MODERN PRIVACY REGULATION, INTERNAL INFORMATION QUALITY, AND OPERATING EFFICIENCY: EVIDENCE FROM THE GENERAL DATA PROTECTION REGULATION

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

---

by
Steven A. Maex
August 2022

Examining Committee Members:

Dr. Jagan Krishnan, Advisory Co-Chair, Department of Accounting
Dr. Jayanthi Krishnan, Advisory Co-Chair, Department of Accounting
Dr. Yi Liang, Department of Accounting
Dr. Hyun Park, Department of Accounting
Dr. Sunil Wattal, External Reader, Department of Management Information Systems

ABSTRACT

In May 2018, the European Union enacted the General Data Protection Regulation (GDPR). I examine its impact on firms' internal information quality (IIQ) and operating efficiency in the United States. Although privacy regulations, such as GDPR, target one subset of firms' information assets (i.e., personal data), academics and practitioners have emphasized the ability of these regulations to drive broad improvements in firms' information management practices resulting in higher quality information available for decision making and, by extension, more efficient operations. At the same time, GDPR's regulatory mandates are likely to burden operations. Using multiple modeling approaches to identify the effect of GDPR on US firms and a variety of IIQ proxies from financial reports and disclosures, I find that (a) GDPR leads to improvements in IIQ for impacted firms and (b) that these improvements in IIQ are beneficial to firm operations. However, the regulatory burden of GDPR has overwhelmed these benefits resulting in a negative net effect on firms' operating efficiency.

To Jess, my parents, and the rest of my family

ACKNOWLEDGMENTS

Over the course of my professional and academic career, I have been blessed with numerous mentors that have contributed significantly to my personal and professional development. I highlight specifically two of my former colleagues at KPMG, Peter Dunbar and Nicole Lauer, who mentored me as a young professional, and my dissertation co-chairs, Dr. Jagan Krishnan and Dr. Jayanthi Krishnan, who have given so much of their time and energy to help me develop as an academic over the past five years. I owe these individuals a significant debt of gratitude. I also thank the other members of my committee (Dr. Yi Liang and Dr. Hyun Park) for their guidance in developing this dissertation, Dr. Sunil Wattal for serving as an external reader, and many other Temple faculty and students who have provided helpful feedback along the way. Any errors are my own.

Additionally, I cannot thank my family enough for their love and encouragement over the course of my PhD journey, particularly my wife, Jess, who has consistently and supportively stood by my side across the ebbs and flows of the past five years. As I enter the next phase of my academic career, I intend to do all that I can to repay those that have helped me reach this point by striving every day to be the best version of myself that I can.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In recent years, many jurisdictions have moved to modernize commercial data privacy regulations, granting individuals greater control over the collection, storage, and processing of their personal information. The foundational example of such regulations, the European Union's General Data Protection Regulation (GDPR), was adopted in April 2016 after a five-year debate over comprehensive reforms to outdated privacy regulations enacted in 1995 (EDPS 2021).[1] As of June 2022, GDPR has resulted in over 1,100 fines totaling close to €1.6 billion levied against global organizations.[2] The Brookings Institute notes that GDPR has "changed the privacy dialogue for businesses and governments around the world" (Chin 2019, para. 2). While the intended goal of such privacy regulations is to improve protections for personal data, practitioners have highlighted GDPR's broad transformative effects on firms' information governance programs resulting in higher quality information available for decision making (ARMA 2017; BDO 2018; Smallwood 2019; Baker McKenzie 2020). Higher quality internal information should, by extension, support greater operating efficiency for firms. However, these benefits to operating efficiency might be overwhelmed by the broader regulatory burden of GDPR.

---

[1] Data privacy refers to the rights of individuals to control their personal data. Data protection refers to the responsibilities of organizations to ensure that the rights of individuals to privacy are instantiated in the organization's controls over the data collected (Forbes 2018a). Although GDPR is focused on data protection, as its title indicates, it is often referred to as a data privacy law even in official EU pages (e.g., https://gdpr.eu/what-is-gdpr/).

[2] Number obtained based on fines publicly reported on EnforcementTracker.com.

In this study, I explore the effects of modern privacy regulation on firms' internal information quality (henceforth, "IIQ") and operating efficiency using GDPR as the earliest and most significant example of such regulations globally, which impacted United States firms to varying degrees.[3] Despite GDPR being an EU regulation, compliance is required for US firms that collect or process data on EU citizens, and a significant number of these firms (over 29% in my sample) reference GDPR in their 10-K "risk factors" (Item 1A) disclosures in the years after the regulation became effective. Further, unlike EU headquartered firms, which were previously required to comply with a broad privacy regulation (the EU's 1995 Data Protection Directive), US firms were generally not subject to significant privacy regulations prior to GDPR except in specific industries.[4] Therefore, the regulatory shock of GDPR to US firms is substantial yet varied based on the extent to which US firms interact with EU citizens, making the US an attractive setting within which to study the effects of GDPR.

I start by studying the effects of the regulation on firms' IIQ, which is defined as "the accessibility, usefulness, reliability, accuracy, quantity, and signal-to-noise ratio of the data and knowledge collected, generated, and consumed within an organization" (Gallemore and Labro 2015, p. 139). GDPR directly requires that firms meet certain obligations in collecting, storing, and using EU citizens' personal data. Among these requirements, firms must conduct data protection impact assessments of systems housing

[3] As I describe later, my focus is on publicly traded United States firms although GDPR applies to both public and private organizations.

[4] As discussed later, the best example of such pre-GDPR privacy regulation affecting US firms was the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule, which became effective in 2003 and focused on protected health information collected by healthcare firms.

personal data (Article 35), implement "data protection by design and default" into these systems (Article 25), implement systems controls and processes to record and track data processing activities (Article 30), and install a Data Protection Officer to a fixed role (Article 37).[5] While the personal data specifically targeted by the regulation represents only one subset of a firm's critical information assets, the governance practices and controls over information that GDPR emphasizes are likely to have broader consequences for the information ecosystem of the organization.[6] The accounting and consulting firm BDO USA LLP notes that GDPR has driven "a paradigm shift on how organizations view and manage their data" (BDO 2018, para. 6). Improvements in these information management practices are likely to support concurrent improvements in the quality and reliability of information assets used for decision making (Smallwood 2019). Therefore, I hypothesize that GDPR will result in improved IIQ for firms more significantly impacted by the regulation.

Turning to GDPR's effects on operating efficiency, past literature in both accounting and information systems has established a positive relationship between IIQ and operating efficiency (e.g., Mithas, Ramasubbu, and Sambamurthy 2011; Cheng, Goh, and Kim 2018). This relationship stems from the theory that management's ability to coordinate operations in a manner that maximizes revenue from a given set of inputs is largely dependent on the quality of information available to make resource allocation decisions. As such, to the extent that GDPR results in higher quality information for

[5] References are to the official Articles of the General Data Protection Regulation available at https://gdpr-info.eu/.

[6] Information assets are defined as "a body of information, defined and managed as a single unit so it can be understood, shared, protected and exploited effectively" (National Archives 2017).

firms, I can also expect to observe a positive effect on operating efficiency. Therefore, I hypothesize that GDPR will have a positive indirect effect on firm operating efficiency by way of improved IIQ.

The operational benefits derived above, however, face other countervailing forces as a consequence of GDPR. Limitations on data sharing resulting from GDPR may prevent the realization of "data synergies," deleteriously affecting decision-making and operations (Gal and Aviv 2020). Further, many of the activities required under the regulation (discussed above) represent compliance costs to firms and will, at least in the short-run, require a diversion of resources away from other revenue-generating activities. By definition, any such diversions will negatively affect firms' operating efficiency. Therefore, I hypothesize that the direct effect of GDPR on operating efficiency (outside of improvements in IIQ) will be negative.

Lastly, I consider whether the net effects of GDPR on operating efficiency are likely to be positive or negative when aggregating the direct and indirect effects of the regulation. To the extent that firms were able to improve IIQ prior to the adoption of GDPR, but did not do so, their revealed preferences suggest that these changes were not worthwhile prior to the regulation. While the passage of GDPR increased the net benefits for firms improving information management practices, these benefits are not likely to have offset the costs of making such changes to comply with the regulation. If they did, firms would have been incentivized to improve information management practices even prior to the passage of GDPR. Therefore, I hypothesize that the net effect of GDPR on operating efficiency will be negative.

Key to studying the hypotheses above is identifying treatment (i.e., US-headquartered firms significantly affected by GDPR) and control firms (i.e., US-headquartered firms not (or less) impacted by GDPR). I identify firms significantly affected by GDPR in two ways. First, I identify firms making a 10-K risk factor disclosure reference to GDPR, which allows me to observe how the effects of GDPR manifest across both time and across firms. I exploit both cross-sectional and intertemporal variation in these disclosures to identify which firms are affected by GDPR and when. Second, as an alternative to the disclosure-based identification strategy above, I identify US-headquartered firms with an EU segment or subsidiary over my sample period as treatment firms assuming that firms with a substantial presence in the EU will be more affected by GDPR than other US-headquartered firms.

As proxies for IIQ and operating efficiency, I rely on past literature studying these outcomes. Starting with IIQ, I use observable proxies for the quality of firms' internal information from publicly available sources (e.g., Gallemore and Labro 2015; Heitzman and Huang 2019). Since no single proxy perfectly captures IIQ, I define an index using attributes from (a) internal control reports under Sarbanes-Oxley Section 404, (b) financial statement filings, and (c) management guidance issuances. My operating efficiency measure, which is based on Demerjian, Lev, and McVay (2012), is a data envelopment analysis (DEA)-based proxy for identifying the extent to which the firm optimally deploys its input resources to generate sales.

