarised in the matrix below (**Table 1**).

Table 1: Literature review

Reference	Title	Summary	Data sources	Details/parameters	Comments
Deane, Goldberg, Rakes and Rees (2019)	The effect of information security certification announcements on the market value of the firm	ESM study of security certification announcements. 111 examples of ISO27001 certification announcements were identified. Stock market reaction both positive and statistically significant and dependent on contingency factors such as industry sector, size and date of certification. All US listed (50% NYSE and 50% NASDAQ) between 2005 and 2015.	PRC, Compustat, CRSP Search for announcements: BSI Group Database PR Newswire Business Wire Yahoo!Finance, PR Web Market Wired Bloomberg, Reuters	Estimation window: 255 days (-300,-46), event window (-1, 0). Market model and FF4FM – note that. Significance testing: t-test recommended: *"The t test is considered to be the best framework for analyzing statistical significance in most event study frameworks and to be relatively robust."* Use SICs for industry sector analyses.	Note the use of a buffer between the estimation and event windows. Filtered for confounding events within the actual event window only so there is some risk of confounding event overlap. FF4FM did not differ significantly from the market model. An example of a favourable information security event study.
Jeong, Lee and Lim (2019)	Information security breaches and IT security investments: Impacts on competitors	ESM used to investigate how a firm's security breaches and IT security investments influence its competitors. Gathered and reviewed 118 information security breaches and 98 IT security investment announcements from 2010 to 2017. "*Substantial*" evidence found that information security breaches have a competition effect: when one firm is breached, its competitors have opportunities to absorb market power. For IT security investment announcements, however, competitors also benefitted. Also observed that the competition effect was higher when breaches occurred after a preceding security investment than when there was no preceding investment.	LexisNexis database (general news topics and business news topics), major newspapers, wire services, and breach related databases such as PRC, DataLossDB, the Heritage Foundation, and Identity Theft Resource Center. CRSP (symbols) and Google Finance for competitors. CRSP for S&P500 data.	Estimation window 180 days with 30 day gap before event windows of (-2,2), (-1,1), (0,1), (0,2). Market model. NAICS used for industry sector analyses and t-test for hypothesis testing.	Useful table of event windows in literature review section. Eventus was the package used. Breach events could be considered favourable or unfavourable depending on whether viewed from the perspective of the breached company or the competitor.

Reference	Title	Summary	Data sources	Details/parameters	Comments
Tweneboah-Kodua, Atsu and Buchanan (2018)	Impact of cyberattacks on stock performance: a comparative study	A dataset of 96 S&P500 firms suffering cyberattacks were analysed via ESM between 2013 and 2017. Financial services sector reacts cumulatively over a three-day period. Technology firms less reactive to data breaches possibly due to improved cyber security posture.	Yahoo!Finance, BLI	Estimation window 250 days immediately prior to the event window. (–1, 1), (–2, 2), (–5, 5), (–10, 10), (–15, 15), (–20, 20), (–30, 30). Market model. For hypothesis testing uses Patell Z cross-sectional T, generalized sign Z, StdCSect Z, generalized rank Z, adjusted Patell Z, generalized rank T and skewness corrected T.	Purely US based. Warn that "*studying the cumulative effects of cyberattacks on prices of listed firms without grouping them into the various sectors may be non-informative*".
Castillo and Falzon (2018)	An analysis of the impact of Wannacry cyberattack on cyber security stock returns	Examines the impact of the WannaCry cyber-attack on stock returns of 43 companies and two ETFs operating in the cyber security industry on the first trading day after the announcement using ESM. Results clearly show that WannaCry had a positive effect on the equity returns of cyber security companies and cyber security investment vehicles. *"Both the size and significance of this finding demonstrate the impact of this worldwide cyber event on ETFs with a specific mandate to invest in the worldwide cybersecurity industry. Having a closer look at the data, it can also be noted that 80% of the companies analysed had positive excess returns on the first trading day after WannaCry. On the other hand, for companies with negative excess returns, none of these were statistically significant even at the 10% level."*	Thomson Reuters DataStream	Estimation window (-244,-6) and event window (0). Market model and Mean Adjusted Returns. Reference indices: NASDAQ for the ETFs and US listed firms. *"Non-US companies were mapped to the main index for that country."*. Uses t-test for hypothesis testing although not explicitly stated.	All US listed except seven EU examples, two Japanese and one South Korean listed. Note: On 10/3/22 the corresponding author confirmed by email that they mainly used *"Excel for significance tests verified by EViews (no changes)."*

Reference	Title	Summary	Data sources	Details/parameters	Comments
Spanos and Angelis (2016)	The impact of information security events to the stock market: A systematic literature review	SLR of the economic consequences (impact on stock price) of security incidents. *"In total, 37 related papers conducting 45 studies were found by the systematic search of bibliographic sources. The majority (75.6%) of these studies report statistical significance of the impact of security events to the stock prices of firms."*	Science Direct, Citeseer, IEEE, Web of Science, Scopus.	Search string: (("Information Security" OR "Computer Security" OR "Network Security" OR "Internet Security" OR "Information System Security" OR "Web Security" OR "Software Security" OR "Application Security") AND ("Market Value" OR "Stock Value" OR "Stock Market" OR "Stock Price" OR "Market Price")) Restricted to title, abstract and keywords only to avoid irrelevant matches.	Useful SLR focussing on ESM, however does not give detail on the parameters used, just compares e.g. number of papers using the Market Model versus Fama-French.
Schatz and Bashroush (2016a)	The impact of repeated data breach events on organisations' market value	Uses ESM to examine the influence of one or more information security breaches on a firm's stock market value. Sample size: 25 firms with 2 events each (50 events total) all S&P500 listed. Although across all 50 events there was a statistically significant negative effect (1.27%) following a breach, could only "weakly conclude" there is a difference in impact between the first and second (lower magnitude abnormal returns) events.	PRC (DatalossDB not used due to copyright issues), Thomson Reuters Datastream	Estimation window (-121,-3). For the event window (-2, 2) due to uncertainty in event date. Market model, OLS. Hypothesis testing using non-parametric tests: GSIGN, BMP.	Solely US based. Comment on lack of available and reliable data and the need to revisit this study in future once new regulations come into force to increase public disclosure of breaches.
Khansa (2015)	M&As and market value creation in the information security industry	Analysis of 787 M&As initiated by 174 public information security firms between 1998 and 2011 using a combination of ESM and regression analyses. On average, M&A events are associated with an increase in stock market value of information security acquirers. Whereas smaller information security acquirers gain more from domestic diversification, larger information security acquirers are better off seeking M&A targets	Thomson SDC Platinum, CRSP	100-day estimation period that ends 15 days prior to the announcement of each M&A. Event windows **(-1,1)**, (-2,2),(-3,3),(-5,5). Longer windows used as *"possible spillover"*. Random effects GLS model. S&P500 used as a market reference. SIC codes (mapped) for sectorial analyses. Hypothesis testing: F-statistic of Wald test. Significance levels are 2-tailed.	Uses Eventus package. Warn against generalising these results to other industry sectors.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		internationally within their line of business, especially during good economic conditions. Analysis also revealed that M&As with identity and access management (IAM) targets are perceived favourably by the stock market regardless of other M&A characteristics.			
Modi, Wiles and Mishra (2015)	Shareholder value implications of service failures in triads: The case of customer information security breaches	ESM used to examine the impact of information security breaches within service triads. A dataset of 146 customer breaches between 2005 and 2010 was used. Of these, 25 were breaches at the front-end service provider (triadic breaches) and 121 were breaches at the buyer firm (dyadic breaches). Service failures due to front-end service providers led to greater losses than such failures within the buyer firms. Also note that "*buyer firm employee productivity can moderate the greater financial penalty associated with such triadic service failures but that buyer firm leverage tends to not have such a mitigating effect."*	ITRC, Factiva, Compustat, CRSP	Estimation window: 255 trading-days ending 10 trading days prior to the event. Event windows: (-2), (-1), (0), (1), (2), (-1,0), (0,1), **(-1,1)**, (-2,2). FF4FM with value weighted index used as a market reference.	Confounding events defined as *"a quarterly earnings release, a merger/ acquisition, a change of a CEO or CFO, a debt restructuring, or an unexpected dividend change – within two trading days of the event date."* SIC codes used for sectorial analyses. Also used non-breached propensity matched firms as a *"control"* and found no significant abnormal returns using FF4FM.
Hinz, Nofer, Schiereck and Trillig (2015)	The influence of data theft on the share prices and systematic risk of consumer electronics companies	Analysis of the share price effect of cyberattacks/data theft on sample of 6 consumer electronics companies between 2007 and 2012. A decrease was observed in both victim and similar companies. Also researched the effects of such events on systematic risk. It appeared that market	Events databases: datalossdb.org and attrition.org. Price data from Thomson Reuters	Estimation window (-200, -30). Event windows: **(-10, 10),** (-3, -1) (0, +0), (0, +1), (0, +2), (0, +3), (0, +4), (0, +5), (0, +20). Market model (OLS). S&P Global 1200 index return used as the market reference.	International study. Note the use of a large buffer between estimation window and event window. Also comment on the market becoming less reactive to information security events over time hence the lack of increase in systematic risk.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		players did not change their evaluations of systematic risk, thus no increase foreseen in cost of capital.			
Hovav and Gray (2014)	The ripple effect of an information security breach event: A stakeholder analysis	Stakeholder analysis of one specific company (TJX) data breach(es) utilising ESM. The impact on stakeholders such as consumers, vendors, banks and hackers are analysed over time and found to vary as new information is released, suggesting a *"wait and see"* attitude by the market. Reports that *"while some stakeholders are losers, other are winners"* (cf. Jeong, Lee and Lim 2019).	Yahoo!Finance used to identify competitors.	Uses the Market Model. Daily abnormal returns reported.	Eventus software used. Useful to gain an understanding of the whole lifecycle of a data privacy breach and associated repeated events (cf. Schatz and Bashroush, 2016a). Also reports that previous studies on the impact of information security events have been "*inconclusive*" and carries out an interesting comparison of cyber events with physical events which are perceived to have a stronger, longer-lasting negative effects on share price.
Goel and Shawky (2014)	The impact of federal and state notification laws on security breach announcements	Uses ESM to examine the impact of federal and state breach notification laws on breached firms before and after the enactment of such laws. Concludes that the negative impact of breach announcements (201 examples between 2001 and 2008) has been reduced significantly after the enactment (0.5% versus 1% before on day of announcement).	Announcements: internet searches, PRC. Announcement dates: LexisNexis, Wall Street Journal, PC Week, Register. Price data and reference: CRSP.	Estimation window 255 days prior to the event. Event window: observed over (-30, 30) but cited AR on event date. FF4FM. CRSP value-weighted index. Hypothesis testing: Patel Z value.	US based study using Eventus package. Confounding events avoided (assumed) by use of short event windows. Also make a comparison of cyber versus physical events (i.a. industrial accidents) and comment these are more consistently negative (cf. Hovav and Gray, 2014). Comment that overapplication of breach notification laws could lead to market desensitisation.
Bose and Leung (2014)	Do phishing alerts impact global	ESM study into the impact of phishing announcements (1942 phishing alerts related to 259 firms in 32 countries between 2003	Data sources: Millersmiles, Websense, and Factiva Stock prices.	Estimation window: 200 trading days ending one month prior to the event. Event windows: (−1), (0), (1), (−1,0), (0,1), **(−1,1)**. FF4FM	International study remarking on the paucity of similar research involving international firms.

Reference	Title	Summary	Data sources	Details/parameters	Comments
	corporations? A firm value analysis	and 2007) on the market value of global firms. Found to be strongly significant for alerts released in 2006-2007 and for financial holding companies. US firms only weakly significant. Loss in market capitalisation estimated to be at least US$411m (based on CAAR of 0.06%; median CAR was more negative).	Trading volume data from Thomson Reuters.	(merged for international) utilised with quantile regression rather than OLS (outliers, non-normality of error terms). FF4FM is compared with CAPM and preferred. Multiple market indices were used due to the international nature of this study (including specialised indices such as S&P banking index) – indices resulting in best adjusted R^2 for the regression model were preferred. Hypothesis testing: Z test, Sign test, Corrado's rank test.	
Pirounias. Mermigas and Patsakis (2014)	The relation between information security events and firm market value, empirical evidence on recent disclosures: An extension of the GLZ study	Impact of security incidents (105 examples) on firm value between 2008 and 2012 employing ESM. Statistically significant negative returns (-0.39%) found with technology firms suffering the most.	Breach announcements: DatalossDB supplemented with ITRC and PRC.	Estimation window (-201, -2). Event windows: (-1, 1), (-1,0), (0,0), (0,1) FF3FM and market model (OLS). Hypothesis testing: t statistic.	Solely US firms. Confounding events only within event window (-1, 1). Useful table of estimation windows and market indices in literature review. Comments on the variability of findings of previous studies in this area and conclude that *"the markets seem to have matured in the way they handle security events".*
Bose and Leung (2013)	The impact of adoption of identity theft countermeasures on firm value	ESM study into the adoption of identity theft countermeasures. 87 announcements (1996-2012) related to US listed companies. *"We show that the news of such adoption increases the short term market value of the announcing firm by 0.63% on an average. Our research also finds that early adopters, adopters of sophisticated identity theft countermeasures, firms with high growth potential, and firms with high credit rating show a strong and positive return in market*	Source of announcements: Factiva, PR Newswire and Business Newswire.	Estimation window (-230, -31). Event window (0,1). Confounding event window (-2, 2). Market model (OLS). Reference market: S&P500 or NASDAQ composite index (depending on highest R^2). Hypothesis testing: Z test and Corrado's rank test.	Favourable information security event study (US based). No industry sector analysis, rather use firm size (market capitalisation) and credit rating.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		value, whereas small firms demonstrate a moderate but positive reaction."			
Khansa, Cook, James and Bruyaka (2012)	Impact of HIPAA provisions on the stock market value of healthcare institutions, and information security and other information technology firms	Example of an ESM study on the impact of legislation, specifically the Health Insurance Portability and Accountability Act (HIPAA) of 1996. This regulation places the onus of storing and transferring healthcare data securely on healthcare providers thereby setting market expectations of a financial burden on the healthcare sector and, conversely, an uplift in demand on IT/IS providers offering products or services related to the HIPAA. The findings are consistent with these expectations with healthcare firms losing up to 2% of market value, whereas IT and IS firms gain 1.5 and 2% respectively. Sample size: 107 healthcare institutions, 735 IT firms and 63 IS firms.	IT and IS firms identified with Yahoo!Finance. Price data from CRSP.	Estimation window of 100 days ending 30 days before the announcement. Event windows: (-1,10), (-2,10), (-3,10), (-5,10), (-7,10), (-10,10) Market model (OLS) using the CRSP value-weighted index as a reference. Hypothesis testing method not specified.	Introduction of HIPAA legislation perceived as positive (favourable) information security event for the IT/IS sectors and negative (unfavourable) for the healthcare sector.
Chen, Li, Yen and Bata (2012)	Did IT consulting firms gain when their clients were breached?	An ESM study on the impact of data breaches on the share price of consulting firms (*"83 breach events affecting a wide range of US firms in various industries in year 2006 and 2007"*). Found that the market value of consulting firms is positively associated with breach announcements (+4.01% during the two days during and after the announcement). Breaches involving a larger number of records, however, resulted in negative returns for consulting firms	Source of breaches: DatalossDB. Price data from CRSP, Compustat.	Estimation window 120 days ending the day before the announcement. Event windows: (0), (1), **(0,1)**. Hypothesis testing: t-value.	US study. The cross-sectional analyses use a different multi-factor regression model involving other parameters such as the number of records breached and article size (Ishiguro et al., 2006).

Reference	Title	Summary	Data sources	Details/parameters	Comments
		particularly for technology intensive firms. "*In other words, generally speaking, the IT consulting firms have similar experiences with the attacked firms."*			
Zafar, Ko and Osei-Bryson (2012)	Financial impact of information security breaches on breached firms and their non-breached competitors	Rather than ESM, investigates the information transfer effect of firms announcing information security breaches (specifically Denial of Service, Website Defacement, Data Theft, and Data Corruption) between 1997 and 2007 by comparing performance with non-breached competitor firms (matched sampling method). Found significant information transfer effects for certain types of breaches and evidence of contagion effects. No similar evidence of a competition effect identified.	Breach announcements: LexisNexis. Financial information: Compustat, EDGAR.	Matching based on 70-130% total assets. Financial KPIs used: ROA, ROS, COGS/S, SGA/S. Hypothesis testing: Wilcoxon matched-paired (Z) test.	Alternative approach to ESM.
Tejay and Shoraka (2011)	Reducing cyber harassment through de jure standards: A study on the lack of the information security management standard adoption in the USA	ESM study on the adoption of ISO Information Security Management System (ISMS) certification announcements on the market value of firms (32 certification examples between 2005 and 2010). No significant economic impact was found.	Event databases: LexisNexis, ProQuest.	Estimation window: 120 days. Event window: (-1), (0), (1), **(-1, 1)** Market model used. Hypothesis testing: t-test.	Favourable event study Also contains useful summary of information security economics.
Yayla and Hu (2011)	The impact of information security events on the stock value of firms: The effect of contingency factors	ESM study of stock market reactions to firm-specific security breaches (123 examples between 1994 and 2006). Reports that breaches have a negative effect with the impact varying with contingency factors such as business type, industry, type of breach,	Google, Yahoo!Finance, LexisNexis for announcements.	Estimation window (-130, 10). Event windows (-1, 1), (-1, 5), (-1, 10). Market model (OLS). Equal-weighted NYSE/AMEX/Nasdaq index used as a market reference. Hypothesis testing: t-value.	US listed firms only. Short summary table in literature review showing ESM parameters of previous studies.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		event year and length of event window. Higher for pure e-commerce firms than traditional bricks-and-mortar. DoS attacks generate greater losses. Also note less negative impact in recent years.		Confounding events removed within (-1, 10).	
Morse, Raval and Wingender (2011)	Market Price Effects of Data Security Breaches	ESM study of data breaches. 306 events between 2000 and 2010 for US listed companies only. Finds evidence of small negative CAARs (-0.3% over two days) but becoming greater over years. Suggests the markets are unsympathetic towards breaches which are clearly avoidable.	DatalossDB for breaches.	Estimation window (-505, -251). Event windows (0), (0, 1), (1, 5), (1, 10). Also (1, 220), (1, 240) and (1, 440), (1, 480). Market model GARCH (1,1) adjusted. Market reference: value-weighted CRSP index. Hypothesis testing: Z test.	Includes an interesting quotation from Warren Buffet: *"Predicting storms doesn't count; building arks does.*". Looks at medium and long-term effects as well (although using CAR rather than BHAR). Also advocates the appointment of CISO (and CRO) on company boards to mitigate risk.
Goldstein, Chernobai and Benaroch (2011)	An event study analysis of the economic impact of IT operational risk and its subcategories	ESM study of 'data' and 'function' related IT operational failures. US financial services firms only, 142 events from 1985 to 2009. Function events yield higher negative returns than data. Firm size and growth potential also affect returns.	Events: Financial Institutions Risk Scenario Trends (FIRST), Factiva, LexisNexis. Price data: CRSP, ComputStat.	Estimation window: (-301, -46). Daily ARs quoted but event window (-1, 2) preferred as most negative CAR. Market model (OLS). Market reference: equal-weighted CRSP index. Hypothesis testing: Patell's one-tailed and standardised Z-statistic.	Split events into Data and Function categories. Uses SIC codes. US financial firms only.
Goel and Shawky (2009)	Estimating the market impact of security breach announcements on firm values	Impact of security breach announcements (reports and news articles) on US firm market value using event study techniques. Data from 2004-2008. Found an impact of around 1% of market value in the days surrounding the event. 205 examples reduced to 168 for which exact dates were known.	Events i.a. LexisNexis and CRSP for price information.	Estimation window 255 days prior to the event period. Event window: (-119,10) with daily ARs quoted. Also graph (-5,5) showing AR & CAR. Uses FF3FM. CRSP market value weighted index.	Negative AR 4 days prior to event which indicates information leakage. Used Eventus package.

Reference	Title	Summary	Data sources	Details/parameters	Comments
Png, Wang and Wang (2008)	The deterrent and displacement effects of information security enforcement: International evidence	ESM study of government enforcement actions regarding attacks on 15 countries between 2004 and 2006. Limited evidence that domestic enforcement deters attacks within that country, but *"compelling"* evidence of a displacement effect that US enforcement increases attacks originating from other countries. Also, a correlation was observed between number of attacks and US unemployment rate.	Events: DShield, Factiva, internet searches. Vulnerabilities: NVD.	Event windows: (-7, 7), (0, 7), (-14, 14), (0, 14). Multifactor regression model (OLS). Hypothesis testing: Wald test.	Notes that *"the motivation of attackers has shifted toward making money"*. Observes that *"the trend is toward attacks for pecuniary gain, rather than to show off technical prowess or gain peer approval"* yet acknowledges that this has not yet been empirically verified.
Kannan, Rees and Sridhar (2007)	Market reactions to information security breach announcements: An empirical analysis	ESM study of security breaches of US listed firms. 102 events (60 companies) between 1997 and 2003. No significant negative returns observed although abnormal returns amplified after the 9/11 attacks. Market also more sensitive to breaches during the dot-com era than before.	Events: New York Times, Wall Street Journal, ZDNet, CNET.	Events windows (-1,2), (-1, 7), (-1, 29). Market reference S&P500 – control firms used as well and qualitatively no difference. Matched firms identified through Hoover's Company Profiles Database.	Uses SIC codes. This study highlights the importance of **long-term confounding events, specifically 9/11 and the dot-com era**. Recommends regulation for transparency in future.
Telang and Wattal (2007)	An empirical analysis of the impact of software vulnerability announcements on firm stock price	ESM study into impact of vulnerability announcements by software vendors *("147 vulnerability announcements pertaining to 18 firms between January 1999 and May 2004")*. Found that a vendor loses ca. 0.6% market value when a vulnerability is reported. More loss of market share if the market is competitive, the vendor is small, the vulnerability is more severe or no patch is provided at announcement time.	Announcements from CERT, BusinessWire, NewsWire, ProQuest, LexisNexis.	Estimation window: (-175, -16). Event window: (0). Uses the Market Model (OLS), Market-Adjusted Model and Mean-Adjusted Model. Hypothesis testing: Sign test and Wilcoxon signed rank test.	Comment on the importance of event studies to corporate policy decision making. *"If the markets are efficient and rational, then event studies should correctly measure the long-term economic impact of an event"* or, to put it another way, *"in the absence of the event, the stock price of the firm at any time would have been higher"*. (cf. MacKinlay, 1997)

Reference	Title	Summary	Data sources	Details/parameters	Comments
Cavusoglu, Mishra, and Raghunathan (2004)	The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers	ESM study to assess the impact of security breaches on the market value of breached firms (66 events between 1996 and 2001). Breached firms lost on average, 2.1 percent within two days of the announcement. Cross-sectional analyses of firm type, firm size, and the year the breach occurred. The information-transfer effect of security breaches (i.e. their effect on the market value of firms that develop security technology) was also studied and gains of 1.36% on average reported.	Lexis/Nexis, and the technology portals CNET and ZDNET. CRSP. NASDAQ, Yahoo!Finance archives. Market value data from Compustat.	Estimation window: 160 days to prior day. Event windows: (0), (1), (0,1). Market model (OLS). Multiple linear regression model used. *"The model regressed the cumulative abnormal returns on hypothesized variables and control variables, namely, firm type, firm size, nature of attack, and time."*.	Frequently cited article. Report both favourable (information transfer to security developers) and unfavourable (breached firm) effects of announcements.
Garg, Curtis and Halper (2003)	The financial impact of it security breaches: What do investors think?	ESM study of cyber-breaches (22 breach announcements between 1996 and 2002). Negative CARs observed of 2.7% on the event day rising to 4.5% two days after (5% and 10% significance levels respectively). It was also noted that internet security vendors reacted positively (4-10%) to announcements prior to 2000 whereas no significant reaction after. For insurance carriers, negative prior (2%) and positive after (0.7-1.7%) *"perhaps reacting favorably in anticipation of increased cyber-insurance sales and the higher premiums as a result of heightened awareness of cyber-insurance"*.	Breach announcements: Bloomberg, Dow Jones Interactive.	Event windows (0), (0, 1), (0, 2). Significance testing: Wilcoxon signed rank test.	Virus attacks excluded because perceived as more market-wide impact. Confounding events: earnings announcements, analyst upgrades and executive resignations. Also consider loss in market capitalisation figures.
Richardson, Smith and Watson (2019)	Much Ado about Nothing: The (Lack of)	Researches the impact of data breaches on four aspects of firms (827 breach disclosures for 417 companies between 2005 and 2018):	PRC. Compustat and CRSP for market	Estimation window (-120, -5). Event windows: (-120, 5), (-1, 3), (-1, 21), (-1, 63), (-1, 126).	Use the Stata 'eventstudy2' routine. Also comment that stolen laptop is not classed as a cyber security incident.

Reference	Title	Summary	Data sources	Details/parameters	Comments
	Economic Impact of Data Privacy Breaches	share price (ESM study), accounting measures, audit and other fees and SOX 404 reporting. ESM showed only -0.3% loss in market value on average, with the exception of a few catastrophic examples. Other effects measured by comparison with propensity matched firms and no differences found. Difficult to justify investment in security based on these findings.	information. Audit Analytics.	FF4FM although comment that the Market Model was similar. Market reference: propensity matched firms. Hypothesis testing: Patell, Boehmer, Corrado, generalised sign test.	*"Companies are unlikely to change their investment patterns unless the cost of breaches increases dramatically or regulatory bodies enforce change."* Frequently cited paper.
Bendovschi, Al-Nemrat and Ionescu (2016)	Statistical Investigation into the Relationship between Cyber-Attacks and the Type of Business Sectors	An initial investigation into the correlation between attack types and industry sectors (4,785 attacks worldwide). Statistically significant correlations found for some industry sectors.	VCDB	Logistic regression used (SAS software). Attack types: Pattern, Action, Actor, Root Cause, Discovery Method.	Uses NAICS. Example findings: payment card skimmer attack most likely in the food industry. Random error most common in the retail industry. For the financial and insurance sector payment card skimmer attacks discovered internally most frequent. As this study uses VCDB, it is not limited to publicly listed companies.
Telang and Wattal (2007)	An Empirical Analysis of the Impact of Software Vulnerability Announcements on Firm Stock Price	An ESM analysis of the impact of vulnerability announcements (147 pertaining to 18 firms January 1999 to May 2004) on the market value of software vendors. Results showed significant impact of -0.63% on stock price when a vulnerability is reported.	Announcements: leading national newspapers, CERT. Compustat for company information.	Estimation window: 160 days (-175, -16). Event windows -1,0, (0,1), (0,2), (0,5), (0,10). Market model, Market Adjusted Model, Mean Adjusted Model.	Also consider loss in market capitalisation figures. Earliest announcement is chosen as event day.
Andoh-Baidoo, Amoako-Gyampah and Osei-Bryson (2010)	How Internet Security Breaches Harm Market Value	ESM study of 41 events/firms between 1997 and 2003. Found a loss of 3.18% on average over a three-day event window. Net firms will more likely suffer damage than non-Net firms. Remark that *"one important insight*	Announcements: LexisNexis. Price data: CRSP.	Estimation window: 120 days up to -2. Event window: (-1, 1). Market model. Indices: NYSE, NASDAQ, AMEX (NASDAQ mainly as mostly tech stocks).	Use both firm and attack characteristics as described by Howard (1997). Also use decision tree induction. Only search for confounding events during the event window.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		our study provides is results refuting an earlier study's claim that only security breach announcements that involve loss of confidential data lead to negative abnormal returns". Results also more negative after February 2000. External attacks more likely than internal.			
Lin, Sapp, Ulmer and Parsa (2020)	Insider trading ahead of cyber breach announcements	An ESM study finding significant evidence of opportunistic insider trading ahead of cyber breach announcements (258 examples). Insiders save an average of $35,009 due to timely selling in the three months before the disclosure. Late filing violations by insiders more likely to occur near the announcement of a cyber breach. Opportunistic (non-routine) trading tends to occur 55–72 days before the announcement. CAAR of -1.44% over a 5- day window (-2, 2).	Compustat, CRSP, PRC for breaches, Thomson Insiders.	Estimation window not specified. Event windows: (-1, 1), (-2, 2), (-2, 18), (-10, 30). Market model verified with FF3FM & FF4FM. Hypothesis testing: t-statistic.	Use 2-digit NAICS for industry. US firms only. Comment that these *"results lend support to the US Security and Exchange Commission's recently announced goal of tightening restrictions on insider trading ahead of cyber breach announcements"*. Also remarks that firms acting on such information asymmetry builds distrust between firm management and the markets.
Rosati, Deeney, Cummins, Van der Werff and Lynn (2019)	Social media and stock price reaction to data breach announcements: Evidence from US listed companies	ESM study on the impact of use of social media (Twitter) in the context of data breach announcements by US firms (87 events from 73 firms 2011-2014). Find that use of social media at the time of a data breach exacerbates negative impact of the breach, whereas as for firms having lower visibility the effect is positive.	PRC, Lexis-Nexis, Social media (Twitter). Thomson Reuters for share prices. CAR source: Datastream Professional	Estimation window: (-125, -6) Event windows: (0,1), (0,2), (0,3), (4,10). Market Model used and compared with FF3FM and found to be consistent.	US only (PRC). Advocates for a contingency model for social media communication in the event of breaches dependent on firm size, visibility and type of breach.
Ettredge, Guo and Li (2018)	Trade secrets and cyber security breaches	Study of the association between firms' disclosures in Forms 10-K of the existence of	ITRC/CyberScout Annual Data Breach Reports,		Use SIC for industry analyses.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		trade secrets, and cyber theft of corporate data (breaches). Firms mentioning the existence of trade secrets have a significantly higher subsequent probability of being breached relative to firms that do not. Results are stronger among younger firms, firms with fewer employees, and firms operating in less competitive industries.	LexisNexis, Compustat, SeekEDGAR.		Not an ESM study, purely looking at probability of a firm being breached if they disclose trade secrets in 10-K by comparing with propensity-matched non-breached firms. Example of potential unfavourable impact of disclosure legislation.
Chen and Jai (2019)	Cyber alarm: Determining the impacts of hotel's data breach messages	Study on the impact of data breach announcements on hospitality companies. Insights are given regarding guests' reactions to cyber-crisis communications through the application of Situational Crisis Communication Theory if the guest received the news through media or directly from the breached hotel and whether the guest was a victim or not.	News media, hotel websites, and trade magazines.	Survey: 255 respondents.	Concludes that method of communication of the breach to guests is important – direct to consumer messaging more effective in rebuilding trust.
Nieuwesteeg and Faure (2018)	An analysis of the effectiveness of the EU data breach notification obligation	Qualitative study of the EU data breach notification obligation (EU DBNO), which is part of the GDPR (Articles 33 and 34). Origins, aims and social benefits are discussed, also the role of national DPAs in inducing data controllers to comply with the regulation.	GDPR, literature review.		Comments that spontaneous privacy breach disclosure by data controllers unlikely without the DBNO. Also highlights the importance of the actions taken by DPAs in making DBNO a success.
Syed (2019)	Enterprise reputation threats on social media: A case of data breach framing	Analysis of social media postings in relation to the 2014 Home Depot data breach. Reputation threats found to vary across crisis stages. Negative emotions such as anger and	Twitter (twitteR package).	Situational crisis communication theory, Crisis Stage Theory.	Provides insights to guide corporate communications strategy post data breach to potentially mitigate reputational damage.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		disgust increase subsequent reputation threats.			
Confente, Siciliano, Gaudenzi and Eickhoff (2019)	Effects of data breaches from user-generated content: A corporate reputation analysis	Investigation into the effects of data breaches (for 35 US firms in 9 industry sectors between 2013 and 2016) on corporate reputation dimensions through user-generated social media content. More dimensions affected for critical incidents. After "*intentional and internal*" type breaches, the "*firm as an employer*" dimension affected negatively due to perceptions of lack of training to avoid such incidents.	PRC for data breaches. SDL SM2 for social media data. Nvivo for analysis.		
Daly (2018)	The introduction of data breach notification legislation in Australia: A comparative view	Review of the impact of the introduction of data breach notification legislation in Australia, the Privacy Amendment (Notifiable Data Breaches) Act 2017. This law is compared with similar legislation in the EU and US and concludes that, although making some inroads into improvement in data security, the need for some reform was identified both in the law and its application as well as lack of consistency in standards to ensure strong cyber security.	(Literature review)		Qualitative study. Useful summary of US/EU regulations included.
Malliouris and Simpson (2020)	Underlying and Consequential Costs of Cyber Security Breaches: Changes in Systematic Risk	Measures changes in upside/downside systematic risk (betas) of organisations (202 events for US firms between 2005 and 2019) suffering security breaches. Found that severe security breaches are associated with	PRC	Risk calculations based on a modified version of the CAPM. Used "*major multi-industry stock market index*" as a reference. (S&P500 cited as an example.)	Excluded events which were supply chain and /or 'unobvious' (non-direct) subsidiaries. Also excluded events not garnering major media attention.

Reference	Title	Summary	Data sources	Details/parameters	Comments
		significantly positive increases in systematic risk and systematic downside risk. Lack of upside risk change indicates the asymmetric nature of the risk impact. Net increase in cost of equity as a result of breaches.		Hypothesis testing: t-test, Wilcoxon signed-rank test.	Sectorial analyses based on S&P Capital IQ industry sector.
Malliouris and Simpson (2019)	The stock market impact of information security investments: The case of security standards	ESM study of the impact of Cyber Essentials (Plus) and ISO/IEC 27001 (re)certifications on the share price of international firms. 145 Cyber Essentials events and 76 ISO/IEC 27001 examples between 2001 and 2018. Found that "*the award of a Cyber Essentials (Plus) certificate is systematically associated with significant and positive market reactions. Surprisingly, our international sample reveals that becoming ISO/IEC 27001-compliant elicits significant negative abnormal stock returns*."	CREST, GCHQ for Cyber Essentials (re)certifications. JAS-ANZ for ISO/IEC 27001 certifications. S&P Capital IQ for financial data.	Estimation window: 252 trading days. Event windows: (-3, 0), (-3, 2), (-2, 0), **(-2, 2).** Market Model (OLS). FTSE350 cited as an example market reference for UK. Hypothesis testing: t-test , Wilcoxon signed-rank test (winsorised).	International study. Example of where an (expected) favourable information security event actually has negative impact.

2.3. Analysis and discussion

Using the data from **Table 1** to answer RQ1 (What is the impact (if any) on share price of a security event, be it favourable or unfavourable and how do these findings compare with the literature?) a good starting point was Spanos and Angelis (2016) which, being an SLR itself, confirmed that research in this area was "*quite limited*" and that "*the majority (75.6%) of these studies report statistical significance of the impact of security events to the stock prices of firms*". Clearly ESM is the preferred approach to such studies (hence it was the focus of their paper) and, although other approaches to measuring economic impact were found to be much rarer, they did, nevertheless, exist as identified through the backward snowball technique e.g. Zafar, Ko and Osei-Bryson (2012) who compared financial KPIs (ROA, ROS, COGS/S, SGA/S) of breached firms with non-breached competitors.

These ESM studies of information security events, however, do seem to vary somewhat regarding the magnitude of the abnormal returns. Earlier papers such as Garg, Curtis and Halper (2003) cite losses of 2.7% on the day of a breach announcement. A frequently cited (and early) source, that of Cavusoglu, Mishra and Raghunathan (2004), report that breached firms lost, on average, 2.1% of market value within two days of the breach announcement. Goel and Shawky (2009) observed an impact of only 1% in the days surrounding a breach event perhaps indicating a trend of reduced market reactions over time, a sentiment echoed even later by Pirounias, Mermigas and Patsakis (2014) who report only -0.39% on average and conclude that "*the markets seem to have matured in the way they handle security events*". This reduced impact over time was again noted by Richardson, Smith and Watson (2019) in another frequently cited paper, noting only a 0.3% reduction in market value and downplay market reaction to data breaches as "*much ado about nothing*". Yayla and Hu (2011) conclude their study by noting that events which *"occurred in recent years were found to have less significant impact than those occurred earlier, suggesting that investors may have become less sensitive to the security events*". That said, Schatz and Bashroush (2016a) quote -1.27%, yet Morse, Raval and Wingender (2011) in a much earlier study report -0.3% on average, just as Richardson et al. (2019). Kannan, Rees and Sridhar (2007) who observe that "*an event study of the effect of such breaches on financial performance found that they do not earn significantly negative abnormal returns*" attribute differences between their findings and that of e.g. Cavusoglu et al. (2004) to their choice of control firms as a reference rather than a composite (one size fits all) index such as the S&P500. Evidently, one has to pay careful

attention to the way in which ESM is applied. Indeed, Hovav and Gray (2014) describe previous studies as "*inconclusive*" and remark that, in contrast with cyber incidents, physical events have a much stronger and longer-lasting impact.

Clearly, there are exceptions to this trend of reduced impact of data breaches over time which requires further investigation. Kannan et al. (2007), somewhat unusually for the time, noted no significant negative impact of security breaches and highlight the importance of long-term confounding events such as the 9-11 attacks and the dot-com era. Clearly macro-economic factors need to be considered as well and any such ESM study here should avoid any overlap with the COVID-19 pandemic. It should also be well noted that by far the majority of these studies focus on US markets. Notable exceptions would be the Bose and Leung (2014) paper[13] on phishing alerts related to 259 firms across 32 countries observing that "*each phishing alert leads to a statistically significant loss of market capitalization that is at least US$ 411 million for a firm*" and Hinz et al. (2015) who investigated cyberattacks and data theft in a small sample of consumer electronics companies worldwide. Although a negative impact was observed in victim companies, an information transfer effect on similar companies was also reported. Castillo and Falzon (2018) analysed the impact of WannaCry on cyber security suppliers: "*event-study methodology is employed for the analysis of this specific event and results clearly indicate that WannaCry had a positive effect on the equity returns of cybersecurity companies and cybersecurity investment vehicles*" – there is a need here to be cognisant of competition effects as well as information transfer. Indeed, Jeong, Lee and Lim (2019) researched the impact of data breaches on competing firms and found "*substantial evidence*" that breaches have a positive impact on competition. The information transfer effect was well noted by Chen et al. (2012) who found that IT consulting firms gained by over 4% on average in the two days following a client breach announcement.

Hinz et al. (2015) also remark on the tendency for the market to become less reactive to data breaches over time, so this effect does not seem to be restricted to the US market. This, and industry specific effects begin to answer RQ2 (Are there any patterns in the data, such as correlations between drop in market value and category of cyber-attack, data breach, industry sector etc.?)

Regarding industry sector, Tweneboah-Kodua, Atsu and Buchanan (2018) analysed breach events for 96 (again) US firms, and although not reporting significant abnormal

[13] The authors also remark on the paucity of international studies in this area.

returns for shorter event windows, do warn that "*studying the cumulative effects of cyberattacks on prices of listed firms without grouping them into the various sectors may be non-informative*". Financial services sector firms, for example, showed more significant abnormal returns over a 3-day event window than those in the technology sector. Yayla and Hu (2011) also provide input to RQ2 from their study stating that "*pure e-commerce firms experienced higher negative market reactions than traditional bricks-and-mortar firms in the event of security breach*". Evidently, industry sector is also a contingency factor in these type of studies with many performing sectorial analyses based on varying standards such as NAICS (Bendovschi, Al-Nemrat & Ionescu, 2016; Jeong et al., 2019; Lin et al., 2020) and SIC codes (Kannan et al., 2007; Goldstein, Chernobai & Benaroch, 2011; Khansa, 2015; Modi, Wiles & Mishra, 2015; Ettredge, Guo & Li, 2018; Deane et al., 2019) whereas Malliouris and Simpson (2020) used that provided by S&P Capital IQ. A different approach was used by Bose and Leung (2013, 2014) who used firm size (market capitalisation) and credit rating for their cross-sectional analyses.

It is acknowledged, however, by Richardson et al. (2019) that exceptional events do occur: "*there is little impact from a data breach except in those rare situations involving massive data exposures*" suggesting a need to categorise data breaches according to severity (RQ2). This was the basis of the study by Bendovschi et al. (2016) who examined the relationship between cyberattack type and industry sector and conclude that "*there is a relation between attacks and some of the business sectors*", a finding underpinned by the work of Yayla and Hu (2011). Their research also found that "*denial of service attacks had higher negative impact than other types of security breaches*". For example. Richardson et al. (2019) use the Breach Level Index (BLI) approach (Stiennon, 2013) to calculate the potential severity of a data breach (**Figure 6**), which considers the number of records breached, the type of data compromised with sensitive personal information ranking higher, the source of the breach and subsequent actions taken. This approach is consistent with Campbell et al. (2003) whose observation that breaches involving unauthorised access to confidential data (data privacy breaches) were more likely to result in significant negative market reaction.

Another factor affecting market perception is the way the breach is communicated. Indeed, Rosati et al. (2019) found that use of social media in reaction to a data breach could exacerbate any negative effect and advocate for a contingency model dependent on firm size and visibility as does Syed (2019) based on research into the Home Depot breach of 2014. Additionally, Amir, Levi and Livne (2018) noted that abnormal returns were

more negative for firms withholding disclosure on breaches themselves and the announcement subsequently being made by a third party instead, or victim firms portraying the breach as more minor than it was. Similar work by Confente et al. (2019) analysed user-generated social media in response to breaches and confirmed that "*intentional and internal*" type events resulted in reputational[14] damages to the victim firm due to perceptions of lack of training to avoid such incidents. Chen and Jai (2019) recommend direct to consumer communication in the event of a data breach as this was found to be more effective in rebuilding trust within the hospitality industry. Malliouris and Simpson (2020) in their ESM study (unusually risk based) actually excluded events from their study which did not garner major media attention.