Using the GDPR impact, IIQ, and operating efficiency proxies described above, I estimate systems of seemingly unrelated regressions to identify (a) the direct effect of GDPR on IIQ, (b) the indirect effect of GDPR on operating efficiency by way of

improved IIQ, and (c) the direct effect of GDPR on operating efficiency. When running these models, three notable findings emerge. First, I find that firms' IIQ improves in the presence of GDPR. Second, GDPR-induced improvements in IIQ have a positive (statistically significant albeit small in magnitude) effect on operating efficiency. Third, GDPR-impacted firms experience significant declines in operating efficiency that overwhelm the benefits stemming from improvements in IIQ. Therefore, while GDPR drives firms to improve IIQ, the resulting operational benefits are engulfed by the regulatory burden of GDPR.

In addition to my main tests, I show that my results vary cross-sectionally across dimensions of industry and firm size. Namely, high technology firms, which process significant volumes of personal data, and small firms, which are less likely to have had the infrastructure in place for GDPR compliance, experience greater improvements in IIQ and larger declines in operating efficiency. As robustness tests, I show that my results are robust to using multiple variants of my operating efficiency measure and that the improvements in IIQ that I observe persist after controlling for differences in boards' attention to cybersecurity. The latter of these analyses is important in light of recent literature (Klein, Manini, and Shi 2022) showing that, around the adoption of GDPR, firms increase board attention to cybersecurity, which could conceivably drive the effects on IIQ that I document.

Collectively, this dissertation contributes to an emerging stream of literature studying the implications of modernized privacy regulations on the global economy. Recent studies have begun to explore the effects of GDPR on topics such as customer web behavior, firm web traffic tracking, and technology venture investment (e.g., Aridor,

Che, and Salz 2020; Goldberg, Johnson, and Shriver 2021; Jia, Jin, Wagman 2021). Other studies have focused on the shift in attention of firms and their boards to cyber risk after the passage of privacy-related regulations such as GDPR (e.g., Ashraf and Sunder 2021; Klein, Manini, and Shi 2022). However, studying the broader effects of such regulations on firms' information assets and operations is important given the movement by many governing bodies globally, including those in the US, toward stronger privacy regulations and cybersecurity reporting requirements.[7] As such, the evidence presented in this paper should be considered important to academics, practitioners, and regulators alike. Lastly, my findings reflecting both GDPR-induced benefits and costs to firms, which are important societal actors, provide empirical support for the notion that "the protection of privacy can both enhance and detract from individual and societal welfare" (Acquisti, Taylor, and Wagman 2016, p. 442).

[7] Gartner, a global information technology research company, predicts that, by 2023, 65% of the world's population will have its personal information covered under modern privacy regulations, up from 10% in 2020 (Moore 2020). Further, the SEC has recently announced a proposal to standardize "disclosures regarding cybersecurity risk management, strategy, governance, and incident reporting by public companies" (SEC 2022, para. 1).

CHAPTER 2

INSTITUTIONAL BACKGROUND

2.1. The EU's General Data Protection Regulation

Starting in 2010, the EU's governing bodies began formally discussing approaches to address concerns over organizations' handling of personal data. At the time, the Data Protection Directive, enacted in October 1995, was the governing regulation over the handling of personal data within the EU.[8] While this directive had been a significant milestone for data protection at the time of its adoption, the decades that followed introduced new challenges for the handling of personal data driven by technological change and globalization. A 2010 report from the European Commission to the European Parliament highlighted these challenges and, further, noted inconsistencies in laws enacted to execute the directive across EU member states (EC 2010). In the years that followed, the EU's three governmental branches collaborated to arrive at a general framework for GDPR. On January 25, 2012, the European Commission proposed a comprehensive reform of the EU's 1995 data protection rules, and on March 12, 2014, a European Parliament plenary vote demonstrated strong support for GDPR. The final

[8] Three branches compose the EU's political institution. The Council of the EU is comprised of national and EU-level leaders and sets the overall political agenda for the EU. The European Parliament is comprised of representatives directly elected by EU citizens and has legislative and budgetary responsibilities. The European Commission is comprised of commissioners appointed by each EU country and is responsible for drawing up proposals for new EU legislation and implementing/enforcing decisions of the European Parliament and the Council of the EU. These three branches work collectively to propose, enact, and enforce regulations applicable to EU member states. A full overview of the EU political apparatus is available at: https://europa.eu/european-union/about-eu/institutions-bodies_en.

version of the regulation was adopted on April 27, 2016. Two years later, on May 25, 2018, the regulation became effective.[9]

GDPR's intent was to clarify the rights of EU citizens to their personal data and require that organizations handling their data implement processes to support these rights.[10] Under the regulation, personal data is defined broadly to include "any information relating to an identified or identifiable natural person ('data subject')" including names, addresses, and even web cookies (Article 4). To protect individual privacy rights, GDPR requires that organizations take numerous steps based on six principles related to the processing of personal data (see Appendix B for an overview of these principles). Generally, organizations processing these data elements are required to collect only data necessary to support "legitimate" commercial purposes, to be transparent regarding the data collected and how it will be used, and to gain "opt-in" consent for its collection from the data subject. Additionally, once collected, organizations must implement controls and processes to ensure that the data remains protected and accurate.

Under these broad principles, there are numerous specific requirements that GDPR places upon organizations. First, organizations are required to develop systems and processes to track data processing activities throughout the organization (Article 30).

---

[9] See Appendix A for a timeline of the primary events leading to the passage and applicability of GDPR with special attention given to the key dates mentioned above.

[10] Articles 12 through 23 of the regulation outline these rights including (1) the right to be informed of what data is collected; (2) the right to access / review data held; (3) the right to rectify incorrect information or request erasure of data; (4) the right to restrict processing on personal data; (5) the right to receive data in a commonly used format; and (6) the right to object to how personal data is used particularly in the context of automated decision making.

Second, organizations are required to include data protection "by design and by default" into systems processing relevant data (Article 25). Third, organizations are required to conduct data protection impact assessments following changes to activities or systems that may come under the requirements of GDPR (Article 35). Fourth, identified data breaches must be reported to supervisory authorities within 72 hours of discovery by organizations (Article 33). Lastly, for large-scale data processors, organizations are required to install a data protection officer that oversees organizational compliance with the regulation (Article 37). As I discuss in Chapter 3, these requirements have the potential to spur firms to improve the programs governing their key information assets.

2.2. Implications of GDPR Outside of the EU

Importantly, the geographic scope of the regulation spans beyond organizations located in the EU. Rather, GDPR applies to any organization handling the personal data of EU citizens regardless of its jurisdiction (Article 3). Additionally, the financial penalties for non-compliance with the regulation can be significant, with a regulatory limit of up to 4% of annual revenues (Article 83). For example, in December 2020, Google and Amazon were fined $120M and $42M, respectively, by French authorities for loading tracking cookies onto users' computers when visiting their websites in a manner that violated GDPR (Lomas 2020).

GDPR is the first privacy law applicable to US firms broadly (albeit to varying degrees).[11] Previous privacy regulations impacting US firms were focused on specific activities (e.g., state-level breach disclosure laws; see Ashraf and Sunder [2021] for an

[11] The official EU site used to educate organizations on compliance with GDPR (https://gdpr.eu) has even published knowledge materials directed at US companies such as "The GDPR Compliance Checklist for US Companies" (available at https://gdpr.eu/compliance-checklist-us-companies/).

overview of these laws) or industries (e.g., the 2003 HIPAA Privacy Rule, which focuses on protected health information collected by healthcare firms). As such, the passage of GDPR reflected a significant shock to US firms' cost-benefit assessments of whether and how to implement policies and procedures addressing data protection.[12]

GDPR has also spurred individual states in the US to consider and implement similar regulations. The first and most notable of these regulations is California's Consumer Privacy Act (CCPA), which became effective in 2020, close to two years after GDPR's effective date. In Appendix C, I compare key features of GDPR and CCPA noting many similarities except for the level of potential fines, which are far smaller under CCPA than under GDPR.[13]

[12] For example, in Appendix C, I compare key features of GDPR and the HIPAA Privacy Rule noting significant differences in terms of scope of organizational coverage and level of potential fines.

[13] Nonetheless, recognizing the potential confounding effects of CCPA on the models I estimate below, in robustness analyses, I show that my results hold when controlling for differences in outcomes between firms headquartered in California (those most likely to be affected by CCPA) and other firms around the adoption of GDPR.

CHAPTER 3

HYPOTHESES DEVELOPMENT

Below, I present hypotheses pertaining to the effects of GDPR on IIQ and operating efficiency, which I have summarized in Figure 1.

Figure 1
Hypotheses Summary

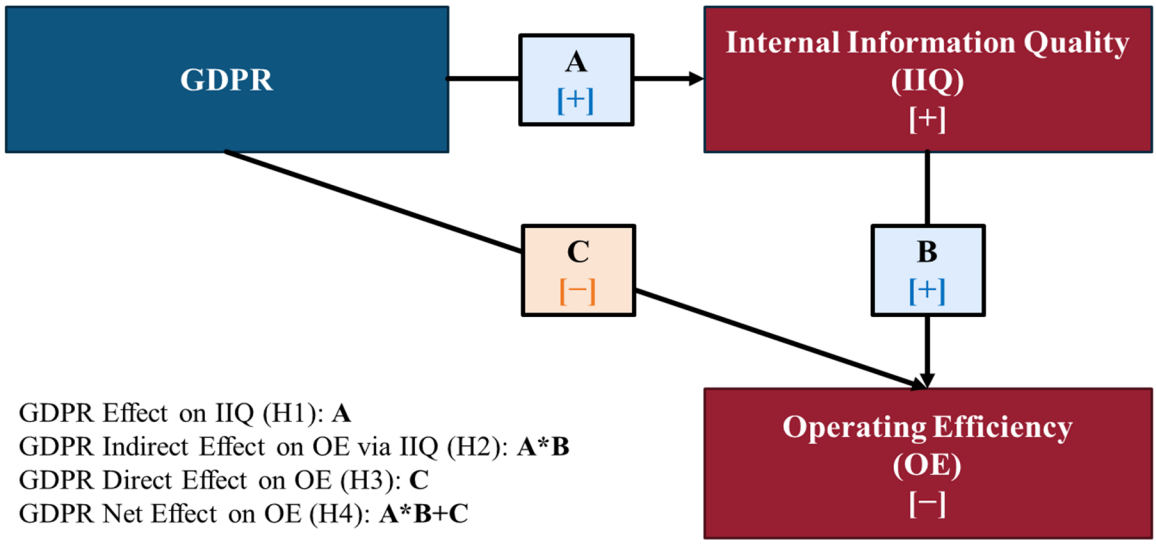


Figure 1 depicts the predicted effects of GDPR on IIQ and operating efficiency.

3.1. GDPR and Internal Information Quality

As stated above, IIQ refers to "the accessibility, usefulness, reliability, accuracy, quantity, and signal-to-noise ratio of the data and knowledge collected, generated, and consumed within an organization" (Gallemore and Labro 2015, p. 139). Numerous recent studies have explored this concept showing it to have implications for areas as varied as operating efficiency, investment efficiency, workplace safety, and tax planning (e.g.,

Cheng et al. 2018; Christensen, Lynch, and Partridge 2021; Laplante, Lynch, and Vernon 2021; Hope, Wang, Yue, and Zhao 2022).

In the broadest sense, the quality of information available for decision making within an organization is based heavily on the processes and systems employed to manage information risks and optimize information value. These objectives and practices fall under the umbrella of "information governance." Information governance is "an organization's coordinated, inter-disciplinary approach to satisfying information compliance requirements and managing information risks while optimizing information value" (Sedona 2019, p. 104). It reflects a collective effort across various departments and organizational roles (e.g., legal, information technology, and business process owners) instead of siloed responsibilities for processing and securing data that have often pervaded organizations. One example of such cross-functional interaction is the classification and valuation of information assets in a coordinated manner between business process owners and legal representatives. The output of this exercise is helpful to IT and other departments responsible for scoping and implementing data protection controls.

The adoption of GDPR, a new and key compliance requirement centered on firms' information management practices, has elevated the importance of information governance for many organizations globally. These improvements are likely to occur due to both (a) certain direct requirements of the regulation as well as (b) the broader emphasis on improved information management that the regulation encourages. First, GDPR directly requires that firms take steps to improve information management practices including conducting protection impact assessments of systems housing

personal data (Article 35) and implementing "data protection by design and default" into these systems (Article 25). While it is conceivable that the systems and practices implemented to support these requirements may be limited in scope to cover only personal data of EU citizens (rather than other key information assets used in making operating decisions), the effects of these improvements are likely to be broader. Since personal information is embedded into the wider information ecosystem of the organization, investments made in systems and technology to better manage one information asset class are likely to spillover to other critical information asset classes.[14] Further, other requirements of the regulation such as implementing systems controls and processes to record and track data processing activities (Article 30) and installing a Data Protection Officer to a fixed role (Article 37) are information-asset-agnostic and are likely to have positive implications for the management of information across the organization.

Even beyond the direct requirements of the regulation, consulting firms have highlighted that organizations are not looking at GDPR as a check-the-box compliance activity. Rather, they see GDPR as an opportunity to make investments in information management practices that enable compliance with GDPR and other similar future regulations and may also achieve broader benefits for the organization. The accounting and consulting firm BDO, for example, stated in March 2018,

> Businesses are using the GDPR as an opportunity to build or in some cases re-establish information governance programs. The GDPR deadline is driving much of this activity with the expectation that this is not a one-

[14] For example, GDPR-driven upgrades/replacements of ERP and CRM systems housing customer data may allow for reductions in the accrued "technical debt" in these systems, improving their broader business value and the ability to extract quality information from them (Banker, Liang, and Ramasubbu 2021).

> time, "check the box" project, but a paradigm shift on how organizations view and manage their data (BDO 2018, para. 6).

Smallwood (2019, p. 3) highlights one specific example of compliance with GDPR spurring improvements to information governance, stating,

> A first step in the GDPR compliance process is to conduct an inventory of an enterprise's information assets to create a data map showing where all the incidences of data are housed. This is commonly the first major implementation step in IG programs, so the discipline and support for IG programs made substantial strides in 2018 with the lead-up to GDPR going into effect.

Supporting the above notion that the effects of GDPR have consequences for information management and quality beyond the information directly subject to the regulation, recent work (i.e., Ashraf and Sunder 2021; Klein et al. 2022) has highlighted that firms take actions such as increasing IT spending and improving board oversight for cyber risk after they become subject to new privacy-related regulations.[15] This evidence is consistent with recent survey data in which 52% of UK firms reported increasing investment in cybersecurity due to GDPR (RSM 2020). Therefore, changes specifically required for GDPR compliance as well as the broader investments in personnel, processes, and systems encouraged by the regulation are likely to lead to improved IIQ for firms. As such, I propose the following hypothesis, which is depicted as Path A of Figure 1:

> *H1: After GDPR's adoption, IIQ improves more for GDPR-impacted firms than other firms.*

[15] In robustness analyses, I find evidence consistent with Klein et al. (2022) that GDPR has led firms to increase the board-level attention to cyber security, which gives credence to the notion that firms are investing in knowledge and expertise over information and technology risks after the adoption of the regulation. I show that the effects of GDPR on IIQ are incremental to these board-level changes.

While I expect GDPR to positively affect IIQ under H1, this hypothesis is not without tension. There are three primary reasons why I may observe no change or a decline in IIQ as a result of GDPR's adoption. First, Jamal, Maier, and Sunder (2005) have shown that the EU privacy regulation pre-dating GDPR (the 1995 Data Protection Directive) had little effect on firms' privacy-related disclosures and practices. While GDPR significantly expands firms' compliance requirements beyond those in the earlier Directive, it could be that GDPR does not represent a significant enough shock to spur changes to US firms' information management practices. Second, even if US firms invest in projects (particularly IT-related projects) to improve information management practices because of GDPR, these projects may not yield expected benefits limiting IIQ improvements (Nelson 2007). Lastly, according to Gallemore and Labro (2015), one of the key components of IIQ is information accessibility. To the extent that GDPR places limitations on the collection and use of certain information assets (i.e., personal information), the expected improvements in IIQ might be muted for these particular information assets. Overall, while these effects may limit the improvement in (or result in the degradation of) the quality of firms' internal information in certain cases, on balance, I expect the broad improvements in information management incentivized by GDPR to lead to improved IIQ.

3.2. GDPR and Operating Efficiency

3.2.1. The Indirect Effect via Improved IIQ

To the extent that GDPR drives improvements in IIQ for firms, it is natural to expect that these improvements will, by extension, enable firms to conduct operations more efficiently. The connection between IIQ and operating efficiency has roots in

management accounting literature, which emphasizes that higher IIQ will lead to improved decision-making (Kinney 1999; Horngren, Datar, Foster, Rajan, and Ittner 2012; Gallemore and Labro 2015). Therefore, it is natural to expect that, as the quality of the information used by management in operational decision making improves, so too will the efficiency of operations. Consistent with this notion, Cheng et al. (2018) find improvements in internal controls over financial reporting, and by extension improvements in management's internal information set, are associated with improved operating efficiency. Outside of academic literature, Iron Mountain (2014), a world-leader in data and records management, highlights that improved information management practices will result in improved operating efficiency. Therefore, I present the following hypothesis, which is depicted as the positive signs on the indirect path between GDPR and firm operating efficiency by way of improved information governance (Path A to B) in Figure 1:

> *H2: The indirect effect of GDPR via improvements in IIQ will have a positive effect on firm operating efficiency.*

3.2.2. The Direct Effect of GDPR's Regulatory Burden on Operating Efficiency

While positive consequences for operating efficiency are likely to manifest through GDPR-induced improvements in IIQ, the regulatory burden of GDPR imposes various costs on impacted firms that will negatively affect firm operations. These costs include both short-term costs of compliance as well as on-going limitations on how firms can collect, share, and use information. According to a 2017 PWC survey of US companies with over 500 employees, 77% intended to spend $1 million or more on GDPR compliance efforts (PWC 2017). These costs are likely much higher for larger organizations, and the price tag for compliance efforts across Fortune's Global 500

companies has been estimated at $7.8 billion (IAPP 2017).[16] To the extent that these compliance costs divert resources from other revenue-generating activities, operating efficiency will be adversely affected. Even beyond the short-term costs of compliance, GDPR may disrupt firm operations in a less quantifiable manner for years into the future. For example, limitations on data sharing may prevent the realization of "data synergies" and inhibit the ability of firms to extract value from their data assets (Gal and Aviv 2020).

Interestingly, both sets of costs (short-term compliance costs and longer-term effects on the use of information) are explicitly highlighted in Microsoft's 2019 10-K risk factors section. First, Microsoft speaks to the expenses incurred to comply with GDPR (emphasis added),

> In May 2018, a new EU law governing data practices and privacy, the General Data Protection Regulation ("GDPR"), became effective. The law…imposes a range of new compliance obligations regarding the handling of personal data. **Engineering efforts to build new capabilities to facilitate compliance with the law have entailed substantial expense and the diversion of engineering resources from other projects and may continue to do so** (Microsoft 2019, p. 26).