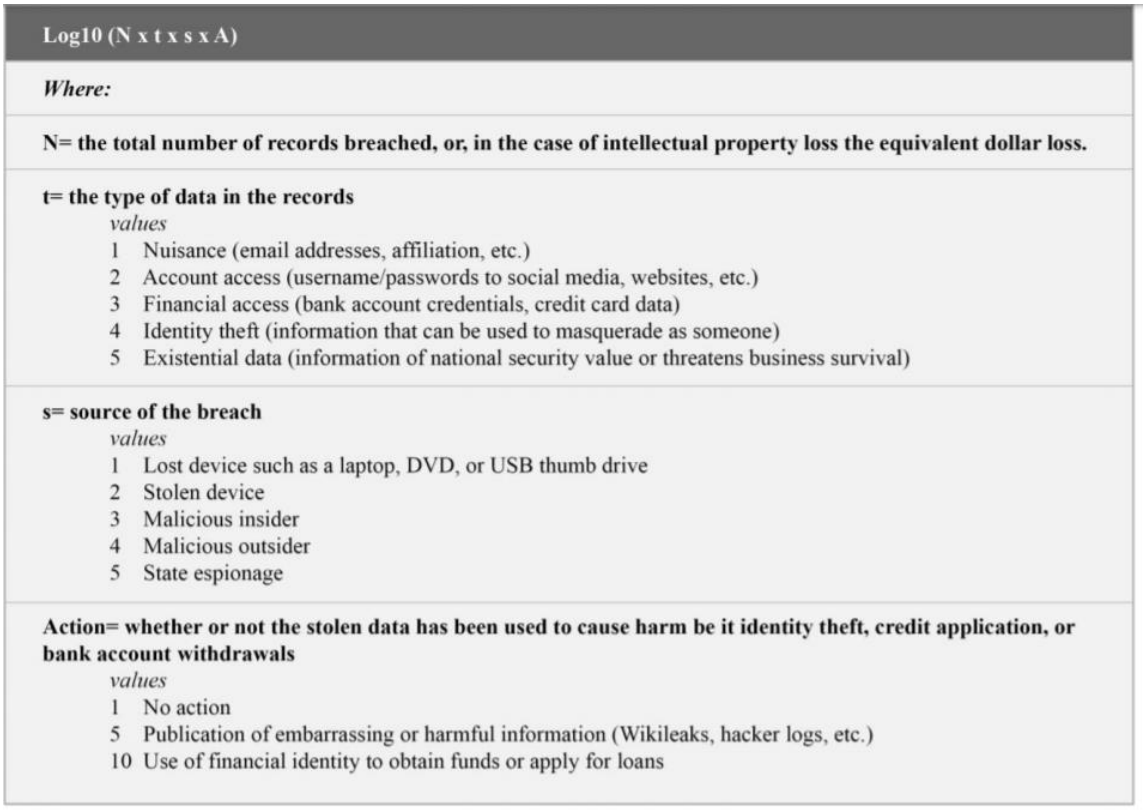
Regarding RQ4 (How can these findings be incorporated into security investment strategies of organisations?), studies such as that of Cavusoglu et al. (2004) and Bose and Leung (2014) reporting high magnitude negative CARs directly as a result of breaches would obviously assist business case justifications for improving an organisation's cyber security posture. The same could be said for studies investigating favourable information security events and finding corresponding positive CARs as a result, such as Deane et al. (2019), Jeong et al. (2019) and Bose and Leung (2013). Although Malliouris and Simpson (2019) found UK companies benefitted from Cyber Essentials (Plus) certification, there was actually a downside found for ISO/IEC27001 certifications worldwide showing that a one-size fits all approach to investment strategies does not work. There are, yet again, other contingency factors to consider as noted by i.a. Bendovschi et al. (2016): "*this study may be the basis of an in-depth analysis with the purpose of providing insights and open the way towards a systematic channelling of the limited security budget towards the right internal controls.*" These challenges are echoed by Schatz and Bashroush (2016b) in their SLR concerning the economic valuation of information security investment and report that "*it remains difficult for practitioners to identify key approaches in current research*". The findings of Richardson et al. (2019) are somewhat disheartening in supporting cyber security investment as reflected in the title of their data breach article "*much ado about nothing*" as are other more recent studies as indicated above. That said, Richardson et al. (2019) argue that "*companies are unlikely to change their investment patterns unless the cost of breaches increases dramatically or regulatory bodies enforce change*" – starting to move towards the introduction of legislation and RQ3 (Regarding the introduction of

[14] Corporate reputation may be defined as "a collective assessment of a company's attractiveness to a specific group of stakeholders relative to a reference group of companies with which the company competes for resources" (Fombrun, 2012).

the GDPR, what is the economic impact of infringement fines on the market value of firms, including those appealed and overturned?).

As the introduction of the GDPR is so recent (2018), literature in this area is rare, however, Goel and Shawky (2014) carried out a similar US based ESM study: "*our results show that the negative impacts of security breach announcements on stock prices have been reduced significantly after the enactment of federal and state security breach notification laws*." This sets the expectation that the introduction of legislation is a favourable (macroeconomic) information security event. Nieuwesteeg and Faure (2018) did carry out a qualitative study of the introduction of the GDPR, however, and opine that the success of this regulation lies very much in the hands of the national DPAs who need to apply it in the most beneficial way. Ettredge et al. (2018) examined the probability of firms mentioning trade secrets in their Form 10-K disclosures being breached and found a higher breach probability for those who did disclose trade secrets over those who did not. Although dependent on contingency factors such as industry sector and firm size this finding was, nevertheless, an example of an unfavourable event due to introduction of legislation. The GDPR was compared with recently introduced legislation in Australia (Daly, 2018) along with equivalent US regulations and areas for potential improvement identified. A study by Khansa et al. (2012) focussed on the healthcare sector and the introduction of the Health Insurance Portability and Accountability Act (HIPAA). The authors found healthcare firms lost 2% of market value as a result of the new regulation placing the onus on them to invest in data storage and transfer whereas service providers showed gains of up to 2%. A tightening of existing US SEC legislation regarding insider

trading is recommended by Lin et al. (2020) who found significant evidence of opportunistic (non-routine) trading in advance of a cyber breach announcement.



Log10 (N x t x s x A)

Where:

N= the total number of records breached, or, in the case of intellectual property loss the equivalent dollar loss.

t= the type of data in the records

values

1 Nuisance (email addresses, affiliation, etc.)
2 Account access (username/passwords to social media, websites, etc.)
3 Financial access (bank account credentials, credit card data)
4 Identity theft (information that can be used to masquerade as someone)
5 Existential data (information of national security value or threatens business survival)

s= source of the breach

values

1 Lost device such as a laptop, DVD, or USB thumb drive
2 Stolen device
3 Malicious insider
4 Malicious outsider
5 State espionage

Action= whether or not the stolen data has been used to cause harm be it identity theft, credit application, or bank account withdrawals

values

1 No action
5 Publication of embarrassing or harmful information (Wikileaks, hacker logs, etc.)
10 Use of financial identity to obtain funds or apply for loans

Figure 6: Breach Level Index (Source: Stiennon, 2013)

2.4. Addendum

Since this initial literature review was carried out, a particularly interesting and thorough study was published (Ali et al., 2021) and is worth a comparison here. The authors present an up to date (samples from 1988 up to 2018) SLR concerning ESM studies of the influence of favourable and unfavourable information security events on the stock market with a view to building on the work of Spanos and Angelis (2016). The similarity of the search query string Ali et al. (2021) used to that employed above is evident, noting the inclusion of the final ESM term: *((“Information Security” OR “Computer Security” OR “Network Security” OR “Internet Security” OR “Information System Security” OR “Web Security” OR “Software Security” OR “Application Security” OR “Cyber Security” OR “Data Privacy” OR “Security”) AND (“Market Value” OR “Stock Value” OR “Stock Market” OR “Stock Price” OR “Market Price” OR “Shareholder Wealth” OR “Firm Value” OR “Market Impact” OR “Share Price” OR “Shareholder Value” OR*

"Market Reactions" OR "Capital Market" OR "Market Securities") AND ("Event Study"))).

Furthermore, Scopus formed the bulk (67%) of the papers of interest identified, again underpinning the decision to use Scopus as the basis for this literature review. Considering the similarity between both the search string and the digital sources it would be natural to expect a great deal of overlap between this literature review and that of Ali et al. (2021) which was indeed the case.

Some key observations identified by Ali et al. (2021) are:

1. The prevalence of PRC as a source of information security events. 55% of papers used this as a data source. LexisNexis was the next most commonly used (29%).
2. Most studies of this type were based on solely US listed firms (76%). Nevertheless, there is a trend more recently for more international studies.
3. The ESM regression model most commonly used was the single-factor market model (79% of studies). FF3FM was the next most common.
4. Sample sizes were not huge, 91% being around 200 events for analysis. Favourable information security events (such as security investments) typically smaller (ca. 100).
5. Estimation windows ranged from 299 days (maximum) down to 50 days.
6. Event windows varied again (0, 1) most common followed by (-1, 1). Many studies used multiple windows to test for significance (46%). Long-term ESM studies are rare.
7. Hypothesis testing: most studies used the t-test (55%), followed by the z-test (37%).
8. Most studies (56%) focus on the negative impact of breaches on the victim firm. 71% found a significant negative impact. Fewer studies exist concerning favourable events such as the introduction of legislation.
9. Stock market reactions to negative events are more volatile (-5.5% to zero with an average of -3.5%) than to favourable events (0.63% to 1.36%).
10. In 54% of studies there is evidence of market reaction in advance of the event (information leakage). This is more prevalent for favourable events.
11. The most significant contingency factors are: "*time frame, industry type, type of breach, and firm size*". The technology industry sector is the most reactive. 3 out

of 5 studies cite a higher magnitude CAR if the breach involves confidential information.

12. Price-based ESMs were by far most common. Only a handful of examples involving trading volume or risk were identified.

Points (1), (2), (3)[15], (7), (8), (9) and (12) are very much consistent with this literature review and quite 'generic'. Regarding (4), (5), (6), (10) and (11), these are rather more variable depending on the exact type of ESM study and thus are discussed in more detail in the relevant chapters.

2.5. Conclusion

Based on the above initial literature review, and reinforced by Ali et al. (2020), it appears as though articles (especially recent ones) in this area are indeed limited and that there is a very strong US bias for what little is out there. As Ali et al. (2020) opine, the US is favoured for ESM studies because of its accessibility (English language, readily available data such as PRC) which in turn arises from the early adoption of breach notification legislation and thus the willingness of US firms to disclose information. Therefore, a UK/EU study, it seems, would be a useful research contribution although it is worth noting this was a search based on English language articles only and so may not cover all EU wide literature[16]. In addition, such a study could also compare and contrast existing US based studies to understand how market reactions differ across nations, thereby explaining some of the seemingly conflicting observations e.g. Richardson et al. (2019) versus Bose and Leung (2014) which Kannan et al. (2007) goes some way toward answering.

The recent change in EU legislation (GDPR) is also of interest – how has this impacted the market reaction of data breach announcements and, again, how does this compare with similar studies in the US? Therefore, it would be necessary to analyse UK/EU breaches both before and after the introduction of the GDPR so this needs to be considered during data selection. The comments of Yayla and Hu (2011) regarding a change in market behaviour over time in becoming less sensitive to breach announcements, also needs to be borne in mind for any analysis.

[15] Concerning the regression model an important decision is the choice of market reference which was not captured by Ali et al. (2021). This is expanded on in Chapter 4.

[16] Note, however, around 80% of articles on Scopus are written entirely in English (https://www.theatlantic.com/science/archive/2015/08/english-universal-language-science-research/400919/. Accessed on: 04/04/23)

The importance of categorising data breaches such as by industry sector (Tweneboah-Kodua et al., 2018) or combination of attack type and industry sector (Bendovschi et al., 2016) for any analysis has also been highlighted providing a solid foundation for RQ2 (and underpinned by Ali et al. 2021).

Initial investigation of RQ4 in the area of investment justification shows that literature in this area is also sparse (Bendovschi et al., 2016) and needs to be revisited as part of this research project once a greater understanding of RQ1 and RQ2 has been established. As noted above, Bendovschi et al. (2016) also link RQ2 and RQ4 suggesting that by gaining such deeper comprehension of the links between industry sector and cyberattack type that more efficient (industry specific) investment frameworks could be developed.

Based on this literature review as well as gaps clearly identified by Ali et al. (2021) the studies which form the core chapters of this thesis are a review of data breaches with a focus on EU markets (Chapter 4), followed by an analysis of GDPR related infringement fines (Chapter 5) and, finally, a positive (favourable) information security event which is the 'CISO Effect' as described in Chapter 6.

The next chapter explains, in detail, the ESM process used in these studies.

Chapter 3. Methodology (General Approach)

3.1. Introduction

In this chapter the general approaches applied in Chapters 4, 5 and 6 are described in detail. Methods which differ between chapters, such as the specifics of data collection, are addressed in the respective chapters.

3.2. Event Study methodology (ESM)

The purpose of an event study is to measure the effect of a corporate event, such as a merger, acquisition or other corporate announcement, on share price on or around the date of the announcement. Such events may be within the (publicly listed) firm's control or outside the firm's control such as a macroeconomic announcement or unexpected cyberattack.

It may be useful to begin with a few examples of corporate events, all of which not only had a significant economic effect in general but also, each one has a personal connection with the author.

In April 1991, Gerald Ratner, then CEO of the UK-listed jewellery company Ratners Group, in a speech at an Institute of Directors conference attended by an audience of 6,000 businesspeople and journalists at the Royal Albert Hall, London decried one of his company's products as being "*absolute crap*". This statement would be reported as having wiped £500m off the value of the company and gave rise to the expression "*doing a Ratner*" (Oxford Reference, 2022).

The share price of the pharmaceutical company Pfizer rallied in 1998 due to market hype about its new anti-impotency drug "Viagra". The increase in market capitalisation to $147bn was sufficient to overtake Merck, the previous market leader in that sector (Washington Post, 1998). Furthermore, the sudden interest in Viagra caused an uplift in other listed companies within the Pfizer supply chain such as UK based chemical manufacturer BTP plc which also had a favourable run in 2000 due to rumours of takeover talks (This is Money, 2000).

Not only can information have an impact on the market value of listed companies, but a broader economic impact across multiple (global) markets. Fellow oenophiles would surely be aware of the "Sideways effect". The film "Sideways" released in late 2004 (and subsequently nominated for five Oscars) derided Merlot and praised Pinot Noir. A study by Cuellar, Karnowsky and Acosta (2009) was able to show a negative impact on Merlot

sales and a corresponding (larger) uptake in Pinot Noir sales and price-point. This data was gathered over a much longer time-period whereas event studies tend to centre on the days immediately surrounding the initial release of information itself. The limitations of long-horizon event studies have been highlighted by e.g. Khotari and Warner (2006) whereas short-horizon were found to be quite reliable given the event date being known precisely – a key requirement for event studies in general.

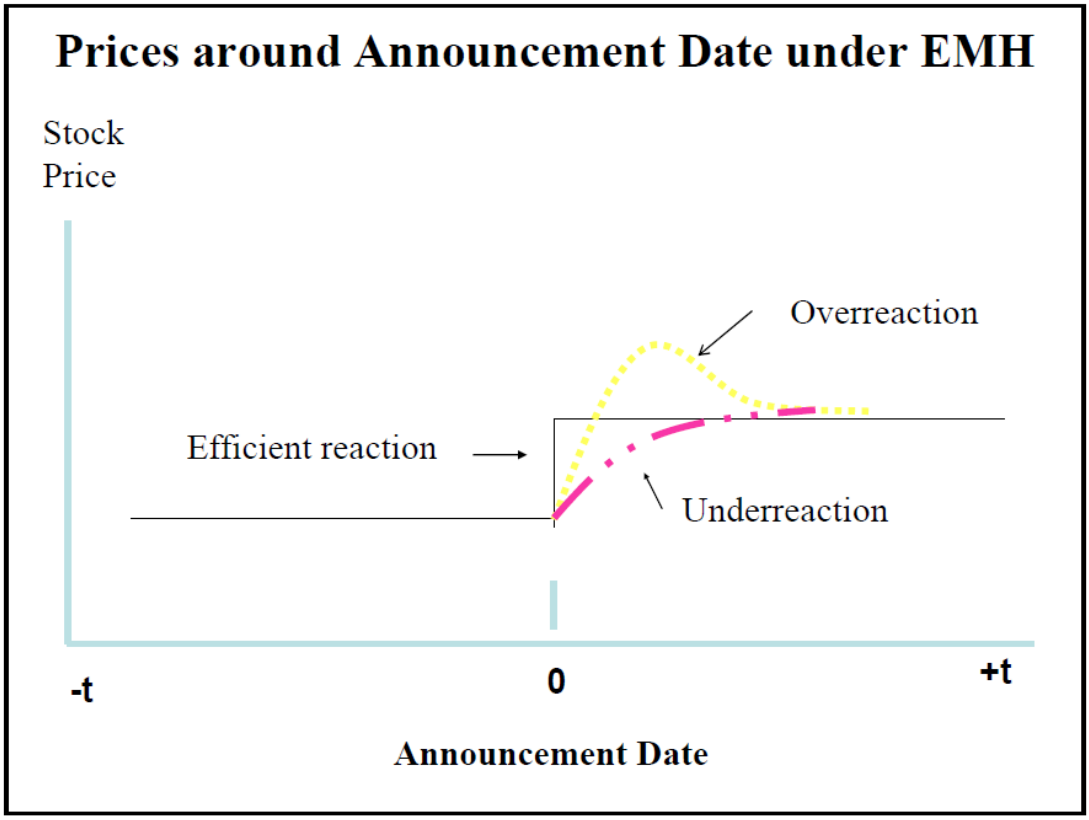


Figure 7: Efficient market hypothesis (Source: Bauer College, Houston, US)

Under the efficient market hypothesis (EMH) (i.a. Fama, 1970), any new (public) information will be immediately reflected in the share price as indicated in **Figure 7**. If all markets were truly efficient, however, it would be impossible to "beat the market" based on publicly available information alone (thus why bother trading unless there is information asymmetry?) which is the basis of the Grossman-Stiglitz paradox (Grossman & Stiglitz, 1980). There are three forms of the EMH: weak, semi-strong and strong. In the strong form, all information both public and private is already included in the share price, in the semi-strong form, public only and in the weak form all historical information is included and, therefore, it would be impossible to 'beat the market'. Potential market inefficiencies have been investigated by i.a. Salameh and AlBahsh (2011) who tested the

Palestinian stock market using event study techniques for abnormal returns following public announcements by listed firms under the mandatory disclosure regulations introduced by the Palestinian Stock Exchange in 2005. The authors found that the Palestinian market reacted gradually (underreaction curve in **Figure 7**), therefore was not perfectly efficient under the semi-strong model. For information security events, a study by Ishiguro et al. (2006) found that the Japanese market tends to react more slowly (ten days versus one) and those having higher intangible assets react more strongly. Clearly, there are variations in efficiencies across markets and, from Chapter 2, it was also noted in e.g. Tweneboah-Kodua et al. (2018) focussing on the US market (all S&P500 firms), that certain industry sectors such as financial services react more rapidly than others to breach events so there are also variations within markets.

This thesis focusses on publicly available information and thus the semi-strong form is most relevant here, although some examples of information leakage before the announcement date (strong form) were also encountered (cf. Lin et al., 2020).

Event studies are, indeed, a widely reported method to assess the impact of a specific event on the share price (market value) of firms - a detailed description may be found in e.g. MacKinlay (1997). Through the observation of share price movements in reaction to information regarding a specific event (such as those described above) over a short time period (the event window), it is possible to ascertain how the market reacted to that specific event, with the caveat of no other confounding events occurring during this period.

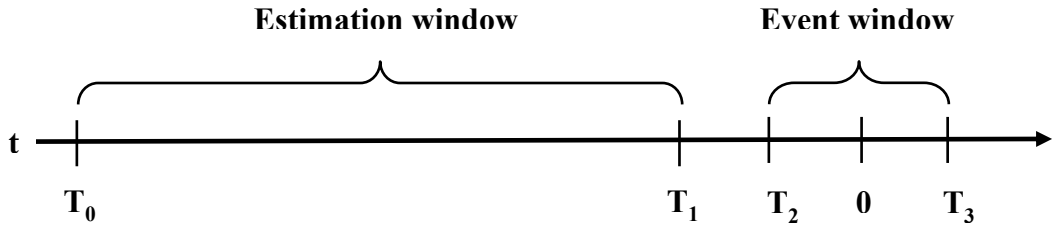


Figure 8: Event study timeline

A frequently used approach for information security related event studies is the market model (e.g. Cavusoglu et al., 2004; Telang & Wattal, 2007; Andoh-Baidoo, Amoako-Gyampah & Osei-Bryson, 2010; Goldstein et al., 2011; Tejay & Shoraka, 2011; Khansa et al., 2012; Bose & Leung, 2013; Hovav and Gray, 2014; Pirounias et al., 2014; Hinz et al., 2015; Schatz & Bashroush, 2016; Castillo & Falzon, 2018; Tweneboah-Kodua et al.,

2018; Deane et al., 2019; Jeong et al., 2019; Malliouris & Simpson, 2019; Rosati et al., 2019; Lin et al., 2020) which, through a regression analysis of a firm's share price over a relatively long estimation window (see **Figure 8**), is able to predict stock returns during the later (usually much shorter) event window. This approach assumes that returns follow a single factor model (**1**) where the return of firm i on day t ($R_{i,t}$) is dependent on the corresponding daily return of the reference market ($R_{m,t}$) and the extent of the security's responsiveness (β_i) offset by its abnormal return (α_i)[17]. The error term $\varepsilon_{i,t}$ is expected to be zero with finite variance. Abnormal returns are calculated for the event window (**2**) and reported as a cumulative abnormal return (CAR) over the whole event window (**3**). For cross-sectional analyses, a cumulative average abnormal return (CAAR) was calculated over N events (**4**).

$$R_{i,t} = \alpha_i + \beta_i \cdot R_{m,t} + \varepsilon_{i,t} \quad (1)$$

$$AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i \cdot R_{m,t}) \quad (2)$$

$$CAR_i = \sum_{t=T_2}^{T_3} AR_{i,t} \quad (3)$$

$$CAAR = \frac{1}{N} \sum_{i=1}^{N} CAR_i \quad (4)$$

Now that the method of calculation of CAAR is explained, it is useful to refer to **Figure 9** to understand the magnitude of abnormal returns from event studies concerning different types of corporate events. It can be seen that the largest impact event is that of director dealing at -36% whereas a board change leads to only a 5% loss on average. Although an acquisition announcement (as seen with the aforementioned BTP plc example) can have a positive effect on the target, the acquiring company sees an overall negative effect of 12% on average (although Khansa (2015) reported a positive effect on the acquiring company for information security mergers and acquisitions). It would certainly be interesting to see where information security events sit within this range.

[17] Compare with the CAPM: $R_{i,t} = R_f + \beta_i(R_{m,t} - R_f) + \varepsilon_{i,t}$

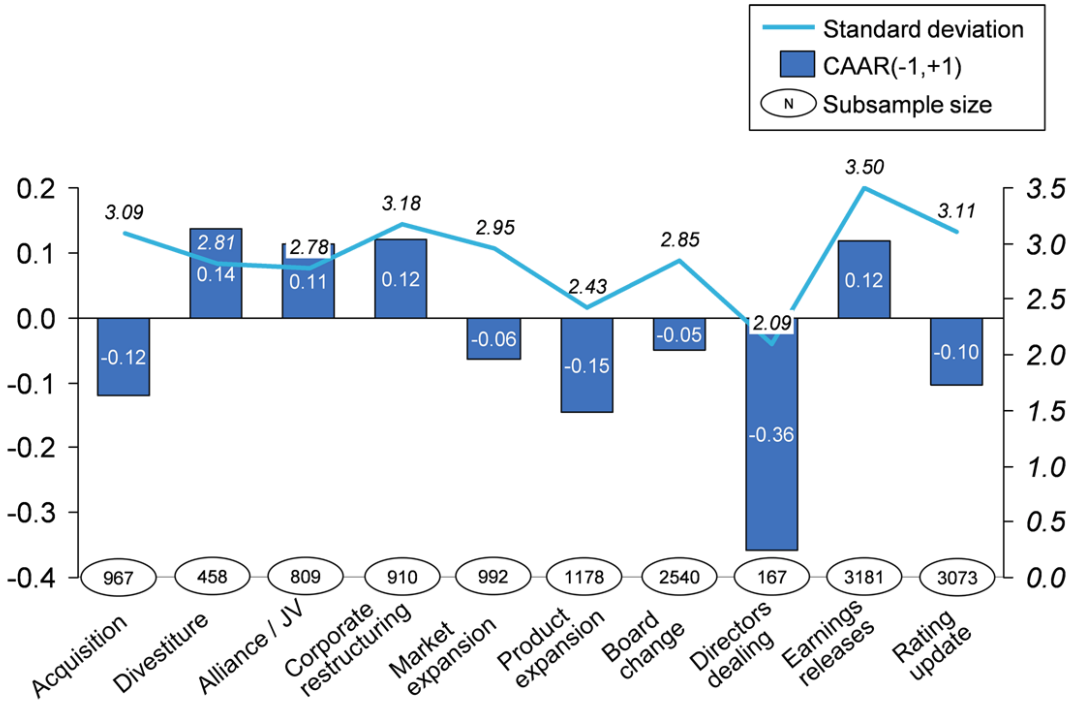


Figure 9: CAARs for different types of corporate events
(Source: https://www.eventstudytools.com. Accessed on: 04/04/23)

3.3. EventStudyTools package

There are various software packages available to perform event studies. The literature review in Chapter 2 showed that Eventus and Stata are two frequent choices. There is also, of course, the possibility to adopt a more manual technique such as Microsoft Excel. The package chosen for this research was EventStudyTools (Schimmer, Levchenko & Müller, 2014). This package was chosen as it was used earlier in an unpublished paper by the author and therein found to be consistent with literature, although further, more detailed validation of the package against literature was carried out and is described later in this chapter. In keeping with the EMH semi-strong (public) nature of this work, it also made sense to choose a public domain (freeware) application if possible. EventStudyTools seemed ideal in these respects, however, it was monetised in September 2020 in order to remain a sustainable proposition. Access can now (at the time of writing) be purchased on a (non-recurring) monthly basis to either the GUI access via the website ($9pm) or R API ($10pm). Furthermore, academic institutions have the right to apply to have their email domains whitelisted and bypass the paywall for the GUI access, at least[18].

[18] University of East London included.

For this research the R API was utilised. The R package was downloaded from the CRAN repository[19] and (once chargeable) an API-key purchased. Both the R API (wrapper) and GUI submit event studies to be processed on remote severs.

3.4. Data collection

The high-level strategy to data collection was to lean towards publicly available data rather than commercial packages to make this research as accessible and reproducible as possible.

Some data sources are common to each of the core chapters (Chapters 4, 5 and 6), such as share and market (reference) price data, industry sector, market capitalisation and currency exchange rates (Yahoo!Finance, 2019) whereas the source of the event announcements themselves varies from chapter to chapter and thus are described in detail therein. It is worth noting here that for the data breaches in Chapter 4, finding any reliable data source for EU markets was challenging and this had to be gathered manually, which was probably the most time-consuming effort, followed by CISO announcements. For Chapter 5 the process was made much easier by having a single reliable source, the Enforcement Tracker (CMS Legal, 2021) with very few gaps in the data to fill. In all cases, Microsoft Excel was used as the initial manual input and filtering tool, then CSV files generated to load into R (R Core Team, 2018). Due to concerns over COVID-19 events confounding the market (e.g He et al., 2020; Alam, Wei & Wahid, 2020) the data was date-capped at 31/12/2019. Nevertheless, some post COVID-19 data was utilised in Chapter 5, specifically a few examples of GPDR infringement fine appeals.

3.5. Data analysis

To execute event studies for each event, a core R function was developed, the outline of which is shown in **Figure 10** (the code is shown in the Appendix).

```
estudy <- function (
  firmSymbol,  # Stock symbol (Yahoo!Finance)
  firmName,     # Just a label somewhat more explicit than the
symbol alone
  indexSymbol, # Index symbol (Yahoo!Finance)
  indexName,   # Just a label for the index
  eventDate,   # "%d.%m.%Y"
  startEvent,  # start of event window
  endEvent,    # end of event window
```

[19] https://cran.r-project.org/web/packages/EventStudy/index.html. Accessed on: 04/04/23

```
  endEst,      # end of estimation window
  lengthEst,   # length of estimation window
  estprice = "adjusted", # use "adjusted" or "close"
  estgroup = "breach",
  estfiletype = "csv",
  estbenchmarkmodel = "mm", # default is market model otherwise
"ff3fm" could be used
  estreturntype = "simple", # use as default rather than "log"
  estnontradingdays = "later", # not relevant if day 0 is trading
day
  estffdata = "" # string containing name of data.frame in global
environment to use for ff3fm
  ) { … }
```

Figure 10: estudy function outline

The function begins by performing some basic validation checks on the data, namely:

1. Both firm and index data should be non-empty
2. Firm and index data should start and end on the same day and would be expected to contain the same number of rows
3. The data has to be historic (today latest) and not in the future – a conservative estimate is made, and a warning flagged if a few days before current time
4. There should be at least enough trading days data to cover the estimation window and event window.

Warnings are generated based on the outcome of these tests for the log file. The required data files for the EventStudyTools package ("*01_requestFile.csv*", "*02_firmData.csv*" and "*03_marketData.csv*") are created automatically in the function by extracting data from Yahoo!Finance directly (using the R tidyquant package) based on the function parameter values provided. These files are then submitted to the EventStudyTools servers and returned as an R object. Finally, the function saves the R objects in both a results table and text files in a uniquely named results directory and moves the log files (stdout, stderr) into the same folder. As running event studies is processor intensive, it makes sense to retain all results to avoid both the additional time and carbon footprint involved in a rerun. The function also returns the name of the results path so this can be added into the results table and referred back to if necessary (EventStudyTools provides a parser function to read back files into an object, if needed).

The parameters input are described in each chapter as there is some variation, default values are indicated in **Figure 10**. Most should be self-explanatory – it is usual to use the adjusted closing price (to take into account stock splits and dividends etc.) to reflect a more accurate value - this was not overridden other than for testing or validation purposes. The estgroup parameter is simply a label to be used if multiple firms are being submitted in the data files and, as the calculations here are for similar events affecting different firms at different times, this was not relevant and, therefore, not changed from "*breach*" after the Chapter 4 analyses were completed. The single factor Market Model (MM) was favoured for all chapters as this was the most widely reported analysis method from the literature review, although other methods are supported by EventStudyTools such as the Market Adjusted Model, Comparison Period Mean Adjusted Model, Market Model with Scholes-Williams Beta Estimation and Fama-French 3-Factor Model (FF3FM). A comparison of the MM with the FF3FM is shown later in this chapter as these are two frequently cited models in ESM.

The selection of the reference index (indexSymbol) is also important as has been seen from the literature review (Kannan et al., 2007) and the logic behind the selection is elaborated in each chapter, as underpinned by evidence from Chapter 4 (SPEUR350 versus market specific indices).

For data analysis and visualisations, separate scripts were developed for each chapter. At the end of the project the resulting R code comprised around 2,000 lines.

3.6. Hypothesis development

The null hypothesis (H_0) for event studies (**5**) maintains that there are no abnormal returns within the event window and the alternative hypothesis (H_1) states the opposite (**6**). Here, μ may represent either AR, CAR or CAAR in the case of cross-sectional studies. The standard deviation of abnormal returns during the estimation window is described by (**7**) where M_i refers to the number of non-missing returns. The t-value for the CAR over the event window for each firm i was then calculated according to (**8**).

$$H_0 : \mu = 0 \qquad (5)$$

$$H_1 : \mu \neq 0 \qquad (6)$$

$$S_{AR_i} = \sqrt{\frac{1}{M_i - 2}\sum_{t=T_0}^{T_1}\left(AR_{i,t}\right)^2} \quad (7)$$

$$t_{CAR_i} = \frac{CAR_i}{\sqrt{(T_3 - T_2 + 1)S_{AR_i}^2}} \quad (8)$$

For cross-sectional analyses the t-statistic (t_{CAAR}) was calculated based on the CAAR (**10**) with S_{CAAR} being the standard deviation of the CARs for each firm i across the sample of size N (**9**).

$$S_{CAAR} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(CAR_i - CAAR)^2} \quad (9)$$

$$t_{CAAR} = \sqrt{N}\frac{CAAR}{S_{CAAR}} \quad (10)$$

Significance testing in this way is consistent with e.g. Castillo and Falzon (2018), Deane et al. (2019) and Jeong et al. (2019). Indeed, Deane at al. (2019: 115) argue that "*the t test is considered to be the best framework for analyzing statistical significance in most event study frameworks and to be relatively robust*".

3.7. Comparison of the Market Model with the Fama-French 3-Factor Model

Two statistical regression models, based on the Capital Asset Pricing Model (CAPM), frequently cited in the literature review (Chapter 2) were the single-factor Market Model (MM) and the Fama-French 3-Factor Model (FF3FM), which introduces two additional factors, specifically size (market capitalisation) and book-to-market ratio (Fama & French, 1992). Some studies (e.g. Rosati et al., 2019; Lin et al. 2020) have reported little variation between the two, and EST offers both methods, thus it was a useful exercise

here to compare results calculated using each of the two methods. Furthermore, Deane et al. (2019) ran the Fama-French Carhart 4-Factor Model, which involves the additional factor of momentum (Fama & French, 1996), alongside the MM and reported that they did not differ significantly, with Richardson et al. (2019) adopting a similar approach also commenting on the similarity of the results.

For the GDPR infringement fine study (Chapter 5), 250 event studies were performed using the MM and summarised in **Table 10**. All 250 studies were rerun using the FF3FM[20] and the results are tabulated in **Table 2**. The two methods both yielded a mean CAAR of ca. -0.4% overall and, at first glance, the results for each event window also look similar, with the CAARs matching exactly for (-2, 2) and (0, 20). That said, the t-test values for the FF3FM seem to be of lower magnitude than those of the MM in nearly all cases which results in the statistical significance of the non-zero abnormal returns effectively dropping a level. For example, the MM showed event windows (0, 2) and (0, 3) having significance at the 1% level, whereas they are only significant at the 5% level for the FF3FM. In a similar vein, the slightly negative CAAR (-0.6%) observed by the MM for the five-day window (0, 4) at the 10% significance level ceases to be recognised at all under the FF3FM as the result is not significant and the null hypothesis of zero abnormal returns cannot be rejected. Notwithstanding these differences, the two methods both show the (0, 3) event window as having the most (and statistically significant) negative abnormal returns and thus this window would have been selected to carry forward for further analyses in either case. The (0, 2) event window shows the second most negative CAAR in using both the MM and the FF3FM as well, and, although this window has a slightly higher magnitude t-test value than the (0, 3) window in the case of the FF3FM, they both still remain significant at the 5% level therefore, again, the (0, 3) window would be the one of interest for this study.

[20] Daily Fama-French factors were downloaded from the website of Kenneth R. French (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french. Accessed on: 04/04/23) using the European (developed market) file for all examples except Marriott and Alphabet Inc. (US listed).

Table 2: Comparison of MM and FF3FM

		MM Results			FF3FM Results		
Event Window	N	CAAR	t_{CAAR}	% Negative CAR	CAAR	t_{CAAR}	% Negative CAR
(-2, 2)	25	-0.0049	-1.6188	56	-0.0049	-1.5057	56
(-1, 1)	25	-0.0041	-1.2112	64	-0.0039	-1.2109	64
(-1, 0)	25	-0.0022	-0.6746	52	-0.0029	-1.0278	52
(0, 1)	25	-0.0064	-2.7453**	72	-0.0050	-2.0247*	68
(0, 2)	25	-0.0072	-3.0748***	80	-0.0057	-2.7281**	80
(0, 3)	25	-0.0096	-3.2341***	76	-0.0077	-2.4381**	68
(0, 4)	25	-0.0064	-2.0190*	72	-0.0050	-1.4381	56
(0, 5)	25	-0.0061	-1.4128	56	-0.0046	-1.1292	48
(0, 10)	25	0.0020	0.2795	48	0.0004	0.0548	44
(0, 20)	25	0.0011	0.0968	40	0.0011	0.1136	44
	250	**-0.0044**		**62**	**-0.0038**		**58**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

3.8. Validation of the method against literature

As mentioned above, although EST was already familiar to the author and previously validated at basic level (Schimmer et al. (2014) also report that they have benchmarked their abnormal return calculators against other applications such as Eventus and Stata) there was a need for more robust, 'independent' testing of the EventStudyTools package and analysis method as a whole, ideally, a comparison with literature.

To begin with, internet searches for event study application were carried out resulting in an interesting reference: "*Performing an event study: An exercise for finance students*" which states "*the accompanying spreadsheet calculates cumulative abnormal returns and cumulative abnormal trading volume and plots them in separate graphs*" (Reese & Robins, 2017). The spreadsheet attachment extracts data directly from Yahoo!Finance and processes it using the Excel Analysis Toolpak add-in. Although not related to cyber security, these example event studies were regarding additions to the S&P 500 index, thus could theoretically be used to benchmark the EST package. On attempting to install it, it was found that the spreadsheet was not functioning correctly and the corresponding author confirmed that Yahoo!Finance have changed the way data is extracted since the tool was developed and, furthermore, this tool would not be maintained going forwards.

The recommendation was simply to calculate the regressions manually in Microsoft Excel. This process is described in detail (with examples online) in e.g. Benninga (2008).

Four examples (from Chapter 5) which ran cleanly[21] in EST were chosen and calculated manually using Excel and the resulting CAR values are tabulated in **Table 3** along with the EST reported figures for comparison. A paired (two-sided) t-test between the CARs showed significance at the 1% level, $t(3) = 1.6634$, $p = 0.1948$, thus the null hypothesis that the difference in means was zero could not be rejected. As these values reported by EST were generated using the API, as a further check they were also all processed using the aforementioned GUI abnormal return calculator accessible through the EST website[22]. No differences in values between the two methods were observed.

Table 3: Comparison of CAR calculation methods

Ultimate Parent Company	Date	Event Window	CAR (EST)	CAR (Excel)
Endesa SA	2019-04-09	(0, 3)	-0.0300	-0.030106986
Marriott	2019-07-09	(0, 3)	-0.0097	-0.009716540
International Airlines	2019-10-01	(0, 3)	-0.0303	-0.030338354
Direct Line Insurance	2019-12-03	(0, 3)	-0.0007	-0.000697273
(Mean)			**-0.0177**	**-0.017714790**

The next step would be to compare with literature reported abnormal returns. Such a comparison was not without its challenges, it was necessary to find a study which reported individual CARs rather than CAARs aggregated across multiple firms or sectors to ensure an easy and direct comparison. Then, of course, once such examples are found, share price data needs to be still accessible for that particular stock symbol on that particular date (no mergers or acquisitions, delistings etc.) which means focussing on more recent studies to reduce the likelihood of such occurrences. Two examples were found in Schatz and Bashroush (2016a) as they mention, firstly, outlier Wyndham Hotels and Resorts Inc. (WH) as being *"-22%*" CAR in response to a data breach announced on 16/2/2009. The authors report using the "*S&P 500 Composite"* index as a reference using Thomson Reuters[23] as a data source. This gave sufficient data to run a similar study in EST for

[21] Meaning no missing data on trading days, or events announced on non-trading days etc.
[22] https://www.eventstudytools.com/arc/upload. Accessed on: 04/04/23
[23] Now Refinitiv.

comparison, however, the returned value from EST was -23.42%. This difference could be attributed to the fact that a different price data source was used (Thomson Reuters versus Yahoo!Finance) and it was also noted that the event date 16/02/2009 was in fact Presidents' Day (a non-trading day). EST was configured for this test to move to the earlier nearest trading day (13th) whereas it is possible that the method used by Schatz and Bashroush (2016a) defaulted to the later day (17th) – the approach was not specified in the paper. By running again in EST and forcing the 17th (later trading day) as the event date, the result changed to -22.52% so this is much closer to that observed by Schatz and Bashroush (2016a). The estudy function shown above was set to 'later' by default for future studies. It is also worth mentioning that the "*S&P 500 Composite"* index cited by the authors is not normally referred to as a composite index and one has to wonder if the market references were, indeed, identical as well. Nevertheless, this (as it was an outlier) was an 'extreme' example (compare with **Figure 9** – a CAR of this magnitude is much higher than expected on average for information security type events) and thus any slight differences in data would surely be amplified in such a case. The second example from this study was a repeat event for WH on 28/02/10. Here, the authors did not specify a precise value in the text, but on the graph displayed the CAR certainly looked slightly negative although very close to zero. The EST returned result was +0.1%. Although these examples show that EST is certainly in the same ballpark as other applications across a broad range of values, it was still well worth an extensive search for a more appropriate paper for comparison. Although studies such as that of Richardson et al. (2019), who use the Stata package eventstudy2 as described by Kaspereit (2015), and Tweneboah et al. (2018), who cite Yahoo!Finance as a data source, would seem ideal for comparison purposes, they again cite summarised cross-sectional CAARs rather than individual firm CAR thus are not suitable candidates (Richardson et al. (2019) also favours FF4FM over the single-factor market model, although comments they are similar). Overall, there is much variation in data source and parameters and packages used for event studies (cf. Ali et al., 2021). One popularly cited package is Eventus (Cowan Research LC)[24] and there are a couple of worked examples available on the internet. The first, (O'Hara & Shaw 1990) is a (now dated) paper regarding an announcement by the US Comptroller of the Currency in 1984. This example researches the market reaction to 22 listed US bank stocks. The problem here would be tracking down the correct listings as the dataset is so out of date (the CUSIPs reported are no longer valid – a name search would be required).

[24] Eventus runs in SAS (from the SAS Institute).

Also, the study uses the CRSP value weighted index as a market reference which is not available from Yahoo!Finance. There is a more recent example (Cowan Research LC, 2001) however, the price data appears to be rolled up by weeks (EST works on daily prices) and CAAR figures are cited across three stocks so, again, this is not really viable even though the Russell 1000 index is used as a market reference which is accessible from Yahoo!Finance. Ideally, a study showing a much higher level of granularity would be preferred and, fortunately the study by Castillo and Falzon (2018) was found to be more appropriate. Here the single-factor market model is favoured and compared with the mean-adjusted model (the comparison here utilises only the market model data). The corresponding author confirmed that Microsoft Excel was used as the primary tool for statistical analysis and was checked by EViews with no changes reported. What makes this paper particularly useful is that only a single market event is reported (Wannacry) and the announcement date (Monday 15th May 2017), estimation window length (-244, -6) and event window are all clearly stated in the paper. Even better, rather than a CAR figure summed over a number of trading days, the paper reports the AR for each individual firm on the event day itself (day 0) and t-tests are calculated using exactly the same method here, thus are ideal for comparison purposes. It was also an opportunity to compare price data from Thomson Reuters versus Yahoo!Finance. The authors analysed ARs (day 0) for 43 companies and two ETFs. They also mentioned that the market reference used for the US listed companies and the ETFs was the Nasdaq Composite index[25]. Other countries were mapped to the "*main index for that country*", therefore it was decided to restrict the comparison to only the US listed symbols to avoid any possible confusion over the market reference. After checking for existing data on Yahoo!Finance, 23 (53% of the original dataset) of the Castillo and Falzon (2018) stocks were identified for comparison which included one ETF (CIBR). The AR values reported by EST[26] are plotted against those from Castillo and Falzon (2018) in **Figure 11**.