Next, Microsoft expresses concerns over the longer-term implications of GDPR for extracting value from information collected,

> The Company's investment in gaining insights from data is becoming central to the value of the services we deliver to customers, to our operational efficiency and key opportunities in monetization, customer perceptions of quality, and operational efficiency. Our ability to use data in this way may be constrained by regulatory developments that impede realizing the expected return from this investment (Microsoft 2019, p. 26).

[16] In comparison, Krishnan, Rama, and Zhang (2008) report that US publicly traded firms spent, on average, $2.2 million to comply with Sarbanes-Oxley Section 404, arguably the most onerous US regulation relating to financial information management to date.

Overall, there is substantial evidence that GDPR will negatively affect firms' operating efficiency. As such, I present the following hypothesis and depict it as the negative sign on the direct path between GDPR and operating efficiency (Path C) in Figure 1:

H3: When controlling for changes in IIQ, the direct effect of GDPR on operating efficiency is negative.

Tension in this expectation comes from considering changes in behavior by privacy-conscious consumers when engaging with firms affected by the regulation. Past information systems literature has found that firms' privacy assurances and policies lead to greater user/customer engagement (Hui, Teo, and Lee 2007; Tsai, Egelman, Cranor, and Acquisti 2011).[17] These changes in user and consumer behavior may allow for increased sales, which in turn, would partially mitigate other forces (discussed above) that negatively affect operating efficiency.

3.2.3. The Net Effect of GDPR on Operating Efficiency

Finally, I consider the net effect of GDPR on operating efficiency for firms. To the extent that GDPR drives firms to make changes that lead to improved information quality, they reveal that these changes were not worthwhile prior to GDPR's passage. As such, I expect that benefits from GDPR-induced improvements in IIQ are unlikely to exceed the negative ramifications of the regulation for operating efficiency. Therefore, I hypothesize that the net effect of GDPR on operating efficiency will be negative.

H4: The net effect of GDPR on operating efficiency is negative.

[17] However, studying crowdfunding platforms, Burtch, Ghose, and Wattal (2015) find that reductions in information control questions lead to increased engagement and net contributions by users. Analogously, changes to websites to require opt-in consent for web cookies (a requirement under GDPR) may reduce engagement by customers.

CHAPTER 4

GDPR IMPACT IDENTIFICATION STRATEGY

The primary challenge in studying the effects of GDPR is delineating firms that are more and less affected by the regulation. Considering the regulation's extraterritorial scope, any global firm that handles relevant data (discussed above) of EU citizens is required to comply with GDPR. I focus on a setting where variation in the effects of GDPR is significant, i.e., public firms headquartered in the US and listed on US exchanges. Unlike EU firms, which were subject to the 1995 Data Protection Directive prior to GDPR, US firms were not subject to a previous omnibus privacy regulation. Therefore, compared to EU firms, it is likely that US firms may experience an even more significant shock as a consequence of GDPR. For example, large US (Fortune 500) firms spent seven times as much as large UK (FTSE 350) firms ahead of the May 2018 effective date for the regulation (Forbes 2018b). Focusing on US firms has the benefit of exploring the effects of GDPR across a sample of firms with a relatively homogeneous external environment (e.g., SEC oversight, litigation pressures, etc.). To identify differences in the extent to which US firms are affected by GDPR, I use two approaches, which I discuss below.

4.1. Identification Strategy #1: Changes in GDPR Risk Factor Disclosures

My first approach to identify US firms more acutely impacted by GDPR is to focus on risk factor disclosures in annual 10-K filings. Within the 10-K filing, Item 1A risk factors include information about the most significant risks that apply to the company and its securities. References to GDPR within this section imply that the firm considers the regulation to be of appropriate significance to report as a risk to investors.

Past studies have found that the market values proactive disclosures regarding information security (e.g., Gordon, Loeb, and Sohail 2010) suggesting that firms are incentivized to be forthcoming with information on the effects of related regulations, such as GDPR, in these disclosures.[18] Therefore, I use these disclosures as a direct proxy to measure differences in the extent to which firms perceive themselves to be affected by GDPR.[19] Specifically, I define *GDPRRISK* as an indicator for the presence of a GDPR-related risk factor disclosure for a given firm-year observation. When used in firm fixed effects regressions with controls for other determinants of these disclosures, *GDPRRISK* identifies how the effects of GDPR manifest for a firm over time.

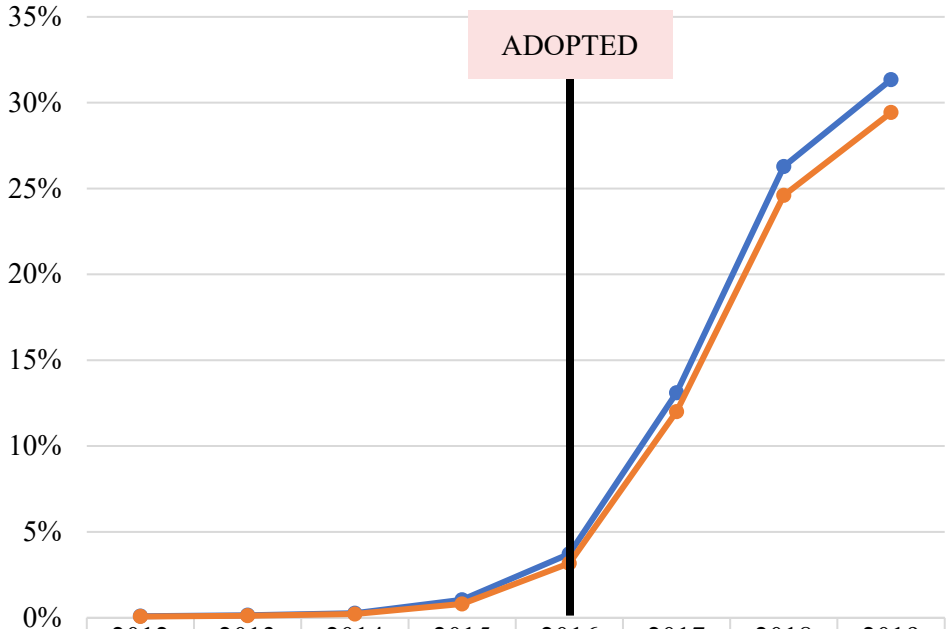
To construct this measure, I extract the risk factors section from 10-K filings. I then search the extracted risk factors section to identify uses of the terms "General Data Protection Regulation" or "GDPR." In Figure 2, I plot the percentage of firm-year observations by year for which these terms appear both within the risk factors section (and, for comparison purposes, the entire 10-K). The strongest upward trend in GDPR-related disclosures occurs between 2016 and 2018 as the regulation was adopted and eventually became enforceable. By 2019, over 31% of firms in my sample referred to the regulation in their annual filing. Of these cases, more than 93% included a GDPR-related risk factor. In Figure 3, I present the percentage of firm-year observations in each of the

[18] Other studies exploring the increasing prevalence in cyber risk factor disclosures have highlighted an increasing trend in these disclosures particularly after SEC guidance published in May 2011 (e.g., Hilary, Segal, and Zhang 2016; Li, No, and Wang 2018).

[19] In common with other studies of risk disclosures, I acknowledge that risk factor disclosures are based on the implicit assumption that firms accurately reveal the most significant risks to their organization. Evidence to support this notion comes from past studies that show risk factor disclosures are informative to the market (Campbell, Chen, Dhaliwal, Lu, and Steele 2014; Hope, Hu, and Lu 2016).

Fama-French 12 industry classifications that make GDPR-related disclosures. Not surprisingly, the three most likely industry classifications to make such disclosures are the Business Equipment/Information Technology, Telecommunications, and Healthcare industries. The first two of these industries are composed of many large data processing entities, and healthcare entities handle significant volumes of individuals' personally identifiable and protected health information.

Figure 2
GDPR References in 10-K Filings by Year



	2012	2013	2014	2015	2016	2017	2018	2019
Full 10-K	0.1%	0.1%	0.3%	1.0%	3.7%	13.1%	26.3%	31.3%
Risk Factors	0.1%	0.1%	0.2%	0.8%	3.2%	12.0%	24.6%	29.4%

Figure 2 provides the percentage of firms in my sample (see Chapter 6 for sample definition) referencing GDPR in their 10-K filing each year.