[25] ^IXIC in Yahoo!Finance

[26] Although EST calculates CARs only over multiple days, the aforementioned EST parser function was utilised to extract the AR on day 0.

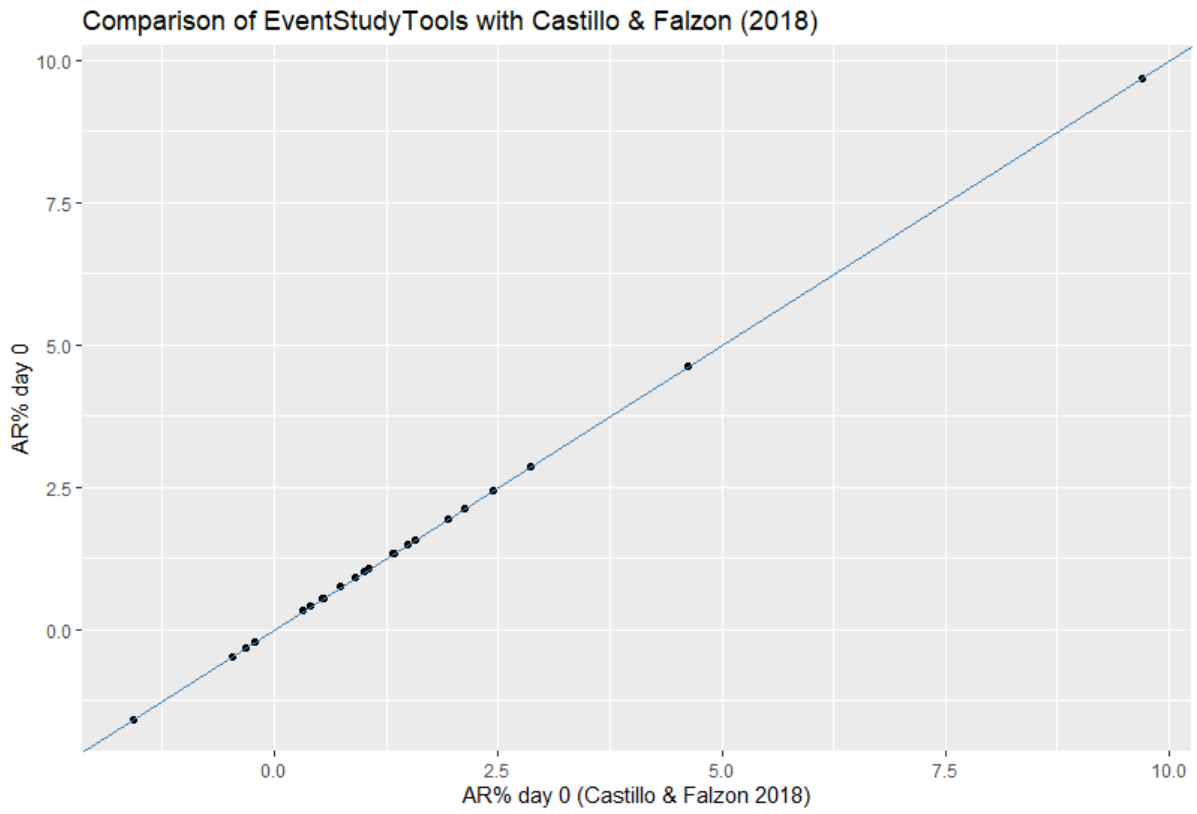


Figure 11: Comparison of EST AR values with Castillo and Falzon (2018)

The strong correlation between the ARs reported by Castillo and Falzon (2018) and those calculated by EST is clearly visible – they match almost exactly. The authors also report a t-test value calculated the same way as described in (**8**) above. A similar plot of the Castillo and Falzon (2018) t-test values versus those calculated by EST is shown in **Figure 12**.

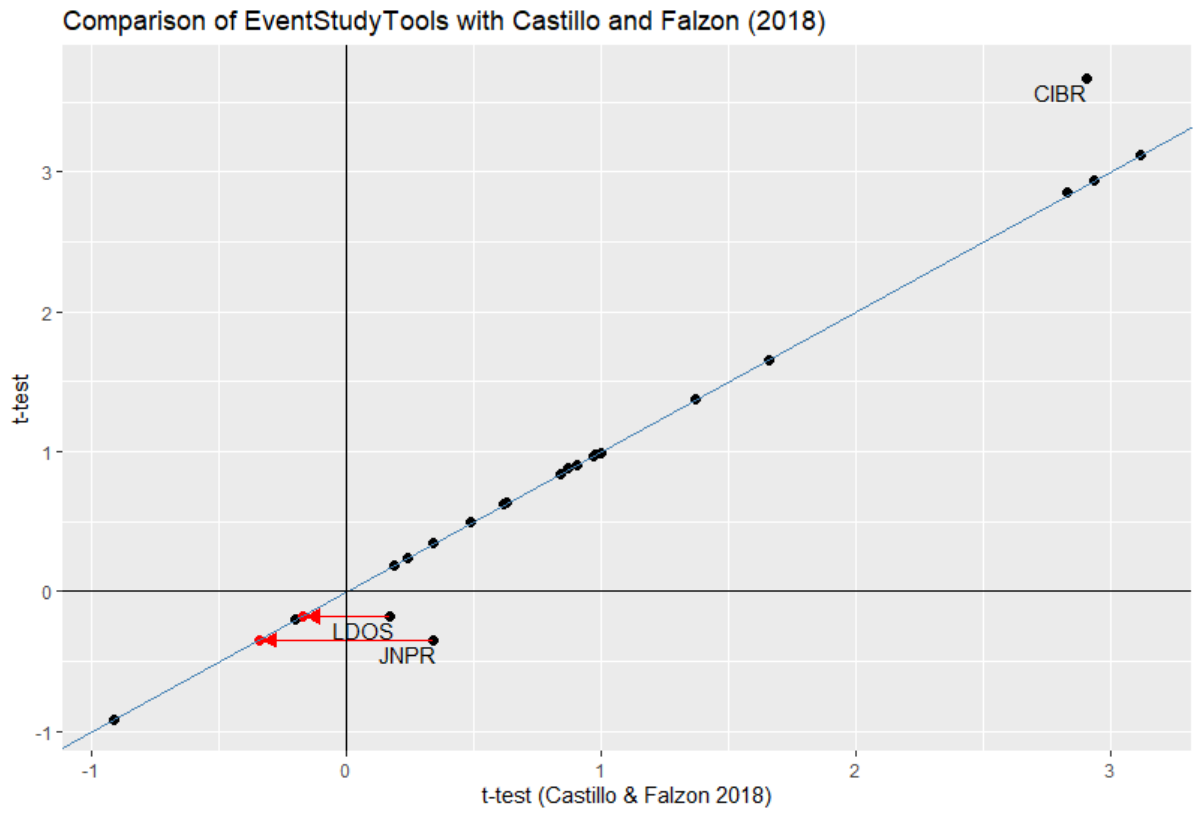


Figure 12: Comparison of EST t-test values with Castillo and Falzon (2018)

Although a strong correlation is clearly visible for the AR values, the t-test values, however, were found to have outliers, namely LDOS, JNPR and CIBR. Upon closer inspection it was observed that the signs of the t-tests appeared to have been misreported in the literature and, indeed, this was confirmed by the corresponding author. Once the incorrect signs were changed for LDOS and JNPR the points fell perfectly on the correlation line as indicated in **Figure 12**. The difference in the t-test values for the ETF (CIBR), however, remains unexplained. This was reported higher by EST than Castillo and Falzon (2018), but even the lower value of the two was high enough to demonstrate statistical significance. As the two AR values match but not the t-test, EST must have calculated a lower standard deviation in the case of CIBR but, as standard deviation figures are not quoted in the Castillo and Falzon (2018) paper. this could not be confirmed.

3.9. Conclusion

In conclusion, the results from EST have been sufficiently validated both against another calculation method (Microsoft Excel) and a commercially available package (EViews) as reported by Castillo and Falzon (2018). As the underlying price and market reference data came from different sources, there is also a high degree of confidence that the

Yahoo!Finance data is as reliable as a commercial source such as Thomson Reuters Datastream.

Furthermore, the similarity of the results calculated using the MM and the FF3FM has been highlighted and is consistent with previous studies. The fact that the MM seems more likely to detect abnormal returns due to the higher probability of rejecting the null hypothesis, along with its lower level of complexity, are surely contributing factors in its prevalence in previous studies over the FF3FM. Indeed Ali et al. (2021) report that 79% of studies in their SLR utilised the MM.

This methodology is applied in the next three (core) chapters. As some ESM parameters vary between studies, such as the choice of event window (cf. Ali et al., 2021) an explanation is given in each chapter as to the why these options were chosen.

Chapter 4. The Impact of Data Breach Announcements on Company Value in European Markets

4.1. Introduction

The Cyber Security Breaches Survey (DCMS, 2022) reports that 72% of large size firms in the UK "*have identified [cyber] breaches or attacks in the last 12 months*". In the year to June 2019, the Office for National Statistics (2022) cites 641,000 incidences of computer misuse for England and Wales alone, a figure including both personal and business-related hacking attempts. With the Data Protection Act (2018) now in force in the UK along with the equivalent 'General Data Protection Regulation' (GDPR) EU wide, firms must disclose any breaches involving personal data within 72 hours and face hefty fines of up to €20m or 4% of turnover (whichever is the greater) for failure to comply.

In light of the above, as well as some recent high profile data breach disclosures such as that of British Airways (2018), it would be reasonable to expect cyber security to be a major concern at board level for not only UK firms, but across Europe. This chapter aims to investigate the impact of data breach announcements on the market value of publicly listed companies with a view to influencing investment in cyber security. Existing literature in this area was found to be somewhat sparse recently and exhibited a strong US bias, hence this chapter focusses on European markets and compares and contrasts with the literature.

By gaining an up to date understanding of any potential negative impact of data breach related announcements on market value, this will highlight the importance of information security to business management as well as the need to invest in cyber security to avoid such incidents. Such insight would also assist practitioners of information security with business case justifications. This research would be of benefit to business management, practitioners of cyber security, investors and shareholders as well as researchers in cyber security or related fields.

The literature review (Chapter 2) showed that the impact of publicly announced data breaches on market value (RQ1) is a topic which has been researched for some years. For example, Cavusoglu et al. (2004) reported that those firms suffering a serious data breach lost, on average, 2.1% of their market value within two days of the announcement, whereas Goel and Shawky (2009) cite a figure of around 1%. A recent literature review carried out by Spanos and Angelis (2016) noted that research in this area, despite its

longevity, was "*quite limited"* although the majority (76%) of studies did show an impact of security events on company market value which was statistically significant. Indeed, even more recently, Lin et al. (2020) cited a loss of 1.44% on average over 5 days. Tweneboah-Kodua et al. (2018: 646), who analysed breach events for 96 S&P500[27] listed firms between 2013 and 2017, however, did not find significant impact over shorter event windows and warn that "*studying the cumulative effects of cyberattacks on prices of listed firms using event study methodology without grouping the firms into various sectors may not be informative*". Financial services sector firms, for example, showed larger abnormal returns over a 3-day event window than those in the technology sector, starting to provide input to RQ2.

Consistent with Tweneboah-Kodua et al. (2018), Richardson et al. (2019) also report a lesser effect on market price, citing an average of less than 0.3% based on an analysis of 827 breach disclosures for 417 companies. Again, this was a US based study, the breach event data sourced from Privacy Rights Clearinghouse[28] (PRC). Richardson et al. (2019) chose propensity matched firms as a reference market rather than the more usual S&P500 composite index which could explain why their findings are so different from Cavusoglu et al. (2004). Indeed, Kannan et al. (2007) in their study found no significant impact either, also using control firms as a reference. An alternative explanation is provided by Yayla and Hu (2011) who note that the market appears to have become less sensitive to breach events in recent years – another factor to be mindful of in any analyses.

Commenting on their findings, Richardson et al. (2019: 249) argue that "*companies are unlikely to change their investment patterns unless the cost of breaches increases dramatically or regulatory bodies enforce change*" – a contribution towards RQ3. It is acknowledged, however, by Richardson et al. (2019) that exceptional events do occur with cases of massive data exposure having potentially catastrophic impact, suggesting a need to categorise data breaches according to their severity (RQ2), such as number of records exposed or level of data sensitivity. Campbell et al. (2003) observed that breaches involving unauthorised access to confidential data were more likely to result in significant negative market reaction.

The above quotation from Richardson et al. (2019) also poses another question – what of the recent change of legislation (GDPR) in UK/EU, has there been any impact? As the

[27] Standard & Poor's index of 500 US stocks representative of US markets in general
[28] https://privacyrights.org/. Accessed on: 04/04/23

introduction of GDPR is so recent (2018), literature in this area is rare, however Goel and Shawky (2014) carried out a similar US based study and observed that negative effects of security breaches were reduced significantly after the enactment of security breach notification laws. In a recent study of the economic impact of GDPR infringement fine disclosures, Ford et al. (2021b) observed negative returns of around 1% up to 3 days after the announcement with this loss of market valuation being far greater than the monetary value of the fine itself in almost all cases. Seemingly minor fines could result in huge losses even for firms having large market capitalisations.

In summary, although there have been differences in results from studies related to the impact of data breach disclosures on market value, there are certainly some common themes such as: event study techniques (described below) are the favoured method for quantitative analyses and the research has a strong US bias, presumably because of readily accessible breach datasets for that market as well as a kind of 'one size fits all' market reference index in the S&P500, with a few notable exceptions such as Bose and Leung (2014). Thus, this research aims to go some way to address the deficit of European centric studies in this area although it should be recognised that literature searches were limited to English language only, thereby possibly excluding some studies of interest.

4.2. Methodology

The high-level approach to this research was to gather a dataset of data breach announcements for European publicly listed companies, then analyse the impact of these announcements on share price using ESM as elucidated in Chapter 3.

4.3. Data collection

The scope for data collection was limited to breach announcements for companies (or their ultimate parents) publicly listed on a European exchange. Ownership of subsidiaries was confirmed through Dun & Bradstreet[29]. To maximise the initial dataset, a broader geographic concept of Europe was used including both continental and trans-continental countries, 52 ISO-3166 country codes in total.

The manual data gathering exercise for European data breaches is described by the following steps (other breaches of relevance identified serendipitously were added, of course):

[29] https://www.dnb.com. Accessed on: 04/04/23

1. Review monthly cyber security blogs (e.g. IT Governance Ltd, 2023; DataBreaches LLC, 2023) for data breach announcements from 01/01/2017 stopping at 31/12/2019 to avoid possible market effects of COVID-19 which could be considered a long-term confounding event in itself. The resulting dataset would be centred roughly around the introduction of GDPR in May 2018 to help with before/after comparisons.
2. Identify breaches of interest, namely those specific to European listed companies or subsidiaries of European listed companies. Breach announcements regarding technology vulnerabilities which applied to multiple companies were disregarded. Privately owned companies were filtered out (e.g. Monzo, Yves-Rocher). Non-European examples filtered out included Everis Spain (Japan) and Three UK (Hong Kong). Also filtered out were cases where the 'breach' was only an allegation and the parent company immediately denied any breach had occurred e.g. Choice Hotels, British Airways.
3. Perform internet searches for the earliest dated public announcement (thereby removing uncertainty around the event date). In each case the announcement was validated against multiple later disclosures.
4. Where possible, gather additional data fields such as the nature of the breach, number of breached records and whether the incident involved personal data.

After completing the above steps, the resulting dataset comprised 33 records. To supplement these, a useful potential data source relevant for Europe mentioned in the literature was the Breach Level Index (BLI) as provided by Gemalto (Thales Group 2017), however since its acquisition by Thales, this data source seems to be no longer publicly available. Instead, the VERIS Community Database (VCDB)[30] was also reviewed as a possible data source, but data here was found to be sparse (only 8,857 records in total worldwide to date) having very little overlap with the hand-gathered dataset (actually only one, the UniCredit SpA incident). The original dataset was augmented by 12 breach disclosures as a result of the VCDB search bringing the total to 49 records. Such a sample size is nowhere near that used by e.g. Richardson et al. (2018) of 827 records but, nevertheless, closer to that of Tweneboah-Kodua et al. (2018) at 96. The difficulty of obtaining a breach database of similar size to these US based studies does, again, underpin favouritism towards this market by researchers due to accessibility

[30] http://veriscommunity.net/vcdb.html. Accessed on: 04/04/23

of data and highlight the need for a European equivalent as there is no reason to believe European companies are not just as susceptible to data breaches as their US counterparts.

Share price and market index data were sourced from Yahoo!Finance (2019) along with firm demographics such as industry sector. For each ultimate parent company, the most appropriate reference index was selected, ideally one of which the candidate firm was a component but adjusted to closest match when data could not be extracted from Yahoo!Finance. This selection of the reference market is important (Kannan et al. 2007; Richardson et al. 2019). Some firms had multiple listings, in which case the primary listing was favoured along with the associated index. Share price data were not available for all of the 49 records and a further four had to be removed, namely Npower, Quickbit and Debenhams (no longer listed) as well as CD Projekt Red (no data currently available pre-2021). This left 45 events going forwards for analysis (see **Table 9**).

4.4. Data analysis

To facilitate the analyses, R (R Core Team 2018)[31] scripts were developed to extract share price and index data directly from Yahoo!Finance for each of the 45 events and then event studies run via an R package (Schimmer, Levchenko & Müller, 2014) [32] using the market model as described above[33]. Non-trading event days were defaulted to the next available trading day. An estimation window of 120 days was chosen consistent with e.g. Goel and Shawky (2009), Andoh-Baidoo et al. (2010), Schatz and Bashroush (2016a), Richardson et al. (2019). In all cases the estimation window ended one trading day before the event window. Tweneboah-Kodua et al. (2018: 641) recommend avoiding overlap of the estimation and event windows in this way to avoid "*parameter contamination*". Although the event window should be broad enough to contain any uncertainty in the date of the event, the longer the window the less likely it is to detect abnormal returns (Dyckman, Philbrick & Stephan 1984). Previous studies have shown market reaction before the event date due to information leakage. For example, using event study techniques, Lin et al. (2020) show significant evidence of opportunistic pre-official announcement insider trading related to data breaches. For this study, a range of event windows were initially chosen starting from up to two days before the event and varying in length from two to

[31] R version 4.0.3 (2020-10-10)
[32] EventStudy package version 0.36.900 (API version 0.374-alpha)
[33] The R code used is included in the Appendix.

fifty[34] trading days in order to give visibility of these effects and others such as sector specific effects reported by e.g. Tweneboah-Kodua et al. (2018).

4.5. Hypothesis development

For hypothesis development, please refer to Chapter 3.

4.6. Results and discussion

To identify any significant CAR (RQ1) an initial visualisation similar to **Figure 13** showed that Travelex was a major outlier (having a CAR of -75% over a 3-day window) and would fall into the category which Richardson et al. (2019: 248) describe as *"those rare situations involving massive data exposures"*. The company has since gone into administration citing both the cyber-attack and COVID-19 effects as contributing factors (The Guardian, 2020). Since this event occurred on 31/12/2019, it was at the limit of the data selection range and the event window would certainly extend into potential 'COVID-19 territory'. Therefore, Travelex was excluded from further analyses leaving 44 breach events remaining. These residual events were re-visualised as shown in **Figure 13**. No obvious evidence of information leakage prior to the announcement date (e.g. Lin et al. 2020) was observed with, in fact, slightly positive CAAR being observed for event windows (-2, 2), (-1, 1) and (-1, 0). Studies where there was uncertainty around the announcement date favoured event windows such as these to ensure abnormal returns were not missed (e.g. Schatz & Bashroush 2016a) but here all dates were validated. Nevertheless, the expectation based on previous studies would be to see market reaction kicking in 1-2 days after the event, growing to a maximum and disappearing over longer event windows. What can be seen here is that the market reaction appears to be much slower overall with no visible negative trend until the (0, 5) window at the earliest, disappearing the following day and subsequently reappearing a month after the event (0, 20). These longer windows operating at the outer limits of event study methodology also seem to be skewed by outliers such as NatWest enjoying, surprisingly, a positive run following their breach announcement and Fox-IT along with The AA falling over 20%. That Fox-IT, a cyber security specialist company itself, suffered such a negative market reaction would certainly come as no surprise, albeit seemingly somewhat late. Clearly there is a need to look more deeply here into the nature of the businesses affected

[34] The limit of this software for CAR event windows was 50 days. For longer windows the buy-and-hold abnormal return (BHAR) approach is recommended.

(beginning to answer RQ2) as recommended by Tweneboah-Kodua et al. (2018) and Bendovschi et al. (2016).

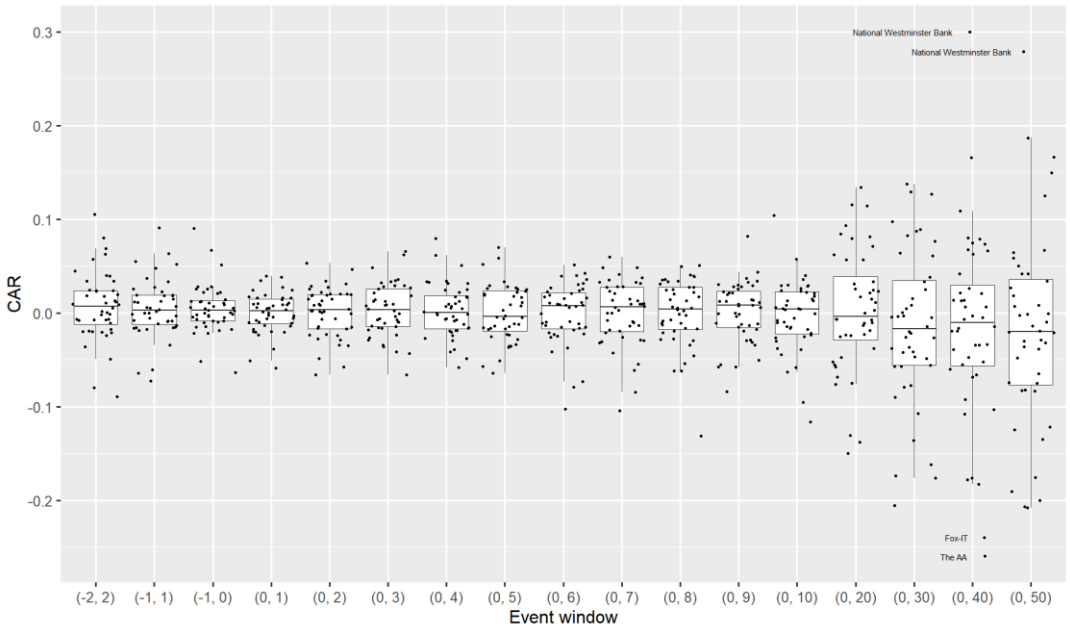


Figure 13: Boxplots of CAR values per event window

For this purpose, a graph of CAAR by sector for each event window is shown in **Figure 14**. It appears that the fastest negative reaction to a breach is, indeed, shown by the technology sector peaking two days after the disclosure in the short term. Financial and communication services companies show little reaction at all over this time period which is somewhat unexpected based on previous studies. The basic materials sector shows the largest short term negative impact after five days, although it should be noted that there was only one company (Norsk Hydro) assigned to this sector, so it made sense to choose the (0, 2) window for a closer look at sector performance. The results are shown in **Table 4**.

Table 4: Analysis of event window (0, 2) by sector

Industry Sector	N	CAAR	S_{CAAR}	t_{CAAR}	Negative CAR %	Total Records Breached	Personal %	GDPR %
Technology	4	-0.0188	0.0390	-0.9656	50	-	75	75
Financial Services	11	-0.0036	0.0250	-0.4730	55	1,360,584,255	100	55
Communication Services	8	-0.0015	0.0193	-0.2124	50	3,117,453	88	75
Industrials	8	-0.0007	0.0345	-0.0574	38	404,700	50	75
Basic Materials	1	0.0098			0	-	0	100
Consumer Cyclical	8	0.0124	0.0250	1.4021	25	358,000	88	75
Consumer Defensive	3	0.0158	0.0034	7.9690**	0	17,295	33	67
Healthcare	1	0.0190			0	-	0	100
	44	**0.0010**			**39**	**1,364,481,703**	**75**	**70**

*,**,*** Indicate statistical significance at the 10%, 5% and 1% levels respectively.

Although 4 sectors show negative CAAR for this event window, the average impact is still slightly positive (0.001) over all 44 events and the negative CAARs are not statistically significant, thus the null hypothesis cannot be rejected for these. However, the CAAR is significant at the 5% level for the consumer defensive sector but in a positive way with the share prices rising over 1% in response to the breach disclosure. Companies in this sector, however, could reasonably be expected to outperform under adversity due to their defensive nature.

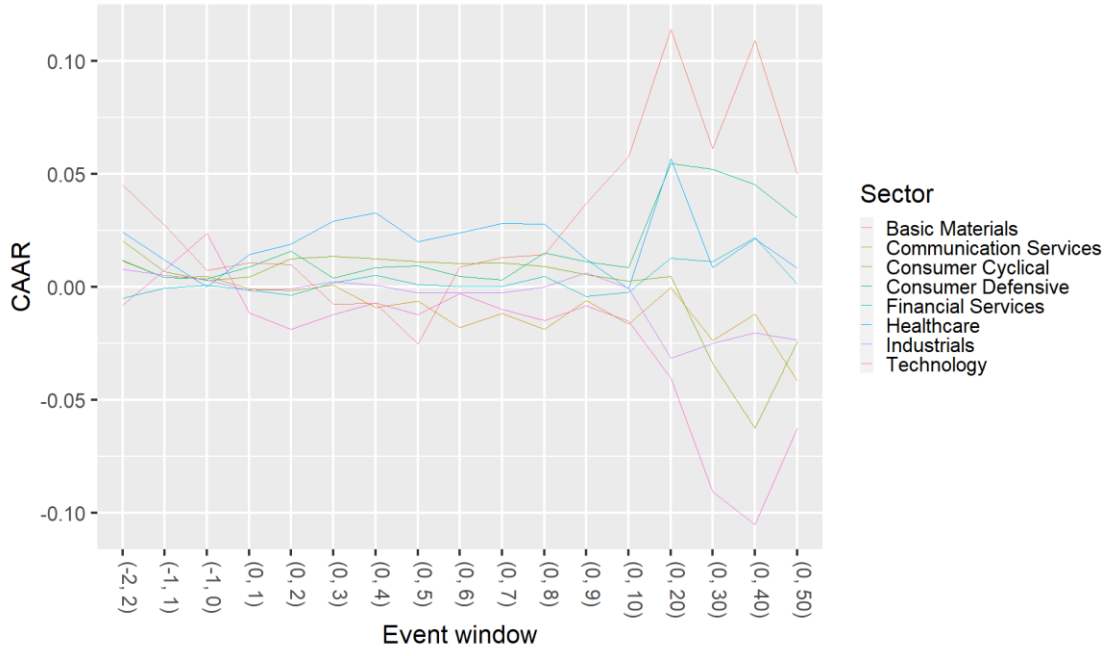


Figure 14: CAAR by industry sector

Not having found evidence of negative impact so far, the observation of Richardson et al. (2019) regarding massive breach volumes leading to more serious effects warrants

investigation. Where it was possible to gather an indication of the number of records breached, this information was added to the dataset (25 examples). It can be seen that the financial services sector was responsible for over 99% of all the records breached, in all cases involving sensitive (personal) data and the majority (55%) being GDPR relevant, thus it seems somewhat surprising the market reaction is not more severe.

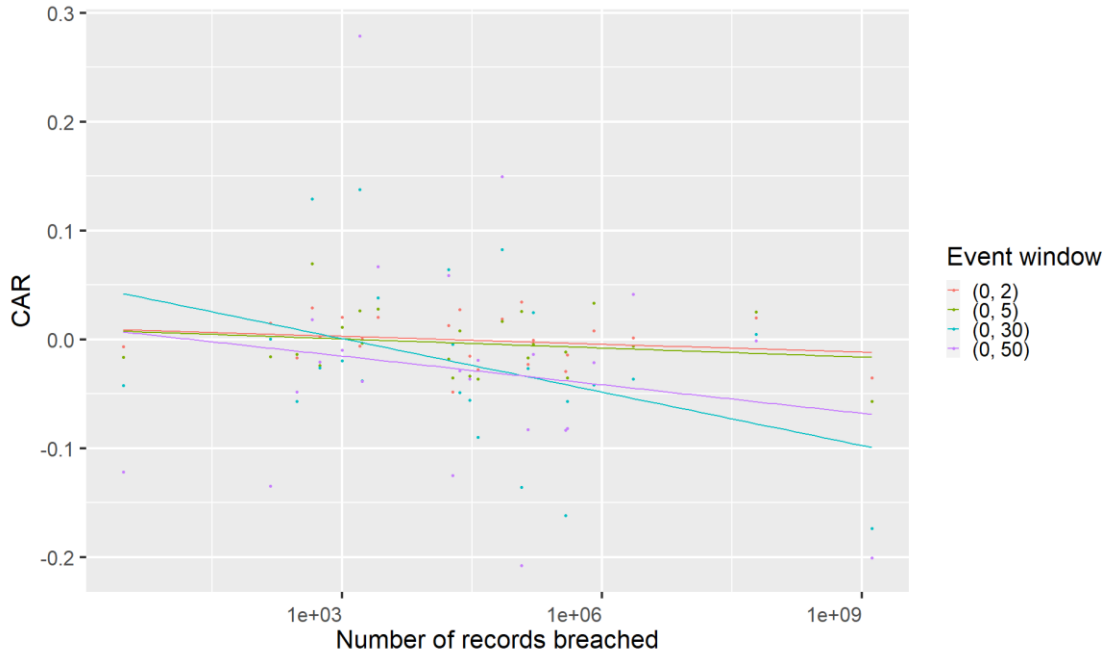


Figure 15: CAR versus records breached

An idea of a correlation between the number of records breached (logarithmic scale) and CAR for a selection of the more interesting event windows is shown in **Figure 15**. There appears to be a weak trend that CARs become more negative the more records there are breached which becomes stronger with the longer windows. As event studies are better suited to the days immediately surrounding, it was decided to focus on other, more major factors. Campbell et al. (2003) noted that breaches involving sensitive personal data led to more negative CARs. For this reason, the sector analysis in **Table 4** was rerun restricting the dataset to only those events involving sensitive (personal) data and the results are shown in **Table 5**. This had the effect of altering the mean CAAR from a slightly positive value to a slightly negative value (0.001 to -0.001). Despite the reduction in events to 33, the financial services sector was unaffected meaning all 11 events involved sensitive data. The technology sector became 37% more negative yet the results are still not statistically significant for any sector. Thus, the null hypothesis of zero abnormal returns still stands.

Table 5: Analysis of event window (0, 2) by sector (personal data)

Industry Sector	N	CAAR	S_{CAAR}	t_{CAAR}	Negative CAR %
Technology	3	-0.0259	0.0446	-1.0040	67
Communication Services	7	-0.0061	0.0152	-1.0623	57
Financial Services	11	-0.0036	0.0250	-0.4730	55
Industrials	4	0.0063	0.0266	0.4725	25
Consumer Defensive	1	0.0130			0
Consumer Cyclical	7	0.0138	0.0266	1.3720	29
	33	**-0.0008**			**45**

,**, Indicate statistical significance at the 10%, 5% and 1% levels respectively.**

Goel and Shawky (2014) observe that the introduction of data breach notification laws led to a reduction in negative market reaction. For this purpose, **Table 6** shows an analysis of abnormal returns for four particularly negative event windows, both before and after the enactment of the GDPR, for the above set of 33 events specifically involving personal data. In three of the four cases, pre-GDPR negative CAAR was turned positive after enactment and, even in the fourth case, the negative CAAR was reduced over 90%. Unfortunately, the results were only statistically significant (at the 10% level) for the longer event windows pre-GDPR and these longer-term event study observations are known to be less reliable.

Table 6: Market effect of GDPR enactment

Event Window	GDPR	N	CAAR	S_{CAAR}	t_{CAAR}	Negative CAR %
(0, 2)	PRE	12	-0.0079	0.0267	-1.0260	50
(0, 2)	POST	21	0.0033	0.0257	0.5953	43
(0, 5)	PRE	12	-0.0114	0.0281	-1.4023	75
(0, 5)	POST	21	0.0039	0.0303	0.5946	48
(0, 30)	PRE	12	-0.0564	0.0916	-2.1330*	83
(0, 30)	POST	21	-0.0047	0.0791	-0.2702	57
(0, 50)	PRE	12	-0.0592	0.1101	-1.8645*	83
(0, 50)	POST	21	0.0022	0.1135	0.0881	62

,**, Indicate statistical significance at the 10%, 5% and 1% levels respectively.**

Finally, an analysis by market reference was carried out (**Table 8**) to give, effectively, a geographic breakdown and see which markets were more sensitive to data breach announcements. Of the shorter event windows, (0, 1) proved to be of particular interest as this was the first real evidence of a statistically significant abnormal return (at the 5% level), specifically related to the Spanish market (IBEX35). Although there were only

four breach events relevant to this market, they spanned three different industry sectors, three out of four were GDPR relevant, and half and half sensitive versus non-sensitive data therefore, it seems, the market itself was the common factor here. One of these breaches was, however, by far the largest (TSB/Sabadell) so volume could have played a part. The most negative impact for this 2-day window was that of the AEX25 (Netherlands) at around -3.8%, but there was only one example here (ING Bank). It is interesting to note that the FTSE350[35] index would effectively cover 17 (39%) of the events so there was a strong UK bias here. As an additional check on the importance of the reference index (Kannan et al. 2007; Richardson et al. 2019), the sector analysis (**Table 4**) was rerun using the SPEUR350[36] as a reference across all events. The resulting abnormal returns are shown in **Table 7**.

Table 7: Analysis of event window (0, 2) by sector (SPEUR350)

Industry Sector	N	CAAR	S_{CAAR}	t_{CAAR}	Negative CAR %
Technology	4	-0.0213	0.0356	-1.1955	50
Financial Services	11	-0.0051	0.0241	-0.7056	45
Communication Services	8	-0.0003	0.0199	-0.0445	38
Industrials	8	0.0000	0.0369	-0.0010	38
Consumer Cyclical	8	0.0107	0.0243	1.2415	38
Consumer Defensive	3	0.0165	0.0080	3.5875*	0
Basic Materials	1	0.0179			0
Healthcare	1	0.0248			0
	44	**0.0008**			**36**

,**, Indicate statistical significance at the 10%, 5% and 1% levels respectively.**

The overall mean CAAR only differs by 0.0002 and the results look very similar, with again only the consumer defensive sector showing statistical significance but this time only at the 10% level. Using market specific indices produced higher t_{CAR} values on average so this was the preferred method (cf. Bose & Leung 2014).

[35] The FTSE100 and FTSE250 combined.
[36] Standard & Poor's index of 350 stocks representative of European markets in general

Table 8:Analysis by market index for event window (0, 1)

Reference Market	Country	N	CAAR	S_{CAAR}	t_{CAAR}	Negative CAR %	Total Records Breached	Personal %	GDPR %
AEX25	NL	1	-0.0383			100	19,055	100	100
OMXH25	FI	1	-0.0237			100	-	100	100
FTSEMIB	IT	1	-0.0207			100	400,000	100	0
ATXPRIME	AT	1	-0.0200			100	-	0	100
IBEX35	ES	4	-0.0098	0.0038	-5.1543**	100	1,300,000,000	50	75
FTSE250	GB	7	-0.0067	0.0366	-0.4858	43	240,000	86	57
CAC40	FR	5	0.0005	0.0163	0.0631	60	181,300	100	80
FTSE100	GB	10	0.0053	0.0136	1.2364	30	398,753	80	70
ISEQ20	IE	2	0.0056	0.0063	1.2472	0	845	50	50
DAX30	DE	7	0.0069	0.0205	0.8980	57	2,440,750	71	86
MOEX50	RU	1	0.0090			0	60,000,000	100	100
OSEAX†	NO	1	0.0107			0	-	0	100
SMI20	CH	1	0.0161			0	800,000	100	0
OMXCBGI‡	DK	2	0.0219	0.0184	1.6846	0	1,000	50	50
		44	**0.0001**			**48**	**1,364,481,703**	**75**	**70**

*,**,*** Indicate statistical significance at the 10%, 5% and 1% levels respectively.
† OBX25 not available
‡ OMXC25 not available

4.7. Conclusion

Overall, no clear impact has been seen on share price of data breach announcements (RQ1) in European companies across all sectors and markets other than Spain. Based on this evidence it is difficult to support business cases for investment in cyber security measures (RQ4), although there could be other approaches as Deane et al. (2019) report a significant uplift in share price for organisations following an announcement related to security certification. Thus, justification for investment would have to depend on other factors such as risk appetite (no company wants to be the next Travelex), industry sector, nature of the data compromised and relevant legislation. These findings are consistent with Richardson et al. (2019) who refer to their observations on the (lack of) economic impact of data breach announcements as "*much ado about nothing*" yet other, mostly earlier, US based research in this area did find significant evidence of negative market reaction supporting the finding of i.a. Yayla and Hu (2011) that markets were becoming less sensitive to data breach disclosure over time. That said, the Spanish market (RQ2) showed statistically significant and rapid sensitivity to data breach announcements, continuing after the enactment of GDPR. Other European markets showed a slight reduction in negative CAR post-GDPR as predicted by Goel and Shawky (2014) but, again, not statistically significant. At the time of writing the Spanish data protection authority (AEPD) has issued more GDPR infringement fines (236 examples) than any

other (CMS Legal, 2021) so perhaps this is a contributing factor to the higher market sensitivity towards data breaches in Spain.

Some differences with US markets were identified, for example, the slower response of the European financial services sector (RQ2). The specific case of Travelex also fits with the observations of Richardson et al. (2019) that in the case of particularly severe breaches, the situation may become irrecoverable, although COVID-19 was cited as a contributing factor in its demise. Following on from this, some evidence of a (weak) correlation between negative CAR and number of records breached was identified, as reported by i.a. Chen at al. (2012), but not really in the short term. Nevertheless, this should be borne in mind for any risk assessment along with the nature of the data itself.

One shortcoming identified as part of this research was the lack of a publicly available breach database like, for example, PRC which features heavily in similar US based studies. Although the VCDB project seems well-intentioned as a global research resource, what is really needed is a much more comprehensive and richer dataset in order to study European and other markets to a depth equivalent to that of US research in this area. Although this study has begun to look at the economic impact of GDPR this is another potential area for future research once the market stabilises and more data becomes available. It must be recognised that these disclosure events are early in the cyber security incident lifecycle and, although appearing no more than a nuisance to the markets generally, there may well be more surprises to follow depending on how effectively they are managed.

The next chapter examines the impact of announcements related to GDPR infringement fines which one would naturally also expect to be of the unfavourable information security event type. This is a logical progression as announcements of breaches in this chapter which are data privacy relevant may well indeed be followed up with such punitive action by DPAs.

Table 9: List of Data Breaches

Company	Symbol	Event Date	Index	Confounding Events	Number of Records Breached	Personal Data?	VCDB Identifier	Comments
Airbus	AIR.PA	30.01.2019	CAC40	NA	NA	TRUE		
Allianz	ALV.DE	27.12.2019	DAX30	NA	160,000	TRUE		
Allied Irish Banks	A5G.IR	12.09.2017	ISEQ20	NA	550	TRUE	47cb7da0-1122-11ea-bd5f-71a8176be0a0	
Avast	AVST.L	21.10.2019	FTSE100	NA	NA	FALSE	5822eae0-ffce-11e9-81dd-7371e517d223	
Axa Insurance	CS.PA	25.02.2019	CAC40	NA	2,600	TRUE	2feb8a20-936b-11e9-96b8-11c597846497	
B&Q	KGF.L	25.01.2019	FTSE250	NA	70,000	TRUE		
Barclays	BARC.L	27.07.2017	FTSE100	NA	NA	TRUE	249c8a80-7c56-11e7-8ec0-e369e9638092	
Bayer	BAYN.DE	04.04.2019	DAX30	NA	NA	FALSE	af9a1ac0-c4f3-11e9-bbcb-6517d645041f	
BMW	BMW.DE	06.12.2019	DAX30	NA	NA	FALSE		
British Airways	IAG.L	06.09.2018	FTSE100	NA	380,000	TRUE		
BT Security	BT-A.L	12.11.2019	FTSE100	NA	150	TRUE		
Cadena SES	PRS.MC	04.11.2019	IBEX35	NA	NA	FALSE		Not a component of IBEX35.
Capita	CPI.L	12.07.2018	FTSE250	NA	23,000	TRUE	ffca5f10-8ab0-11e8-874b-dd84fb61b260	

Computacenter	CCC.L	22.05.2019	FTSE250	NA	NA	TRUE		
Costa Coffee	WTB.L	02.07.2018	FTSE100	NA	NA	TRUE		
Deutsche Bank	DBK.DE	28.07.2019	DAX30	NA	450	TRUE	117a9da0-c4f9-11e9-8a5a-abdcbb867af8	
Eir	ILD.PA	22.08.2018	CAC40	NA	37,000	TRUE		CACMID60 not available.
Fox-IT	NCC.L	14.12.2017	FTSE250	NA	NA	TRUE		
Gekko Group	AC.PA	20.11.2019	CAC40	NA	140,000	TRUE		
HSBC	HSBA.L	05.06.2017	FTSE100	NA	NA	TRUE		
ING Bank	INGA.AS	02.03.2019	AEX25	NA	19,055	TRUE	02a64550-0a3f-11ea-9821-617d71ee6253	
James Fisher	FSJ.L	05.11.2019	FTSE250	NA	NA	FALSE		
Jewson	SGO.PA	14.11.2017	CAC40	NA	1,700	TRUE	8bd20230-d91b-11e7-be42-df02fe7390b5	
Kerry Group	KRZ.IR	11.07.2018	ISEQ20	NA	295	FALSE	f1178b00-972b-11e8-8342-0f8f10fcd812	
Maersk	MAERSK-A.CO	27.06.2017	OMXCBGI	NA	NA	FALSE		OMXC25 not available.
Mercedes Benz	DAI.DE	19.10.2019	DAX30	NA	NA	TRUE		
National Westminster Bank	NWG.L	19.08.2019	FTSE100	NA	1,600	TRUE	cef428d0-cfec-11e9-aa40-01794b8efffe	
Nokia	NOKIA.HE	21.03.2019	OMXH25	NA	NA	TRUE	c1a7f4b0-023a-11ea-ad43-7952870528d4	
Norsk Hydro	NHY.OL	19.03.2019	OSEAX	NA	NA	FALSE		OBX not available.