Figure 3
GDPR-Related Risk Factors by Industry

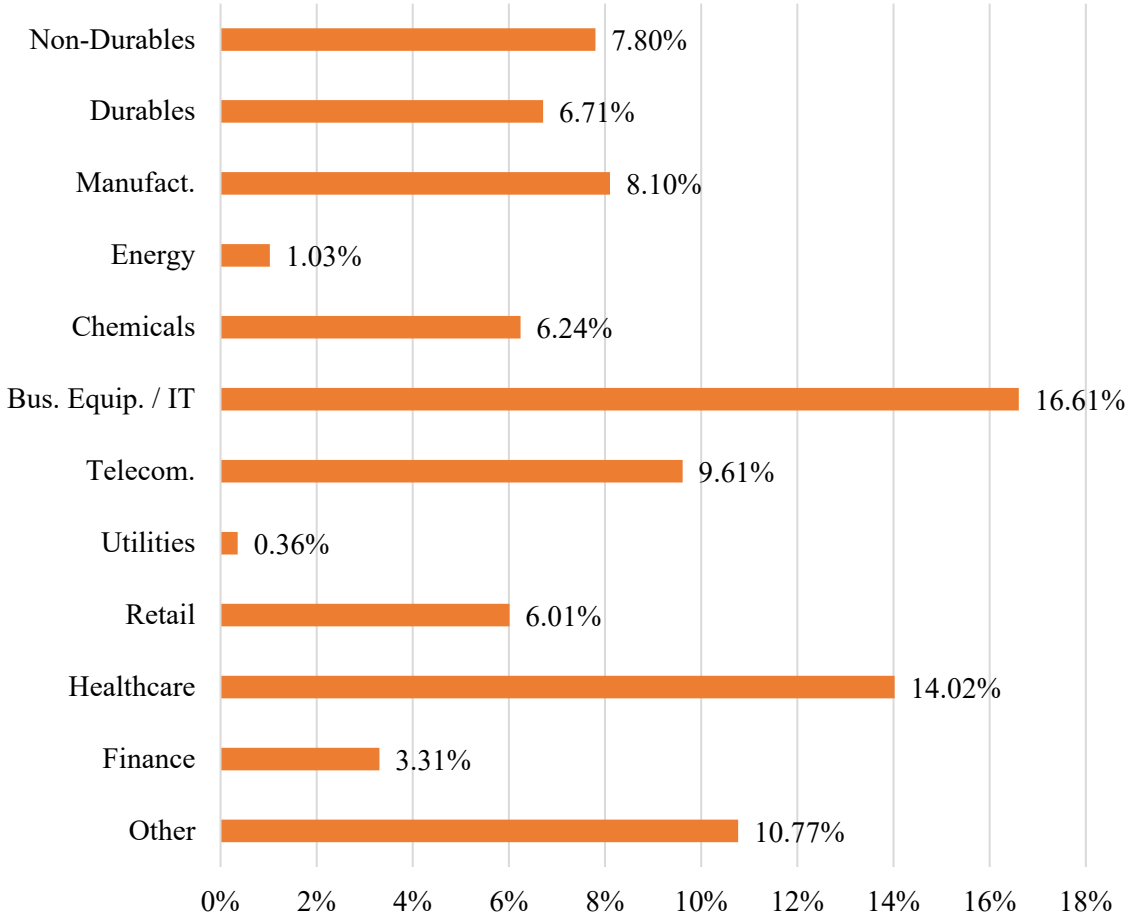


Figure 3 provides the percentage of firm-year observations in my sample (see Chapter 6 for sample definition) referencing GDPR in the risk factors section of their annual 10-K filing by industry (based on the Fama-French 12 industry classifications).

4.2. Identification Strategy #2: Firms with European Operations

GDPR is likely to have a more significant effect on firms that interact explicitly with EU citizens as customers, vendors, or employees. For this reason, I define an indicator coded 1 for firms reporting an EU segment (from Compustat) or EU subsidiary

(from WRDS Subsidiary data) over the sample period (*EUOPS*).[20] *EUOPS* identifies *cross-sectional* variation in the impact of GDPR across firms. This variation can be exploited in difference-in-differences models to identify how my outcomes change for firms with European operations compared to other firms around the year of GDPR adoption.

[20] The subsidiaries in the WRDS Subsidiary database are obtained from Exhibit 21 disclosures required to accompany Form 10-K by the SEC. Dyreng, Hoopes, Langtieg, and Wilde (2020, p. 643) study these disclosures and conclude that they "provide a reasonable proxy for locations of significant subsidiaries." Since WRDS Subsidiary data is only available through a portion of 2019 (and my sample period covers fiscal years 2012 to 2019), I base my *EUOPS* measure on firm observations from 2012 to 2018. As I describe below, when running my difference-in-differences models using this variable as a treatment indicator, I require firms to have observations in both the pre (2012 to 2015) and post (2016 to 2019) periods. Nonetheless, there is a chance that firms who have not reported an EU segment or subsidiary from 2012 to 2018 report one in 2019. These firms would be in my control group and, since they are likely to be more similar to those that I identify as *EUOPS*=1 firms, their presence in the control group should bias against me finding a result.

CHAPTER 5

RESEARCH DESIGN

5.1. GDPR and IIQ (H1)

5.1.1. Measurement of IIQ

Past studies have highlighted that there is no single, publicly-available proxy that perfectly captures the quality of firms' internal information (e.g., Gallemore and Labro 2015; Heitzman and Huang 2019; Hope et al. 2022). Therefore, as my primary proxy for IIQ, I define an index (*IIQ*) comprising observable traits of firms used to capture components of IIQ in prior literature. Specifically, I define *IIQ* as the average of the standardized values of variables reflecting firms' internal control quality (*NOICMW* and *NOITMW*), financial reporting accuracy and speed (*NORESTATE* and *NOLATEFILER*), and management guidance issuances (*GUIDANCE* and *GUIDANCEACC*). I discuss each of these variables in the sections below.

Internal control quality. The first two variables used in the construction of *IIQ* are indicators coded one for firms that (a) maintain effective ICFR (*NOICMW*) and (b) maintain effective IT-related ICFR (*NOITMW*). These variables reflect the quality of firms' internal controls over managing financial information, which when deficient can lead to errors in management's internal reports (Feng, Li, and McVay 2009). I specifically include an indicator related to maintaining effective IT-related ICFR (incremental to *NOICMW*) considering the expected effects of GDPR on systems and technologies employed to manage information. IT-related material weaknesses have also been shown to have more significant negative effects relative to other types of material

weaknesses (e.g., Kim, Richardson, and Watson 2018; Ashraf, Michas, and Russomanno 2020).

Financial reporting accuracy and speed. The next two variables used in the construction of *IIQ* are indicators coded one for firms that do **not** subsequently restate their current year financials (*NORESTATE*) and firms that make an on-time annual 10-K filing (*NOLATEFILER*), respectively. *NORESTATE* captures the firms' ability to prepare accurate external financial reports, which is driven by the accuracy of internal reporting (Gallemore and Labro 2015).[21] *NOLATEFILER* captures differences in external financial report filing speed. Since the quality of internal information is a key input into external financial filings, firms with low quality information will take longer to prepare these reports (Dorantes, Li, Peters, and Richardson 2013).

Management guidance issuance and accuracy. The final two variables used in the construction of *IIQ* are focused on management's issuance of sales guidance.[22] *GUIDANCE* is an indicator coded one for firms that issue annual or quarterly sales guidance during the fiscal year. In an environment where guidance issuances expose firms to increased litigation risk, managers with higher information quality will be more likely to issue guidance (Heitzman and Huang 2019). *GUIDANCEACC* measures the

[21] In my main analyses, I define *NORESTATE* based on all restatements reported in Audit Analytics. However, certain restatements due to fraudulent reporting by management (coded as "Financial Fraud, Irregularities and Misrepresentations" in Audit Analytics) may not be indicative of weak internal information quality. If I exclude these restatements from consideration in constructing *NORESTATE*, my results are consistent.

[22] While other studies tend to focus on earnings guidance as a measure of IIQ, I focus on the issuance of sales guidance since GDPR is likely to be most closely connected to customer information, which can have implications for sales forecasting. For example, GDPR requirements over the handling of web traffic data (by way of web cookies) could induce firms to apply greater governance over demand data that could be useful in predicting future sales (Yang, Pan, and Song 2014).

accuracy of guidance issuances made by taking the negative value of the average difference between the guidance issuances made and actual sales reported (scaled by the actual sales reported) (e.g., Gallemore and Labro 2015; Heitzman and Huang 2019).

Summarization of IIQ components into *IIQ*. To arrive at the value for *IIQ* included in models below, I standardize (i.e., demean and divide by the standard deviation) each of the six IIQ proxies above and take the average of these standardized values.[23] I further standardize the resulting average to allow for easier interpretation in regressions.

5.1.2. Modeling the Effects of GDPR on IIQ

Under H1, I predict that GDPR will lead to improvements in IIQ for firms. To study this effect, I employ the following model:

$$IIQ_{i,t} = \alpha + \beta_1 GDPRIMPACT_{i,t} + \gamma CONTROLS_{i,t} + Fixed\ Effects\ +\ e_{i,t} \quad [1]$$

The dependent variable, *IIQ*, is described above. My test variable, *GDPRIMPACT*, which has two variants, captures the impact of GDPR on IIQ. H1 predicts a positive sign for its coefficient (β_1). The two variants of *GDPRIMPACT* correspond to each of the GDPR impact identification strategies discussed above. The first is *GDPRRISK*, which is a time-varying indicator for firms making GDPR-related risk factor disclosures. The second is

[23] To construct *IIQ,* I require that data be available for at least my internal control (*NOICMW* and *NOITMW*) and financial reporting (*NORESTATE* and *NOLATEFILER*) input variables. For firms covered by I/B/E/S (and for firms with available sales guidance and comparable actuals reported), I also include *GUIDANCE* (and *GUIDANCEACC*) in my construction of IIQ. My main results are robust to limiting my analyses to only firms covered by I/B/E/S as presented in Chapter 9.

the interaction of *EUOPS,* an indicator coded 1 (0 otherwise) for firms with European operations, and *POST,* an indicator coded 1 (0 otherwise) for the year of GDPR adoption (2016) and onward.[24] When employing *EUOPS*POST* as *GDPRIMPACT*, model [1] takes the form of a difference-in-differences regression.[25] In addition to a vector of controls discussed below, I include cross-sectional (either industry – based on the Fama-French 48 classifications – or firm) and year fixed effects to capture unobservable differences across industries/firms and time. These fixed effects lead to the suppression of the standalone *EUOPS* (in firm fixed effects models) and *POST* indicators in models employing *EUOPS*POST* as *GDPRIMPACT*.[26]

As controls, I include a vector of time-varying firm attributes that may be correlated with both my GDPR impact proxies as well as the quality of management's internal information. Particularly important to my setting are controls for the cyber risk exposure of the firm, which may drive improvements in information quality agnostic of GDPR. Specifically, I include (a) an indicator of whether the firm has disclosed a breach in the two years leading up to its fiscal year end (*BREACHFIRM*; Sheneman 2022); (b) the natural log of the number of cyber breaches in the firm's Fama-French 12 industry classification over the two years leading up to the firm's fiscal year end (*BREACHIND*;

[24] I focus on the year of GDPR adoption (2016), rather than the first year of enforcement (2018), as the start of my post period, recognizing that firms will begin to make changes to comply with the regulation once it has been adopted and prior to it being enforced. In additional analyses, I perform a hedonic adaptation analysis where I consider the effects of GDPR on my outcomes of interest in both the years following GDPR adoption (2016 and 2017) and after enforcement (2018 and 2019).

[25] When running these difference-in-differences models, I require that each firm have at least one observation with available data both before and after the adoption of GDPR.

[26] Recall that *EUOPS* is time-invariant for a given firm. Therefore, the base *EUOPS* term is identified and reported in models employing industry-fixed effects, but not when employing firm-fixed effects.

Ashraf 2022); and (c) the scaled count of cyber-related keywords in the firm's 10-K risk factors section (*CYBERRISK*). I also control for the natural log of the word count of the risk factors section (*WRDCNTRISK*) to control for differences in the portion of the 10-K from which *GDPRRISK* is derived.

Remaining controls are drawn from past literature on determinants of internal control issues (e.g., Ashbaugh-Skaife, Collins, and Kinney 2007; Rice and Weber 2012; Ge, Koester, and McVay 2017) along with other factors that may shape management's IIQ and/or external reporting incentives. These include differences in firm fundamentals (size (*SIZE*), leverage (*LEV*), and return on assets (*ROA*)) as well as differences in accounting measurement and information management risk proxied by sales growth (*GROWTH*), inventory (*INVENTORY*), the log of the number of firm segments (*SEG*), as well as indicators for foreign income (*FOREIGN*), merger and acquisition activity (*MA*), and restructuring activity (*REST*). I also include the log of the firm's age (*AGE*) and an indicator for whether the firm employs a Big Four auditor (*AUDBIG4*). I further control for whether the firm is an accelerated filer (*ACCEL*) since these firms typically receive audit opinions on ICFR and have shorter filing deadlines. Lastly, the percentage of institutional ownership for the firm (*INSTOWN*) and the log of the number of analysts following the firm (*ANALYSTS*) capture investor demand for information.

5.2. GDPR and Operating Efficiency (H2 through H4)

5.2.1. Measurement of Operating Efficiency

I use the frontier-based measure of operating efficiency introduced by Demerjian et al. (2012) (*OPEREFF*). This measure uses data envelopment analysis (DEA) to identify the extent to which firms are efficient at converting inputs (cost of sales; selling,

general, and administrative expenses; property, plant, and equipment; operating leases; research and development; goodwill; and other intangible assets) to output (revenue).[27] The output of the DEA model takes the value 0 to 1 with 1 being firms at the frontier of efficiency. Following Cheng et al. (2018), I take the percentile rank of the continuous output from the DEA model by year and Fama-French 48 industry to make my measure more comparable across industries and time.

5.2.2. Modeling the Effects of GDPR on Operating Efficiency

H2 through H4 require that I estimate (a) the direct effect of GDPR on IIQ, (b) the direct effect of GDPR on operating efficiency, and (c) the indirect effect of GDPR on operating efficiency via IIQ. As such, I employ systems of seemingly unrelated regressions to allow hypothesis testing of parameters across equations estimating changes in IIQ and operating efficiency. This set of equations (henceforth, referred to as model [2] composed of equations [2A] and [2B]) takes the following general form:

$$IIQ_{i,t} = \alpha + \beta_{11} GDPRIMPACT_{i,t} + \gamma CONTROLS_{i,t} + Fixed\ Effects + e_{i,t} \quad [2A]$$

$$OPEREFF_{i,t} = \alpha + \beta_{21} IIQ_{i,t} + \beta_{22} GDPRIMPACT_{i,t} + \gamma CONTROLS_{i,t} + Fixed\ Effects + e_{i,t} \quad [2B]$$

[27] Specifically, the measure is derived by solving the following optimization problem as specified in Demerjian et al. (2012). The definition of each variable is described in detail in Demerjian et al. (2012).

$$\max_{v} \theta = (Sales) \cdot (v_1 CoGS + v_2 SG\&A + v_3 PPE + v_4 OpsLease + v_5 R\&D + v_6 Goodwill + v_7 OtherIntan)^{-1}$$

Equation [2A], used to identify the direct effect of GDPR on IIQ, takes a similar form to model [1].[28] Equation [2B], which is used to identify the direct effects of IIQ and GDPR on operating efficiency, takes *OPEREFF* as the dependent variable and includes *IIQ* and *GDPRIMPACT* as key test variables. In both equations, I include the vector of controls employed in running IIQ models above (except for *ROA*, which itself is driven by operating efficiency). In equation [2B], I further include the firm's sales market share in its industry (*MKTSHARE*) and the firm's business segment concentration (*CONCEN*), which have been shown to be associated with operating efficiency in prior literature (Demerjian et al. 2012; Cheng et al. 2018).

To test each of my hypotheses, I rely on coefficients (or products of coefficients) from the set of seemingly unrelated regressions defined in model [2]. Specifically, under H2 (studying the indirect effect of GDPR on operating efficiency via improved IIQ), I expect $\beta_{11}*\beta_{21}$ to be positive and significant. Under H3 (studying the direct effect of GDPR on operating efficiency), I expect β_{22} to be negative and significant. Lastly, under H4 (studying the net effect of GDPR on operating efficiency), I expect $\beta_{22} + (\beta_{11}*\beta_{21})$ to be negative and significant.

[28] Although the equations in model [1] and equation [2A] of model [2] are equivalent, they are implemented differently in STATA owing to the fact that model [1] is specified as a standalone regression while model [2] is specified as a system of seemingly unrelated regressions. Model [1] results are obtained using STATA's standard REG and AREG commands. Model [2] is implemented using the SUREG command to estimate equations [2A] and [2B] together as one system of seemingly unrelated regressions. Since SUREG is unable to efficiently handle large numbers of firm fixed effects indicators (as is the case with AREG for model [1]), I simulate firm fixed effects by demeaning all variables and year fixed effects indicators in equations [2A] and [2B] within the firm (CIK). In all cases, standard errors reported are clustered by firm.

CHAPTER 6

SAMPLE AND GDPR IMPACT PROFILE

In this section, I present the sample used to execute the models outlined above. Before estimating these models, I perform preliminary analyses to examine the correlation between my GDPR impact proxies, *GDPRRISK* and *EUOPS*.

6.1. Sample Selection

Table 1 presents the sample selection for my analyses. I start by identifying firm-year observations for firms headquartered in the US with an annual 10-K filing, positive assets, and a management internal control (i.e., SOX 404a) report in Audit Analytics with fiscal years ending from 2012 to 2019. 2012 is the year in which the European Commission proposed a comprehensive reform of the EU's 1995 data protection rules (see Appendix A), and a small number of firms (0.1% of the sample per Figure 2) began to make GDPR-related risk factor disclosures. I end my sample in 2019 to avoid any confounding effects of the COVID-19 pandemic, which began to affect firm operations significantly in early 2020. Collectively, this eight-year period includes four years prior to (2012 to 2015) and four years after (2016 to 2019) the adoption of the regulation. After removing firm-year observations (a) with missing values for variables necessary to construct model [1] studying IIQ, (b) for which a risk factors section of at least 250 words cannot be programmatically extracted from the 10-K, and (c) with 10-K filing lags (i.e., number of days from fiscal year end date to filing date) greater than 360 days, I am left with 31,630 observations to conduct analyses aimed at (1) understanding the determinants of GDPR-related risk factor disclosures and (2) studying IIQ under H1. For analyses incorporating effects on operating efficiency, my sample is restricted to non-

financial and non-utilities firms for which there is available data to construct my operating efficiency models (model [2]) leaving a total of 20,225 observations.

Table 1
Sample Selection

Full Sample (Determinants of GDPR Risk Factor Disclosures / H1 Analyses)	
Number of firm-year observations from 2012 to 2019 meeting the following criteria: • 10-K filers headquartered in the United States • Positive assets in Compustat • Availability of an audit opinion and SOX 404a internal control report in Audit Analytics	37,862
Less: Firm-year observations without required Audit Analytics or Compustat data to construct variables necessary for H1 (IIQ) analyses	(3,917)
Less: Firm-year observations without a system-identifiable risk factors section of 250 words or more	(2,236)
Less: Firm-year observations for which their 10-K is filed more than 360 days after their fiscal year end	(79)
Full Sample	***31,630***
Efficiency Subsample (H2 through H4 Analyses)	
Sample for Determinants of GDPR Risk Factor Disclosure and H1 Analyses	31,630
Less: Financial firms (SIC codes 6000 to 6999) and utilities firms (SIC codes 4900 to 4999)	(9,095)
Less: Firm-year observations for which data necessary to construct operating efficiency models is not available	(2,310)
Efficiency Subsample	***20,225***

Table 1 reports the sample selection for observations used in both the IIQ (H1) and operating efficiency (H2 through H4) analyses.