Porr	POS.VI	02.05.2019	ATXPRIME	NA	NA	FALSE		Not a component of ATX.
Prosegur	PSG.MC	27.11.2019	IBEX35	NA	NA	FALSE		Not a component of IBEX35.
Sberbank	SBER.ME	03.10.2019	MOEX50	NA	60,000,000	TRUE	ca136300-06ef-11ea-8714-b997dc99273b	
Sports Direct	FRAS.L	08.02.2017	FTSE250	NA	30,000	TRUE		
Swisscom	SCMN.SW	07.02.2018	SMI20	NA	800,000	TRUE		
T-Mobile	DTE.DE	20.08.2018	DAX30	NA	2280,000	TRUE		
T-Mobile	DTE.DE	11.10.2017	DAX30	NA	300	TRUE		
Telefonica	TEF.MC	17.07.2018	IBEX35	NA	NA	TRUE		
Tesco	TSCO.L	20.09.2019	FTSE100	NA	NA	FALSE		
Tesco Travelex	TSCO.L	13.03.2018	FTSE100	NA	17,000	TRUE		
The AA	AA.L	03.07.2017	FTSE250	NA	117,000	TRUE		
Tivoli	TIV.CO	18.08.2019	OMXCBGI	NA	1,000	TRUE		Tivoli not included in C20. OMXC25 not available.
Travelex	FIN.L	31.12.2019	FTSE250	NA	NA	FALSE		Major outlier – removed from analysis.
TSB	SAB.MC	23.04.2018	IBEX35	NA	1,300,000,000	TRUE		
UniCredit SpA	UCG.MI	26.07.2017	FTSEMIB	NA	400,000	TRUE	ce142850-093e-11e8-bbae-ed39c73f1398	Could use EWI instead.
Vodafone	VOD.L	25.09.2019	FTSE100	NA	3	TRUE		

Chapter 5. The Impact of GDPR Infringement Fines on the Market Value of Firms

5.1. Introduction

The European Union Agency for Cybersecurity (ENISA, 2020) reported a "*54% increase in the total number of* [data] *breaches by midyear 2019 compared with 2018*". ENISA (2020) also remarks that "*55% of the responders to a Eurobarometer survey responded that they are concern*[sic] *about their data being accessed by criminals and fraudsters*". Clearly, there is major concern out there in the fields of data security and privacy. Such was the driver for the General Data Protection Regulation (GDPR) which, as Proton AG put it, *"is the toughest privacy and security law in the world. Though it was drafted and passed by the European Union (EU), it imposes obligations onto organizations anywhere, so long as they target or collect data related to people in the EU. The regulation was put into effect on May 25, 2018.",* going on to add that *"the GDPR will levy harsh fines against those who violate its privacy and security standards, with penalties reaching into the tens of millions of euros"* (Proton AG, 2023).

A primary objective of the GDPR is to protect "*fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data*" (Data Protection Act, 2018). The requirement therein, to notify data breaches to the relevant supervisory authority within 72 hours of becoming aware (where feasible), could reasonably be expected to increase visibility of non-compliance. For example, in the UK, before the introduction of the GDPR as the Data Protection Act (2018), the preceding Data Protection Act (1998), according to the Information Commissioner's Office (ICO)[37], stated "*although there is no legal obligation on data controllers to report breaches of security which result in loss, release or corruption of personal data, the Information Commissioner believes serious breaches should be brought to the attention of his Office*". Prior to 2010, the ICO was limited to serving enforcement notices for contraventions of the DPA (1998), however in April 2010 the ICO was granted the power to issue fines of up to £500,000 on its own authority. For example, Sony Computer Entertainment Europe were fined £250,000 in January 2013 for a "*serious breach*" when their PlayStation Network was hacked (BBC, 2013) and in 2016, TalkTalk were fined £400,000 for leaking personal data of almost 157,000 customers due to poor website security (BBC, 2016). Serious infringements under the GDPR, those violating the fundamental principles of the

[37] The supervisory (data protection) authority of the UK (https://ico.org.uk. Accessed on: 04/04/23)

right to privacy and the right to be forgotten, could result in a fine of up to €20 million or 4% of the firm's worldwide annual revenue from the preceding financial year (whichever amount is higher), a clear deterrent against carelessness concerning data privacy and security. Indeed, total fines issued by data protection authorities since the introduction of the GDPR currently stand at over €275m (CMS Legal, 2021).

This research is concerned with the impact the announcement of these GDPR fines has on the market value of publicly listed companies. Spanos and Angelis (2016) report that data breach announcements are associated with a negative impact on market value. Could it be that, since the introduction of the GDPR, a firm's share price may suffer a 'double whammy' of both initial breach notification and subsequent punitive action? This chapter aims to assess the economic impact of the introduction of the GDPR on publicly listed companies through the application of fiscal penalties levied by its supervisory authorities on those firms which have suffered a data privacy breach. Such an understanding would better inform the cyber security investment strategies of companies. To achieve this objective, the research questions as detailed in Chapter 1 were considered.

This research would be of benefit to business management, practitioners of cyber security, investors and shareholders as well as researchers in cyber security or related fields. It could also be of value to data protection authorities to increase their understanding of the impact and enforcement of legislation on the economy. Another benefit of this study would be the European focus, thereby beginning to offset the strong US bias of the existing literature in this area.

The initial literature review (Chapter 2) identified an SLR concerning the impact of data breach events on the stock market carried out by Spanos and Angelis (2016), who report that, although research in this area was "*quite limited*", the majority of studies (76%) found a statistically significant negative impact. For example, Lin et al. (2020) report a loss of 1.44% on average over a 5-day window. Andoh-Baidoo, Amoako-Gyampah and Osei-Bryson (2010) report -3.18% abnormal returns over a 3-day period. Cavusoglu, Mishra and Raghunathan (2004) cite -2.1% on average within two days after the announcement. Goel and Shawky (2009) quote -1% in the days surrounding the event. These studies also note some correlations between these negative returns and, for example, industry sector. Tweneboah-Kodua et al. (2018), warn that "*studying the cumulative effects of cyberattacks on prices of listed firms without grouping them into the various sectors may be non-informative*". They noted that financial services firms reacted

more rapidly and more significantly than those in the technology sector. It was also observed by Campbell et al. (2003) that those breaches involving unauthorised access to confidential data were more likely to result in significant negative market reaction, which one would reasonably expect to apply across the board for this study.

Such observations would support the idea of governments introducing legislation to not only counter this negative economic impact but also to help protect data subjects who are effectively innocent victims of such breaches of confidentiality. Indeed, the right to privacy is a component of the European Convention on Human Rights (1950) and the EU has sought to protect this right through legislation ever since with, firstly, the introduction of the European Data Protection Directive (1995), then the Privacy and Electronic Communications Directive (2002) and, in response to ever-evolving technology and increases in data transfers, the GDPR in 2018 along with the (delayed) ePrivacy Regulation due to repeal the 2002 Directive (European Commission, 2021).

These technological advances are reflected in the extended definition of 'personal data' in the GDPR which includes, for example, geographic location and 'online identifiers' such as website tracking cookies which are not stipulated in the DPA (2018). Apart from the hefty monetary penalties for failure to comply and the obligation to notify breaches already highlighted above, another key difference between the GDPR and the DPA (1998) is the concept of 'consent'. Whereas the DPA (1998) deemed a simple 'opt out' of marketing communications adequate, the GDPR has more specific and granular requirements in this area. Under the GDPR, organisations must obtain specific 'opt ins' to marketing communications via different channels (email, telephone etc.) and offer individuals the ability to withdraw that consent at any time. A full and detailed comparison between the two regulations is out of scope here, but hopefully this gives a flavour of how data privacy legislation in the UK (and EU) has evolved over the years.

The relatively recent introduction of the GDPR naturally limits the availability of research on its impact, so it is necessary to look elsewhere. The introduction of data breach notification laws in the US was found to reduce identity theft by over 6% on average (Romanosky, Telang & Acquisti, 2011). Clearly, if data subjects are rapidly made aware that their personal data has been compromised, and which data, they should be better positioned to take preventative action. There are already, however, some criticisms of the effectiveness of the GPDR in this area as notification to data subjects is only required in certain *"high risk"* cases and where it would not place too onerous a burden on the

reporting organisation (Nieuwesteeg & Faure, 2018). Data breach notification laws have been widely adopted in the US, albeit not centrally – federal law in this area only covers certain specific sectors. Nevertheless, 47 jurisdictions have implemented their own notification legislation. In fact, the US could be considered an early adopter. By contrast the EU GDPR model is central and adopted by member states and includes the notification requirement within the data protection law itself unlike, for example, Australia (Daly, 2018) where a separate law was introduced in early 2017. Goel and Shawky (2014) carried out a US based study examining the impact of data breach announcements on share price and found a significant reduction in negative returns after the enactment of both federal and state laws. The continuing introduction of such legislation could explain why Yayla and Hu (2011) observed a general trend of reduction in the market impact of information security related events over time. Murciano-Goroff (2019) researched Californian company investment in web server security following the introduction of state data breach notification law yet only noted a modest effect with server software being, at most, 2.8% newer. Indeed, Richardson et al. (2019) argue that "*companies are unlikely to change their investment patterns unless the cost of breaches increases dramatically or regulatory bodies enforce change*" underpinning the need for an understanding of the impact and effectiveness of the GDPR on cyber security investment – an area which this research aims to inform as well as bringing an EU specific perspective to offset the strong US bias of previous studies. This US bias was also observed by Ali et al. (2021) who revisited and expanded the work of Spanos and Angelis (2016), reporting that 76% of papers reviewed were based solely on US data although note a growth in non-US based studies (up to 40%) since 2017. They attribute this to the increasing adoption of regulation outside the US for disclosure of cyber security events to investors, the GDPR being an example of this, at least in those cases involving personal data. Lack of disclosure would, naturally, result in lack of breach data as highlighted in the previous chapter.

5.2. Methodology

The high-level approach to this research was to download a list of publicly announced GDPR infringement fines from the Enforcement Tracker (CMS Legal, 2021), filter this dataset for those cases involving publicly listed companies and analyse the impact of these announcements on share price using ESM as elucidated in Chapter 3.

5.3. Data collection

The base dataset used to identify fine announcements was from the GDPR Enforcement Tracker (CMS Legal, 2021). Although not professing to be an exhaustive list, the initial data download resulted in 277 records. Manually filtering these records for those involving publicly listed companies (or a subsidiary of a publicly listed company[38]) resulted in 71 rows. Some announcement dates were found to be missing and were instead found from press reports and official data protection authority publications where applicable. It was necessary to exclude certain records due to a missing date such as Facebook (Germany) and Unicredit (Czech Republic/Slovakia). Events on the same day were consolidated into one e.g. Eni Gas e Luca, EDP Spain. Entries which had potentially overlapping (confounding) event windows were also filtered e.g. Vodafone (2 events). Share price and market index data were extracted from Yahoo!Finance (2019) along with firm demographics such as annual revenue, market capitalisation and industry sector. Information was not available for all the events on Yahoo!Finance e.g. Louis Group (Cyprus), Xfera (now privately owned) and Avon Cosmetics (event was pre-public), thus these events had to be filtered out also, leaving 48 records. The most appropriate market index was chosen as a reference in each case (Kannan et al. (2007) highlighted the importance of the market reference), ideally one which included the candidate company itself but adjusted, if needed, due to lack of availability of data in Yahoo!Finance. Some firms had multiple listings, in which case the primary listing and associated index were used. The date range was limited, naturally, from the earliest fine since the introduction of the GDPR in 2018 (actually, January 2019) until the date of download but it was decided to cap the data at 31/12/2019 in order to avoid market uncertainties due to COVID-19, that being a long-term confounding event in itself[39] – for example, He at al. (2020) report on the impact of COVID-19 on Chinese markets in general using event study techniques citing the closure of Wuhan in January 2020 as the start of the outbreak with Alam et al. (2020) making similar observations on the Australian stock market commencing February 2020 through a similar approach[40]. This COVID-19 date capping reduced the dataset from 48 to 25 events going forwards for analysis (see **Table 16**). Internet searches were carried out for other confounding events near the announcement date such as financial results, dividends, changes of CEO or CFO and mergers or

[38] Ultimate parent companies were identified from Dun & Bradstreet (https://www.dnb.com. Accessed on: 04/04/23)

[39] Examples of other such long-term confounding events are the dot-com era and the 9/11 attacks (Kannan, Rees and Sridhar, 2007)

[40] Interestingly, both COVID-19 event study papers yet again favour the single-factor model as used here.

acquisitions, consistent with prior studies of this type (i.a. Garg et al., 2003; Modi et al., 2015). It was noted that the nearest confounding events were at least four trading days away from the announcement (two examples, Telefónica and Vodafone), yet on average there was a gap of 16 days.

5.4. Data analysis

To facilitate the analyses, R (R Core Team, 2018)[41] scripts were developed to pull share price and index data directly from Yahoo!Finance for each data record and then event studies run via an R package (Schimmer, Levchenko & Müller, 2014)[42] using the market model as described above[43]. Non-trading event days were defaulted to the next available trading day. An estimation window of 120 days was chosen consistent with e.g. Goel and Shawky (2009), Andoh-Baidoo et al. (2010), Schatz and Bashroush (2016a), Richardson et al. (2019). In all cases the estimation window ended one trading day before the event window. Tweneboah-Kodua et al. (2018) recommend avoiding overlap of the estimation and event windows in this way to avoid "*parameter contamination*". Although the event window should be broad enough to contain any uncertainty in the date of the event, the longer the window the less likely it is to detect abnormal returns (Dyckman et al., 1984). Previous studies have shown market reaction before the event date due to information leakage. For example, using event study techniques, Lin et al. (2020) show significant evidence of opportunistic pre-official announcement insider trading related to data breaches. For this study, a range of event windows was initially chosen starting from up to two days before the event and varying in length from 2 up to 20 trading days to give visibility of these effects and others such as sector specific effects reported by e.g. Tweneboah-Kodua et al. (2018) who observed more rapid response from the financial services sector, for instance.

5.5. Hypothesis development

For hypothesis development, please refer to Chapter 3.

5.6. Results and discussion

Event studies were carried out as described above for 10 event windows of varying length across all 25 GDPR fine events. A visualisation of the overall results is shown in **Figure 16**. It appears at first glance that the most negative impact is seen around the 4-day event window (0, 3) with the market value gradually recovering over longer windows and

[41] R version 4.0.3 (2020-10-10)
[42] EventStudy package version 0.36.900 (API version 0.374-alpha)
[43] The R code used is included in the Appendix.

beginning to see positive recovery 10 days after the event. After 20 days, for IAG (Vueling) and EDF (Madrileña Red de Gas) the abnormal returns had grown to over 10% either way yet the median CAR remained much closer to zero.

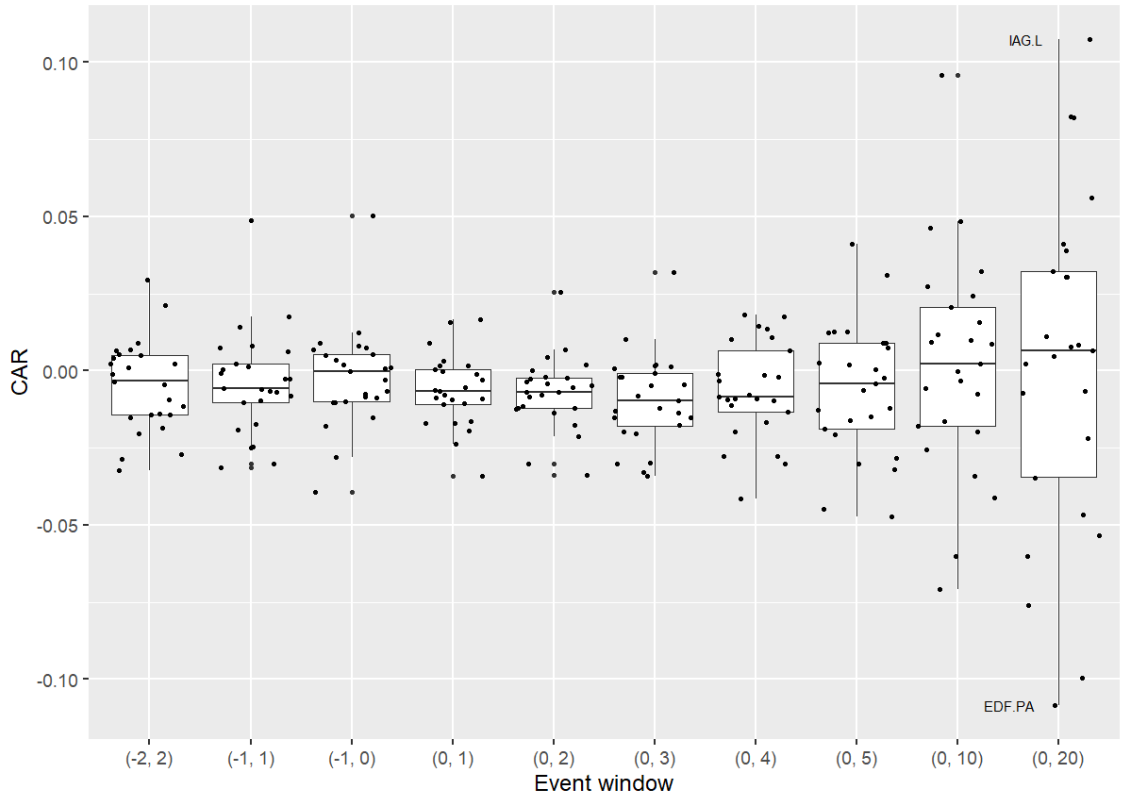


Figure 16: Comparison of event windows

A CAAR was calculated for multiple firms across each window and is shown in **Table 10**. Here the 3 and 4-day event windows (0, 2), (0, 3) show the most negative abnormal returns and are statistically significant at the 1% level. It is interesting to note that the null hypothesis cannot be rejected for the three earlier event windows involving pre-event days, thereby indicating no information leakage prior to the fine announcements and consistent with the lack of uncertainty in the event dates for this exercise. As above, there is also lack of statistical significance for the longer windows indicative of a tendency of market recovery towards zero abnormal returns over time as reported by Dyckman et al. (1984). The event window (0, 3) showed the most negative (almost 1%) CAAR, consistent with the findings of Goel and Shawky (2009). Within this window, 19 of the 25 events (76%) had abnormal returns of less than zero, therefore this window was chosen as the basis for further analyses. Usage of this event window (0, 3) has been previously reported in studies of this type e.g. Hinz et al. (2015), Rosati et al. (2019), although the majority tend to see a slightly faster market reaction (Ali et al. 2021) indicating perhaps

less information salience here (e.g. Ramos, Latoeiro & Veiga, 2020). Another benefit of this choice of event window was the nearest confounding events (Telefónica and Vodafone) fell on day 4, just outside this window and other studies of this type (e.g. Deane et al., 2019) exclude only confounding events falling within the event window itself. In fact, Goel and Shawky (2014) comment that the shorter the event window, the less chance there is of finding a confounding event and, for a larger sample size, did not filter for confounding events at all.

Table 10: CAAR by event window

Event Window	N	CAAR	t_{CAAR}	% Negative CAR
(-2, 2)	25	-0.0049	-1.6188	56
(-1, 1)	25	-0.0041	-1.2112	64
(-1, 0)	25	-0.0022	-0.6746	52
(0, 1)	25	-0.0064	-2.7453**	72
(0, 2)	25	-0.0072	-3.0748***	80
(0, 3)	25	-0.0096	-3.2341***	76
(0, 4)	25	-0.0064	-2.0190*	72
(0, 5)	25	-0.0061	-1.4128	56
(0, 10)	25	0.0020	0.2795	48
(0, 20)	25	0.0011	0.0968	40
	250	**-0.0044**		**62**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

An analysis by ultimate parent company of CAAR is shown in **Table 11**. It can be seen that four firms suffered more than one fine under GDPR, but no more than two during the date range of this study. The firm suffering the most negative abnormal return is listed first and the most positive last. The overall average fine levied was found to be almost €17m and it appears that the supervisory authorities have been relatively lenient so far with the average penalty sitting at around 0.15% of previous year's annual revenue (the greatest being just over 1%) and nowhere near the possible maximum of 4% for more serious GDPR infringements[44]. That said, the average loss in market capitalisation based on the CAAR was estimated to be of the order of nearly 29,000 times that at €1.2bn.

[44] Note that percentages were calculated based on ultimate parent revenues and not necessarily that of the infringing legal entity.

Clearly this figure is heavily skewed by the €19bn loss Alphabet Inc. experienced following their €50m fine. It seems that a huge market value is little protection against abnormal returns with the smallest company in the sample, Österreichische Post, having a slightly positive return. Also noteworthy was the seemingly innocuous €2k fine for BNP Paribas precipitating a market value fall of nearly €1bn. It was also noted that there was only one case (Österreichische Post) out of all 25 where the ratio of change in market capitalisation to fine was less than one, so firms need to recognise that the overall financial impact of a GDPR penalty is likely to be much greater than the value of the actual fine itself.

Table 11: Analysis by ultimate parent company

Ultimate Parent Company	N	CAAR	Average Revenue† € 000,000	Average Fine € 000	Fine as % of Revenue	Market Capitalisation‡ € 000,000	Δ Market Capitalisation € 000	Δ MC to Fine Ratio
United Internet	1	-0.0342	5,131	9,550	0.1861	7,104	242,957	25
Endesa SA	1	-0.0300	19,555	60	0.0003	22,634	679,020	11,317
Iberdrola	2	-0.0253	35,076	42	0.0001	63,221	1,602,652	38,618
UniCredit	1	-0.0204	20,674	130	0.0006	18,639	380,236	2,925
Delivery Hero	1	-0.0198	665	195	0.0294	23,691	469,082	2,401
Alphabet Inc	1	-0.0153	120,380	50,000	0.0415	1,245,280	19,052,788	381
BNP Paribas	1	-0.0152	52,030	2	0.0000	61,513	934,998	467,499
International Airlines	2	-0.0148	24,406	102,315	0.4192	10,354	153,246	1
Vodafone	1	-0.0130	43,666	60	0.0001	40,960	532,482	8,875
Eni SpA	1	-0.0123	75,822	11,500	0.0152	33,157	407,831	35
Deutsche Telekom	2	-0.0110	75,351	21	0.0000	70,219	768,898	36,614
Marriott	1	-0.0097	18,507	110,390	0.5965	41,340	400,995	4
Enel SpA	1	-0.0049	74,221	6	0.0000	82,095	402,266	67,044
ING Group	1	-0.0046	18,304	80	0.0004	34,953	160,784	2,010
OTP Bank	1	-0.0019	2,955	511	0.0173	10,979	20,861	41
Direct Line Insurance	1	-0.0007	3,937	5	0.0001	4,954	3,468	694
Électricité de France	1	0.0014	68,976	12	0.0000	31,142	43,599	3,633
Engie SA	1	0.0016	60,596	60	0.0001	30,778	49,245	821
Österreichische Post	1	0.0019	1,958	18,000	0.9191	2,320	4,408	0
Telefónica	2	0.0042	48,693	39	0.0001	20,019	84,080	2,156
Deutsche Wohnen	1	0.0320	1,438	14,500	1.0086	13,665	437,280	30
	25	**-0.0096**	**38,235**	**16,796**	**0.1462**	**81,313**	**1,177,602**	**28,901**

† Revenue of fiscal year prior to the event (consistent with GPDR penalties). Currencies converted based on rate at time of event.

‡ Current market capitalisation (Feb-21). Currencies converted based on rate at 31/12/2019.

Noting that of the top four negative CAAR events in **Table 11**, three of them are related to electricity companies it would certainly be interesting to look at industry sector analysis as recommended by e.g. Tweneboah-Kodua et al. (2018). A breakdown by sector is

shown in **Table 12**. Here it can be seen that the most reactive industry sector was *Consumer Cyclical* (-1.5%), however, only *Utilities, Communication Services* and *Financial Services* showed statistical significance of non-zero (negative) abnormal returns albeit only at the 10% level.

Table 12: CAAR by industry sector

Industry Sector	N	CAAR	t_{CAAR}	% Negative CAR
Consumer Cyclical	2	-0.0148	-2.9208	100
Utilities	6	-0.0138	-2.1852 *	67
Energy	1	-0.0123		100
Communication Services	7	-0.0109	-2.1098 *	86
Industrials	3	-0.0092	-0.8761	33
Financial Services	5	-0.0086	-2.1881 *	100
Real Estate	1	0.0320		0
	25	**-0.0096**		**76**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

A geographical analysis is shown in (**Table 13**). Although France shows the most negative CAR, there is only one example. Interestingly, the majority of fines (15 out of 25 = 60%) came from the Spanish and Romanian data protection authorities, both exhibiting negative CAARs which are statistically significant at the 5% level. These appear to be low value fines overall (combined only 0.14% of total) so there does not seem to be any obvious correlation between CAAR and value of fines – the UK being responsible for 75% of the total fine value yet having a negative CAAR of less than half the overall mean. It would appear that the markets in Spain and Romania are more sensitive to GDPR fine announcements despite the low fine values. At the time of writing, according to CMS Legal (2021), the Spanish data protection authority has issued 342 fines since the advent of the GDPR which is over three times more than its nearest rival, Italy, with 101. As there was no (statistically significant) result from Italy in the dataset here, the next most prolific fine issuer was actually Romania with 68 which seems consistent with this dataset and would appear to indicate that it is the number of fines issued which is the major factor in market nervousness rather than their monetary value.

Table 13: Analysis by country

Country	N	CAAR	t_{CAAR}	% Negative CAR	Total fines (€000)
FRANCE	1	-0.0153		100	50,000
SLOVAKIA	1	-0.0137		100	40
ITALY	1	-0.0123		100	11,500
SPAIN	10	-0.0113	-2.2826**	70	388
ROMANIA	5	-0.0107	-3.4456**	100	220
GERMANY	3	-0.0073	-0.3648	67	24,245
UK	2	-0.0045	-0.8654	50	314,990
BULGARIA	1	-0.0019		100	511
AUSTRIA	1	0.0019		0	18,000
	25	**-0.0096**		**76**	**419,894**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

During the data collection exercise, it was noted that some of the larger GDPR fines had been appealed and the results of the appeals formally announced. This enabled an additional dataset to be built (**Table 14**) and analysed in the same way as the initial announcements.

Table 14: Summary of GDPR fine appeals

Ultimate Parent	Date	Original fine	Result of appeal
Alphabet Inc	12/06/2020	€50m	Rejected
International Airlines	16/10/2020	£190m	Reduced to £20m
Marriott	30/10/2020	£99.2m	Reduced to £18.4m
United Internet	12/11/2020	€9.55m	Reduced to €900k

The expected outcome of these appeal announcements would be negative market price impact for the unsuccessful appeal by Alphabet Inc and positive for the other three examples where the fines were massively reduced. The results are shown in **Table 15**. It appears there is indeed, a negative trend for Alphabet beginning on the announcement day itself and not disappearing until 20 days after the event. International Airlines has a strongly increasing positive return after the event whereas, although positive, United

Internet remains fairly constant. Marriott however, experienced some negative market sentiment after the event. One has to be mindful of market conditions and volatility due to the COVID-19 pandemic and its effect on (especially the hospitality) industry here. That was the reason the original dataset was capped at 31/12/2019 and, in analysing these more recent events, the results were not found to be statistically significant thus the null hypothesis of zero abnormal returns still stands.

Table 15: CAR by event window of fines appealed

Event Window	N	Alphabet Inc		International Airlines		Marriott		United Internet	
		CAR	t_{CAR}	CAR	t_{CAR}	CAR	t_{CAR}	CAR	t_{CAR}
(-2, 2)	1	0.0164	0.5686	0.1459	1.1842	0.0455	0.7426	0.1039	1.9689
(-1, 1)	1	0.0026	0.1164	0.0499	0.5229	0.0143	0.3013	0.0563	1.3715
(-1, 0)	1	0.0054	0.2960	-0.0110	-0.1412	0.0346	0.8929	0.0431	1.2859
(0, 1)	1	-0.0076	-0.4166	0.0345	0.4427	-0.0045	-0.1179	0.0598	1.7917
(0, 2)	1	-0.0075	-0.3357	0.1059	1.1096	-0.0009	-0.0192	0.0812	1.9865
(0, 3)	1	-0.0008	-0.0310	0.0899	0.8158	-0.0187	-0.3463	0.0839	1.7775
(0, 4)	1	-0.0148	-0.5131	0.1349	1.0949	-0.0230	-0.3810	0.0753	1.4269
(0, 5)	1	-0.0171	-0.5412	0.1523	1.1284	0.0073	0.1104	0.0796	1.3770
(0, 10)	1	-0.0379	-0.8858	0.1596	0.8733	0.1250	1.3959	0.0827	1.0566
(0, 20)	1	0.0160	0.2707	0.3824	1.5145	0.1686	1.3626	0.0902	0.8340

5.7. Conclusion

It has been seen how the announcement of monetary penalties related to GDPR infringement can result in statistically significant negative CARs of around 1% up to three days after the event. It was also observed that the economic impact on the market value of a publicly listed firm far outweighs the monetary value of the fine itself in almost all cases, and that a very small fine can have huge impact on market value (cf. BNP Paribas). It is also known from the literature that CARs of a similar magnitude are generated at the time of the initial announcement of a breach (and as seen in the previous chapter). Considering all these negative factors, the need for firms to invest in cyber security to protect data privacy is clearly underpinned by this research, as well as showing a clear economic impact of the introduction of the GDPR itself. Significant negative market reactions to particularly punitive data protection authorities have also been highlighted, as in the case of Spain and Romania, despite their relatively low monetary penalties.

In light of the recent introduction of the GDPR, the dataset for this study was (necessarily) limited. Once more data becomes available and the market recovers from the COVID-19

pandemic, future research is expected to give a better idea of the impact of GDPR infringement fines on publicly listed firm value. Although four examples of GDPR fine appeals were identified and positive returns were observed where those appeals were successful (and the reverse), the results were not statistically significant, and the null hypothesis of zero abnormal returns could not be rejected. Future research is needed in this area also as recently there has been news of Deutsche Wohnen successfully appealing their €14.5m fine. Considering the high-profile reductions of the fines for International Airlines (British Airways) and Marriott, a precedent appears to have been set with the ICO clearly recognising the need to encourage infringing firms to use their available funds in these difficult economic times to invest in cyber security measures (Macfarlanes, 2020). Future studies may, therefore, reveal more about the positive impact of the GDPR on cyber security investment following its introduction and subsequent punitive actions. In this study, only 2 out of 21 (10% of) ultimate parent firms were US based with the balance being European, therefore this work also begins to offset the strong US bias of these types of studies in the literature as predicted by Ali et al. (2021).

The next chapter focusses on information security events of a more positive nature (which could, in theory, have been driven by an unfavourable event such as major data privacy breach), that of CISO appointment announcements ('The CISO Effect').

Table 16: List of GDPR Infringement Fine Announcements

Company	Symbol	Index	Date	Country	Amount (€)	Type	Comments
Austrian Post	POST.VI	ATX20	23/10/2019	AUSTRIA	18,000,000	Insufficient legal basis for data processing	
BNP Paribas Personal Finance SA	BNP.PA	CAC40	22/11/2019	ROMANIA	2,000	Insufficient fulfilment of data subjects rights	
British Airways	IAG.L	FTSE100	08/07/2019	UK	204,600,000	Insufficient technical and organisational measures to ensure information security	
Curenergía Comercializador de último recurso	IBE.MC	IBEX35	28/11/2019	SPAIN	75,000	Insufficient legal basis for data processing	
Delivery Hero	DHER.DE	MDAX60	19/09/2019	GERMANY	195,407	Insufficient fulfilment of data subjects rights	
Deutsche Wohnen SE	DWNI.DE	MDAX60	30/10/2019	GERMANY	14,500,000	Non-compliance with general data processing principles	
DSK Bank	OTP.BD	CEE	28/08/2019	BULGARIA	511,000	Insufficient technical and organisational measures to ensure information security	^BUX.BD not available.
ENDESA (energy supplier)	ELE.MC	IBEX35	09/04/2019	SPAIN	60,000	Insufficient legal basis for data processing	Date added.
Eni Gas e Luce	ENI.MI	FTSEMIB	11/12/2019	ITALY	11,500,000	Insufficient legal basis for data processing	Could use EWI instead.
Google Inc	GOOGL	SP500	21/01/2019	FRANCE	50,000,000	Insufficient legal basis for data processing	
Iberdrola Clientes	IBE.MC	IBEX35	16/10/2019	SPAIN	8,000	Insufficient cooperation with supervisory authority	
ING Bank NV	INGA.AS	AEX25	28/11/2019	ROMANIA	80,000	Insufficient technical and organisational measures to ensure information security	
Linea Directa Aseguradora	DLG.L	FTSE100	03/12/2019	SPAIN	5,000	Insufficient legal basis for data processing	

Madrileña Red de Gas	EDF.PA	CACN20	21/01/2019	SPAIN	12,000	Insufficient technical and organisational measures to ensure information security	Minority stake only. Date added.
Marriott International	MAR	SP500	09/07/2019	UK	110,390,200	Insufficient technical and organisational measures to ensure information security	
SC Enel Energie SA (Electricity Distributor)	ENEL.MI	FTSEMIB	16/12/2019	ROMANIA	6,000	Insufficient legal basis for data processing	Could use EWI instead.
Slovak Telekom	DTE.DE	DAX30	27/09/2019	SLOVAKIA	40,000	Insufficient technical and organisational measures to ensure information security	Date added.
Telecoms provider (1&1 Telecom GmbH)	UTDI.DE	TECDAX	09/12/2019	GERMANY	9550,000	Insufficient technical and organisational measures to ensure information security	
Telefonica Moviles España SAU	TEF.MC	IBEX35	06/05/2019	SPAIN	48,000	Non-compliance with general data processing principles	Date added.
Telefónica SA	TEF.MC	IBEX35	14/11/2019	SPAIN	30,000	Non-compliance with general data processing principles	
Telekom Romania Mobile Communications SA	DTE.DE	DAX30	18/12/2019	ROMANIA	2,000	Insufficient technical and organisational measures to ensure information security	Also traded as OTE in Athens.
Unicredit Bank SA	UCG.MI	FTSEMIB	27/06/2019	ROMANIA	130,000	Insufficient technical and organisational measures to ensure information security	Could use EWI instead.
Viaqua Xestión Integral Augas de Galicia	ENGI.PA	CAC40	21/11/2019	SPAIN	60,000	Insufficient legal basis for data processing	
Vodafone España SAU	VOD.L	FTSE100	06/11/2019	SPAIN	60,000	Insufficient legal basis for data processing	
Vueling Airlines	IAG.L	FTSE100	01/10/2019	SPAIN	30,000	Insufficient legal basis for data processing	

Chapter 6. (The CISO Effect:) The Impact of CISO Appointment Announcements on the Market Value of Firms

6.1. Introduction

The US Federal Bureau of Investigation (IC3, 2020) reported a 69% increase in internet crime related complaints from 2019. With cybercrime so rife, along with damaging high-profile data breaches such as that of Marriott and British Airways (Ford et al., 2021a), cyber security should be very much a concern for organisations globally. Indeed, the UK Department for Digital Culture, Media and Sport report that 77% of businesses view cyber security as a high priority at board level (DCMS, 2021) and Gartner Inc. (2021) predict that by 2025, 40% of boards will have cyber security committees established and overseen by a *"suitably qualified"* executive. Therefore, given the importance of cyber security, in the case of publicly listed companies one could reasonably expect markets to react in a positive way to news of investment in this area. This study is concerned with announcements of investment in human capital, specifically heads of security at executive level such as chief information security officer (CISO), chief security officer (CSO) or similar roles[45], the primary research question being what is the impact (if any) of this information on the market value of firms? The ability to measure and clearly identify any positive impact would surely encourage organisations to both invest in, as well as publicise establishment or improvement of their security functions and thereby their overall cyber security posture[46].

Existing literature relating specifically to CISOs is rather sparse (Karanja & Rosso, 2017) so this small-scale initial study would also begin to fill a knowledge gap. Such research is expected to be of interest to business management, cyber security practitioners, investors and shareholders as well as researchers in cyber security or related fields. Looking elsewhere, due to the dearth of literature on CISOs as reported by Karanja and Rosso (2017), an interesting study by Banker and Feng (2019) showed that data breaches caused by system deficiencies (but not criminal fraud or human error) increased the likelihood of CIO turnover by 72%. Chatterjee, Richardson and Zmud (2001) examined the economic impact of the creation of new chief information officer (CIO) positions using event study techniques and observed positive market reactions of 1.16% on average, and even higher (almost 3%) for heavily information technology (IT) driven industry

[45] For convenience, hereinafter the acronym 'CISO' is used as a generic term to refer to roles of this type.
[46] Cyber security posture is a term encompassing training and awareness in addition to governance and technical solutions.

sectors. This variation between sectors was also reported by Tweneboah-Kodua et al. (2018) who observed that financial services firms reacted more rapidly and more significantly than those in the technology sector. Their study focussed on the negative impact of cyberattacks on stock returns and a systematic literature review by Spanos and Angelis (2016) found that 76% of studies in this area reported an impact which was both negative and statistically significant. For instance, Lin et al. (2020) observed losses of 1.44% on average over a five-day window. Andoh-Baidoo, Amoako-Gyampah and Osei-Bryson (2010) found -3.18% abnormal returns over a three-day period. Cavusoglu, Mishra and Raghunathan (2004) observe -2.1% on average over the two days following the announcement. Goel and Shawky (2009) cite -1% in the days immediately surrounding the event.

Moving back to positive economic impacts, literature in this area was found to be lacking. Cavusolgu et al. (2004) found a positive impact on security developers after the disclosure of security breaches by affected firms, consistent with Chen et al. (2012) who found IT consulting firms gained market share following a client security breach (with the caveat the breach was not too severe). Deane et al. (2019) studied ISO 27001 certification announcements and found the abnormal market returns both positive and statistically significant (0.72% on average over two days). A more recent SLR (Ali et al., 2021) updating the work of Spanos and Angelis (2016), acknowledges the relative lack of studies concerning the positive impact of favourable information security events such as regulation, certification or investment. The authors also note that such favourable events generate lower magnitude abnormal returns (in the range 0.63% to 1.36%) than unfavourable events which could lose up to 10% of market value.

This study, therefore, not only increases the CISO research knowledge base but also helps to address the shortfall in studies on favourable security events. Ali et al. (2021) also note that 76% of existing similar studies were based solely on US data, so the aim here was also to look globally as far as possible (given English language restrictions) in an attempt to offset this US bias (cf. Ford et al., 2021a).

6.2. Methodology

From a high-level perspective, the approach to this research was to hand-gather a dataset of CISO appointment announcements then filter this dataset keeping only those examples related to publicly listed companies. The final step was to identify any impact of these announcements on the share price of each firm through ESM as elucidated in Chapter 3.

6.3. Data collection

The base dataset was hand-gathered from internet searches for CISO appointments[47] building up a Microsoft Excel spreadsheet. Many results were clearly governmental organisations or not-for-profits or private companies, so were discarded. It was decided to cap the date range at 31/12/2019 in order to avoid market uncertainties due to COVID-19, that in itself being a long-term confounding event. Once the residual set seemed to be mostly publicly listed companies, each data record was carefully reviewed, ensuring that the stock was still listed and share price data available. Data fields extracted included company name, announcement date, job title, position in the organisation (reporting line) and gender. Some announcements made it clear this was a newly created CISO position and so these were marked as such whereas others merely implied this – these were flagged separately due to the lesser information salience (e.g. Ramos et al., 2020). Other records had to be filtered out as they were not listed before the start of the estimation window. A few examples (Bridgestone Americas, Santander UK and CareerBuilder) were related to subsidiary companies, but these were kept as there would be a contribution, at least, to any market reaction despite less salience, also to maintain as large a dataset as possible. Once the filtering was complete, a dataset of 41 records remained (see **Table 23**). The final step was to look for confounding events. The dates of the nearest events (before/after day zero) were also recorded so these could be filtered out dynamically as needed. Confounding events were considered to be earnings/dividend announcements or another executive appointment (consistent with e.g. Chatterjee et al., 2001). There were three examples (Digital Realty, Wells Fargo, Axon) where joint announcements were made with other positions and these were included, again to maximise the dataset despite the lesser salience.

The importance of the choice of market index to use as a reference has been highlighted by e.g. Kannan et al. (2007). Consistent with Ford et al. (2021a), the most appropriate market index was chosen in each case being, ideally, one in which the stock in question was included and favouring narrower, more focussed indices rather than a one-size-fits-all approach given, of course, that this index data was available from Yahoo!Finance (otherwise the next best was selected). Some firms had multiple listings, in which case the primary listing and associated index were used.

[47] Sources included ProQuest (CSO Online), Google and Bing

6.4. Data analysis

To facilitate the analyses, R (R Core Team, 2018)[48] code was developed to extract share price and index data directly from Yahoo!Finance for each data record. Event studies were then carried out through an R package (Schimmer, Levchenko & Müller, 2014) [49] applying the market model as described above[50]. Announcements falling on non-trading days were defaulted to the next available trading day. An estimation window of 120 days was chosen, consistent with previous studies (e.g. Goel & Shawky, 2009; Andoh-Baidoo et al., 2010; Schatz & Bashroush, 2016a; Richardson et al., 2019) ending one trading day before the event window in all cases. Tweneboah-Kodua et al. (2018) recommend avoiding overlap of the estimation and event windows in this way to avoid "*parameter contamination*". Although the event window should be broad enough to contain any uncertainty in the date of the event, the longer the window, the less likely it is to detect abnormal returns (Dyckman et al., 1984). Some studies observed a market reaction in advance of the event date due to information leakage, such as Lin et al. (2020) who found significant evidence of opportunistic pre-official announcement insider trading related to data breaches using event studies. For this exercise, a range of event windows was initially chosen starting from up to two days before the event and varying in length from 2 up to 20 trading days. This approach was used to catch any pre-event effects as well as others, for instance, sector specific effects reported by i.a. Tweneboah-Kodua et al. (2018) who observed more rapid responses to information security events from the financial services sector, for example, thereby justifying longer event windows for sectorial comparison purposes.

6.5. Hypothesis development

For hypothesis development, please refer to Chapter 3.

6.6. Results and discussion

Event studies were carried out as above for 10 event windows of differing length across all 41 CISO appointment announcements. A visualisation of the overall results showed that event windows (-1, 1) and (-1, 0) were of interest and thus records with confounding events ±2 days were filtered out[51] leaving 37 records. The revised visualisation is shown

[48] R version 4.1.2 (2021-11-01)
[49] EventStudy package version 0.36.900 (API version 1.059)
[50] The R code used is included in the Appendix.
[51] This is a quite stringent approach, similar studies have only filtered out confounding events within the event window, this approach leaves a gap of one day in case of any advance information leakage of the confounding event.

in **Figure 17**.The most positive impact is seen around the three or two-day event windows (-1, 1), (-1, 0) with the market value reverting back to normal by day 4. There was also a slight peak at day 5, but of course, this may be due to confounding events. By day 20 it can be seen that Fortinet is the best performing stock with a CAR of almost 20%, and MGIC the worst.

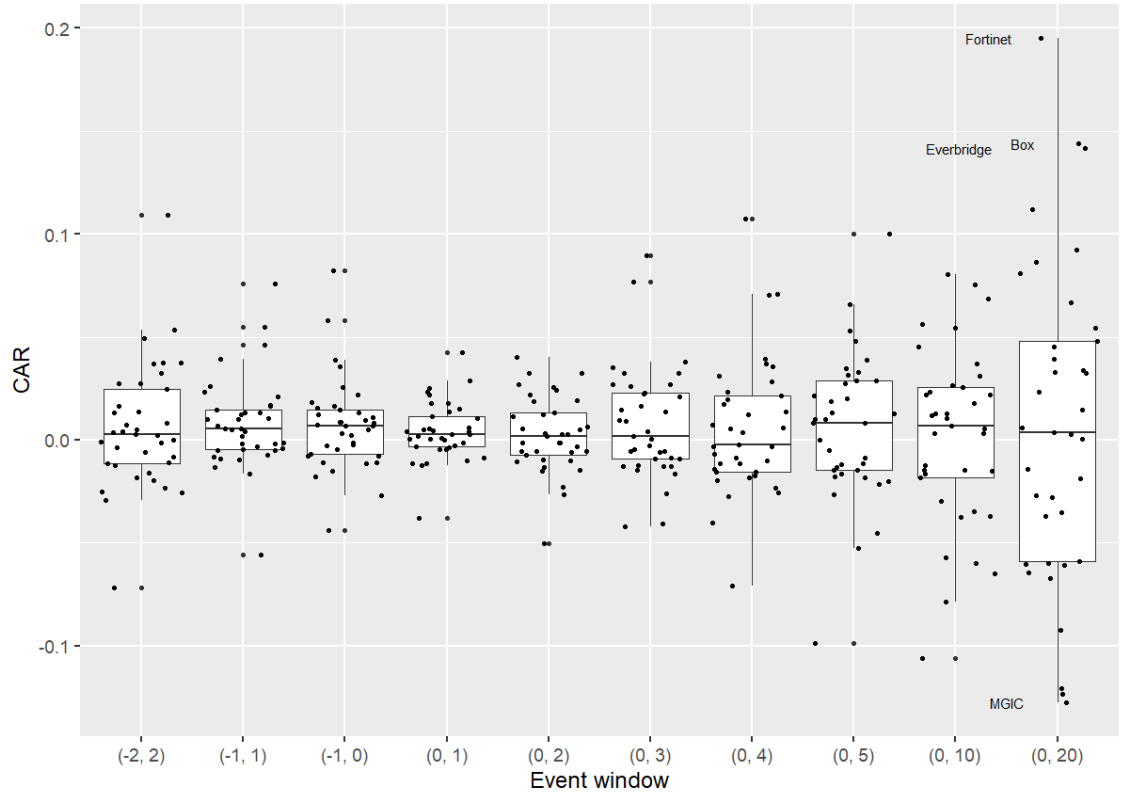


Figure 17: Comparison of event windows

A CAAR was calculated for multiple firms across each window and the results shown in **Table 17**. The three-day event window (-1, 1) showed the most positive CAAR of almost 0.8% (significant at the 5% level), similar to the findings of e.g. Chatterjee et al. (2001) for CIO appointments with their dataset being of a comparable size as well (1.16% over 96 announcements). The fact that this includes a pre-event day indicates a little information leakage prior, as the date accuracy was carefully verified (cf. Lin et al. 2020). Within this window, 22 of the 37 events (59%) had abnormal returns of greater than zero, therefore this window was chosen as the basis for further analyses. Usage of this event window (-1, 1) has been previously reported in studies of this type (e.g. Chatterjee et al., 2001; Andoh-Baidoo et al., 2010; Bose & Leung, 2014; Khansa, 2015; Modi et al., 2015). A lack of statistical significance for the longer windows is also seen, indicative of a

tendency of market recovery towards zero abnormal returns over time as reported by Dyckman et al. (1984).

Table 17: CAAR by event window

Event Window	N	CAAR	t_{CAAR}	% Positive CAR
(-2, 2)	37	0.0062	1.2356	54
(-1, 1)	37	0.0077	2.1091**	59
(-1, 0)	37	0.0070	1.8997*	62
(0, 1)	37	0.0043	1.8334*	65
(0, 2)	37	0.0025	0.8411	51
(0, 3)	37	0.0075	1.6717	54
(0, 4)	37	0.0055	1.0091	49
(0, 5)	37	0.0059	1.0054	57
(0, 10)	37	0.0015	0.2074	59
(0, 20)	37	0.0070	0.5529	57
	370	**0.0055**		**57**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

A visualisation of how the data changes over time is shown in **Figure 18**. Out of the 37 total events in the dataset, by far the majority (92%) occurred between 2017 and 2019 but the overall number of CISO announcements does not seem to be increasing year-on-year - it actually dropped from 18 to 5 in the last two years sampled, so there seems to be a lack of awareness of the potential benefits of sharing this information with the market. A relative measure of the CAAR is also shown flipping from positive in 2012 to negative in 2015 and remaining positive for the bulk of the dataset in the final three years.

Regarding the origin of the appointment, internal appointments were low (27%) in 2017 dropping to 11% in 2018 and subsequently to zero in 2019. So, ignoring the sparse data in previous years, a trend toward solely external recruitment is evidenced. Interestingly the internal appointments, although fewer at 6 out of 37 (16%), generated a CAAR of 1.28% which was almost double that of the much more frequent external appointments (CAAR=0.67%), although these results were not statistically significant. Chatterjee et al. (2001) suggest that the market may respond better to internal appointments because the appointee would 'hit the ground running' due to existing in-depth knowledge of the business and established relationships with management.

Karanja and Rosso (2017) observed in their dataset of 55 CISOs spanning 2010-2014 that only 11% were female. Although here it is difficult to identify any clear trend over time, it can be noted there were no such examples in the data sample prior to 2017 (although only three data points) yet in the final three years the percentage of women was fairly constant (18, 22, 20%). Despite the lack of any obvious trend in the data here, comparing with Karanja and Rosso (2017) there is almost double the percentage of women CISOs – a refreshing increase in diversity. That said, the 30 examples (82%) of male appointees yielded a CAAR of +0.85% significant at the 5% level whereas the remaining 7 female appointments generated only +0.38% which was, unfortunately, not statistically significant.

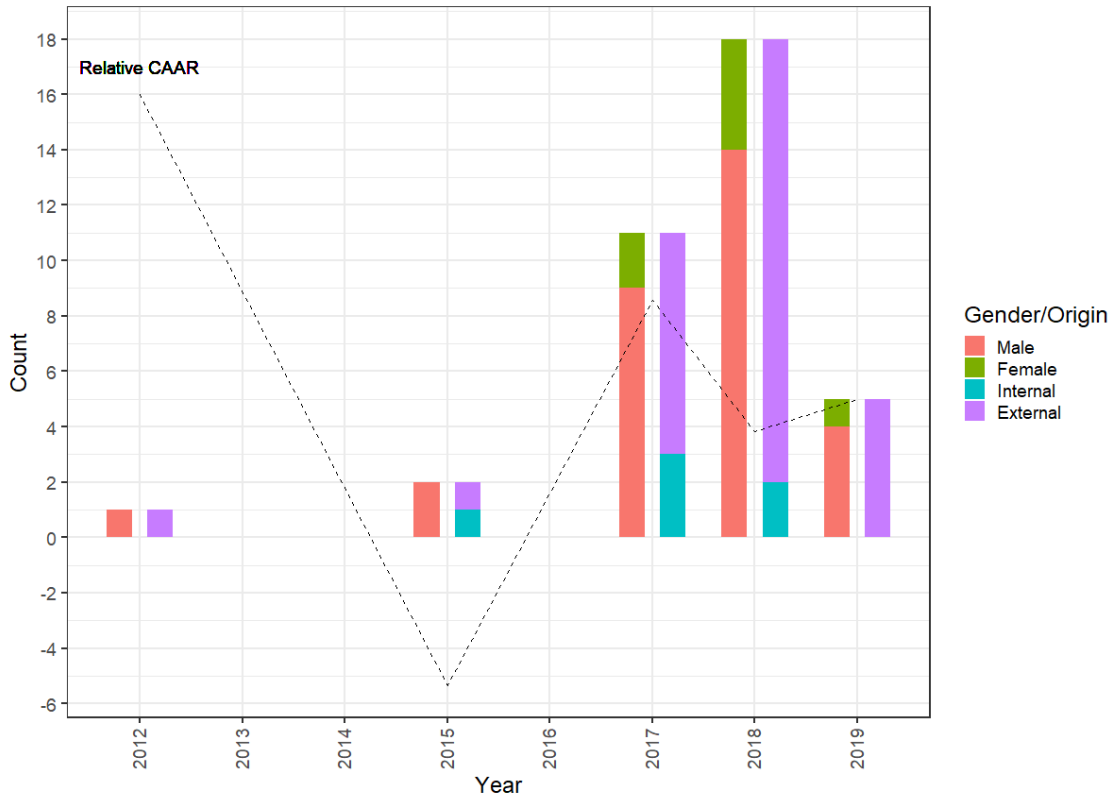


Figure 18: Breakdown by year

A breakdown of CAAR by sector is shown in **Table 18**. The financial services sector was the clear winner showing positive CARs of almost 1.8% on average and statistical significance at the 1% level. Such sector specific behaviour has been highlighted in e.g. Tweneboah-Kodua et al. (2018). It is interesting to note that 17 out of 37 announcements sampled (46%) belonged to this sector, perhaps indicating more willingness (or regulation, of course) for transparency within this industry. It does seem surprising that negative market sentiment was identified for 10 examples spanning 4 sectors especially

one as sensitive as communication services which came in at more than 1% negative. Nevertheless, there is a deficiency of data here and no statistical significance.

Table 18: CAAR by industry sector

Industry Sector	N	CAAR	t_{CAAR}	% Positive CAR
Financial Services	17	0.0178	3.3027***	76
Healthcare	1	0.0169		100
Consumer Cyclical	2	0.0077	3.2553	100
Utilities	1	0.0066		100
Technology	6	0.0046	0.5974	50
Real Estate	1	-0.0047		0
Consumer Defensive	1	-0.0082		0
Industrials	5	-0.0084	-0.6592	40
Communication Services	3	-0.0101	-2.8456	0
	37	**0.0077**		**59**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

Looking more closely at the content of the announcements, information on job title and position in the organisation are shown in **Table 19** and **Table 20** respectively. The most positive returns seem to occur when the CISO title is combined with a VP/SVP position – clearly this is recognised by the market as having more influence within the organisation, although due to lack of data it is not statistically significant. The title CSO (a broader role not restricted to only information security) does show some significance but only at the 10% level. Even the single example of appointment of a Deputy CISO role (occurring shortly after the CISO was recruited) showed positive CAR at nearly 0.5%. Interestingly, combination of the security role with the CIO title results in the least positive CAR, or with that of "*trust*" (CTSO) yields negative of 1.7%. It appears that the market is preferring a focussed role with a high level of influence.

Table 19: CAAR by job title

Title	N	CAAR	t_{CAAR}	% Positive CAR
VP/CISO	3	0.0358	2.4170	100
SVP/CISO	1	0.0234		100
CSO	6	0.0098	2.0353*	83
Global Head of Cyber Risk	1	0.0053		100
Deputy CISO	1	0.0048		100
CISO	23	0.0043	0.8788	43
CIO/CISO	1	0.0017		100
CTSO	1	-0.0165		0
	37	**0.0077**		**59**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

Carrying forward the argument of influence, it is possible to get an idea of the position in the organisation from some (16 out of 37 were not specified)[52] of the announcements where the reporting line was cited in **Table 20**. First of all, the results where data were available were not significant, so the null hypothesis cannot be rejected, although it is interesting to note that the expected result of direct CEO reporting (and therefore greatest influence) being the highest CAAR is not the case – it seems that COO or CFO reporting is more well received by the market. If the concept of operations is combined with technology (CTO) or EVP though, the result is negative. The observation that a reporting line into the CIO yields negative CAAR seems consistent with Williams (2007) and the conflict-of-interest argument – would there be reluctance on the part of such a CISO to call out security flaws in the CIO's IT environment?

[52] In some cases, it could be assumed from the announcement, however unless explicit it was classified as unspecified.

Table 20: CAAR by CISO reporting line

Reporting line	N	CAAR	t_{CAAR}	% Positive CAR
COO	2	0.0215	0.8737	50
CFO	1	0.0162		100
CEO	4	0.0148	1.3565	75
Corporate Secretary	1	0.0134		100
CTO	1	0.0130		100
(not specified)	16	0.0107	1.7592*	62
CISO	1	0.0048		100
CTO/COO	1	-0.0013		0
CIO	8	-0.0033	-0.3978	50
EVP Operations	1	-0.0047		0
SVP	1	-0.0050		0
	37	0.0077		59

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

The Chatterjee et al. (2001) study was restricted to the establishment of newly created CIO positions only. If the CISO announcements here contained a clear or implied indication that this was a new CISO position this was captured in the data record and analysed as shown in **Table 21**. The results were almost the exact opposite of expected, that the market would react better if an organisation was currently lacking an established security function. Perhaps the market expectation is that there should already be such a function in existence? The results are not significant except those cases where new or established was not specified which was a clear positive CAAR approaching 1.2%.

Table 21: Analysis of new or established CISO roles

New or established?	N	CAAR	t_{CAAR}	% Positive CAR
(not specified)	25	0.0115	2.6205**	64
New (implied)	4	0.0015	0.2505	50
New (specified)	8	-0.0013	-0.1501	50
	37	**0.0077**		**59**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

In view of the intended international nature of this study, an analysis by different markets is shown in **Table 22** using market currency as a primary key (to be less granular due to data deficiency). Here, India shows by far the highest CAR of 7.6% but with only one example. It is the UK and US who show significantly positive CAARs at the 5% level (with the UK almost double that of the US). Only Israel and Australia are displaying negative CAARs, but again lacking data. Despite no intended US focus, it can be seen that 28 out of 37 (76%) announcements originated from US markets so there is little opportunity to analyse other markets in detail. This percentage matches closely the observation of Ali et al. (2021) on US dominance in such research who attribute this to regulatory effects encouraging transparency in addition to language restrictions. Over time there is an expectation that transparency (and thus availability of data) will increase in other markets as new regulations come into force.

Table 22: Analysis by market currency

Market Currency	N	CAAR	t_{CAAR}	% Positive CAR
INR	1	0.0757		100
GBP	2	0.0158	14.3636**	100
USD	28	0.0081	2.4346**	61
JPY	1	0.0053		100
EUR	2	0.0052	0.6242	50
ILS	1	-0.0013		0
AUD	2	-0.0320	-1.3453	0
	37	**0.0077**		**59**

*,**,*** Represent statistical significance at the 10%, 5% and 1% levels respectively.

6.7. Conclusion

It has been shown in this introductory study that the announcement by a publicly listed company of a CISO appointment does indeed induce a positive market reaction, particularly within the financial services sector (+1.8% in the three days surrounding the event). In the data sample of 37 announcements analysed, 17 were financial services companies indicating a sector specific willingness to report, or regulatory effect.

Based on the above analyses, the following content for a CISO appointment announcement to deliver maximum positive abnormal market return would be advised, with the caveat that not all of the findings were statistically significant, so these are indicative trends only (ideally the hiring organisation would be a US or UK listed company in the financial services sector): job title CISO with VP or SVP responsibility stipulated (no mention whether this is a new or established role), internal placement, male, reporting to the COO, CFO, or CEO but definitely not the CIO.

Chatterjee et al. (2001) calculated an approximate range of US$7.5m (median approach) to US$76m (mean approach) increase in market capitalisation for CIO appointments thereby easily justifying "*the trend in escalated executive salaries*". Repeating this estimation across this whole CISO dataset gives a range of US$94m to US$318m, so it seems that this trend is very much continuing and clearly applies to CISOs as well!

This research should highlight the clear economic benefit of CISO appointments, as well as the advantage of transparency in this area, to business management through the value the market places on the CISO role. There is an opportunity for other sectors to follow the lead of financial services and for other markets to adopt US practices and get ahead of the curve before new regulations come into effect. Once more data becomes available this exercise could be revisited for more in-depth analyses hopefully revealing less of a bias towards male CISO appointments as observed here.

The next chapter considers the overall impact of multiple events as there is some overlap in sample firms between this chapter and/or the previous two.

Table 23: List of CISO Appointment Announcements

Firm	Symbol	Index	Date	Nearest Confounding Event Date	Position	Reports to	Origin	Gender	First appointment?	Comments
AIG	AIG	SP100	15/07/2019	22/07/2019	Deputy CISO	CISO	External	Male		
AIG	AIG	SP100	01/04/2019	05/04/2019	CISO	CIO	External	Male		From Wells Fargo.
Avnet	AVT	SP400	08/12/2015	04/01/2016	CISO	CIO	Internal	Male	First	Changed to NASDAQ in 2018.
Axon	AXON	SP400	21/12/2017	13/02/2018	CISO		Internal	Male		Joint announcement.
Barclays	BARC.L	FTSE100	11/04/2018	26/04/2018	CSO		External	Male		Came from JP Morgan.
Box	BOX	RUSSELL2000	14/01/2019	27/02/2019	CISO		External	Female	Implied	Came from SAP Ariba.
Bridgestone Americas	5108.T	TOPIX	15/05/2018	16/07/2018	CISO	CIO	External	Male		Subsidiary
CareerBuilder	APO	RUSSELL1000	12/12/2017	26/10/2017	CIO/CISO		External	Male		Subsidiary.
CenturyLink	LUMN	SP500	18/09/2018	08/11/2018	CSO		External	Male		Ticker previously CTL.
Comerica	CMA	SP500	10/10/2012	17/10/2012	SVP/CISO		External	Male		Came from Morgan Stanley.

Customers Bank	CUBI	RUSSELL2000	01/08/2017	04/08/2017	VP/CISO	COO	Internal	Female		Promotion
Digital Realty	DLR	SP500	12/09/2018	24/09/2018	CISO	EVP Operations	External	Male		Joint CIO/CISO announcement.
Dominion Energy	D	DJUA	20/11/2018	26/11/2018	VP/CISO		External	Male	First	ex FBI.
Equifax	EFX	SP500	12/02/2018	26/02/2018	CISO	CEO	External	Male		Previously Home Depot.
Everbridge	EVBG	RUSSELL1000	30/04/2018	07/05/2018	CSO	CEO	External	Male		Previously Fannie Mae.
F5	FFIV	SP500	11/07/2018	20/07/2018	CISO		External	Female	Implied	
Factset	FDS	SP400	20/06/2018	26/06/2018	CISO		External	Male		Previously Dell.
Flipkart	WMT	SP100	15/11/2019	14/11/2019	Head of Information Security		External	Male		Subsidiary. Came from WiPro.
Fortinet	FTNT	SP500	10/01/2017	02/02/2017	CISO	CEO	External	Male		
GlaxoSmithKline	GSK.L	FTSE100	11/12/2018	06/02/2019	CISO		External	Female		Pharma, R&D only.
Grainger	GWW	SP500	15/12/2017	12/12/2017	CISO		Internal	Male	Implied	Promotion
Home Depot	HD	SP100	08/08/2018	14/08/2018	CISO	CIO	External	Male		Tweeted by ciodive.
Huntington Ingalls Industries	HII	SP500	26/06/2017	25/07/2017	CISO	CFO	External	Male	First	From Vencore.

IDFC First Bank	IDFCFIRSTB.NS	NIFTY500	01/05/2017	25/04/2017	CISO		External	Male		
Jetstar	QAN.AX	ASX200	01/02/2017	23/02/2017	CISO	CIO	External	Female	First	From Asciano IT.
Leumi Bank	LUMI.TA	TA35	17/10/2017	21/11/2017	CISO	CTO/COO	External	Male	First	From SMBC.
McDonalds	MCD	SP100	19/09/2017	21/09/2017	CISO	VP Operations	External	Male		Lack of clarity on date.
MGIC	MTG	SP400	22/03/2018	12/03/2018	VP/CISO		External	Male		From Avnet.
Moody's	MCO	SP500	17/10/2018	23/10/2018	Global Head of Cyber Risk		Internal	Male		MIS division only. Formerly CISO.
Palo Alto	PANW	RUSSELL1000	24/11/2015	23/11/2015	CISO	CFO	External	Male		First (but not obvious).
Popular Bank	BPOP	RUSSELL1000	19/04/2018	24/04/2018	CSO	CEO	Internal	Female	First	Promotion
Santander UK	SAN.MC	IBEX35	24/09/2018	08/10/2018	CISO	COO	External	Female	Implied	Subsidiary
Silicon Valley bank	SIVB	SP500	07/02/2019	24/01/2019	CSO	CIO	External	Male		
Société Générale	GLE.PA	CAC40	28/09/2017	04/10/2017	CSO	Corporate Secretary	External	Male	First	ex Air Force

Southwest Airlines	LUV	SP500	28/01/2019	30/01/2019	MD Technology/CISO		Internal	Male		Promotion - announced with others.
Twilio	TWLO	RUSSELL1000	16/08/2018	06/08/2018	CTSO		External	Male		Previously an adviser.
Unisys	UIS	SP600	16/04/2018	01/05/2018	CISO	SVP	External	Male		ex IBM.
Vonage	VG	SP600	11/04/2017	09/05/2017	CISO		External	Male		From hosting.com.
Voya Financial	VOYA	RUSSELL1000	19/07/2018	26/07/2018	CISO	CIO	External	Male		
Wells Fargo	WFC	SP100	28/05/2019	06/06/2019	CISO	CTO	External	Male		Joint announcement.
Woolworths	WOW.AX	ASX200	25/05/2015	06/05/2015	CISO	CIO	External	Male	First	From KPMG.

Chapter 7. The Impact of Repeated Information Security Events on Market Value

7.1. Introduction

A closer look at the datasets from Chapters 4, 5 and 6 reveals some examples of ultimate parent companies being the subject of repeated information security events either within each study or between different ones. In this chapter (Chapter 7) the overall impact of such repeated events is examined. It would be interesting to see if the markets react more or less strongly to, for example, the first or subsequent data breach announcements for a specific ultimate parent firm (unfavourable event). One might expect a stronger 'not again' type reaction to a second breach, but if the breached firm has put in place security measures following the first, such as hiring a CISO perhaps, then the markets could well have expectations of mitigated effects the second time around. The same would apply to other unfavourable events such as GDPR infringement fines. Considering a favourable security event such as that of a CISO recruitment announcement, one would naturally expect the initial impact on market value to be greater, reflecting the recognition by the organisation of the importance of security and willingness to invest. A routine replacement of a CISO would not naturally be expected to generate as much excitement in the market. A successful infringement fine appeal would certainly be expected to generate positive returns (favourable event), but exactly how those would compare with the market reaction to the initial data privacy violation (unfavourable) seems difficult to predict.

For convenience, **Table 24** shows a summary of the repeated events involved. The most repeat events seen within a single study was that of GDPR infringement fines with four companies suffering two fines. Only two companies reported more than one data breach and AIG was the only company reporting a CISO appointment more than once. Due to the US centric nature of the CISO appointment dataset, there is, naturally, very little overlap with other primarily European studies with only Barclays (BARC.L) being the exception in suffering a breach as well. It should be noted that data gathering for these studies was carried out independently and, therefore, there is not necessarily a direct relation or causality between the events and nor is the order in the table (by chapter) necessarily the order in which the events happened - this is explained in more detail later. Deutsche Telekom (DTE.DE) and International Airlines Group (IAG.L) both showed four events, however IAG was unique in spanning three different datasets due to their

initial data breach, GDPR infringement fine and subsequent GDPR fine appeal with three out of four events being related in this particular case (the Vueling fine was not directly connected to the British Airways events). The expectation here is that data breaches and infringement fines would be perceived by the market as negative (unfavourable) events and (successful) fine appeals and CISO appointments as positive (favourable). Of the total of 31 repeated examples identified, by far the majority (24 = 77%) would fall into the unfavourable category, therefore, with the possibility of adding in the Alphabet Inc (GOOGL) fine appeal example as well as that proved to be unsuccessful. On this basis, then, it would seem reasonable to expect an overall negative CAAR.

Table 24: Summary of repeated events by company (stock symbol)

Company	Chapter 4 (Breaches)	Chapter 5 (GDPR fines)	Chapter 5 (GDPR Fine appeals)	Chapter 6 CISO Appointments	Total
AIG	0	0	0	2	**2**
BARC.L	1	0	0	1	**2**
GOOGL	0	1	1	0	**2**
IBE.MC	0	2	0	0	**2**
INGA.AS	1	1	0	0	**2**
MAR	0	1	1	0	**2**
TSCO.L	2	0	0	0	**2**
UCG.MI	1	1	0	0	**2**
UTDI.DE	0	1	1	0	**2**
VOD.L	1	1	0	0	**2**
TEF.MC	1	2	0	0	**3**
DTE.DE	2	2	0	0	**4**
IAG.L	1	2	1	0	**4**
Total	**10**	**14**	**4**	**3**	**31**

From the literature review in Chapter 2, there was only one paper identified which focussed solely on the economic impact of repeated events[53] and that was Schatz and Bashroush (2016a) who built a dataset (using the PRC database) of 25 organisations

[53] Hovav and Gray (2014) also investigated repeated information security events, but these were multiple announcements all related to one original data breach.

suffering two data breach events. The events could not overlap i.e. they had to be far enough apart not to be confounding events and, actually, one example (CVS) involved a gap of almost seven years in between. The study found a tendency towards statistically significant negative returns for the earlier reported events (Group 1), although the dataset suffered from non-normality. For the later reported events (Group 2), the CAAR was much closer to zero (-0.16%) and not showing significance. Although it appeared that Group 1 showed a noticeably stronger negative reaction than Group 2, it could not be shown that it was significantly different i.e. "*we found merely weak statistical evidence in this study that the market reacts differently to a subsequent breach event affecting the same organisation*" (Schatz & Bashroush, 2016a).

Having such a small dataset was acknowledged as a limitation by the authors, who recommended revisiting the methodology in future once more data becomes available. Other limitations cited are date accuracy, potential unrelated confounding events and measures taken after the first event, such as a change of CISO, for example. Although not stipulated in the paper, it also has to be asked if Group 1 events were actually the first ever reported breaches for each firm. Additionally, it should be noted that no sectorial analyses were carried out, nor other contingency factors considered, such a type of breach which, as known from previous chapters, could affect CAR values. It should also be borne in mind that a trend was observed by i.a. Yayla and Hu (2011) that the market becomes less sensitive to breaches over time as confirmed by Ali et al. (2021) who comment on the volatility of unfavourable information security events (such as data breaches) versus favourable.

7.2. Results and discussion

Using data from Chapter 4 (data breach announcements) it was possible to build two sets as shown below (**Table 25**) using the same 5-day event window as Schatz and Bashroush (2016a).

Table 25: Data breaches repeated events (-2, 2)

Firm	Event Date	Event Window	CAR	t-value	Negative CAR
T-Mobile	2017-10-11	(-2, 2)	-0.0176	-0.8844	1
T-Mobile	2018-08-20	(-2, 2)	-0.0020	-0.1162	1
Tesco	2018-03-13	(-2, 2)	0.0185	0.6515	0
Tesco	2019-09-20	(-2, 2)	0.0237	0.9137	0

Only two example stock symbols were found where there were one or more breaches within the dataset, namely, Deutsche Telekom AG (T-Mobile) and Tesco. Clearly there is not enough data here to show any statistical significance, but T-Mobile seems to follow the expected pattern, as per Schatz and Bashroush (2016a), of a second breach having much reduced impact whereas Tesco seems to be on the rise during both breach announcements.

Table 26: Data breaches repeated events (0, 4)

Firm	Event Date	Event Window	CAR	t-value	Negative CAR
T-Mobile	2017-10-11	(0, 4)	-0.0185	-0.9296	1
T-Mobile	2018-08-20	(0, 4)	-0.0020	-0.1260	1
Tesco	2018-03-13	(0, 4)	-0.0164	-0.5775	1
Tesco	2019-09-20	(0, 4)	0.0175	0.6747	0

The aforementioned uncertainty in date caused Schatz and Bashroush (2016a) to use the 5-day window (-2, 2). The data used in Chapter 4 was already validated (including event date) thus it would not be essential to use the same window (no uncertainty in the date of announcement). Choosing a 5-day window beginning on the event date resulted in **Table 26**.

It can now be seen that despite there being no major change in T-Mobile, evidence of a negative impact has now been found in the first event for Tesco, but not the second, although note that the overall impact of the second is, although positive, still of greater magnitude than the first. Again, this is only referencing the data from Chapter 4 on data breaches and it should also be noted that Group 1 in this dataset was not necessarily the very first example of a data breach in that company.

Comparing the CAARs of Group 1 and Group 2 for these two event windows results in **Table 27**. It is evident that the choice of event window can make a major difference to the CAR, but there is, potentially, an overall trend of becoming more positive in Group 2.

Table 27: Comparison of CAAR for repeated breach events

Event window:	(-2, 2)	(0, 4)
Group 1	+0.0005	-0.0175
Group 2	+0.0109	+0.0078

There were also four instances of ultimate parent companies suffering multiple GDPR infringement fines from the dataset in Chapter 5. For this study, the 4-day event window was used (most negative on average) and the results are shown in **Table 28**. Note that these are different types of events from Schatz and Bashroush (2016a) who only studied breach announcements, therefore different market behaviour could be expected.

Table 28: Repeated events (GDPR infringement fines) event window (0,3)

Firm	Event Date	Window	CAR	Total	Group
IBE.MC	2019-10-16	(0, 3)	-0.0178		1
IBE.MC	2019-11-28	(0, 3)	-0.0329	-0.0507	2
IAG.L	2019-07-08	(0, 3)	0.0007		1
IAG.L	2019-10-01	(0, 3)	-0.0303	-0.0296	2
DTE.DE	2019-09-27	(0, 3)	-0.0137		1
DTE.DE	2019-12-18	(0, 3)	-0.0082	-0.0219	2
TEF.MC	2019-05-06	(0, 3)	0.0103		1
TEF.MC	2019-11-14	(0, 3)	-0.0019	0.0084	2
CAAR			-0.0117		
Group1			-0.0051		
Group2			-0.0183		

With the exception of Deutsche Telekom (DTE.DE), in all cases the Group 2 event is more negative, a mean CAR of -1.8% for Group 2 versus only -0.5% for Group 1.

For comparison purposes, the Schatz and Bashroush (2016a) methodology of event window (-2, 2) rather than that used in Chapter 5, was applied (**Table 29**) and the results do not alter the overall conclusion that Group 2 is more negative. Deutsche Telekom is showing positive, thus the prior 2 days to event were not relevant here due to the accuracy of the date for GDPR infringement fine announcements. This event window is not as effective in detecting abnormal returns, yet both windows seem to show that GDPR

infringement fine announcements are having a greater economic impact second time around, unlike breach events.

Table 29: Repeated events (GDPR infringement fines) event window (-2, 2)

Firm	Event Date	Window	CAR	Total	Group
IBE.MC	2019-10-16	(-2, 2)	-0.0186		1
IBE.MC	2019-11-28	(-2, 2)	-0.0322	-0.0508	2
TEF.MC	2019-05-06	(-2, 2)	0.0023		1
TEF.MC	2019-11-14	(-2, 2)	-0.0153	-0.0130	2
IAG.L	2019-07-08	(-2, 2)	0.0066		1
IAG.L	2019-10-01	(-2, 2)	-0.0034	0.0032	2
DTE.DE	2019-09-27	(-2, 2)	0.0009		1
DTE.DE	2019-12-18	(-2, 2)	0.0041	0.0050	2
CAAR			-0.0069		
Group1			-0.0022		
Group2			-0.0117		

Before investigating repeated examples across different studies, there was also one repeated example within the CISO appointment study (Chapter 6) so this is displayed in **Table 30**.

Table 30: Repeated events (CISO appointments) event window (-1, 1)

Firm	Event Date	Event Window	CAR	t-value	Negative CAR
AIG	2019-04-01	(-1, 1)	-0.0018	-0.0615	1
AIG	2019-07-15	(-1, 1)	0.0048	0.1848	0

Here, using the 3-day window as per Chapter 6, a negative return is seen for the first and positive for the second. Using the 5-day window as per Schatz and Bashroush (2016a) the results are shown in **Table 31**.

Table 31: Repeated events (CISO appointments) event window (-,2 2)

Firm	Event Date	Event Window	CAR	t-value	Negative CAR
AIG	2019-04-01	(-2, 2)	0.0026	0.0688	0
AIG	2019-07-15	(-2, 2)	-0.0111	-0.3332	1

By changing the event window, the signs have switched signalling a complete reversal in the CAR values. Naturally, looking at one single company over two events alone it is challenging to identify any clear trend in CAR.

Now that repeated events within each study have been examined, it would be interesting to study an overall impact on share price across multiple studies for all thirteen example firms (**Table 32**). Here the 5-day event window as per Schatz and Bashroush (2016a) is used. The first point to note here is that the overall CAAR is positive at 0.43% despite the introductory expectation of negative impact on market value due to the prevalence of unfavourable events. A closer look shows that the fine appeal announcements from Chapter 5 are having a major uplift overall on share price, offsetting negative bias due to unfavourable events. It was seen in Chapter 5 that even though the fine appeal from Alphabet Inc (GOOGL) was unsuccessful (expected negative), there was still positive CAR exhibited (ca. 2%). For the successful appeals of Marriott, UTDI.DE and IAG there are much larger CARs at ~5, 10 and 15% respectively. If those figures are compared with **Figure 9** for event study methodology in general, these examples appear to be of above average magnitude, and may be explained, of course, by COVID-19 market effects as they were after the original data date cap of 31/12/2019. Nevertheless, with such high magnitudes, GDPR fine appeals surely warrant more detailed investigation in future. On removal from the dataset, the CAAR changed to -0.66% which shows what a huge difference these four examples make.

Table 32: Repeated events (all) event window (-2, 2)

Symbol	Event Date	Window	CAR	Total	Origin
INGA.AS	2019-03-02	(-2, 2)	-0.0894		Breach
INGA.AS	2019-11-28	(-2, 2)	0.0023	-0.0871	GDPR
IBE.MC	2019-10-16	(-2, 2)	-0.0186		GDPR
IBE.MC	2019-11-28	(-2, 2)	-0.0322	-0.0508	GDPR
TEF.MC	2018-07-17	(-2, 2)	-0.0201		Breach
TEF.MC	2019-05-06	(-2, 2)	0.0023		GDPR
TEF.MC	2019-11-14	(-2, 2)	-0.0153	-0.0331	GDPR
DTE.DE	2017-10-11	(-2, 2)	-0.0176		Breach
DTE.DE	2018-08-20	(-2, 2)	-0.0020		Breach
DTE.DE	2019-09-27	(-2, 2)	0.0009		GDPR

DTE.DE	2019-12-18	(-2, 2)	0.0041	-0.0146	GDPR
AIG	2019-04-01	(-2, 2)	0.0026		CISO
AIG	2019-07-15	(-2, 2)	-0.0111	-0.0085	CISO
BARC.L	2017-07-27	(-2, 2)	-0.0069		Breach
BARC.L	2018-04-11	(-2, 2)	0.0073	0.0004	CISO
GOOGL	2019-01-21	(-2, 2)	-0.0144		GDPR
GOOGL	2020-06-12	(-2, 2)	0.0164	0.0020	Appeal
VOD.L	2019-09-25	(-2, 2)	0.0003		Breach
VOD.L	2019-11-06	(-2, 2)	0.0051	0.0054	GDPR
UCG.MI	2017-07-26	(-2, 2)	-0.0077		Breach
UCG.MI	2019-06-27	(-2, 2)	0.0294	0.0217	GDPR
MAR	2019-07-09	(-2, 2)	-0.0045		GDPR
MAR	2020-10-30	(-2, 2)	0.0455	0.0410	Appeal
TSCO.L	2018-03-13	(-2, 2)	0.0185		Breach
TSCO.L	2019-09-20	(-2, 2)	0.0237	0.0422	Breach
UTDI.DE	2019-12-09	(-2, 2)	-0.0144		GDPR
UTDI.DE	2020-11-12	(-2, 2)	0.1039	0.0895	Appeal
IAG.L	2018-09-06	(-2, 2)	-0.0238		Breach
IAG.L	2019-07-08	(-2, 2)	0.0066		GDPR
IAG.L	2019-10-01	(-2, 2)	-0.0034		GDPR
IAG.L	2020-10-16	(-2, 2)	0.1459	0.1253	Appeal
CAAR			0.0043		

The 'biggest loser' in the list (**Table 32**) is ING (INGA.AS) – again demonstrating the propensity for greater market reactions to financial services sector companies, although it is surprising the infringement fine was received slightly positively to offset the breach a little. Closely following ING is Iberdrola (IBE.MC) – an energy utilities company – another highly regulated sector. Interestingly, Barclays (BARC.L) managed to offset all losses from their data breach by subsequently hiring a CISO and were the only company appearing in both the Chapter 4 and Chapter 6 datasets.

The overlap between data breaches (Chapter 4) and GDPR events (Chapter 5) comprised six firms: Vodafone (VOD.L), IAG, Deutsche Telekom (DTE.DE), Telefonica (TEF.MC), UniCredit (UCG.MI) and ING Group. Three companies exhibited negative

CAAR (ING, Telefonica and Deutsche Telekom) overall as expected. Vodafone and UniCredit showed (unexpected) positive CAAR over this event window as did IAG, although IAG was massively offset by the positive market reaction (ca. 15%) to the subsequent successful GDPR infringement fine appeal, this being the only example appearing in all three datasets (Chapters 4, 5 and 6). Disregarding the appeal (due to COVID-19 market volatility), IAG would also be negative CAAR overall. For clarity, a visualisation is shown in **Figure 19**.

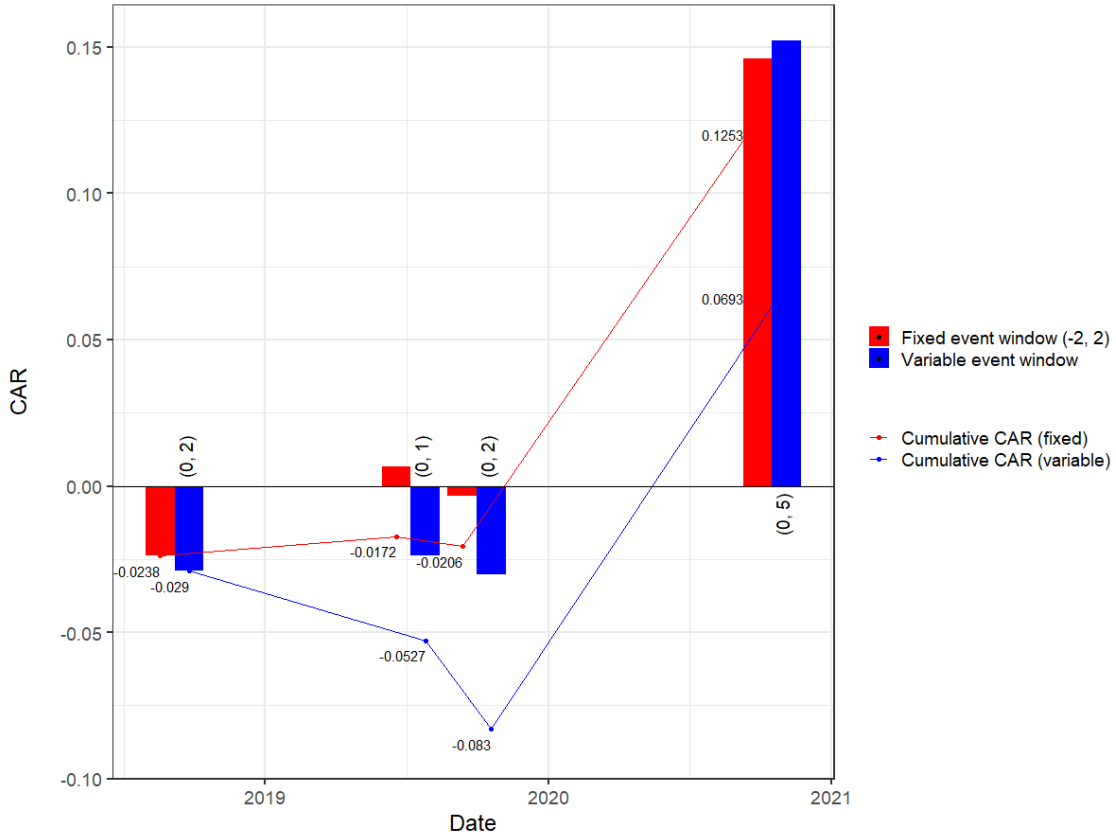


Figure 19: Repeated events for IAG

As the information security events here vary in nature, not only are the results displayed for the Schatz and Bashroush (2016a) choice of event window (-2, 2) but also a variable selection of event window which showed the highest magnitude CAR over all event windows analysed[54], arbitrarily of length less than 10 days on the basis that ESM is more effective at detecting abnormal returns over shorter time periods. Indeed, Ali et al. (2021) in their SLR reported that *"most event windows extending beyond two days of an event are insignificant"* and that 55% of studies used event windows (-1, 1), (0, 1) and (0, 2)

[54] In the event multiple windows showed the same CAR, the earlier value was taken to correctly reflect the market efficiency (speed of reaction). For example, for the Vueling fine announcement of 01/10/19 a CAR of -0.303 was calculated for both the (0, 2) and (0, 3) event windows.

which seems to be consistent with the findings here, if the highly positive reaction to the (successful) British Airways GDPR fine appeal post-COVID (announcement date 16/10/20) with the six-day window (0, 5) being the greatest is excluded. Comparing the variable and fixed event window approaches, there results are visibly similar for both the initial British Airways data breach (06/09/18) and the GDPR fine appeal, however, in the case of the two infringement fine announcements (British Airways on 08/07/19 and Vueling on 01/10/19)[55] it can be seen that the fixed window approach was less effective in detecting negative market reaction. Whereas the Schatz and Bashroush (2016a) displays a cumulative effect which stays around the -2% mark after three events, the variable window approach shows cumulative losses of over 8% in market capitalisation which, of course, makes a much larger dent in the positive market effect of the later successful fine appeal. Usage of the (-2, 2) window seems to have picked up on pre-event market optimism or, perhaps, other confounding events. This finding does highlight the importance of the choice of event window in ESM studies as highlighted by i.a. Tweneboah-Kodua et al. (2018) and Ali et al. (2021), especially here where three different types of information security events are being analysed. It is also worth noting here that the three British Airways events were all related (data privacy breach, subsequent GDPR infringement fine and later appeal) whereas the Vueling fine was for a non-compliant cookie policy on its website (CMS Legal, 2021) and so did not appear in the Chapter 4 (breach) dataset.