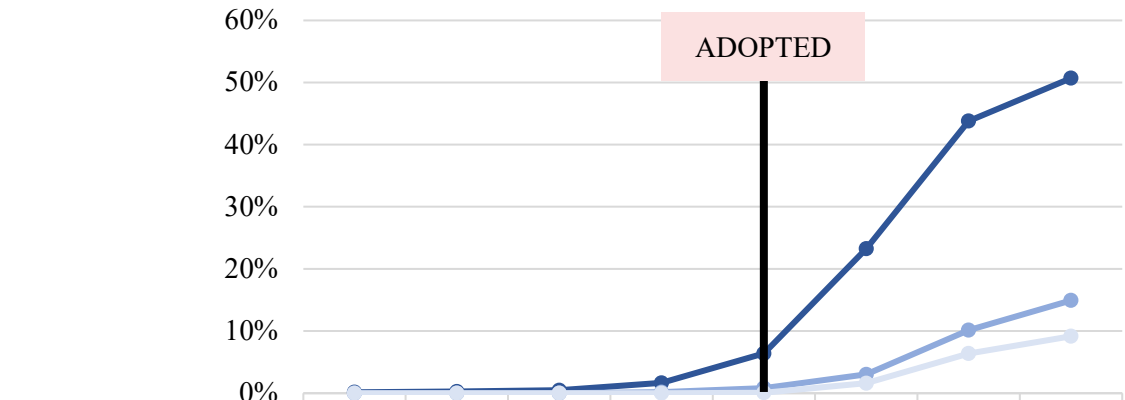
6.2 Association between GDPR Impact Proxies

Next, I explore the association between *GDPRRISK* and *EUOPS* as the two proxies that I use to identify the effects of GDPR on firms. As reported above, by 2019, over 29% of firms in my sample make a GDPR-related risk factor disclosure. I decompose this overall trend in GDPR-related risk factor disclosures into three classifications of firms: (a) those with European operations (*EUOPS* = 1), (b) those without European operations but with other foreign operations (non-US segments and subsidiaries), and (c) those with only domestic (US) operations (no non-US segments or subsidiaries). I report these trends in Figure 4. As expected, the positive slope of the trendline for firms with European operations is much steeper compared to the other two firm classifications and weakest for firms without foreign segments or subsidiaries. By 2019, over 50% of firms with European operations reported a GDPR-related risk factor disclosure compared to less than 15% of firms without European operations.

To provide further empirical evidence in support of the above relationship, I model *GDPRRISK* as a function of *EUOPS* and a series of controls that could explain differences in firms' reported risk factors. I include indicators for firms in high-tech industries (*HITECH*; based on the definition in Barron, Byard, Kile, and Reidl [2002]), which are likely to give extra attention to privacy regulations such as GDPR, differences in audit quality (*AUDBIG4*), litigation risk (*LITIGATION*), institutional ownership (*INSTOWN*), analyst following (*ANALYSTS*), firm fundamentals (*SIZE*, *LEV*, and *ROA*), and the firm's attention to cybersecurity (*BREACHFIRM*, *BREACHIND*, and *CYBERRISK*). Lastly, I control for *WRDCNTRISK* consistent with previous analyses and

include year fixed effects indicators to control for the increasing trend in GDPR-related disclosures highlighted in Figure 2.

Figure 4
GDPR-Related Risk Factors by Operating Geography and Year



	2012	2013	2014	2015	2016	2017	2018	2019
EU Ops.	0.2%	0.3%	0.5%	1.7%	6.4%	23.2%	43.8%	50.7%
Other Foreign Ops.	0.0%	0.0%	0.0%	0.2%	0.8%	3.0%	10.1%	14.9%
US Ops. Only	0.0%	0.0%	0.0%	0.0%	0.1%	1.6%	6.4%	9.2%

Figure 4 presents the percentage of firms making GDPR-related risk factor disclosures by year based on whether the firm has EU operations, foreign (but only non-EU) operations, or domestic only operations.

In Table 2, I report descriptive statistics for variables used in estimating the model described above (Panel A) as well as the results of estimating this model (Panel B). As noted in Panel A, 8.3% of firm-year observations report a GDPR-related risk factor. In Panel B, *EUOPS* is positively and significantly associated with *GDPRRISK* both over the entire sample period and when limiting my sample to the post-GDPR adoption years (2016 to 2019). The strength of this association provides further evidence that (a) firms with European operations are differentially affected by GDPR and (b) that GDPR-related

risk factor disclosures are reported by firms that, by the nature of the geographic scope of their operations, would be more affected by GDPR than others.

Table 2
Determinants of GDPR-Related Risk Factor Disclosures

Panel A: Descriptive Statistics

	Firm-Year Sample = 31,630				
Variable	**Mean**	**SD**	**Q1**	**Median**	**Q3**
GDPR-Impact Proxies					
GDPRRISK	0.083	0.276	0.000	0.000	0.000
EUOPS	0.471	0.499	0.000	0.000	1.000
Privacy-Sensitive Industries					
HITECH	0.261	0.439	0.000	0.000	1.000
External Factors					
AUDBIG4	0.647	0.478	0.000	1.000	1.000
ACCEL	0.728	0.445	0.000	1.000	1.000
LITIGATION	0.223	0.416	0.000	0.000	0.000
INSTOWN	0.511	0.367	0.101	0.590	0.856
Analysts (#)	7.098	7.846	1.000	5.000	10.000
ANALYSTS	1.573	1.092	0.693	1.792	2.398
Fundamentals					
Total Assets (mil)	6,916.487	19,422.330	174.549	1,052.807	4,264.915
SIZE	6.634	2.525	5.162	6.959	8.358
LEV	0.724	0.857	0.406	0.612	0.840
ROA	-0.156	0.764	-0.043	0.013	0.054
Cyber Risk					
BREACHFIRM	0.022	0.146	0.000	0.000	0.000
BREACHIND	2.377	1.026	1.609	2.565	3.219
CYBERRISK	1.654	1.636	0.397	1.252	2.414
Risk Factors Length					
WRDCNTRISK	8.302	0.631	7.900	8.322	8.745

Panel B: Model Estimation

Dependent Variable	**Pr (*GDPRRISK* = 1)**			
Years	**2012 – 2019**		**2016 – 2019**	
	(1)		**(2)**	
	Coefficient	**z-statistic**	**Coefficient**	**z-statistic**
GDPR-Impact Proxy				
EUOPS	2.229***	*(23.65)*	2.199***	*(23.52)*
Privacy-Sensitive Industries				
HITECH	0.650***	*(5.29)*	0.687***	*(5.57)*
External Factors				
AUDBIG4	0.602***	*(4.96)*	0.617***	*(5.09)*
ACCEL	0.258*	*(1.71)*	0.276*	*(1.82)*
LITIGATION	-0.157	*(-1.31)*	-0.149	*(-1.22)*
INSTOWN	0.331**	*(2.29)*	0.322**	*(2.22)*
ANALYSTS	0.421***	*(6.09)*	0.403***	*(5.87)*
Fundamentals				
SIZE	-0.196***	*(-5.97)*	-0.189***	*(-5.76)*
LEV	-0.041	*(-0.45)*	-0.021	*(-0.23)*
ROA	0.072	*(0.54)*	0.077	*(0.57)*
Cyber Risk				
BREACHFIRM	0.336**	*(2.09)*	0.302*	*(1.87)*
BREACHIND	0.293***	*(6.51)*	0.281***	*(6.28)*
CYBERRISK	0.261***	*(9.68)*	0.253***	*(9.43)*
Risk Factors Length				
WRDCNTRISK	1.489***	*(16.38)*	1.436***	*(16.01)*
Observations	31,630		15,199	
Fixed Effects	Year		Year	
Pseudo R^2	0.491		0.368	

Table 2 reports the descriptive statistics for variables used in estimating my GDPR risk factor disclosure determinants model (Panel A) as well as the results of estimating the model (Panel B). Panel B presents the results of estimating this model both over the full sample period (2012 to 2019; column 1) as well as a sample period limited to the period after GDPR adoption (2016 to 2019; column 2). All variables are defined in Appendix D. z-statistics calculated based on firm-clustered standard errors are presented in parentheses next to coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

CHAPTER 7

DESCRIPTIVE STATISTICS AND RESULTS

7.1. Descriptive Statistics

In Table 3, I report descriptive statistics for the variables used in models testing my four hypotheses. Approximately 8.3% of firm-year observations include a GDPR-related risk factor disclosure, and 47.1% of firms in my sample have either a geographic segment or a subsidiary in the EU. Turning to the variables used in defining *IIQ*, 90.1% (93.5%) of firm-year observations maintain effective internal controls (IT-related internal controls). 92.2% of firm-year observations **are not** associated with subsequent restatements, and 94.0% of firm-year observations are associated with on-time 10-K filings. 43.4% of firm-year observations have at least one sales guidance issuance during the fiscal year, and the average difference between sales guidance and actual sales is approximately 5%. *OPEREFF*, which is the percentile ranking by industry and year of firm efficiency (0.01 to 1.00), has a mean of 0.53.