7.3. Conclusion

Although Schatz and Bashroush (2016a) reported weak evidence that a second breach announcement had a different (lesser) impact on market value and recommend revisiting once a larger dataset of breaches becomes available, this chapter analyses a much smaller sample size within each study (versus 25 in each group), so more solid statistical evidence would surely be challenging here. Nevertheless, in the case of breaches, there is a tendency to be less negative in Group 2. For the GDPR fine announcements, there seems to be different market behaviour. These appear to react more strongly to second infringements and even more strongly again to a GDPR fine appeal announcement, be that a positive or negative outcome, with the caveat that these appeal market gains were observed during a period of market instability due to COVID-19. Nevertheless, this is a clear pointer to future research and these results seem strong enough to warrant further investigation. Analysis of the single repeat CISO appointment announcement was

[55] Both subsidiaries of International Airlines Group (IAG).

inconclusive and shows that the choice of event window is important and choosing a one-size-fits-all approach when different industry sectors and event types are involved may not be ideal.

The special case of IAG experiencing four information security events spanning all three datasets was an interesting one. After suffering a data breach followed by two infringement fines, the cumulative effect was a drop of 8.3% in share price (using the variable window approach). Based on the market capitalisation of IAG around that time (see **Table 11**) that would correspond to a loss in market value of €860m. The total value of fines levied was €205m giving a total negative financial impact of ca. €1.1bn – an amount certain to raise some eyebrows at board level. Although the subsequent fine appeal market response showed as positive (although not statistically significant) and appeared to more than offset these losses as the share price rallied, it is worth recalling the EMH and the words of Telang and Wattal (2007): *"in the absence of the event, the stock price of the firm at any time would have been higher".*

Regarding the successful GDPR infringement fine appeal of British Airways, a reason cited by the ICO for the reduction in monetary penalty from £183m to £20m was *"that BA has also implemented a number of remedial technical measures so as to reduce the risk of a similar attack in future, and has indicated that expenditure on IT security will not be reduced as a result of the impact of COVID-19"* (ICO, 2020: 73). There were other factors at play, of course, such as the representation by BA that *"the* [original] *amount of the fine is not 'effective' because issuing large fines is likely to be counterproductive"* (ICO, 2020: 78), which all point to the importance of investment in information security measures to maintain the CIA triad.

The next chapter (Chapter 8) presents an overview of literature in this area with the aim of giving pointers to firms looking to improve their cyber security posture.

Chapter 8. Investment in Information Security

8.1. Introduction

The core chapters (Chapters 4, 5 and 6) of this work have investigated the impact of both favourable and unfavourable events on the market value of organisations. This chapter now focusses on RQ4, concerning how organisations might incorporate the findings from these chapters into their investment strategies with a view to improving their cyber security posture, thereby avoiding unfavourable information security events such as data breaches or infringement fines. A good starting point is to revisit the literature review from Chapter 2 through the lens of security investment and supplement this with additional material in this subject area.

8.2. Related work and discussion

The study of the economic impact of data breaches in Chapter 4 revealed no clear evidence of a decrease of share price in European markets following a breach announcement, with the exception of Spain. That said, many examples were cited in Chapter 2 where statistically significant negative impacts were identified and, indeed, Spanos and Angelis (2016) in their SLR quote a figure of 76% falling into that category, mostly US based. It appears though, that markets have become less sensitive to breach announcements over time as confirmed by Ali et al. (2021) in their (later) SLR which may go some way towards explaining the difference observed between European and US markets. Indeed, Richardson et al. (2019) question whether their US market-based findings (of a *"lack of"* economic impact) can support business cases for security investment, although they do recognise certain extreme, catastrophic cases of major data leaks where the breach could result, ultimately, in the demise of the company as was seen with Travelex in Chapter 4. This should send a powerful message to CEOs and CFOs of the need to improve their security, remembering the Warren Buffet quotation from Chapter 2 that *"Predicting storms doesn't count; building arks does."* (Morse et al., 2011).

Although Campbell et al. (2003) reported that breaches involving sensitive data were more likely to invoke negative market reactions, the investigation in Chapter 4 was not able to confirm this; even though the CAAR of data privacy relevant breaches was lower than that of the whole dataset, the results were not statistically significant. What is more of a help in investment justification though are the findings from Chapter 5 regarding infringement fines following a data privacy breach. Here, clear evidence, statistically

significant at the 1% level, was found of a negative CAAR of 1% in the few days around a GDPR infringement fine announcement. This loss of market value was found to far outweigh the monetary value of the fine, on average being 29,000 times costlier. The case of multiple (repeated) unfavourable events experienced by IAG reported in the previous chapter (Chapter 7) resulted, potentially, in huge losses totalling €1.1bn – a figure sure to grab the attention again of the CEO and CFO and certainly helpful in building business cases for investment for organisations processing personal data.

If negative evidence alone is insufficient to be persuasive, then an alternative approach would be to utilise evidence of positive returns from favourable information security events to justify investment. It has been seen from Chapter 2 that, for example, investment in security certification (Bose & Leung, 2013; Deane et al., 2019) can result in significant positive abnormal returns for companies, which would seem to be a more direct method of gaining business case approval over multiple 'horror stories' of what happened to other organisations who failed to protect their crown jewels[56]. Indeed, Moore, Dynes & Chang (2015) report that this ROI approach is feasible and clear evidence was found in Chapter 6 ('The CISO Effect') of a rise in market value of 0.8% for companies announcing the appointment of a CISO type role. The effect was more significant within the financial services sector, and, with the expectation of reaping benefits in the range US$94m to US$318m, it seems easy to justify the cost of such a hire based on ROI alone. Nevertheless, there are examples of where 'favourable' security events yielded unexpected results such as that of ISO/IEC27001 certification announcements reported by Malliouris and Simpson (2020) who found that this news was not well received by the market and, in actual fact, had a slightly negative effect. Perhaps the market perception here was that of overinvestment in security, a concept supported by Srinidhi, Yan and Tayi (2015) who argue that, due to misalignment between managers and investors, there is a tendency for managers to overinvest in security measures to avoid serious security incidents during their limited tenure (cf. Banker and Feng (2019) who showed evidence that data breaches involving system deficiencies led to increased CIO turnover). One cynical security professional disagrees with this, however, and viewed the route to being a successful CISO as having incidents and managing them well – *"if you have a clean sheet, nobody's interested"* (Schatz & Bashroush, 2018: 9).

[56] *"Data Is the New Oil of the Digital Economy"* (Wired, 2014)

The concept of overinvestment nicely segues into the next question which is how much to invest – referred to as the "*cyber security investment challenge*" by Fielder et al. (2016: 13) – identifying an effective decision-making strategy. An SLR in the area of economic valuation for security investment was carried out by Schatz and Bashroush (2016b) who report the challenges of security investment in general, and find, in contrast to Moore et al. (2015), more focus around reduction of risk rather than revenue generation, stating that *"security measures aim to reduce loss and not commonly generate revenue"* (Schatz & Bashroush, 2016b: 2).

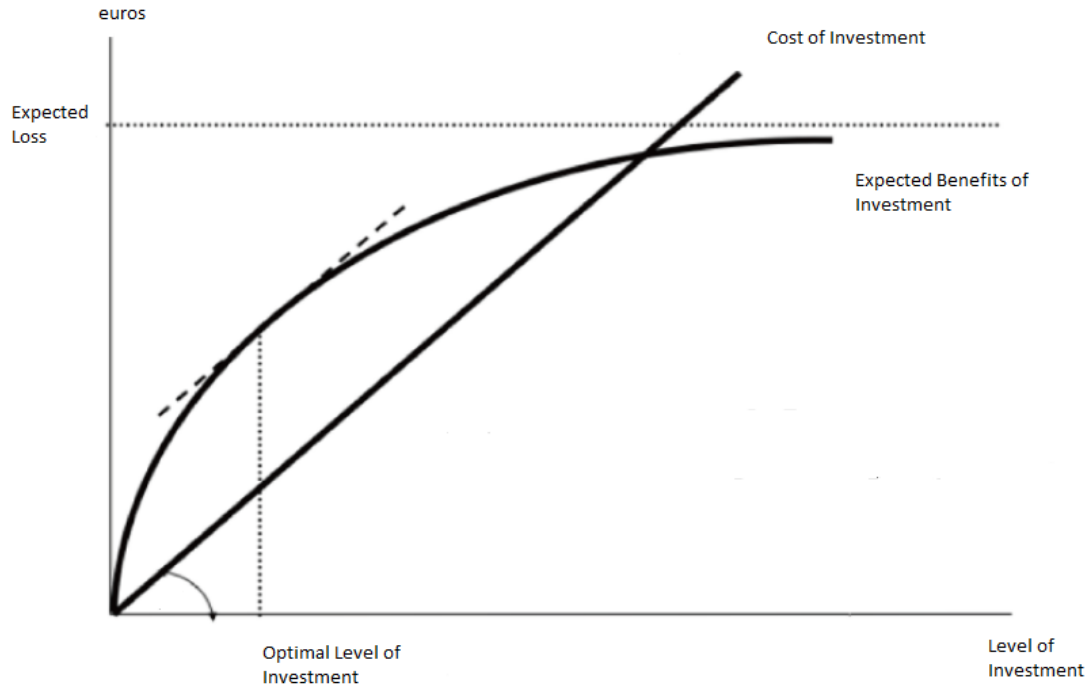


Figure 20: Gordon-Loeb investment model
(Source: https://en.wikipedia.org/wiki/Gordon-Loeb_model. Accessed on: 04/04/23)

Gordon and Loeb (2002) developed a mathematical model (**Figure 20**) for information security investment which recommends the maximum amount that a risk-neutral firm should spend should not exceed 37% (1/e) of the expected losses. This work has since been extended (Gordon, Loeb & Zhou, 2016) to give a more practical perspective. A further study (Gordon et al., 2018) indicated positive findings on private firms investing in cyber security, citing the reasons as being: incorporation of their security investment into financial reporting, internal control systems and the (reduced) risk of loss.

The question "*How much is enough?*" was also asked by Hoo (2000). This thesis proposes a quantitative decision model in which parameters, such as the existing security posture and costs to implement, could be input to calculate a value for optimal investment.

Furthermore, Hoo (2000) shows a worked example using this model where the optimal investment figure for a large (10,000+ employees, US$500m+ turnover) high-tech company was calculated as US$327,500 based on the avoidance of possible losses of US$2.811m. The optimal spend to expected loss ratio is, therefore, 12% which is lower than the 37% predicted by the Gordon and Loeb (2002) model, although there is no suggestion that the caveat of a risk neutral firm has been met here. Indeed, the author recognises that the model would need to (and could) be adjusted to reflect other factors such as industry sector with the optimal security posture for a financial services firm, for example, requiring a very different investment profile, a sentiment echoed by Schatz and Bashroush (2018: 6) who opine that *"An online retail business cares more about the availability of their web services than a brick and mortar business would. Likewise, such a business would be more concerned with potential reputational impact should a breach occur, which further impacts the way information security spending is prioritised"*. Hoo (2000) also states that cost estimates for implementation in this example were at the lower bounds, thus the figure of US$327,500 is optimistic.

To complement these primarily theoretically based studies (Hoo, 2000; Gordon & Loeb, 2002) a recent, more practical, qualitative study into security investment was carried out by Schatz and Bashroush (2018). Here the authors interviewed a (primarily UK based) sample of security professionals to understand existing practices in investment decision making. A key finding was that *"it is uncommon for security practitioners to apply accounting performance metrics such as NPV, ROI, IRR etc. Rather, investments tend to be pre-allocated through means of annually assigned budgets attached to risk-based performance metrics without further hurdle rate requirements. Notable exceptions to this practice where ad-hoc requirements arise from incidents or specific business demands were found"* (Schatz & Bashroush, 2018: 16). This tendency for an anecdotal type of approach to investment decision making was also well noted much earlier by Hoo (2000), predicting a return to a more quantitative risk management type approach in future based on such driving forces as the cyber insurance market. Based on the findings of Schatz and Bashroush (2018) then, it appears there has not been much progress in this prediction of Hoo (2000) over the last two decades, with budgets continuing to primarily be driven by the previous year, industry best practices or a must-do approach (Gordon & Loeb, 2006; Schatz and Bashroush, 2016b).

The aforementioned Hoo (2000) model also went a stage further in suggesting how the optimal budget might be allocated to specific security measures, in other words, not just

'how much?' but 'what to buy?' as well. A total of twelve security measures (or "*safeguards*") are considered for adoption and of those twelve, only three were recommended as part of the optimal (US$327,500) investment strategy, specifically Screen Locking Software, Communications Content Screening, and Intrusion Detection System.

A decision support tool to inform security budget allocation was also developed by Fielder et al. (2016) through a combination of game theory, combinatorial optimisation and a hybrid of the two. The authors found that the advice given by this tool was consistent with the UK Government's Cyber Essentials scheme (NCSC, 2022b)[57] which recommends five controls, namely Firewalls, Secure Configuration, User Access Control, Malware Protection and Security Update Management.

Evidently, the output of these two decision support tools is different, despite some degree of overlap (Communications Content Screening included in Firewall, Screen Locking Software is part of Secure Configuration). The difference may be attributed to, in part at least, the fact that, whereas the Hoo (2000) example was cited as a 'large' US based high-tech company, the Fielder et al. (2016) study was based on the network design of a typical SME, thereby implying an expected much lower level of maturity in its existing cyber security posture and, consequentially, a greater number of security measures needed to reach the optimum.

The nine remaining safeguards not selected in the Hoo (2000) optimal model example are: Security Awareness, HW/SW Network Upgrade, Response Team, Nightly Back-ups, Encryption, Central Access Control, Security Management Team, Anti-Virus Software and Intrusion Detection System. So how do these controls compare with those of the Cyber Essentials (NCSC, 2022b)?

Certainly, the Cyber Essentials control of Security Update Management includes HW/SW Network Upgrade (commonly referred to as 'patch management'). The Cyber Essentials stance on nightly back-ups is as follows: *"Backing up your data is not a technical requirement of Cyber Essentials; however we highly recommend implementing an appropriate backup solution."* (NCSC, 2022b: 14). This recommendation is actually documented in the *"Further Guidance"* section, which seems surprising considering the increased threat of i.a. ransomware. However, as NCSC (2022c) themselves state: *"We're*

[57] A later version than that cited by Fielder et al. (2015) but the core five controls remain the same.

frequently asked about backups and why we don't include them in the Cyber Essentials controls. It's certainly not because we don't think they are important. The main reason is that we don't want to overload organisations in the certification process. Implementing the controls properly makes your organisation a harder target for common types of cyber attack, such as ransomware, and therefore reduces the criticality of backups". The other Hoo (2000) safeguards of Central Access Control, Anti-Virus Software and Intrusion Detection System are all in scope for Malware Protection and Firewalls under Cyber Essentials. Although Hoo (2000) also mentions Encryption as the final technical control, this is not directly reflected in Cyber Essentials, although reference is made to VPN in the context of home working (a COVID-19 driven revision). Presumably, it is assumed that the VPN in question encrypts the traffic (in transit) although this is not explicitly stated, nor is there any requirement for 'at rest' data encryption.

The only Hoo (2000) safeguards which now remain to be discussed are Security Awareness, Response Team and Security Management Team. All of these three controls are focussed on the 'human factor' whereas Cyber Essentials professes to be a list of technical control themes. Nevertheless, User Access Control under Cyber Essentials does include a human element, that of user education: *"Educating people on how to avoid common or discoverable passwords"* (NCSC, 2022b: 11). Hoo (2000) positions this as a more comprehensive user awareness training programme, yet it still remains at the lower end of the cost scale to implement (it was not incorporated in the optimal case study due to the level of maturity of cyber security and high-tech nature of the organisation). Such end-user training is referred to as building a *"human firewall"* (see i.a. Whitman et al., 2005). Incorporating such a (more detailed) end-user awareness training programme would certainly seem to be an attractive option for any, especially limited, security budget. Indeed, employee training and awareness forms part of e.g. the ICO's Accountability Framework (ICO, 2022), an example for how legislation (in this case the GDPR) and regulators can drive investment decision making. That said, relying on such regulatory input alone may not yield the best results - as Schatz and Bashroush (2018) observe: *"Conventional budgeting approaches cause information security departments to direct their funds towards a 'minimum protection/maximum compliance' strategy rather than initiatives that contribute the most value to the organisation"*. Regarding awareness training it should also be borne in mind that *"after a while, a security poster, no matter how well designed, will be ignored; it will, in effect, simply blend into the*

environment. For this reason, awareness techniques should be creative and frequently changed" (NIST, 1995).

Finally, the remaining two safeguards cited by Hoo (2000), Response Team and Security Management Team have been covered by Chapter 6 which shows clear ROI on the appointment of a CISO type function, despite this being one of the higher cost items (and thus not included in the optimal strategy) of the worked example (Hoo, 2000) – nor is it included in the optimal list of controls of Fielder et al. (2016) who were targeting SMEs (unlikely to be large enough to justify dedicated security teams) and basing the list of controls on Cyber Essentials which, as stated above, included technical measures only.

8.3. Conclusion

It appears the subject area of information security investment is not straightforward and, as Schatz and Bashroush (2018: 16) remark, *"security investments are made in context of a highly complex organisational system relying on a wide range of unique business environment factors"*. This topic is certainly worthy of an entire thesis in its own right – this chapter barely scratches the surface, so to speak, but is included to show how the studies described in Chapters 4, 5, 6 (and 7) are capable of aiding security practitioners by adding value to investment decision-making processes in organisations. Chapters 4 and 5 can inform the risk-based approach, whereas Chapter 6 is more aligned with ROI-based justifications. Although Chapter 4 results were somewhat inconclusive, Chapter 5 very much stresses the importance of investing in security where sensitive data is being processed, especially, since the advent of the GDPR. When building business cases, practitioners should not only consider the monetary value of an infringement fine itself, but also the much (ca. 29,000 times) larger potential drop in company market value. These findings would surely be an aid to security investment approval as would those of Chapter 6 which showed a clear ROI on recruiting a CISO.

In Chapter 7, the specific case of IAG was highlighted which identified a potential downside (risk) of the order of €1.1bn following a data breach and two subsequent infringement fines under the GDPR. Applying the Gordon and Loeb (2002) model in its most basic form[58], the optimal level of investment would be 1/e of this figure, or €405m which is ca. £350m at current exchange rates. Although this figure may seem high, the subsequent appeal to ICO for a reduction was successful and partly justified[59] due to

[58] Assuming IAG is a risk-neutral firm.
[59] Aside from other representations made by British Airways, COVID-19 also played a significant part.

evidence of commitment from British Airways to a continuing investment programme in information security. The reduction was around the 90% mark from £183m to £20m. It is not known here what the level of investment was that British Airways committed to, but the £163m difference alone would sit at the kind of optimal level (12%) which Hoo (2000) was suggesting in the worked example of a larger high-tech company with a more mature cyber security posture. One of the reasons cited by the ICO for the initial punitive action was the failure of British Airways to address the basics – clearly their cyber security posture was not as mature as one would expect – so the advice of Fielder et al. (2016) to focus on Cyber Essentials, despite being targeted at SMEs, would certainly have been of benefit in this case, in hindsight. That said, further good advice would be to consider investing in back-ups and an enhanced, frequently refreshed end-user security awareness training programme.

Another key message here is how DPAs can influence company information security investment strategies for the greater good by applying new legislation mindfully, not just levying huge fines as a deterrent, but instead ensuring the money is put to good use, as suggested also by i.a. Nieuwesteeg and Faure (2018) and underpinned by the British Airways case study in Chapter 7.

Chapter 9. Overall Conclusion and Contribution to Knowledge

9.1. Introduction

The final chapter of this thesis begins, for convenience, with a summary of the chapter content, followed by a reflection of the research questions and how these have been answered, along with the contribution to knowledge. Finally, challenges encountered during the research (limitations) are described before concluding on pointers to future research.

9.2. Summary

Chapter 1 began by giving some background on the importance of information security and why it should be a major concern at board level. Research aims and objectives were developed into research questions and ESM was introduced as the methodology. The expected contribution to the knowledgebase was also mentioned along with an explanation of the thesis structure and a list of publications arising from this research.

The literature review is detailed in Chapter 2 beginning with definitions of information and cyber security related terms as well as econometric and financial definitions for convenience. From these definitions, search strings were developed over multiple iterations, and it was found that there was not a huge body of research in existence regarding the economic impact of information security events in general, and what was available was very US centric. It was also identified that the most frequently used method for quantitative analysis of such events was ESM, hence this has been the focus here, along with a UK/EU bias in data selection, particularly concerning data breach announcements (see Chapter 4). Due to the relatively recent introduction of the GDPR there was also little literature in existence regarding this legislation in general, another gap identified (see Chapter 5). As Ali et al. (2021) reported, another deficit in the literature was that of favourable information security events, which led to the research in Chapter 6 (“The CISO Effect”).

A detailed review of ESM, including mathematical models, is given in Chapter 3 and the software package of choice, EST, is introduced along with the data gathering approach, the R code used for the analyses and hypothesis development. Also, some comparisons were run of the favoured market model versus FF3FM as well as EST versus literature and, in both cases, the results were found to be very similar, in fact, virtually identical for the EST test case.

In Chapter 4, the impact on company market value of data breach announcements (a hand-gathered data set of 45 examples) in UK/EU markets was investigated using ESM and the results found to be inconclusive overall, with the notable exception of Spain.

A similar approach was employed in Chapter 5, this time analysing a dataset of 25 GDPR infringement fine announcements downloaded from the GDPR Enforcement Tracker (CMS Legal, 2021). Statistically significant CARs of 1% were found up to three days after the event with the Spanish and Romanian markets being particularly sensitive. It was also found that the drop in market value was way greater than the monetary value of the fine itself, actually ca. 29,000 times larger on average. GDPR fine appeals were also investigated, and although the results were not statistically significant, this is an area which certainly warrants future research.

As Chapters 4 and 5 had focussed on unfavourable security events, Chapter 6 used a similar approach to investigate the (anticipated positive) impact of CISO appointment announcements on market value ('The CISO Effect'). A dataset of 37 events was analysed and, indeed, the effect was found to be positive, actually a CAR of +0.8% on average over the three days surrounding the event, with stronger market reaction (+1.8%) for the financial services sector, being statistically significant at the 1% level.

Chapter 7 examined the impact of repeated events existing between the datasets of Chapters 4, 5 and 6. Due to the small sample size, no statistical significance could be shown, but the general trends identified were reduced market reaction for a second data breach announcement whereas GPDR infringement fines exhibited stronger market reactions for subsequent events, including a fine appeal be it successful or not.

The previous chapter (Chapter 8) aimed to give security investment advice to organisations based on the output of the preceding chapters and incorporates a supplementary literature review of the topic.

9.3. Reflection on research questions

The first research question (RQ1) asked *"What is the impact (if any) on share price of a security event, be it favourable or unfavourable and how do these findings compare with the literature?"*. The results from the analysis of data breach announcement in Chapter 4 were generally inconclusive with the notable exception of the Spanish market. In light of the strong US bias in existing literature, does this mean the European markets behave

differently? Well, not necessarily, as recent US based studies such as Richardson et al. (2019) have also found the market to have become less sensitive to data breaches over time as also observed by e.g. Yayla and Hu (2011), with examples from both US and UK/EU ranging from catastrophic failures to market indifference. Findings from Chapters 5 (GDPR infringement fines) and 6 (CISO recruitment announcements) both showed statistically significant abnormal returns, although infringement fines will be covered in more detail under RQ3 below (a drop in market value of around 1% on average). The favourable 'CISO Effect' reported in Chapter 6 showed an uplift of 0.8% of share price on average, a similar order of magnitude to that observed by Chatterjee et al. (2001) for newly created CIO positions (+1.16%).

Secondly (RQ2), the question *"Are there any patterns in the data, such as correlations between drop in market value and category of cyber-attack, data breach, industry sector etc.?"* was asked. Due to the inconclusive nature of the data breach study in Chapter 4, it was of course, challenging to identify anything other than indicative trends through cross-sectional analyses, although it was noted that the European financial services sector seemed to respond less rapidly to breach announcements than its US equivalent. Sectorial analyses for infringement fines were only weakly significant, although geographically it was the Spanish and Romanian markets which were shown to be less tolerant of GDPR fines than others. The 'CISO Effect' was more marked for the financial services sector, showing an uplift of 1.8% and statistical significance at the 1% level. Again, due to the small dataset, other sectorial analyses were mostly inconclusive with just indicative trends for future research.

In response to the third question (RQ3) *"Regarding the introduction of the GDPR, what is the economic impact of infringement fines on the market value of firms, including those appealed and overturned?"* from Chapter 5 it was seen above that a drop in share price of around 1% was observed overall in the three days following the event. What is of particular interest here is how the effect on market value far outweighs the monetary value of the fine itself, typically by a factor of around 29,000, so this should be the major concern for organisations. In the study of repeated information security events (Chapter 7), it was seen that in the rather unfortunate case of British Airways, an initial data breach followed by two infringement fines resulted in a loss of 8.3% of market value. Coupled with the initial monetary value of the fines themselves, the resulting losses were of the order of €1.1bn. Although some investigation was done into fine appeals, there were only four examples all occurring during a period of COVID-19 market instability. That said,

there was some evidence found in Chapter 7 of the markets reacting more strongly to subsequent events including the positive impact of a successful appeal, certainly something to re-examine in future once more data becomes available.

The final question (RQ4) asked *"How can these findings be incorporated into the security investment strategies of organisations?"*. Such security investment advice is discussed in Chapter 8. Although the material in Chapter 4 (data breaches) was mostly inconclusive, there is clear business case support from the GDPR infringement fine findings and, in particular, the British Airways case as highlighted in Chapter 7. The positive 'CISO Effect' is also supportive of investment in human capital as covered in Chapter 6. Regarding what to invest in, practitioners are reminded not to lose sight of the basics (e.g. Cyber Essentials) and to consider data back-ups and a comprehensive end-user training programme.

9.4. Contribution to knowledge

The literature review in Chapter 2 showed that there was not a huge knowledgebase concerning the economic impact of information security events in general, and what did exist tended to be very US centric. This research goes some way to offsetting that bias with the UK/EU centric study on data breaches in Chapter 4 and the GDPR infringement fine impact study in Chapter 5. Due to the relatively recent introduction of the GDPR (2018), studies in this area were, naturally, few and far between at the time of writing. Another gap in literature identified was that concerning CISOs (and similar roles) in general (Karanja & Rosso, 2017) which is addressed by Chapter 6, along with the recognised lack of studies focussing on favourable information security events (Ali et al., 2021) as CISO recruitment was found to yield clear positive returns. Studies regarding the impact of repeated data breaches were also found to be lacking (Schatz & Bashroush, 2016a) and Chapter 7 contributes to this area. The paucity of literature in the area of company security investment strategy and budget allocation is identified and addressed in Chapter 8.

Not only have these studies contributed to the economics of information security events knowledgebase as above, but also there is some useful input here into ESM in general, such as highlighting the importance of the choice of event window and market reference, the approach to filtering confounding events and comparison of the MM and FF3FM. During the literature review (Chapter 2), no existing studies in this area were identified which explicitly stated the use of EST as the analysis software package, so this approach

seems novel also, thus a comparison of EST with literature was carried out in advance as shown in Chapter 3. The R code used for the EST analyses is shown in the Appendix thus could easily be re-engineered for future studies of this type.

This research should encourage firms not only to invest in information security but also to invest in an optimal manner. Publicly listed firms are also encouraged to be transparent about their investment in security measures through key findings here such as highlighting the potential negative effects of security breaches and data privacy fines whilst reinforcing the benefits of improving cyber security postures. Regulators should also be influenced to mandate disclosure of such information in statutory reports.

As mentioned previously, this research would be of benefit to business management, managers and practitioners of cyber security, investors and shareholders and policy makers as well as researchers in cyber security or related fields.

9.5. Research limitations

Although some idea of the challenges encountered have been highlighted in the previous section, it is worth reiterating the major limitations here, mainly the COVID-19 pandemic and the lack of a comprehensive breach database for Europe. The market instability caused by the pandemic reduced the amount of, already limited, data available as date cut-offs of 31/12/2019 had to be observed. Fortunately, as this study was mostly internet based, lockdown restrictions did not directly hinder progress. The lack of a comprehensive breach database for Europe led to time-consuming hand-gathering of announcements (Chapter 4) and is surely a contributing factor to the observed propensity for US based studies of this type, along with the greater maturity of the US markets regarding data breach notification legislation.

Other limitations encountered (Chapter 2) included the recent introduction of the GDPR (2018) which, naturally, reduced the knowledgebase in this area, and perhaps to some degree the use of English language only literature searches, although the set of matches (papers) returned compared well with other studies of this type and any restriction was effectively mitigated through use of the backward snowball technique.

It should also be remembered that the scope of this thesis is naturally limited to publicly listed companies, to facilitate the measurement of market reactions. Of course, privately owned enterprises, SMEs, partnerships, cooperatives, not-for-profit organisations, charities, educational institutions, governments, NGOs, armed forces and others, also

experience security incidents and, although not needing to be concerned about the impact on their share price of a data breach, for example, nevertheless incur costs identifying and containing the breach, notifying the relevant authorities, paying any infringement fine and carrying out any post-breach mitigating actions concerning lost business and reputational damage. Such costs would include both internal effort as well as that of external consultants where needed. A report by IBM Security (2022: 5) cites the average cost of a data breach to an organisation as US$4.35m through such an activity-based costing approach. A publicly listed company has not only these costs to contend with, but also any impact on market value over and above, and that is the focus of this research, additional evidence to support security investment business cases.

9.6. Pointers to future research

In Chapter 4, the lack of a comprehensive breach database for Europe was identified. Such resources are readily available in the US, most likely due to the maturity of disclosure legislation in this area and the benefits of information sharing to avoid future occurrences. Further research on this topic is surely justified. Certain markets, specifically Spain and Romania, were found to react more strongly to information security events (see Chapters 4 and 5), possibly due to having particularly active DPAs. Again, further research is warranted once more data becomes available.

Some (weak) evidence was found of increasing magnitude of abnormal returns concerning GDPR fine appeals (Chapter 5) which, although occurring during a period of market instability (COVID-19), seems again to warrant more investigation in future.

In fact, as documented above, these studies were all hampered, to some degree, by market effects of the pandemic and thus would all benefit from being revisited in future once the markets re-stabilise after the pandemic.

References

Aiken, M. (2017), *The cyber effect: An expert in cyberpsychology explains how technology is shaping our children, our behavior, and our values--and what we can do about it*, Random House

Alam, M.M., Wei, H., Wahid, A.N.M. (2020), *COVID-19 outbreak and sectoral performance of the Australian stock market: A event study analysis*, Australian Economic Papers, **60**(3), https://doi.org/10.1111/1467-8454.12215. Accessed on: 24/02/22

Ali, S.E.A., Lai, F.-W., Dominic, P.D.D., Brown, N., Lowry, P.B., Ali, R.F. (2021), *Stock market reactions to favorable and unfavorable information security events: A systematic literature review*, Computers & Security, 110, 102451

Amir, E., Levi, S., Livne, T. (2018), *Do firms underreport information on cyber-attacks? Evidence from capital markets,* Review of Accounting Studies, **23**(3), 1177-1206.

Andoh-Baidoo F.K., Amoako-Gyampah K., Osei-Bryson K.M. (2010), *How Internet security breaches harm market value*, IEEE Security and Privacy **8**(1), 36–42

Banker, R., Feng, C. (2019), *The Impact of Information Security Breach Incidents on CIO Turnover*, Journal of Information Systems, **33**(3), 309–329

BBC (2013), *Sony fined over 'preventable' PlayStation data hack*, https://www.bbc.co.uk/news/technology-21160818. Accessed on: 30/03/21

BBC (2016), *TalkTalk fined £400,000 for theft of customer details*, https://www.bbc.co.uk/news/business-37565367. Accessed on: 26/04/21

Bendovschi, A., Al-Nemrat, A., Ionescu, B. (2016), *Statistical Investigation into the Relationship between Cyber-Attacks and the Type of Business Sectors*, International Journal of Business, Humanities and Technology, (6)1, 49-61

Benninga, S. (2008), *Financial modeling (3 ed.)*, Boston, MA, MIT Press

Bloomberg (2018), *IAG Share Price Hit After British Airways Data Hack*, https://www.bloomberg.com/news/articles/2018-09-06/british-airways-says-hackers-stole-customers-credit-card-data. Accessed on: 16/09/22

Bose, I., Leung, A.C.M. (2013), *The impact of adoption of identity theft countermeasures on firm value*, Decision Support Systems, **55**(3), 753-763.

Bose, I., Leung, A.C.M. (2014), *Do phishing alerts impact global corporations? A firm value analysis*, Decision Support Systems, **64**, 67-78.

British Airways (2018), *Customer Data Theft*. Available on: https://www.britishairways.com/en-gb/information/incident/data-theft/latest-information. Accessed on: 29/03/19

Cabinet Office (2022), *National Cyber Strategy 2022*, https://www.gov.uk/government/publications/national-cyber-strategy-2022/national-cyber-security-strategy-2022. Accessed on: 12/08/22

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

Castillo, D., Falzon, J. (2018), *An analysis of the impact of Wannacry cyberattack on cybersecurity stock returns*, Review of Economics and Finance, **13**(3), 93-100

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**, 69-104

Chatterjee, D., Richardson V.J., Zmud, R.W. (2001), *Examining the Shareholder Wealth Effects of Announcements of Newly Created CIO positions*, MIS Quarterly, **25**(1), 43-70, https://www.jstor.org/stable/3250958. Accessed on: 04/04/23

Chen, H.S., Jai, T.M.C. (2019), *Cyber alarm: Determining the impacts of hotel's data breach messages*, International Journal of Hospitality Management, **82**, 326-334

Chen, J.V., Li, H.C., Yen, D.C., Bata, K.V. (2012), *Did IT consulting firms gain when their clients were breached?*, Computers in Human Behavior, **28**(2), 456-464

CMS Legal (2021), *GDPR Enforcement Tracker*, https://www.enforcementtracker.com/. Accessed on: 26/02/21

Confente, I., Siciliano, G.G., Gaudenzi, B., Eickhoff, M. (2019), *Effects of data breaches from user-generated content: A corporate reputation analysis*, European Management Journal, **37**(4), 492-504, ISSN 0263-2373, https://doi.org/10.1016/j.emj.2019.01.007. Accessed on: 04/04/23

Corbet, S., Gurdgiev, C. (2020), *An Incentives-Based Mechanism for Corporate Cyber Governance Enforcement and Regulation*. In: Walker, T., Gramlich, D., Bitar, M., Fardnia, P. (eds) Ecological, Societal, and Technological Risks and the Financial Sector, Palgrave Studies in Sustainable Business In Association with Future Earth. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-38858-4_13. Accessed on: 04/04/23

Cowan Research LC (2001), *Eventus® Example: Using Non-CRSP Data in Eventus 8*, http://www.eventstudy.com/NonCRSPEventus8.pdf. Accessed on: 06/09/22

CSO (2020), *What is information security? Definition, principles, and jobs*, https://www.csoonline.com/article/3513899/what-is-information-security-definition-principles-and-jobs.html. Accessed on: 12/08/22

Cuellar S.S, Karnowsky, D., Acosta, F. (2009), *The Sideways Effect: A Test for Changes in the Demand for Merlot and Pinot Noir Wines*, Journal of Wine Economics, Volume 4, Issue 2, Winter 2009, Pages 1–14

CyBOK (2022), About CyBOK, https://www.cybok.org/about/. Accessed on: 11/07/22

Daly A. (2018), *The introduction of data breach notification legislation in Australia: A comparative view*, Computer Law & Security Review, **34**, 477–495

DataBreaches LLC (2023), *DataBreaches.net – The Office of Inadequate Security*, https://www.databreaches.net/. Accessed on: 25/01/23

Data Protection Act (1998), https://www.legislation.gov.uk/ukpga/1998/29/contents. Accessed on: 30/04/21

Data Protection Act (2018), http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted. Accessed on: 11/06/21

Data Protection Directive (1995), https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri= CELEX:31995L0046: en:HTML. Accessed on: 30/04/21

DCMS (2020), Impact of the GDPR on Cyber Security Outcomes, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/906691/Impact_of_GDPR_on_cyber_security_outcomes.pdf. Accessed on: 06/09/22

DCMS (2021), *Cyber Security Breaches Survey 2021*, https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2021/cyber-security-breaches-survey-2021. Accessed on: 06/09/22

DCMS (2022), *Cyber Security Breaches Survey 2022*, https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2022. Accessed on: 21/12/22

Deane J.K., Goldberg D.M., Rakes T.R., Rees L.P. (2019), *The effect of information security certification announcements on the market value of the firm. Information Technology & Management*, **20**(3), 107-121

Dyckman, T., Philbrick, D., Stephan, J. (1984), *A Comparison of Event Study Methodologies Using Daily Stock Returns: A Simulation Approach*, Journal of Accounting Research, **22**, (Supplement)

Edgar, T. Manz, D. (2017), *Research methods for cyber security*, Syngress

ENISA (2020), *ETL2020 The Year in Review*, https://www.enisa.europa.eu/publications/year-in-review. Accessed on: 16/09/22

Ettredge, M., Guo, F., Li, Y. (2018), *Trade secrets and cyber security breaches*, Journal of Accounting and Public Policy, **37**(6), 564-585, https://doi.org/10.1016/j.jaccpubpol.2018.10.006. Accessed on: 04/04/23

European Commission (2021), *Proposal for an ePrivacy Regulation*, https://digital-strategy.ec.europa.eu/en/policies/eprivacy-regulation. Accessed on: 16/09/22

European Convention on Human Rights (1950), https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/005. Accessed on: 30/04/21

Fama, E.F. (1970), *Efficient Capital Markets: A Review of Theory and Empirical Work*, The Journal of Finance, **25**(2), 383–417

Fama, E.F., French, K.R. (1992), *The cross-section of expected stock returns*, Journal of Finance, **47**(2), 427-465, https://onlinelibrary.wiley.com/doi/pdfdirect/10.1111/j.1540-6261.1992.tb04398.x. Accessed on: 22/09/22

Fama, E.F., French, K.R. (1996), *Multifactor explanations of asset pricing anomalies*, Journal of Finance **51**(1), 55–84, https://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1996.tb05202.x. Accessed on: 22/09/22

Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., Smeraldi, F. (2016), *Decision support approaches for cyber security investment*, Decision support systems, **86**, 13-23

Fombrun, C.J. (2012), *The building blocks of corporate reputation: Definitions, antecedents, consequences*

Forbes (2018), *Marriott Breach -- What Happened, How Serious Is It And Who Is Impacted?*, https://www.forbes.com/sites/kateoflahertyuk/2018/11/30/marriott-breach-what-happened-how-serious-is-it-and-who-is-impacted/. Accessed on: 11/07/22

Ford, A., Al-Nemrat, A., Ghorashi, S.A., Davidson, J. (2021a), *The Impact of Data Breach Announcements on Company Value in European Markets*, WEIS 2021: The 20th Annual Workshop on the Economics of Information Security, https://weis2021.econinfosec.org/wp-content/uploads/sites/9/2021/06/weis21-ford.pdf. Accessed on: 21/12/21

Ford, A., Al-Nemrat, A., Ghorashi, S.A., Davidson, J. (2021b), *The Impact of GDPR Infringement Fines on the Market Value of Firms*, Proceedings of the 20th European Conference on Cyber Warfare and Security, https://doi.org/10.34190/EWS.21.088. Accessed on: 04/04/23

Ford, A., Al-Nemrat A., Ghorashi, S.A., Davidson, J. (2022a), *The Impact of CISO Appointment Announcements on the Market Value of Firms*, 17th International Conference on Cyber Warfare and Security, https://doi.org/10.34190/iccws.17.1.49. Accessed on: 06/09/22

Ford, A., Al-Nemrat A., Ghorashi, S.A., Davidson, J. (2022b), *The Impact of GDPR Infringement Fines on the Market Value of Firms,* Information and Computer Security, **31**(1), https://doi.org/10.1108/ICS-03-2022-0049. Accessed on: 12/09/22

Fortune (2021), *Massive data leak exposes 700 million LinkedIn users' information*, https://fortune.com/2021/06/30/linkedin-data-theft-700-million-users-personal-information-cybersecurity/. Accessed on: 11/07/22

Garg, A., Curtis, J., Halper, H. (2003), *The Financial Impact of IT Security Breaches: What Do Investors Think?*, Information Systems Security, 12:1, 22-33, DOI: 10.1201/1086/43325.12.1.20030301/41478.5

Gartner Inc. (2021), *Gartner Predicts 40% of Boards Will Have a Dedicated Cybersecurity Committee by 2025*, https://www.gartner.com/en/newsroom/press-releases/2021-01-28-gartner-predicts-40--of-boards-will-have-a-dedicated-. Accessed on: 12/09/22

Gartner Inc. (2022), *What is a Metaverse?*, https://www.gartner.com/en/articles/what-is-a-metaverse. Accessed on: 12/09/22

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

Goel, S., Shawky, H.A. (2014), *The Impact of Federal and State Notification Laws on Security Breach Announcements,* Communications of the Association for Information Systems, **34**, 37-50

Goldstein, J., Chernobai, A., Benaroch, M. (2011), *An Event Study Analysis of the Economic Impact of IT Operational Risk and its Subcategories*, Journal of the Association for Information Systems, **12**(9), DOI: 10.17705/1jais.00275

Gordon, S., Ford, R. (2006), *On the Definition and Classification of Cybercrime*, J. Comput. Virol, **2**, 13–20

Gordon, L.A., Loeb, M.P. (2002), *The economics of information security investment*, ACM Transactions on Information and System Security (TISSEC), **5**(4), 438-457

Gordon, L.A., Loeb, M.P. (2006), *Budgeting process for information security expenditures. Communications of the ACM*, **49**(1), 121-125

Gordon, L.A., Loeb, M.P., Zhou, L. (2016), *Investing in cybersecurity: Insights from the Gordon-Loeb model. Journal of Information Security*, **7**(02), 49

Gordon, L.A., Loeb, M.P., Lucyshyn, W., Zhou, L. (2018), *Empirical evidence on the determinants of cybersecurity investments in private sector firms*, Journal of Information Security, **9**(02), 133

Grossman, S.J., Stiglitz, J.E. (1980), *On the Impossibility of Informationally Efficient Markets*, American Economic Review, **70**(3), 393–408

He, P., Sun, Y., Zhang, Y., Li, T. (2020) *COVID-19's Impact on Stock Prices Across Different Sectors - An Event Study Based on the Chinese Stock Market*, Emerging Markets Finance and Trade, **56**(10), 2198-2212, https://doi.org/10.1080/1540496X.2020.1785865. Accessed on: 24/02/22

Hinz, O., Nofer, M., Schiereck, D., Trillig, J. (2015), *The influence of data theft on the share prices and systematic risk of consumer electronics companies*, Information & Management, **52**(3), 337-347

Hoo, K.J.S. (2000), *How much is enough: a risk management approach to computer security*, Stanford University

Hovav, A., Gray, P. (2014), *The Ripple Effect of an Information Security Breach Event: A Stakeholder Analysis,* Communications of the Association for Information Systems, **34**(50), DOI: 10.17705/1CAIS.03450

Howard, J. (1997), *An Analysis of Security Incidents on the Internet 1989–1995*, PhD thesis, Dept. of Engineering and Public Policy, Carnegie Mellon University

IBM Security (2022), *Cost of a Data Breach Report*, https://www.ibm.com/uk-en/reports/data-breach. Accessed on: 31/01/23

ICO – Information Commissioner's Office (2020), *Penalty Notice: Section 155, Data Protection Act 2018, Case ref: COM0783542, British Airways plc*, https://ico.org.uk/media/action-weve-taken/mpns/2618421/ba-penalty-20201016.pdf. Accessed on: 12/09/22

ICO – Information Commissioner's Office (2022), *Key concepts and definitions*, https://ico.org.uk/for-organisations/the-guide-to-nis/key-concepts-and-definitions/. Accessed on: 12/08/22

IC3 – Internet Crime Complaint Center (2020), *Internet Crime Report 2020*, https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf. Accessed on: 12/09/22

Interpol (2022), *COVID-19 Cyberthreats*, https://www.interpol.int/en/Crimes/Cybercrime/COVID-19-cyberthreats. Accessed on: 06/09/22

Ishiguro, M., Tanaka, H., Matsuura, K., Murase, I. (2006), *The effect of information security incidents on corporate values in the Japanese stock market*, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.212.4556. Accessed on: 12/09/22

ISO/IEC (2009), *27000:2009, (E) Information technology, Security techniques - Information security management systems - Overview and vocabulary*

IT Governance Ltd (2023), *IT Governance Blog En – Protect – Comply - Thrive*, https://www.itgovernance.eu/blog/en. Accessed on: 25/01/23

Jeong, C., Lee, S., Lim, J. (2019), *Information security breaches and IT security investments: Impacts on competitors*, Information & Management, **56**(5), 681-695

Johnson, M.E., Goetz, E. (2007), *Embedding information security into the organization*, IEEE Security & Privacy, **5**(3), 16-24

Kannan, K., Rees, J., Sridhar, S. (2007), *Market Reactions to Information Security Breach Announcements: An Empirical Analysis*, International Journal of Electronic Commerce, 01 September **12**(1), 69-91

Karanja, E., Rosso, M.A. (2017), *The Chief Information Security Officer: An Exploratory Study, Journal of International Technology and Information Management*, **26**(2), 23-47

Kaspereit, T. (2015), *EVENTSTUDY2: Stata module to perform event studies with complex test statistics*, Statistical Software Components

Khansa L. (2015), *M&As and market value creation in the information security industry*, Journal of Economics and Business, **82**, 113–134

Khansa, L., Cook, D.F., James, T., Bruyaka, O. (2012), *Impact of HIPAA provisions on the stock market value of healthcare institutions, and information security and other information technology firms*, Computers & Security, **31**(6), 750-770.

Khotari, S.P., Warner, J.B. (2006), *Econometrics of Event Studies* (Working Paper), https://www.bu.edu/econ/files/2011/01/KothariWarner2.pdf. Accessed on: 12/09/22

Lin, Z., Sapp, T.R., Ulmer, J.R., Parsa, R. (2020) *Insider trading ahead of cyber breach announcements*, Journal of Financial Markets, **50**, 100527

Macfarlanes (2020), *Lessons from the ICO's decisions to reduce the BA and Marriott GDPR fines*, https://www.macfarlanes.com/what-we-think/in-depth/2020/lessons-from-the-ico-s-decisions-to-reduce-the-ba-and-marriott-gdpr-fines/. Accessed on: 26/02/21

MacKinlay, A. C. (1997), *Event Studies in Economics and Finance*, Journal of Economic Literature **35**(1) (March)

Makridis, C., Dean, B. (2018), *Measuring the Economic Effects of Data Breaches on Firm Outcomes: Challenges and Opportunities*, **1**, 59 – 83.

Malliouris, D.D., Simpson, A.C. (2019), *The stock market impact of information security investments: The case of security standards*, Workshop on the Economics of Information Security, https://weis2019.econinfosec.org/wp-content/uploads/sites/6/2019/05/WEIS_2019_paper_22.pdf. Accessed on: 12/09/22

Malliouris, D.D., Simpson, A.C. (2020), *Underlying and consequential costs of cyber security breaches: Changes in systematic risk*, Workshop on the Economics of Information Security, https://weis2020.econinfosec.org/wp-content/uploads/sites/8/2020/06/weis20-final14.pdf. Accessed on: 12/09/22

Merriam-Webster (2022), https://www.merriam-webster.com/dictionary/cyberspace. Accessed on: 12/08/22

Modi, S.B., Wiles, M.A., Mishra, S. (2015), *Shareholder value implications of service failures in triads: The case of customer information security breaches*, Journal of Operations Management, 35, 21–39

Moore, T., Dynes, S., Chang, F.R. (2015)*, Identifying how firms manage cybersecurity investment*, http://blog.smu.edu/research/files/2015/10/SMU-IBM. pdf. Accessed on: 12/08/22

Morse, E.A., Raval, V., Wingender Jr, J.R. (2011), *Market price effects of data security breaches*, Information Security Journal: A Global Perspective, **20**(6), 263-273

Murciano-Goroff (2019), *Do Data Breach Disclosure Laws Increase Firms' Investment in Securing Their Digital Infrastructure?,* WEIS 2019, https://weis2021.econinfosec.org/wp-content/uploads/sites/6/2019/05/WEIS_2019_paper_33.pdf. Accessed on: 12/09/22

NCSC (2022a), *What is a cyber incident*, https://www.ncsc.gov.uk/information/what-cyber-incident. Accessed on: 15/09/22

NCSC (2022b), *Cyber Essentials: Requirements for IT infrastructure v3.0*, https://www.ncsc.gov.uk/files/Cyber-Essentials-Requirements-for-Infrastructure-v3-0-January-2022.pdf. Accessed on: 15/09/22

NCSC (2022c), *We think Cyber Essentials is, well, still essential …*, https://www.ncsc.gov.uk/blog-post/we-think-cyber-essentials-is-well-still-essential. Accessed on: 15/09/22

Neumann, A., Statland, N., Webb, R. (1977) *Post-processing audit tools and techniques*, in Proceedings of the NBS Invitational Workshop. Miami Beach, Florida: US Department of Commerce, National Bureau of Standards, pp. 11–3; 11–4. Available at: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nbsspecialpublication500-19.pdf. Accessed on: 12/09/22

Neville-Rolfe, L.J. (2020), *PrivSec London*, Queen Elizabeth II Centre, London, 04/02/20

Nieuwesteeg, B., Faure, M. (2018), *An analysis of the effectiveness of the EU data breach notification obligation*, Computer Law & Security Review: **34**(6), 1232-1246

NIST (1995), *Special Publication 800-12: An Introduction to Computer Security: The NIST Handbook*, https://csrc.nist.rip/publications/nistpubs/800-12/800-12-html/chapter13.html. Accessed on: 15/09/22

NIST (2022a), *Supplemental Information for the Interagency Report on Strategic U.S. Government Engagement in International Standardization to Achieve U.S. Objectives for Cybersecurity*, https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8074v2.pdf. Accessed on: 12/08/22

NIST (2022b), *Security Posture*, https://csrc.nist.gov/glossary/term/security_posture. Accessed on: 12/08/22

OED – Oxford English Dictionary (2022), https://www.oed.com. Accessed on: 12/09/22

Office for National Statistics (2022), *Crime Survey England & Wales 2022*, https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingjune2022#computer-misuse. Accessed on: 21/12/22

O'Hara, M, Shaw, W. (1990), *An Event Study Example Using Eventus*, http://www.eventstudy.com/ohara.pdf. Accessed on: 06/09/22

Oxford Reference (2022), *Doing a Ratner*, https://www.oxfordreference.com/view/10.1093/acref/9780199916108.001.0001/acref-9780199916108-e-2235. Accessed on: 06/09/22

Parker, D. B. (1998), *Fighting Computer Crime: A new Framework for Protecting Information New York*, Wiley Computer Publishing, John Wiley & Sons, Inc.

Phillips, K., Davidson, J.C., Farr, R.R., Burkhardt, C., Caneppele, S., Aiken, M. (2022), *Conceptualizing Cybercrime: Definitions, Typologies and Taxonomies*, Forensic Sci. 2, 379–398. https://doi.org/10.3390/forensicsci2020028. Accessed on: 04/04/23

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, http://dx.doi.org/10.1016/j.jisa.2014.07.001. Accessed on: 04/04/23

Png, I.P., Wang, C.Y., Wang, Q.H. (2008), *The deterrent and displacement effects of information security enforcement: International evidence. Journal of Management Information Systems*, **25**(2), 125-144

Privacy and Electronic Communications Directive (2002), https://eur-lex.europa.eu/legal-content/EN/ALL/? uri=CELEX%3A32002L0058. Accessed on: 30/04/21

Proton AG (2023), *What is GDPR, the EU's new data protection law?*, https://gdpr.eu/what-is-gdpr/. Accessed on: 23/01/23

R Core Team (2018), *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/. Accessed on: 12/09/22

Ramos S.B., Latoeiro P., Veiga, H. (2020), *Limited attention, salience of information and stock market activity*, Economic Modelling, **87**, 92-108

Reese, W.A., Robins, R.P. (2017) *Performing an event study: An exercise for finance students*, Journal of Economic Education, **48**(3), 206-215, DOI: 10.1080/00220485.2017.1320603

Richardson, V.J., Smith, R.E, Watson, M.W. (2019) *Much Ado about Nothing: The (Lack of) Economic Impact of Data Privacy Breaches*, Journal of Information Systems: **33**(3), 227-265

Romanosky, S., Telang, R., Acquisti, A. (2011), *Do Data Breach Disclosure Laws Reduce Identity Theft?*, Journal of Policy Analysis and Management, **30**(2), 256-286

Rosati, P., Deeney, P., Cummins, M., Van der Werff, L., Lynn, T. (2019), *Social media and stock price reaction to data breach announcements: Evidence from US listed companies*, Research in International Business and Finance, **47**, 458-469

Roškot, M., Wanasika, I., Kroupova, Z.K. (2020), *Cybercrime in Europe: surprising results of an expensive lapse*, Journal of Business Strategy

Salameh, H., AlBahsh, R. (2011), *Testing the Efficient Market Hypothesis at the Semi Strong Level in Palestine Stock Exchange – Event Study of the Mandatory Disclosure*, International Research Journal of Finance and Economics, ISSN 1450-2887, **69**

SANS (2022), *Information Security Resources*, https://www.sans.org/information-security/. Accessed on: 12/08/22

Schatz, D. (2018), *Towards a Comprehensive Evidence-Based Approach for Information Security Value Assessment,* (Doctoral dissertation, University of East London), https://repository.uel.ac.uk/item/84590. Accessed on: 12/09/22

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

Schatz, D., Bashroush, R. (2016b), *Economic Valuation for Information Security Investment: A Systematic Literature Review*, Information Systems Frontiers, **19**(5), 1205-1228

Schatz, D., Bashroush, R. (2018), *Corporate Information Security Investment Decisions: A Qualitative Data Analysis Approach*, International Journal of Enterprise Information Systems, **14**(2), 1-20

Schimmer, M., Levchenko, A., and Müller, S. (2014), *EventStudyTools (Research Apps), St.Gallen*, http://www.eventstudytools.com. Accessed on: 26/02/21

Sharpe, W.F. (1964), Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, Journal of Finance, **19**(3), 425–442

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

Srinidhi, B., Yan, J., Tayi, G.K. (2015), *Allocation of resources to cyber-security: The effect of misalignment of interest between managers and investors*, Decision Support Systems, **75**, 49-62.

Stiennon, R. (2013), *Categorizing data breach severity with a breach level index*, https://breachlevelindex. com/pdf/Breach-Level-Index-WP. pdf. Accessed on: 11/08/22

Syed, R. (2019), *Enterprise reputation threats on social media: A case of data breach framing*, Journal of Strategic Information Systems, **28**, 257–274

Tejay, G.P., Shoraka, B. (2011), *Reducing cyber harassment through de jure standards: a study on the lack of the information security management standard adoption in the USA*, International Journal of Management and Decision Making, **11**(5-6), 324-343

Telang, R., Wattal, S. (2007), *An empirical analysis of the impact of software vulnerability announcements on firm stock price*, IEEE Transactions on Software engineering, **33**(8), 544-557

Thales Group (2017), *First half 2017 Breach Level Index report*, https://www.thalesgroup.com/en/markets/digital-identity-and-security/press-release/first-half-2017-breach-level-index-report-identity-theft-and-poor-internal-security-practices-take-a-toll. Accessed on: 10/06/21

The Guardian (2020), *Travelex falls into administration, with loss of 1,300 jobs*, https://www.theguardian.com/business/2020/aug/06/travelex-falls-into-administration-shedding-1300-jobs. Accessed on: 10/06/21

The Independent (2021), *Classified MoD documents found at bus stop in Kent*, https://www.independent.co.uk/news/uk/politics/classified-documents-bus-stop-kent-b1873492.html. Accessed on: 12/08/22

This is Money (2000), *BTP shares soar on Swiss bid*, https://www.thisismoney.co.uk/money/news/article-1574940/BTP-shares-soar-on-Swiss-bid.html. Accessed on: 06/09/22

Thomas, D., Loader, B. (2000), *Introduction-Cybercrime: Law Enforcement, Security and Surveillance in the Information Age*, In Cybercrime: Law Enforcement, Security and Surveillance in the Information Age; Thomas, D., Loader, B., Eds.; Routledge: London, UK

Tweneboah-Kodua, S., Atsu, F. and Buchanan, W. (2018), *Impact of cyberattacks on stock performance: a comparative study*, Information and Computer Security, **26**(5), 637-652

Verizon (2022), *2022 Data Breach Investigations Report*, https://www.verizon.com/business/resources/reports/2022/dbir/2022-dbir-data-breach-investigations-report.pdf. Accessed on: 12/08/22

Von Solms, R., Van Niekerk, J. (2013), *From information security to cyber security*, Computers & Security, **38**, 97-102

Washington Post (1998), *Pfizer's Stock Soars on Success of Drug*, https://www.washingtonpost.com/wp-srv/national/longterm/viagra/stories/pfizer26.htm. Accessed on: 06/09/22

Whitman, M.E, Fendler, P., Caylor, J., Baker, D. (2005), *Rebuilding the human firewall*, In Proceedings of the 2nd annual conference on Information security curriculum development (InfoSecCD '05), Association for Computing Machinery, New York, NY, USA, 104–106, https://doi.org/10.1145/1107622.1107646. Accessed on: 04/04/23

Wiener, N. (1948), *Cybernetics; or control and communication in the animal and the machine*, John Wiley

Williams, P. (2007), *Executive and board roles in information security, Network Security*, **8**, 11-14

Wired (2014), *Data Is the New Oil of the Digital Economy*, https://www.wired.com/insights/2014/07/data-new-oil-digital-economy/. Accessed on: 20/09/22

Yahoo!Finance (2019), *Historical Data*, https://finance.yahoo.com/quote. Accessed on: 11/06/21

Yayla, A. A., Hu, Q. (2011), *The Impact of Information Security Events on the Stock Value of Firms: The Effect of Contingency Factors*, Journal of Information Technology, **26**(1), 60–77

Zafar, H., Ko, M., Osei-Bryson, K.M. (2012), *Financial impact of information security breaches on breached firms and their non-breached competitors*, Information Resources Management Journal (IRMJ), **25**(1), 21-37

Appendix (R Code)

```
estudy <- function (
  firmSymbol,  # Stock symbol (Yahoo!Finance)
  firmName,    # Just a label somewhat more explicit than the symbol alone
  indexSymbol, # Index symbol (Yahoo!Finance)
  indexName,   # Just a label for the index
  eventDate,   # "%d.%m.%Y"
  startEvent,  # start of event window
  endEvent,    # end of event window
  endEst,      # end of estimation window
  lengthEst,   # length of estimation window
  estprice = "adjusted", # use "adjusted" or "close"
  estgroup = "breach",
  estfiletype = "csv",
  estbenchmarkmodel = "mm", # default is market model otherwise "ff3fm" could be used
  estreturntype = "simple", # use as default rather than "log"
  estnontradingdays = "later", # not relevant if day 0 is trading day
  estffdata = "" # string containing name of data.frame in global environment to use for ff3fm
  ) {
    #
    # Prerequisites:
    # install.packages("devtools")
    #devtools::install_github("EventStudyTools/api-wrapper.r")
    require(tidyquant)
```

```
require(dplyr)
require(readr)
#
# open a logfile and capture errors/warnings
options(warn = 1) # output warnings as they happen
logfilename = "estudylog.txt"
logfile <- file(logfilename, open = "wt")
sink(logfile, type = "message")
sink(logfile, type = "output")
#
# calculate earliest and latest dates (give a month flex either way)
# (event study will fail if event window runs into future)
t1 <- min(startEvent, 0, (endEst - lengthEst + 1)) # earliest point in time (expect -ve)
t2 <- max(endEvent, 0, endEst) # latest point in time (expect +ve)
startDate <- as.Date(eventDate, format("%d.%m.%Y")) + (t1*7/5) - 30
endDate   <- as.Date(eventDate, format("%d.%m.%Y")) + (t2*7/5) + 30
# use default names for data files
dataFiles <- c("request_file" = "01_requestFile.csv",
               "firm_data" = "02_firmData.csv",
               "market_data" = "03_marketData.csv")
#
# output directory
resultPath = paste("data", firmSymbol, basename(tempfile("")), sep = "\\")
# Get Firm Data
```

```
firmSymbol %>%
  tidyquant::tq_get(from = startDate, to = endDate) %>%
  dplyr::mutate(date = format(date, "%d.%m.%Y")) -> firmData
firmData$symbol <- firmName # make the name more explicit and avoid special characters (.)
# knitr::kable(head(firmData), pad=0)
# Get Index Data
indexSymbol %>%
  tidyquant::tq_get(from = startDate, to = endDate) %>%
  dplyr::mutate(date = format(date, "%d.%m.%Y")) -> indexData
indexData$symbol <- indexName # avoid use of special characters (^)
# knitr::kable(head(indexData), pad=0)
#
# Perform basic validation checks on firm and index data:
warnstr = ""
# 1. both firm and index data should be non-empty
if (is.null(firmData) | is.null(indexData)) {
  warning(firmSymbol, indexSymbol, " is empty?")
  warnstr = paste(warnstr, "Empty?")
}
# 2. firm and index data should start and end on the same day and be same # rows
if ( (min(firmData$date) != min(indexData$date)) |
     (max(firmData$date) != max(indexData$date)) ) {
     warning(firmSymbol, indexSymbol, " start/end dates do not match?")
     warnstr = paste(warnstr, "Start/end?")
```

```
}
if ( nrow(firmData) != nrow(indexData) ) {
  warning(firmSymbol, indexSymbol, " mismatch in data rows?")
  warnstr = paste(warnstr, "Mismatch?")
}
# 3. data has to be historic - see if we are starting to get close to today
if (endDate >= today()) { # endDate is a conservative estimate!
  warning(firmSymbol, indexSymbol, " dates running into future?")
  warnstr = paste(warnstr, "Future?")
}
# 4. should be at least enough trading days data to cover estimation window and event window
if (nrow(firmData) < (t2 - t1 + 1)) {
  warning(firmSymbol, " firm data missing rows?")
  warnstr = paste(warnstr, "Missing firm?")
}
if (nrow(indexData)< (t2 - t1 + 1)) {
  warning(indexSymbol, " index data missing rows?")
  warnstr = paste(warnstr, "Missing index?")
}
#
# Price files for firms and market
firmData %>%
  dplyr::select(symbol, date, estprice) %>%
  readr::write_delim(file      = dataFiles["firm_data"],
```

```
                    delim     = ";",
                    col_names = F)
# for market data, need to add Fama-French data columns if needed
if (estbenchmarkmodel == "mm") {
        indexData %>%
          dplyr::select(symbol, date, estprice) %>%
          readr::write_delim(file      = dataFiles["market_data"],
                             delim     = ";",
                             col_names = F)
}
else {  # assume this is "ff3fm"
        if (estbenchmarkmodel != "ff3fm") {
            warning(firmSymbol, " assuming ff3fm")
            warnstr = paste(warnstr, "assuming ff3fm")
        }
        ffdata <- get(estffdata) # expect to error if does not exist?
        ffdata$date <- paste(substr(ffdata$X, 1, 4), substr(ffdata$X, 5, 6), substr(ffdata$X, 7, 8), sep = "-")
        ffdata$date <- format(as.Date(ffdata$date), "%d.%m.%Y")
        # now perform the join (could do this in a loop for real time error checking)
        indexData <- left_join(indexData, ffdata, by = "date")
        # check for missing rows in ffdata
        if (sum(is.na(indexData$X))> 0) {
            warning(sum(is.na(indexData$X)), " missing rows in ffdata")
            warnstr = paste(warnstr, sum(is.na(indexData$X)), "missing rows in ffdata")
```

```
        }
        indexData %>%
          dplyr::select(symbol, date, estprice, RF, SMB, HML) %>%
          readr::write_delim(file      = dataFiles["market_data"],
                             delim     = ";",
                             col_names = F)
}
# finally create request file
request <- cbind(1, firmName, indexName, eventDate, estgroup, startEvent, endEvent, endEst, lengthEst)
request %>%
  as.data.frame() %>%
  readr::write_delim(dataFiles["request_file"], delim = ";", col_names = F)
# now the files have been created, perform the event study
#
# initialise
#
library(EventStudy)
apiUrl <- "http://api.eventstudytools.com"
apiKey <- "573e58c665fcc08cc6e5a660beaad0cb" # old key not in use after Sept 2020
#
# from above URL
#
# Setup API Connection
estSetup <- EventStudyAPI$new(apiUrl)
```

```
estSetup$authentication(apiKey)
#
EventStudy::checkFiles(dataFiles) # check the format of data files just created is okay
# set parameters for the event study (some are default but specified here for clarity)
estPar <- EventStudy::ARCApplicationInput$new()
estPar$setResultFileType(estfiletype)
estPar$setBenchmarkModel(estbenchmarkmodel)
estPar$setReturnType(estreturntype)
estPar$setNonTradingDays(estnontradingdays)
#
# perform the event study:
estResults <- estSetup$performEventStudy(estParams = estPar,
                                         dataFiles = dataFiles,
                                         destDir = resultPath
                                         )
# now copy the dataFiles into the results folder for retention
file.copy(dataFiles, resultPath, overwrite = T)
file.remove(dataFiles) # clean up
sink(type = "message", split = F) # close logfile
sink(type = "output",  split = F) # close logfile
closeAllConnections() #ensure all files are closed
file.copy(logfilename, resultPath, overwrite = T)
file.remove(logfilename) # clean up
# build return vector
```

```
    # Since v0.39 need to parse results first for analysis report and car results
    estParser <- ResultParser$new()
    retval = c( firmSymbol, # this along with event date is primary key
                estParser$get_analysis_report(paste(resultPath, "analysis_report.csv", sep = "\\")), # includes event
date
                estParser$get_analysis_report(paste(resultPath, "car_results.csv", sep = "\\")),
                startEvent, # included to avoid need to parse window string in results
                endEvent,   # included to avoid need to parse window string in results
                resultPath,
                warnstr,
                estbenchmarkmodel, # added 14/3/22 - record calculation method
                estffdata # Fama-French data.frame used
                ) # also included in logfile
    return(retval)
}
get_name <- function ( firmSymbol = "") {
  require("stringr")
  url=str_c("https://uk.finance.yahoo.com/quote/", firmSymbol, "/profile?p=", firmSymbol)
  tmpdf=as.data.frame(readLines(url, warn = FALSE))
  names(tmpdf)="text"
  sec=str_match(tmpdf$text[6],'Sustain.*Sector\\(s\\)')
  sec=str_replace_all(sec, '<[^>]*>', '#')
  sec=str_replace_all(sec, '#+', '#')
  sec=str_replace_all(sec, '[^ -~]', '') # remove non-printables
```

```
  sec=str_replace(sec, '\\{.*\\}', '')
  sec=str_replace(sec, '^.*?#.*?#', '')
  sec=str_replace(sec, '#.*$', '')
  sec=str_replace(sec,'&amp;','&')
  return(sec)
}
get_sector <- function ( firmSymbol = "") {
  require("stringr")
  url=str_c("https://uk.finance.yahoo.com/quote/", firmSymbol, "/profile?p=", firmSymbol)
  tmpdf=as.data.frame(readLines(url, warn = FALSE))
  names(tmpdf)="text"
  sec=str_match(tmpdf$text[6],'Sector\\(s\\).*Industry')
  sec=str_replace_all(sec, '<[^>]*>', '')
  sec=str_replace_all(sec, '[^ -~]', '') # remove non-printables
  sec=str_replace(sec, 'Sector\\(s\\):', '')
  sec=str_replace(sec, 'Industry', '')
  sec=str_replace(sec,'&amp;','&')
  return(sec)
}
get_industry <- function ( firmSymbol = "") {
  require("stringr")
  url=str_c("https://uk.finance.yahoo.com/quote/", firmSymbol, "/profile?p=", firmSymbol)
  tmpdf=as.data.frame(readLines(url, warn = FALSE))
  names(tmpdf)="text"
```

```
  ind=str_match(tmpdf$text[6],'Industry</span>.*Full-time employees')
  ind=str_replace_all(ind, '<[^>]*>', '')
  ind=str_replace_all(ind, '[^ -~]', '') # remove non-printables
  ind=str_replace(ind, 'Industry:', '')
  ind=str_replace(ind, 'Full-time employees', '')
  ind=str_replace(ind,'&amp;','&')
  return(ind)
}
get_market_cap <- function ( firmSymbol = "") {
  require("stringr")
  url=str_c("https://uk.finance.yahoo.com/quote/", firmSymbol, "?p=", firmSymbol)
  tmpdf=as.data.frame(readLines(url, warn = FALSE))
  names(tmpdf)="text"
  mc=str_match(tmpdf$text[6],'MARKET_CAP-value.*Beta')
  mc=str_replace_all(mc, '<[^>]*>', '')
  mc=str_replace_all(mc, '[^ -~]', '') # remove non-printables
  mc=str_replace(mc, 'MARKET_CAP-value.*>', '')
  mc=str_replace(mc, 'Beta', '')
  mc=str_replace(mc,'&amp;','&')
  mc=str_replace(mc,'T','e+12')
  mc=str_replace(mc,'B', 'e+9')
  mc=str_replace(mc,'M', 'e+6')
  return(as.numeric(mc))
}
```

```
get_market_cap_cur <- function ( firmSymbol = "") {
  require("stringr")
  url=str_c("https://uk.finance.yahoo.com/quote/", firmSymbol, "?p=", firmSymbol)
  tmpdf=as.data.frame(readLines(url, warn = FALSE))
  names(tmpdf)="text"
  mc=str_match(tmpdf$text[6],'Currency in.*Add to watchlist')
  mc=str_replace_all(mc, '<[^>]*>', '')
  mc=str_replace_all(mc, '[^ -~]', '') # remove non-printables
  mc=str_replace_all(mc, 'Currency in', '')
  mc=str_replace_all(mc, 'Add to watchlist', '')
  mc=str_replace(mc,'&amp;','&')
  return(toupper(mc))
}
# this function needs more work - not really any benefit over looking up manually!
get_revenue <- function ( firmSymbol = "") {
  require("stringr")
  url=str_c("https://uk.finance.yahoo.com/quote/", firmSymbol, "/financials?p=", firmSymbol)
  tmpdf=as.data.frame(readLines(url, warn = FALSE))
  names(tmpdf)="text"
  mc=str_match(tmpdf$text[6],'Income.*Cost of revenue')
  mc=str_replace_all(mc, '<[^>]*>', ';')
  mc=str_replace_all(mc, ';+', ';')
  mc=str_replace_all(mc, '[^ -~]', '') # remove non-printables
  mc=str_replace_all(mc, 'Income.*Currency in ', '')
```

```
  mc=str_replace_all(mc, 'Cost of revenue', '')
  mc=str_replace_all(mc, 'Add to watchlist', '')
  mc=str_replace(mc,'&amp;','&')
  return(mc)
}
# this function can easily be adjusted to average over extra days in case cdate is a non-trading day
get_eur_rate <- function (curr, cdate) {
  rv <- 1
  if (curr != "EUR") {
    require(tidyquant)
    cdate <- format(as.Date(cdate, "%d/%m/%Y"),"%Y-%m-%d")
    rv <- tq_get(paste(curr,"EUR=X", sep = ""), from = cdate, to = format(as.Date(cdate, "%Y-%m-%d")+0, "%Y-%m-%d")) # +x
days
    rv <- mean(rv$close)
  }
  return(rv)
}
# this function can easily be adjusted to average over extra days in case cdate is a non-trading day
get_usd_rate <- function (curr, cdate) {
  rv <- 1
  if (curr != "USD") {
    require(tidyquant)
    cdate <- format(as.Date(cdate, "%d/%m/%Y"),"%Y-%m-%d")
```

```
    rv <- tq_get(paste(curr,"USD=X", sep = ""), from = cdate, to = format(as.Date(cdate, "%Y-%m-%d")+0, "%Y-%m-%d")) # +x
days
    rv <- mean(rv$close)
  }
  return(rv)
}
# A couple of example event study function calls
# BA data breach
ba1 <- estudy("IAG.L", "British Airways", "^FTSE", "FTSE100", "06.09.2018", -2, 2, -3, 120, estprice = "close")
ba2 <- estudy("IAG.L", "British Airways", "^FTMC", "FTSE250", "06.09.2018", -2, 2, -3, 120, estprice = "close")
ba <- data.frame(rbind(ba1,ba2))
rm(ba1, ba2)
#
# Data breach events pre Jan 2020 - confirmed meeting 22/05/20 to cap at Dec-19 to avoid COVID-19 effects
# now load from Excel export
events <- read.delim("data\\events.txt", stringsAsFactors = F) # tab delimited
#
write.csv(sapply(events, unlist), file="data\\events.bak") # for backup purposes
#
# initialise results
for (row in 1:nrow(events)) {
  cat("Processing row", row, events[row,1], "\n")
  p <- as.list(c(events[row,1:5])) # start parameter list
  #p[3] <- "^SPEUP" # try S&P Euro 350
```

```
#p[4] <- "SPEUR350"
names(p) <- NULL # to avoid any name matching in function call
results[nrow(results)+1,] <- do.call(estudy, c(p, -2,  2, -3, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p, -1,  1, -2, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p, -1,  0, -2, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  0, -1, 120)) # not an option
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  1, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  2, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  3, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  4, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  5, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  6, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  7, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  8, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0,  9, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0, 10, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0, 20, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0, 30, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0, 40, -1, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p,  0, 50, -1, 120)) # event window must be in range (-50, 50)
results[nrow(results)+1,] <- do.call(estudy, c(p, -1, 1, -2, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p, -2, 2, -3, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p, -5, 5, -6, 120))
results[nrow(results)+1,] <- do.call(estudy, c(p, -10, 10, -11, 120))
```

```
  results[nrow(results)+1,] <- do.call(estudy, c(p, -15, 15, -16, 120))
  results[nrow(results)+1,] <- do.call(estudy, c(p, -20, 20, -21, 120))
  }
#
write.csv(sapply(results, unlist), file="data\\results.csv") # for backup purposes
#
# Coerce data types ready for processing
attach(df) # actually make a new df to convert types
results %>% filter(Reference.Market != "SPEUR350") -> df
results %>% filter(Reference.Market == "SPEUR350") -> dfe
df$Window <- as.character(Window)
df$Window <- factor(df$Window, levels = c("(-2, 2)", "(-1, 1)", "(-1, 0)",
  "(0, 1)", "(0, 2)", "(0, 3)", "(0, 4)", "(0, 5)",
  "(0, 6)", "(0, 7)", "(0, 8)", "(0, 9)", # added
  "(0, 10)", "(0, 20)", "(0, 30)", "(0, 40)", "(0, 50)"))
df$End.of.Estimation.Window <- as.numeric(End.of.Estimation.Window)
df$CAR.Value <- as.double(CAR.Value)
df$CAR.t.test <- as.double(CAR.t.test)
df$Firm.1 = factor(as.character(Firm.1))
df$Event.Date <- as.Date(as.numeric(Event.Date))
df$Reference.Market <- factor(as.character(df$Reference.Market))
df$V27 <- as.character(V27)
df$V1 <- as.character(V1)
df$Analysis.Report <- as.character(Analysis.Report)
```

```
df$sector <- as.character(firmographics$sector[match(df$V1, firmographics$symbol)])
df$records <- events$Records.breached[match(paste(df$V1, df$Event.Date), paste(events$Symbol, as.Date(events$Date,
"%d.%m.%Y")))]
df$gdpr <- if_else(df$Event.Date < "2018-05-25", F, T)
# now check if there were warnings:
# - manually review results?
df %>% count(V27)
df %>% filter(grepl("Future", V27)) %>% count(Event.Date, V27)
df %>% filter(grepl("Mismatch", V27))
df %>% filter(grepl("warning", Analysis.Report)) %>% count(Analysis.Report) %>% knitr::kable()
df %>% filter(grepl("zero", Analysis.Report)) %>% count(Firm.1, Analysis.Report)
#
df <- subset(df, Firm.1 != "Travelex") # Travelex is major outlier
#
# add in new field for personal data
df$personal <- events$personal[match(paste(df$V1, df$Event.Date), paste(events$Symbol, as.Date(events$Date,
"%d.%m.%Y")))]
#
# now check SP Euro 350 results
#
attach(dfe)
dfe$Window <- as.character(dfe$Window)
dfe$Window <- factor(dfe$Window, levels = c("(-2, 2)", "(-1, 1)", "(-1, 0)",
```

```
                                        "(0, 1)", "(0, 2)", "(0, 3)", "(0, 4)", "(0, 5)", "(0, 10)", "(0, 20)", "(0,
30)", "(0, 40)", "(0, 50)"))
dfe$End.of.Estimation.Window <- as.numeric(End.of.Estimation.Window)
dfe$CAR.Value <- as.double(CAR.Value)
dfe$CAR.t.test <- as.double(CAR.t.test)
dfe$Firm.1 = factor(as.character(Firm.1))
dfe$Event.Date <- as.Date(as.numeric(Event.Date))
dfe$Reference.Market <- factor(as.character(dfe$Reference.Market))
dfe$V27 <- as.character(dfe$V27)
dfe$Analysis.Report <- as.character(Analysis.Report)
dfe$sector <- as.character(firmographics$sector[match(dfe$V1, firmographics$symbol)])
dfe$records <- events$Records.breached[match(paste(dfe$V1, dfe$Event.Date), paste(events$Symbol, as.Date(events$Date,
"%d.%m.%Y")))]
dfe$gdpr <- if_else(dfe$Event.Date < "2018-05-25", F, T)
# now check if there were warnings:
# - manually review results?
dfe %>% count(V27)
dfe %>% filter(grepl("Future", V27)) %>% count(Event.Date, V27)
dfe %>% filter(grepl("Mismatch", V27))
dfe %>% filter(grepl("warning", Analysis.Report)) %>% count(Analysis.Report) %>% knitr::kable()
dfe %>% filter(grepl("zero", Analysis.Report)) %>% count(Firm.1, Analysis.Report)
#
#df <- subset(df, Firm.1 != "Travelex") # Travelex is major outlier
#dfe <- subset(dfe, Firm.1 != "Travelex") # Travelex is major outlier
```

```
#
t1 <- as.list(tapply(df$CAR.Value, df$Window, mean)) # a workaround due to not being able to do weighted averages in
add_row
t2 <- df %>% select(Sector = sector, Window, CAR = CAR.Value) %>%
    group_by(Sector) %>%
  pivot_table(
      .rows = c(Sector, ~ COUNT(Sector) / 17),
      .columns = Window,
      .values = ~ AVERAGE(CAR)
  )
  mutate(t2, Mean = rowMeans(t2[3:15])) %>%
  rename(N = 2) %>%
  arrange(Mean) %>%

  add_row(Sector = "", N = sum(.$N),
          `(-2, 2)`    = t1$`(-2, 2)`,
          `(-1, 1)`    = t1$`(-1, 1)`,
          `(-1, 0)`    = t1$`(-1, 0)`,
          `(0, 1)`     = t1$`(0, 1)`,
          `(0, 2)`     = t1$`(0, 2)`,
          `(0, 3)`     = t1$`(0, 3)`,
          `(0, 4)`     = t1$`(0, 4)`,
          `(0, 5)`     = t1$`(0, 5)`,
          `(0, 6)`     = t1$`(0, 6)`,
```

```
          `(0, 7)`     = t1$`(0, 7)`,
          `(0, 8)`     = t1$`(0, 8)`,
          `(0, 9)`     = t1$`(0, 9)`,
          `(0, 10)`    = t1$`(0, 10)`,
          `(0, 20)`    = t1$`(0, 20)`,
          `(0, 30)`    = t1$`(0, 30)`,
          `(0, 40)`    = t1$`(0, 40)`,
          `(0, 50)`    = t1$`(0, 50)`,
          Mean         = mean(as.numeric(t1)) # as test this should == mean(df$CAR.Value)
          ) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put last
  #formattable(align = c("l", "r", "r", "r", "r", "r", "r", "r", "r", "r", "r", "r", "r", "r", "r"))
rm(t1, t2)
#
# another table format
# aim for window, N, CAAR, t-test, p(***), % negative CARs
subset(df, Window !="(0, x)" & records > 1000 & personal) %>% # update based on decision which to use
  #filter(is.element(sector, c("Technology"))) %>%
  select(Window, CAR = CAR.Value, t.value = CAR.t.test) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Window) %>%
  summarise(N = n(), CAAR = mean(CAR), SCAAR = sd(CAR), t.caar = sqrt(n())*(CAAR/SCAAR), pv = 2*pt(-abs(t.caar), n()-1),
    sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
    `NegCAR %` = sum(NegCAR)/n()*100) %>%
```

```
  #arrange(Window) %>%
  add_row(Window = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
    `NegCAR %` = weighted.mean(.$`NegCAR %`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable()
# impact of GDPR
subset(df, is.element(Window, c("(0, 2)", "(0, 5)", "(0, 30)", "(0, 50)")) & personal) %>% # update based on decision
which to use
  select(Window, gdpr, CAR = CAR.Value, t.value = CAR.t.test) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Window, gdpr) %>%
  summarise(N = n(), CAAR = mean(CAR), SCAAR = sd(CAR), t.caar = sqrt(n())*(CAAR/SCAAR), pv = 2*pt(-abs(t.caar), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `NegCAR %` = sum(NegCAR)/n()*100) %>%
  #arrange(Window) %>%
  select(-pv) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  formattable()
# sector analysis for a particular window
subset(df, Firm != "Travelex" & Window == "(0, 2)") %>% # update based on decision which to use
  select(sector, CAR = CAR.Value, t.value = CAR.t.test, records, personal, gdpr) %>%
```

```
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(sector) %>%
  summarise(N = n(), CAAR = mean(CAR), SCAAR = sd(CAR), t.caar = sqrt(n())*(CAAR/SCAAR), pv = 2*pt(-abs(t.caar), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `NegCAR %` = sum(NegCAR)/n()*100,
            Records = sum(records, na.rm = TRUE),
            `Personal %` = sum(personal)/n()*100,
            `GDPR %` = sum(gdpr)/n()*100,
            t.car = mean(abs(t.value))) %>%
  arrange(CAAR) %>%
  add_row(sector = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `NegCAR %` = weighted.mean(.$`NegCAR %`, .$N),
          Records = sum(.$Records),
          `Personal %` = weighted.mean(.$`Personal %`, .$N),
          `GDPR %` = weighted.mean(.$`GDPR %`, .$N),
          t.car = weighted.mean(.$t.car, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  select(-pv, -t.car) %>%
  formattable()
# market analysis for a particular window
subset(df, Firm != "Travelex" & Window ==  "(0, 1)") %>% # update based on decision which to use
  select(Reference.Market, CAR = CAR.Value, t.value = CAR.t.test, records, personal, gdpr) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
```

```
  group_by(Reference.Market) %>%
  summarise(N = n(), CAAR = mean(CAR), SCAAR = sd(CAR), t.caar = sqrt(n())*(CAAR/SCAAR), pv = 2*pt(-abs(t.caar), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `NegCAR %` = sum(NegCAR)/n()*100,
            Records = sum(records, na.rm = TRUE),
            `Personal %` = sum(personal)/n()*100,
            `GDPR %` = sum(gdpr)/n()*100) %>%
  arrange(CAAR) %>%
  select(-pv) %>%
  add_row(Reference.Market = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `NegCAR %` = weighted.mean(.$`NegCAR %`, .$N), Records = sum(.$Records),
          `Personal %` = weighted.mean(.$`Personal %`, .$N),
          `GDPR %` = weighted.mean(.$`GDPR %`, .$N)
          ) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  formattable()
#
# check t.test function for example in above
#subset(df, (Window =="(-5, 5)") & (sector == "Financial Services")) %>%
#  select(CAR = CAR.Value) %>% t.test(alternative = "two.sided")
# Visualisations:
require(ggplot2)
#
```

```
# boxplot
ggplot(subset(df, 1==1)) +
  aes(x = Window , y = CAR.Value) +
  geom_boxplot(outlier.shape = NA) + # otherwise outliers will be duplicated by jitter
  geom_jitter() +
  geom_text(aes(label=if_else(abs(CAR.Value) > 0.21, Firm.1, NULL)), hjust = 1.2) + # to identify outliers
  #stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme_grey(base_size = 24) +
  labs(title=NULL, x = "Event window", y = "CAR") +
  theme(plot.title = element_text(hjust = 0.5)) # centre title
# analysis by industry
ggplot(subset(df, 1==1)) +
  aes(x = reorder(sector, CAR.Value), y= CAR.Value, colour = as.factor(sector)) +
  geom_point() +
  #geom_jitter() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1), base_size = 32) +
  labs(title = "Data breach events", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme_grey(base_size = 18) +
  theme(plot.title = element_text(hjust = 0.5))
# analysis by industry per window
ggplot(subset(df, 1==1)) +
  aes(group = sector, x = Window, y = CAR.Value, colour = sector) +
  theme_grey(base_size = 32) +
```

```
  theme(axis.text.x = element_text(angle = -90, vjust = 0, hjust = 0)) +
  labs(title = NULL, x = "Event window", y = "CAAR", colour = "Sector") +
  stat_summary(aes(group=sector, colour = sector), fun.y=mean, geom="line", linetype = "solid") +
  #stat_summary(aes(group=1), fun.y=mean, geom="line", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
# financial services
ggplot(subset(df, sector == "Industrials")) +
  aes(x = Window , y = CAR.Value) +
  geom_boxplot(outlier.shape = NA) + # otherwise outliers will be duplicated by jitter
  geom_jitter() +
  #geom_text(aes(label=if_else(CAR.Value < -0.3, Firm.1, NULL)), hjust = -0.0) + # to identify outliers
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  labs(title="Breach events", x = "Event window", y = "CAR") +
  theme(plot.title = element_text(hjust = 0.5))
# analysis by firm
ggplot(subset(df, Firm.1 != "Travelex")) +
  aes(x = reorder(Firm.1, CAR.Value), y= CAR.Value, colour = as.factor(Window)) +
  geom_point() +
  geom_jitter() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "Data breach events (excluding Travelex)", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
    theme(plot.title = element_text(hjust = 0.5))
# by date?
```

```
ggplot(subset(df, Firm.1 != "Travelex")) +
  aes(x = Event.Date, y= CAR.Value, colour = as.factor(Window)) +
  scale_x_date(date_breaks = "1 month", date_labels = "%m-%Y") +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "Data breach events (excluding Travelex)", x = NULL, y = "CAR", colour = "Event window") +
  theme(plot.title = element_text(hjust = 0.5)) +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  geom_vline(xintercept = as.Date("2018-05-25"), linetype = "dashed") # GDPR
# per market?
ggplot(subset(df, Firm.1 != "Travelex")) +
  aes(x = reorder(Reference.Market, CAR.Value), y= CAR.Value, colour = as.factor(Window)) +
  geom_point() +
  geom_jitter() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "Data breach events (excluding Travelex)", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
#
ggplot(subset(df, is.element(Window, c("(0, 2)", "(0, 5)", "(0, 30)", "(0, 50)")) & personal)) +
  aes(x = (records), y = CAR.Value, colour = as.factor(Window)) +
  geom_point() +
  #scale_x_continuous() +
  scale_x_log10() +
```

```
  geom_smooth(method='lm', se = F) +
  #stat_summary(aes(group=1), fun=mean, geom="line", colour="steelblue", linetype = "dashed") + # can we add regression
line for this?
  theme_grey(base_size = 32) +
  theme(axis.text.x = element_text(angle = 0, vjust = 0.4, hjust = 1)) +
  labs(title = NULL, x = "Number of records breached", y = "CAR", colour = "Event window") +
  theme(plot.title = element_text(hjust = 0.5))
#
############################
# GDPR fine announcements #
############################
#
# read in CSV file from enforcementtracker.com
#
fines <- read.delim("data\\gdprfines.txt", stringsAsFactors = F) # tab delimited
#fines <- subset(fines, as.Date(Date, "%d/%m/%Y") <= "2019-12-31") # agreed in meeting 22/05/20 to stop at Dec to avoid
COVID-19 effects
# change of plan 30/06/20 do the event studies for now and subset later
#
results.gdpr <- results[0,] # initialise a new results table based on previous
for (row in 1:nrow(fines)) {
  #if(as.Date(fines[row, 'Date'], "%d/%m/%Y") > "2019-12-31") {
  cat("Processing row", row, fines[row,'Symbol'], "\n")
  p <- as.list(c(fines[row, c('Symbol', 'Controller.Processor', 'Index', 'IndexDesc') ])) # start parameter list
```

```
  p <- c(p, format(as.Date(fines[row, 'Date'], "%d/%m/%Y"), "%d.%m.%Y"))
  names(p) <- NULL # to avoid any name matching in function call
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p, -2,  2, -3, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p, -1,  1, -2, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p, -1,  0, -2, 120))
  #results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0,  0, -1, 120)) # not an option
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0,  1, -1, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0,  2, -1, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0,  3, -1, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0,  4, -1, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0,  5, -1, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0, 10, -1, 120))
  results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  0, 20, -1, 120))
  #
  #results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p,  -5,  5, -6, 120))
  #results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p, -10, 10, -11, 120))
  #results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p, -15, 15, -16, 120))
  #results.gdpr[nrow(results.gdpr)+1,] <- do.call(estudy, c(p, -20, 20, -21, 120))
  #}
}
#
write.csv(sapply(results.gdpr, unlist), file="data\\results.gdpr.csv") # for backup purposes
#
# Coerce data types ready for processing
```

```
dfg <- results.gdpr
attach(dfg)
dfg$Event.Date <- as.Date(as.numeric(Event.Date))
dfg <- subset(dfg, Event.Date <= "2019-12-31") # agreed in meeting 22/05/20 to stop at Dec to avoid COVID-19 effects
dfg$Window <- as.character(dfg$Window)
dfg$Window <- factor(dfg$Window, levels = c("(-2, 2)", "(-1, 1)", "(-1, 0)",
  "(0, 1)", "(0, 2)", "(0, 3)", "(0, 4)", "(0, 5)", "(0, 10)", "(0, 20)"))
dfg$End.of.Estimation.Window <- as.numeric(End.of.Estimation.Window)
dfg$CAR.Value <- as.double(CAR.Value)
dfg$CAR.t.test <- as.double(CAR.t.test)
dfg$Firm.1 = factor(as.character(Firm.1))
dfg$Reference.Market <- factor(as.character(dfg$Reference.Market))
dfg$V1 <- as.character(V1)
dfg$V27 <- as.character(V27)
dfg$Analysis.Report <- as.character(Analysis.Report)
#fines$Fine <- as.numeric(fines$Fine)
dfg$Fine <- as.numeric(fines$Fine[match(paste(dfg$V1, as.Date(dfg$Event.Date)), paste(fines$Symbol, as.Date(fines$Date,
"%d/%m/%Y")))])
dfg$Country <- fines$Country[match(paste(dfg$V1, as.Date(dfg$Event.Date)), paste(fines$Symbol, as.Date(fines$Date,
"%d/%m/%Y")))]
#
# now add firmographics - need to modify to read firmographics database
#dfg$sector   <- lapply(dfg$V1, get_sector)    # takes a while!
#dfg$industry <- lapply(dfg$V1, get_industry)  # takes a while!
```

```
dfg$V1name <- firmographics$name[match(dfg$V1, firmographics$symbol)]
dfg$market_cap_eur <- firmographics$market_cap_eur[match(dfg$V1, firmographics$symbol)]
dfg$revenue_eur <- revenues$Last_yr_rev_eur[match(paste(dfg$V1, as.Date(dfg$Event.Date)),
  paste(revenues$Symbol, as.Date(revenues$EventDate, "%d/%m/%Y")))]
dfg$Country_f <- as.character(firmographics$country[match(dfg$V1, symbol)]) # also add firm country for comparison
# now check if there were warnings:
# - manually review results?
dfg %>% count(V27)
dfg %>% filter(grepl("Future", V27)) %>% count(Event.Date, V27)
dfg %>% filter(grepl("Mismatch", V27))
dfg %>% filter(grepl("warning", Analysis.Report)) %>% count(Analysis.Report) %>% knitr::kable()
dfg %>% filter(grepl("zero", Analysis.Report)) %>% count(Firm.1, Analysis.Report)
#
#
# tables of results
#
# Window
subset(dfg, Window !="(x, x)") %>% # update based on decision which to use
  select(Window, CAR = CAR.Value, t.value = CAR.t.test) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Window) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Negative CAR` = sum(NegCAR)/n()*100) %>%
```

```
  #arrange(CAAR) %>%
  add_row(Window = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
    `% Negative CAR` = weighted.mean(.$`% Negative CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
  #kbl(booktabs = T, format = "latex")
#
# validate
  tapply(dfg$Fine, dfg$Window, mean)
# Country (as defined by Enforcement Tracker)
subset(dfg, Window =="(0, 3)") %>% # update based on decision which to use
  select(Country, CAR = CAR.Value, t.value = CAR.t.test, Fine) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Country) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Negative CAR` = sum(NegCAR)/n()*100, sum(Fine) / 1000) %>%
  arrange(CAAR) %>%
  add_row(Country = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Negative CAR` = weighted.mean(.$`% Negative CAR`, .$N), sum(Fine) / 1000) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
```

```
  mutate(across(ends_with("0"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r", "r"))
# Country (from firmographics)
subset(dfg, Window =="(0, 3)") %>% # update based on decision which to use
  select(Country_f, CAR = CAR.Value, t.value = CAR.t.test, Fine) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Country_f) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Negative CAR` = sum(NegCAR)/n()*100, sum(Fine) / 1000) %>%
  arrange(CAAR) %>%
  add_row(Country_f = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Negative CAR` = weighted.mean(.$`% Negative CAR`, .$N), sum(Fine) / 1000) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  mutate(across(ends_with("0"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r", "r"))
#
# Market
subset(dfg, Window =="(0, 3)") %>% # update based on decision which to use
  select(Reference.Market, CAR = CAR.Value, t.value = CAR.t.test) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
```

```
  group_by(Reference.Market) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Negative CAR` = sum(NegCAR)/n()*100) %>%
  arrange(CAAR) %>%
  add_row(Reference.Market = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Negative CAR` = weighted.mean(.$`% Negative CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
# Firm
subset(dfg, Window =="(0, 3)") %>% # update based on decision which to use
  select(V1name, sector, CAR = CAR.Value, Fine, revenue_eur, market_cap_eur) %>%
  group_by(V1name) %>%
  summarise(N = n(), CAAR = mean(CAR), Revenue = mean(revenue_eur) / 1000, Fines = mean(Fine) / 1000, `Fines %` = (Fines
/ Revenue) * 100,
            `Market Capitalisation` = mean(market_cap_eur) / 1000, DeltaMC = abs(`Market Capitalisation` * CAAR),
            Ratio = DeltaMC / Fines) %>%
  arrange(CAAR) %>%
  add_row(V1name = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N), Revenue = weighted.mean(.$Revenue, .$N),
          Fines = weighted.mean(.$Fines, .$N),
          `Fines %` = weighted.mean(.$`Fines %`, .$N),
          `Market Capitalisation` = weighted.mean(.$`Market Capitalisation`, .$N),
```

```
          DeltaMC = weighted.mean(.$DeltaMC, .$N),
          Ratio = weighted.mean(.$Ratio, .$N)
          ) %>%
  mutate(across(.cols = c("CAAR", "Fines %"), round, 4)) %>% # should put rounding last
  mutate(across(.cols = c("Fines", "DeltaMC", "Ratio"), round, 0)) %>%
  mutate(across(.cols = c("Revenue", `Market Capitalisation`, "DeltaMC", "Ratio"), format, big.mark = ",")) %>%
  formattable(align = c("l", "r", "r", "r", "r", "r", "r", "r", "r"))
# Sector
subset(dfg, Window =="(0, 3)") %>% # update based on decision which to use
  select(Sector = sector, CAR = CAR.Value, t.value = CAR.t.test) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Sector) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Negative CAR` = sum(NegCAR)/n()*100) %>%
  arrange(CAAR) %>%
  add_row(N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Negative CAR` = weighted.mean(.$`% Negative CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "r", "r"))
#
# boxplots
```

```
#
ggplot(dfg) +
  aes(x = Window , y = CAR.Value) +
  geom_boxplot() +
  geom_jitter() +
  geom_text(aes(label=if_else(abs(CAR.Value) > 0.1, V1, NULL)), hjust = 1.5) + # to identify outliers
  #stat_summary(aes(group=1), fun.y=sum, geom="line", colour="steelblue", linetype = "dashed") +
  labs(title="", x = "Event window", y = "CAR") +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme_grey(base_size = 18)
with(dfg, aggregate(CAR.Value, by = list(Window), FUN=mean))
# analysis by firm
ggplot(dfg) +
  aes(x = reorder(V1, CAR.Value), y= CAR.Value, colour = as.factor(Window)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "GDPR fines", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
# by date?
ggplot(subset(dfg, Window == "(0, 3)")) +
  aes(x = Event.Date, y= CAR.Value, colour = as.factor(Window)) +
  scale_x_date(date_breaks = "1 month", date_labels = "%m-%Y") +
  geom_point() +
```

```
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "GDPR fines by date", x = NULL, y = "CAR", colour = "Event window") +
  theme(plot.title = element_text(hjust = 0.5)) +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed")
  #geom_vline(xintercept = as.Date("2018-05-25"), linetype = "dashed") # GDPR
# per market?
ggplot(subset(dfg, Window == "(0, 3)")) +
  aes(x = reorder(Reference.Market, CAR.Value), y= CAR.Value, colour = as.factor(Window)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "GDPR fines by market", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
#
# Is there a correlation between CAR and value of the fine?
ggplot(subset(dfg, Window == "(0, 3)")) +
  aes(x = (Fine / revenue_eur), y = (CAR.Value), colour = as.factor(Window)) +
  geom_point() +
  #scale_x_continuous=(breaks = seq(0,30,5)) +
  scale_x_log10() +
  geom_smooth(method='lm', formula= y~x) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "GDPR fines / CAR plot", x = "Fine (€)", y = "CAR", colour = "Event window") +
  theme(plot.title = element_text(hjust = 0.5))
```

```
# Country?
ggplot(subset(dfg, Window == "(0, 3)")) +
  aes(x = reorder(Country, CAR.Value), y= CAR.Value, colour = as.factor(Window)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "GDPR fines by country (authority)", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
# summary stats
tapply(dfg$CAR.Value, dfg$Window, mean)
#
#####################
# GDPR fine appeals #
#####################
#
#fines.app <- subset(fines, Sort %in% c(224, 175, 174, 97))
attach(fines.app)
fines.app$Date[Sort == 224] <- "12/06/2020" # GOOGL appeal failed
fines.app$Date[Sort == 175] <- "16/10/2020" # IAG reduced to £20m
fines.app$Date[Sort == 174] <- "30/10/2020" # MAR reduced to £18.4m
fines.app$Date[Sort == 97 ] <- "12/11/2020" # 1&1 fine reduced 90% to €900,000 https://www.corderycompliance.com/1and1-
gdpr-fine-reduced/
#
results.gdpr.app <- results[0,] # initialise a new results table based on previous
```

```
for (row in 1:nrow(fines.app)) {
  cat("Processing row", row, fines.app[row,'Symbol'], "\n")
  p <- as.list(c(fines.app[row, c('Symbol', 'Controller.Processor', 'Index', 'IndexDesc') ])) # start parameter list
  p <- c(p, format(as.Date(fines.app[row, 'Date'], "%d/%m/%Y"), "%d.%m.%Y"))
  names(p) <- NULL # to avoid any name matching in function call
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p, -2,  2, -3, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p, -1,  1, -2, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p, -1,  0, -2, 120))
  #results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0,  0, -1, 120)) # not an option
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0,  1, -1, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0,  2, -1, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0,  3, -1, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0,  4, -1, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0,  5, -1, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0, 10, -1, 120))
  results.gdpr.app[nrow(results.gdpr.app)+1,] <- do.call(estudy, c(p,  0, 20, -1, 120))
}
write.csv(sapply(results.gdpr.app, unlist), file="data\\results.gdpr.app.csv") # for backup purposes
#
# Coerce data types ready for processing
dfga <- results.gdpr.app
attach(dfga)
dfga$Event.Date <- as.Date(as.numeric(Event.Date))
dfga$Window <- as.character(dfga$Window)
```

```
dfga$Window <- factor(dfga$Window, levels = c("(-2, 2)", "(-1, 1)", "(-1, 0)",
                                              "(0, 1)", "(0, 2)", "(0, 3)", "(0, 4)", "(0, 5)", "(0, 10)", "(0, 20)"))
dfga$End.of.Estimation.Window <- as.numeric(End.of.Estimation.Window)
dfga$CAR.Value <- as.double(CAR.Value)
dfga$CAR.t.test <- as.double(CAR.t.test)
dfga$Firm.1 = factor(as.character(Firm.1))
dfga$Reference.Market <- factor(as.character(dfga$Reference.Market))
dfga$V1 <- as.character(V1)
dfga$V27 <- as.character(V27)
dfga$Analysis.Report <- as.character(Analysis.Report)
#fines.app$Fine <- as.numeric(fines.app$Fine)
#dfga$Fine <- as.numeric(fines.app$Fine[match(paste(dfga$V1, as.Date(dfga$Event.Date)), paste(fines.app$Symbol,
as.Date(fines.app$Date, "%d/%m/%Y")))])
#dfga$Country <- fines.app$Country[match(paste(dfga$V1, as.Date(dfga$Event.Date)), paste(fines.app$Symbol,
as.Date(fines.app$Date, "%d/%m/%Y")))]
#
# now add firmographics - need to modify to read firmographics database
dfga$V1name <- firmographics$name[match(dfga$V1, firmographics$symbol)]
dfga$market_cap_eur <- firmographics$market_cap_eur[match(dfga$V1, firmographics$symbol)]
dfga$revenue_eur <- revenues$Last_yr_rev_eur[match(paste(dfga$V1, as.Date(dfga$Event.Date)),
                                                   paste(revenues$Symbol, as.Date(revenues$EventDate, "%d/%m/%Y")))]
#
subset(dfga, Window !="(x, x)") %>% # update based on decision which to use
  select(Window, CAR = CAR.Value, t.value = CAR.t.test) %>%
```

```
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(Window) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Negative CAR` = sum(NegCAR)/n()*100) %>%
  add_row(Window = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Negative CAR` = weighted.mean(.$`% Negative CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
tapply(dfga$CAR.Value, dfga$V1, mean)
#
dfga %>% select(V1, Window, CAR = CAR.Value, tv = CAR.t.test) %>%
  group_by(Window, V1) %>%
  pivot_table(
    .rows = c(Window, ~COUNT(Window)),
    .columns = V1,
    .values = c( ~ AVERAGE(CAR), ~ AVERAGE(tv))
  ) %>%
  select(1,2, 3,7, 4,8, 5,9, 6,10) %>%
  rename(N = 2) %>%
  rename_with(~gsub("AVERAGE\\(", "", .)) %>%
  rename_with(~gsub(")_", "_", .)) %>%
```

```
  mutate(days = c(5,3,2,2,3,4,5,6,11,21)) %>% # very lazy!! should calculate from V24, V25
  # now can calculate p values
  mutate(pv_GOOGL = 2*pt(-abs(tv_GOOGL), days-1),
         pv_IAG.L = 2*pt(-abs(tv_IAG.L), days-1),
         pv_MAR   = 2*pt(-abs(tv_MAR),   days-1)
         ) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put last
  select(1:10) %>% # no significance
  formattable(align = c("l", "r", "r", "r", "r", "r", "r", "r", "r"))
ggplot(subset(dfga, Window == "(0, 2)")) +
  aes(x = V1 , y = CAR.Value) +
  geom_boxplot() +
  geom_jitter() +
  geom_text(aes(label=if_else(abs(CAR.Value) > 0.075, V1, NULL)), hjust = -0.1) + # to identify outliers
  #stat_summary(aes(group=1), fun.y=sum, geom="line", colour="steelblue", linetype = "dashed") +
  labs(title="GDPR fines: comparison of event windows", x = "Event window", y = "CAR") +
  theme(plot.title = element_text(hjust = 0.5))
#
###################################
# Repeated data breach/fine events #
###################################
#
# Check for any repeated events
# ============================
```

```
#
df %>% filter(is.element(V1, events$Symbol[duplicated(events$Symbol)])) %>% select(V1, Event.Date, Window, CAR.Value)
intersect(results$V1, results.gdpr$V1)
intersect(results$V1, results.ciso$V1)
intersect(intersect(results$V1, results.gdpr$V1), results.ciso$V1)
intersect(intersect(results$V1, results.gdpr$V1), fines.app$Symbol) # only IAG
intersect(results.gdpr$V1, results.ciso$V1)
intersect(fines.app$Symbol, results.gdpr$V1) # just a test - should be all 4!
#
dfrep <- events$Symbol[duplicated(events$Symbol)] # strictly speaking should read this from df!
subset(df, is.element(V1, dfrep) & is.element(Window, c("(0, 4)"))) %>% # update based on decision which to use
  select(V1, Event.Date, Window, CAR = CAR.Value, t.value = CAR.t.test) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  arrange(V1, Event.Date, Window) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  formattable()
#
# GDPR fines
dfgrep <- dfg %>% select(V1) %>% group_by(V1) %>% filter(n() > 10) %>% count()
dfgrep <- as.character(dfgrep$V1)
subset(dfg, is.element(V1, dfgrep) & is.element(Window, c("(-2, 2)"))) %>% # update based on decision which to use
  select(V1, Event.Date, Window, CAR = CAR.Value, t.value = CAR.t.test) %>%
  group_by(V1) %>% mutate(grp = row_number()) %>%
```

```
  summarise(V1, Event.Date, Window, CAR, CAAR = sum(CAR), grp) %>%
  arrange(CAAR, Event.Date) %>% ungroup() %>%
  add_row(V1 = "CAAR", Event.Date = NA, Window = "", CAR = mean(.$CAR), CAAR = mean(.$CAAR)/2) %>%
  add_row(V1 = "Group1", Event.Date = NA, Window = "", CAR = mean(.$CAR[which(.$grp == 1)]), CAAR = NA) %>%
  add_row(V1 = "Group2", Event.Date = NA, Window = "", CAR = mean(.$CAR[which(.$grp == 2)]), CAAR = NA) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  formattable()
#
# CISO
#
dfcrep <- ciso$Symbol[duplicated(ciso$Symbol)]
subset(dfc, is.element(V1, dfcrep) & is.element(Window, c("(-2, 2)"))) %>% ## )) %>% # update based on decision which to
use
  filter(V26 < 10) %>% # exclude larger windows
  arrange(Event.Date, CAR.value) %>%
  #slice(n=5)  %>% # max and min - won't work as not grouped by event
  select(V1, Event.Date, Window, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  arrange(V1, Event.Date, CAR, Window) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  #select(-pv) %>%
  formattable()
```

```
#
#
# these are firms with repeated events:
union_all(df$V1[df$Window == "(0, 1)"], dfg$V1[dfg$Window == "(0, 1)"],
          dfga$V1[dfga$Window == "(0, 1)"], dfc$V1[dfc$Window == "(0, 1)"]) %>% as.data.frame() -> allrep
colnames(allrep) <- "V1"
allrep <- allrep$V1[duplicated(allrep$V1)] %>% unique()
df   %>% filter(is.element(V1, allrep)) %>% select(V1) %>% count(V1)
dfg  %>% filter(is.element(V1, allrep)) %>% select(V1) %>% count(V1)
dfga %>% filter(is.element(V1, allrep)) %>% select(V1) %>% count(V1)
dfc  %>% filter(is.element(V1, allrep)) %>% select(V1) %>% count(V1)
#
# IAG repeated events (specific event window only)
filter(df, V1 == "IAG.L", Window == "(-2, 2)") %>% select(V1, Firm, Event.Date, Window, CAR.Value, CAR.t.test) %>%
mutate(abscar= abs(CAR.Value), gp = "1") %>%
  rbind(filter(dfg , V1 == "IAG.L", Window == "(-2, 2)") %>% select(V1, Firm, Event.Date, Window, CAR.Value, CAR.t.test)
%>% mutate(abscar=abs(CAR.Value), gp = "1")) %>%
  rbind(filter(dfga, V1 == "IAG.L", Window == "(-2, 2)") %>% select(V1, Firm, Event.Date, Window, CAR.Value, CAR.t.test)
%>% mutate(abscar=abs(CAR.Value), gp = "1")) %>%
# try and find (max) IAG repeated events - should use ABS to ensure capturing the largest magnitude
  rbind(filter(df  , V1 == "IAG.L" & V25 < 10) %>% select(V1, Firm, Event.Date, Window, CAR.Value, CAR.t.test) %>%
mutate(abscar=abs(CAR.Value), gp = "2") %>% slice_max(abscar)) %>%
  rbind(filter(dfg , Firm == "British Airways" & V25 < 10) %>% select(V1, Firm, Event.Date, Window, CAR.Value,
CAR.t.test) %>% mutate(abscar=abs(CAR.Value), gp = "2") %>% slice_max(abscar)) %>%
```

```
  rbind(filter(dfg , Firm == "Vueling Airlines" & V25 < 10) %>% select(V1, Firm, Event.Date, Window, CAR.Value,
CAR.t.test) %>% mutate(abscar=abs(CAR.Value), gp = "2") %>% slice_max(abscar)) %>%
  rbind(filter(dfga, V1 == "IAG.L" & V25 < 10) %>% select(V1, Firm, Event.Date, Window, CAR.Value, CAR.t.test) %>%
mutate(abscar=abs(CAR.Value), gp = "2") %>% slice_max(abscar)) -> tdfba
#
# need to remove duplicate value for GDPR fine (two identical CAR.Values!)
tdfba <- filter(tdfba, !(Firm == "Vueling Airlines" & Window == "(0, 3)"))
# Visualisation of IAG
ggplot(tdfba, aes(Event.Date, CAR.Value, fill = gp)) + geom_bar(stat="identity", position = "dodge", width = 75) +
  geom_text(data = subset(tdfba, gp == "2"), aes(label=Window, y = ifelse(CAR.Value >0 , -0.01, 0.01)), vjust = 1.2, size
= 5, angle = 90) +
  geom_line(data = subset(tdfba, gp == "1"), aes(x = Event.Date - 75/4, y = cumsum(CAR.Value), colour = gp)) +
  geom_point(data = subset(tdfba, gp == "1"), aes(x = Event.Date - 75/4, y = cumsum(CAR.Value), colour = gp)) +
  geom_line(data = subset(tdfba, gp == "2"), aes(x = Event.Date + 75/4, y = cumsum(CAR.Value), colour = gp)) +
  geom_point(data = subset(tdfba, gp == "2"), aes(x = Event.Date + 75/4, y = cumsum(CAR.Value), colour = gp)) +
  geom_text(data = subset(tdfba, gp == "1"), aes(x = as.Date(ifelse(cumsum(CAR.Value) < 0, Event.Date - 75/4, Event.Date
- 75/2)), y = cumsum(CAR.Value)-0.005, label = cumsum(CAR.Value)), hjust = 1) +
  geom_text(data = subset(tdfba, gp == "2"), aes(x = as.Date(ifelse(cumsum(CAR.Value) < 0, Event.Date + 75/4, Event.Date
- 75/2)), y = cumsum(CAR.Value)-0.005, label = cumsum(CAR.Value)), hjust = 1) +
  scale_colour_manual(name=NULL, values=c("red", "blue"), labels = c("Cumulative CAR (fixed)", "Cumulative CAR
(variable)")) +
  geom_hline(yintercept = 0) +
  scale_fill_manual(NULL, values = c("red","blue"), labels = c("Fixed event window (-2, 2)", "Variable event window")) +
  labs(x="Date",y="CAR\n") +
```

```
  theme_bw(base_size = 18)
# All repeated events summarised:
#
subset(df  , is.element(V1, allrep)) %>% select(V1, Event.Date, Window, CAR = CAR.Value, t.value = CAR.t.test)  %>%
mutate(across(Window, as.character), origin = "Breach") -> tdf1
subset(dfg , is.element(V1, allrep)) %>% select(V1, Event.Date, Window, CAR = CAR.Value, t.value = CAR.t.test)  %>%
mutate(across(Window, as.character), origin = "GDPR")   -> tdf2
subset(dfga, is.element(V1, allrep)) %>% select(V1, Event.Date, Window, CAR = CAR.Value, t.value = CAR.t.test)  %>%
mutate(across(Window, as.character), origin = "Appeal") -> tdf3
subset(dfc , is.element(V1, allrep)) %>% select(V1, Event.Date, Window, CAR = CAR.value, t.value = CAR.t.value) %>%
mutate(across(Window, as.character), origin = "CISO")   -> tdf4
rbind(tdf1, tdf2, tdf3, tdf4) %>% # add appeals back in? tdf3
  filter(Window == "(-2, 2)") %>% # Schatz & Bashroush method
  mutate(NegCAR = if_else(CAR < 0, 1, 0)) %>%
  group_by(V1) %>%
  summarise(V1, Event.Date, Window, CAR, CAAR = sum(CAR), origin)  %>%
  arrange(CAAR, Event.Date) %>% ungroup() %>%
  add_row(V1 = "CAAR", Event.Date = NA, Window = "", CAR = mean(.$CAR), CAAR = NA, origin = "") %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(ends_with("%"), round, 0)) %>%
  formattable()
#
################################
# Build firmographics database #
```

```
################################
#
attach(firmographics)
# any rows to be added?
for (ticker in setdiff(c(df$V1, dfg$V1, ciso$Symbol), firmographics$symbol)) {
  ts = as.character(ticker)
  add_row(firmographics, symbol = ts, sector = as.list(get_sector(ts)), industry = as.list(get_industry(ts)),
          name = get_name(ts)) -> firmographics
  cat(ticker, " added\n")
}
for (ticker in firmographics$symbol[as.character(firmographics$market_cap_cur) == "NULL"]) {
  ts = as.character(ticker)
  firmographics$market_cap[symbol == ts] <- get_market_cap(ts)
  firmographics$market_cap_cur[symbol == ts] <- get_market_cap_cur(ts)
  cat(ticker, " updated\n")
}
############################
# CISO hire announcements #
############################
#
# read in CSV file from Excel sheet
#
ciso <- read.delim("data\\ciso.txt", stringsAsFactors = F) # tab delimited
rownames(ciso) <- 1:nrow(ciso)
```

```
results.ciso <- results[0,] # initialise a new results table based on previous
for (row in 44:nrow(ciso)) {
  cat("Processing row", row, ciso[row,'Symbol'], "\n")
  p <- as.list(c(ciso[row, c('Symbol', 'Firm', 'Index', 'IndexDesc') ])) # start parameter list
  p <- c(p, format(as.Date(ciso[row, 'Date'], "%d/%m/%Y"), "%d.%m.%Y"))
  names(p) <- NULL # to avoid any name matching in function call
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p, -2,  2, -3, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p, -1,  1, -2, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p, -1,  0, -2, 120))
  #results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0,  0, -1, 120)) # not an option
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0,  1, -1, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0,  2, -1, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0,  3, -1, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0,  4, -1, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0,  5, -1, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0, 10, -1, 120))
  results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  0, 20, -1, 120))
  #
  #results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p,  -5,  5, -6, 120))
  #results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p, -10, 10, -11, 120))
  #results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p, -15, 15, -16, 120))
  #results.ciso[nrow(results.ciso)+1,] <- do.call(estudy, c(p, -20, 20, -21, 120))
  }
write.csv(sapply(results.ciso, unlist), file="data\\results.ciso.csv") # for backup purposes
```

```
#
# Coerce data types ready for processing
dfc <- results.ciso
attach(dfc)
dfc$Event.Date <- as.Date(as.numeric(Event.Date))
#dfc <- subset(dfc, Event.Date <= "2019-12-31") # agreed in meeting 22/05/20 to stop at Dec to avoid COVID-19 effects
dfc$Window <- as.character(dfc$Window)
dfc$Window <- factor(dfc$Window, levels = c("(-2, 2)", "(-1, 1)", "(-1, 0)",
                                            "(0, 1)", "(0, 2)", "(0, 3)", "(0, 4)", "(0, 5)", "(0, 10)", "(0, 20)"))
dfc$End.of.Estimation.Window <- as.numeric(End.of.Estimation.Window)
dfc$CAR.value <- as.double(CAR.value)
dfc$CAR.t.value <- as.double(CAR.t.value)
dfc$CAR.p.value <- as.double(CAR.p.value)
dfc$Firm.1 = factor(as.character(Firm.1))
dfc$Reference.Market <- factor(as.character(dfc$Reference.Market))
dfc$V1 <- as.character(V1)
dfc$V27 <- as.character(V27)
dfc$Analysis.Report <- as.character(Analysis.Report)
dfc$timediff <- as.numeric(ciso$timediff[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol,
as.Date(ciso$Date, "%d/%m/%Y")))])
dfc$Name <- as.character(ciso$Name[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol, as.Date(ciso$Date,
"%d/%m/%Y")))])
dfc$Title <- as.character(ciso$Posn[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol, as.Date(ciso$Date,
"%d/%m/%Y")))])
```

```
dfc$Reportsto <- as.character(ciso$Reports.to[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol,
as.Date(ciso$Date, "%d/%m/%Y")))])
dfc$Origin <- as.character(ciso$I.E[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol, as.Date(ciso$Date,
"%d/%m/%Y")))])
dfc$Gender <- as.character(ciso$m.f[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol, as.Date(ciso$Date,
"%d/%m/%Y")))])
dfc$First <- as.character(ciso$First.[match(paste(dfc$V1, as.Date(dfc$Event.Date)), paste(ciso$Symbol, as.Date(ciso$Date,
"%d/%m/%Y")))])
# now add firmographics - need to modify to read firmographics database
dfc$Currency <- as.character(firmographics$market_cap_cur[match(dfc$V1, symbol)])
dfc$Sector <- as.character(firmographics$sector[match(dfc$V1, symbol)])
dfc$market_cap_usd <- as.numeric(firmographics$market_cap_usd[match(dfc$V1, symbol)])
dfc$Country <- as.character(firmographics$country[match(dfc$V1, symbol)])
# boxplots
ggplot(subset(dfc, abs(timediff) > 2)) +
  aes(x = Window , y = CAR.value) +
  geom_boxplot() +
  geom_jitter() +
  geom_text(aes(label=if_else(abs(CAR.value) > 0.125, Firm.1, NULL)), hjust = 2) + # to identify outliers
  #stat_summary(aes(group=1), fun.y=sum, geom="line", colour="steelblue", linetype = "dashed") +
  labs(title="", x = "Event window", y = "CAR") +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme_grey(base_size = 18)
with(dfc, aggregate(CAR.Value, by = list(Window), FUN=mean))
```

```
# analysis by firm
ggplot(dfc) +
  aes(x = reorder(V1, CAR.value), y= CAR.value, colour = as.factor(Window)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
# by date?
ggplot(subset(dfc, Window == "(-1, 1)")) +
  aes(x = Event.Date, y= CAR.value, colour = as.factor(Window)) +
  scale_x_date(date_breaks = "1 month", date_labels = "%m-%Y") +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "CAR by date", x = NULL, y = "CAR", colour = "Event window") +
  theme(plot.title = element_text(hjust = 0.5)) +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed")
  #geom_vline(xintercept = as.Date("2018-05-25"), linetype = "dashed") # GDPR
# per market?
ggplot(subset(dfc, Window == "(-1, 1)")) +
  aes(x = reorder(Reference.Market, CAR.value), y= CAR.value, colour = as.factor(Window)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "CAR by market", x = NULL, y = "CAR", colour = "Event window") +
```

```
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
#
# Is there a correlation between CAR and market_cap_eur?
ggplot(subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2) & Sector == "Financial Services")) +
  aes(x = market_cap_usd, y = (CAR.value), colour = as.factor(Window)) +
  geom_point() +
  #scale_x_continuous=(breaks = seq(0,30,5)) +
  scale_x_log10() +
  geom_smooth(method='lm', formula= y~x) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "Market cap / CAR plot", x = "Market cap (€)", y = "CAR", colour = "Event window") +
  theme(plot.title = element_text(hjust = 0.5))
# Currency
ggplot(subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2))) +
  aes(x = reorder(Currency, CAR.value), y= CAR.value, colour = as.factor(Window)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "CAR by currency", x = NULL, y = "CAR", colour = "Event window") +
  stat_summary(aes(group=1), fun.y=mean, geom="line", colour="steelblue", linetype = "dashed") +
  theme(plot.title = element_text(hjust = 0.5))
# Show geography
subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2)) %>%
  group_by(region=Country) %>%
```

```
  summarise(n=n()) -> mapdata
wmap <- map_data("world") # could use a subset here to speed up?
wmap <- left_join(wmap, mapdata, by = "region")
ggplot(wmap, aes(map_id = region, fill = n)) +
  geom_map(map = wmap,  colour = "white") +
  expand_limits(x = wmap$long, y = wmap$lat) +
  scale_fill_viridis_c(option = "C", na.value = 0) +
  theme_grey()
# show how characteristics change by year
# summary stats
subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2)) %>%
  group_by(dy=year(Event.Date)) %>%
  summarise(c=mean(CAR.value)) -> dfcs # doesn't seem to work without summarising seperately!
subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2)) %>%
  group_by(dy=year(Event.Date), Gender, Origin) %>%
  summarise(c=mean(CAR.value), n=n()) %>%
  ggplot() +
  geom_bar(position="stack", stat="identity", aes(fill=Gender, x=dy-0.2, y=n), width = 0.25) +
  geom_bar(position="stack", stat="identity", aes(fill=Origin, x=dy+0.2, y=n), width = 0.25) +
  geom_line(data=dfcs, aes(x=dy, y=c/max(c)*16), geom="line", colour="black", linetype = "dashed") +
  scale_fill_discrete(limits = c("Male", "Female", "Internal", "External")) + # change to _discrete for colour or _grey
  theme_bw(base_size = 18) +
  geom_text(x=2012, y = 17, label = "Relative CAAR", colour = "black", size = 5, font = "plain") +
  scale_x_continuous(minor_breaks = NULL, breaks = c(2012:2019)) +
```

```
  scale_y_continuous(minor_breaks = seq(-6,20,1), breaks = seq(-6,20,2)) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.4, hjust = 1)) +
  labs(title = "", x = "Year", y = "Count", fill = "Gender/Origin") +
  theme(plot.title = element_text(hjust = 0.5))
# summary stats
subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2)) %>%
  group_by(dy=year(Event.Date)) %>%
  summarise(c=mean(CAR.value)) -> dfcs
subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2)) %>%
  group_by(dy=year(Event.Date), Origin, c=n()) %>%
  summarise(dy, Origin, n())
# benefits calculation
subset(dfc, Window == "(-1, 1)" & (abs(timediff) > 2 & Sector == "Financial Services")) %>%
  summarise(mean(CAR.value)*mean(market_cap_usd)/1e+6, median(CAR.value)*median(market_cap_usd)/1e+6)
#
# tables of results
#
# Window
subset(dfc, abs(timediff) > 2) %>% # update based on decision which to use
  select(Window, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Window) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
```

```
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  #arrange(CAAR) %>%
  add_row(Window = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
#
# Position / Job title
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Title, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Title) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Title = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
```

```
#
# Position reports to
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Reportsto, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Reportsto) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Reportsto = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
#
# Name
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Name, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Name) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
```

```
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Name = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
#
# Gender
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Gender, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Gender) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Gender = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
```

```
#
# Origin
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Origin, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Origin) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Origin = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
#
# First?
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(First, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(First) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
```

```
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(First = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
#
# Index (this could be country?)
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Reference.Market, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Reference.Market) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Reference.Market = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
```

```
# Currency (use as country?)
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Currency, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Currency) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
  arrange(-CAAR) %>%
  add_row(Currency = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
          `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
  mutate(across(where(is.double), round, 4)) %>% # should put rounding last
  mutate(across(starts_with("%"), round, 0)) %>%
  select(-pv) %>%
  formattable(align = c("l", "r", "r", "r", "l", "r"))
#
# Sector
subset(dfc, abs(timediff) > 2 & Window == "(-1, 1)") %>% # update based on decision which to use
  select(Sector, CAR = CAR.value, t.value = CAR.t.value) %>%
  mutate(PosCAR = if_else(CAR > 0, 1, 0)) %>%
  group_by(Sector) %>%
  summarise(N = n(), CAAR = mean(CAR), t.CAAR = sqrt(n())*(CAAR/sd(CAR)), pv = 2*pt(-abs(t.CAAR), n()-1),
            sig = if_else(pv <= 0.01, "***", if_else(pv <= 0.05, "**", if_else(pv <= 0.10, "*", ""))),
            `% Positive CAR` = sum(PosCAR)/n()*100) %>%
```

```
arrange(-CAAR) %>%
add_row(Sector = "", N = sum(.$N), CAAR = weighted.mean(.$CAAR, .$N),
        `% Positive CAR` = weighted.mean(.$`% Positive CAR`, .$N)) %>%
mutate(across(where(is.double), round, 4)) %>% # should put rounding last
mutate(across(starts_with("%"), round, 0)) %>%
select(-pv) %>%
formattable(align = c("l", "r", "r", "r", "l", "r"))
```