Table 3
Descriptive Statistics

Firm-Year Sample = 31,630

	Mean	SD	Q1	Median	Q3
GDPR-Impact Proxies					
GDPRRISK	0.083	0.276	0.000	0.000	0.000
EUOPS	0.471	0.499	0.000	0.000	1.000
IIQ Proxies					
IIQ	0.000	1.000	0.119	0.119	0.482
NOICMW	0.901	0.298	1.000	1.000	1.000
NOITMW	0.935	0.246	1.000	1.000	1.000
NORESTATE	0.922	0.268	1.000	1.000	1.000
NOLATEFILER	0.940	0.238	1.000	1.000	1.000
GUIDANCE[b]	0.434	0.496	0.000	0.000	1.000
GUIDANCEACC[b]	-0.052	0.082	-0.054	-0.027	-0.014
Operating Efficiency Variable					
OPEREFF[a]	0.528	0.270	0.310	0.540	0.760
Controls					
Total Assets (mil)	6,916.487	19,422.330	174.549	1,052.807	4,264.915
SIZE	6.634	2.525	5.162	6.959	8.358
LEV	0.724	0.857	0.406	0.612	0.840
ROA	-0.156	0.764	-0.043	0.013	0.054
GROWTH	0.162	0.710	-0.034	0.051	0.164
INVENTORY	0.079	0.124	0.000	0.016	0.116
FOREIGN	0.402	0.490	0.000	0.000	1.000
SEG	1.176	0.538	0.693	1.099	1.609
MA	0.332	0.471	0.000	0.000	1.000
REST	0.276	0.447	0.000	0.000	1.000
Age (#)	22.953	16.855	9.000	19.000	30.000
AGE	2.848	0.792	2.197	2.944	3.401
AUDBIG4	0.647	0.478	0.000	1.000	1.000
ACCEL	0.728	0.445	0.000	1.000	1.000
INSTOWN	0.511	0.367	0.101	0.590	0.856
Analysts (#)	7.098	7.846	1.000	5.000	10.000
ANALYSTS	1.573	1.092	0.693	1.792	2.398
BREACHFIRM	0.022	0.146	0.000	0.000	0.000
BREACHIND	2.377	1.026	1.609	2.565	3.219
CYBERRISK	1.654	1.636	0.397	1.252	2.414
WRDCNTRISK	8.302	0.631	7.900	8.322	8.745
MKTSHARE[a]	0.008	0.019	0.000	0.001	0.005
CONCEN[a]	0.959	0.154	1.000	1.000	1.000

[a] N for variables used in efficiency analysis is 20,225 per Table 1.
[b] N for *GUIDANCE* is 22,437, which reflects firms covered by I/B/E/S Guidance database.
[c] N for *GUIDANCEACC* is 9,360, which is the number of firm-year observations with at least one sales guidance issuance with an actual reported in I/B/E/S.

Table 3 reports the descriptive statistics for variables used in estimating my primary models studying the effects of GDPR on IIQ and operating efficiency.

7.2. GDPR and IIQ (H1)

In Table 4, I start by exploring H1, which predicts that GDPR will have a positive effect on IIQ. In Panel A, I report the results of estimating model [1] using *GDPRRISK* as the proxy for GDPR impact. The coefficient on *GDPRRISK* is positive and significant (p-value < 0.001) in both industry (column 1) and firm (column 2) fixed effects models implying a significant increase in *IIQ* as firms make GDPR-related risk factor disclosures. Since *IIQ* is standardized, the coefficients on *GDPRRISK* reflect an increase on the order of 8.3% to 12.8% of a standard deviation in *IIQ* for firms reporting GDPR as a material risk. In considering the coefficients on control variables in column 1, larger firms (*SIZE*), more mature firms (*AGE*), and firms with greater analyst following (*ANALYSTS)* and institutional ownership (*INSTOWN*) have higher IIQ while more distributed firms (*SEG*) as well as those engaging in merger and acquisition (*MA*) activity have lower IIQ.

In Panel B, I report difference-in-differences models using *EUOPS*POST* as my GDPR impact proxy. Again, I find that GDPR leads to improvements in IIQ as evidenced by the coefficient on *EUOPS*POST*, which is positive and significant (p-value < 0.001) in both columns 1 and 2. In other words, firms with European operations exhibit greater improvements in IIQ around the adoption of GDPR compared to other firms.

Collectively, these results support H1's prediction that GDPR-impacted firms improve IIQ more than other firms around the regulation's adoption.

Table 4
GDPR and IIQ

Panel A: GDPR-Related Risk Factor Disclosures and IIQ

Dependent Var.		***IIQ***			
Fixed Effects		**Industry / Year**		**Firm / Year**	
		(1)		**(2)**	
	Pred	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
GDPRRISK	+	0.128***	*(5.56)*	0.083***	*(3.51)*
SIZE		0.047***	*(6.31)*	-0.015	*(-0.63)*
LEV		-0.111***	*(-5.04)*	-0.074***	*(-2.89)*
ROA		0.227***	*(8.80)*	0.066**	*(2.50)*
GROWTH		-0.074***	*(-6.12)*	-0.012	*(-1.01)*
INVENTORY		-0.005	*(-0.05)*	0.146	*(0.61)*
FOREIGN		0.054**	*(2.46)*	0.041	*(0.87)*
SEG		-0.098***	*(-4.89)*	-0.052	*(-1.19)*
MA		-0.080***	*(-5.55)*	-0.040***	*(-2.81)*
REST		0.015	*(0.95)*	0.003	*(0.19)*
AGE		0.071***	*(4.79)*	0.097	*(1.31)*
AUDBIG4		0.107***	*(4.73)*	-0.075	*(-1.28)*
ACCEL		0.020	*(0.72)*	-0.073**	*(-1.97)*
INSTOWN		0.161***	*(5.77)*	0.074**	*(1.96)*
ANALYSTS		0.048***	*(3.80)*	0.079***	*(3.26)*
BREACHFIRM		-0.040	*(-1.29)*	0.000	*(0.00)*
BREACHIND		-0.023	*(-1.58)*	0.004	*(0.29)*
CYBERRISK		0.022***	*(4.19)*	0.007	*(1.17)*
WRDCNTRISK		-0.052***	*(-2.63)*	-0.285***	*(-6.34)*
Observations		31,630		31,630	
Adjusted R^2		0.193		0.555	

Panel B: Change in IIQ for Firms with EU Operations

Dependent Var.		***IIQ***			
Fixed Effects		**Industry / Year**		**Firm / Year**	
		(1)		**(2)**	
	Pred	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
*EUOPS*POST*	+	0.097***	*(4.37)*	0.103***	*(4.30)*
EUOPS		-0.019	*(-0.63)*		
Observations		26,778		26,778	
Model [1] Controls		Included		Included	
Adjusted R^2		0.147		0.491	

Table 4 reports the results of estimating model [1] studying the effects of GDPR on IIQ when using *GDPRRISK* (in Panel A) as well as *EUOPS*POST* (in Panel B) as GDPR impact proxies. The sample sizes in Panels A and B differ since I require firms to have at least one observation with available data in each of the pre-adoption and post-adoption periods when using *EUOPS*POST* as my GDPR impact proxy. All variables are defined in Appendix D. t-statistics based on firm-clustered standard errors are presented in parentheses next to coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

7.3. GDPR and Operating Efficiency (H2 through H4)

Next, in Table 5, I explore H2 through H4, which articulate my expectations regarding the effects of GDPR on operating efficiency indirectly via improved IIQ (H2), directly (H3), and in net (H4). In Panel A (B), I report the results of estimating model [2] using *GDPRRISK* (*EUOPS*POST*) as my GDPR impact proxy. In each panel, I report versions of the model including either industry and year (columns 1 and 2) or firm and year (columns 3 and 4) fixed effects. In each case, I report the results obtained for estimating equations [2A] and [2B] side-by-side and include the estimates for the indirect (via improved IIQ) and net effects of GDPR on operating efficiency at the bottom of each panel.

Starting with Panel A, I find results consistent with my hypotheses. First, in both industry and firm fixed effects models, *IIQ* from equation [2B] (β_{21}; columns 2 and 4) is significantly and positively associated with operating efficiency. When multiplied with the corresponding (positive and significant) coefficient on *GDPRRISK* from equation [2A] (β_{11}; columns 1 and 3), I observe a positive and significant indirect effect of GDPR on operating efficiency via improved IIQ consistent with H2 ($\beta_{11}*\beta_{21} > 0$). Second, in both industry and firm fixed effects models, the direct effect of GDPR on operating efficiency is significantly negative consistent with H3 ($\beta_{22} < 0$). Further, in each case, the negative direct effect dominates the positive indirect effect ($(\beta_{22} + (\beta_{11}*\beta_{21})) < 0$) consistent with H4. The magnitude of the net effect reported at the bottom of Panel A implies that GDPR has led to a decline in relative operating efficiency by 2.2 (1.5) percentiles in industry (firm) fixed effects models.

In Table 5 Panel B, I obtain similar results when using *EUOPS*POST* as my GDPR impact proxy. First, GDPR's indirect effect via improved IIQ ($\beta_{11}*\beta_{21}$) remains positive and statistically significant. Second, the direct (β_{22}) and net effects ($\beta_{22} + (\beta_{11}*\beta_{21})$) of GDPR on operating efficiency appear even stronger in magnitude than those reported above and reflect a decline in relative operating efficiency by 2.6 (2.7) percentiles in industry (firm) fixed effects models. Collectively, these analyses suggest that, although improvements in IIQ driven by GDPR have had positive consequences for operating efficiency, the regulatory burden of GDPR has overwhelmed these benefits. Overall, GDPR appears to have been costly to firm operations.

Table 5
GDPR and Operating Efficiency

Panel A: GDPR-Related Risk Factor Disclosures and Operating Efficiency

Fixed Effects			**Industry / Year**		**Firm / Year**	
Dependent Var.			***IIQ***	***OPEREFF***	***IIQ***	***OPEREFF***
Model [2] Eq.			**2A**	**2B**	**2A**	**2B**
	Pred.		**(1)**	**(2)**	**(3)**	**(4)**
	Eq. 2A	**Eq. 2B**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**
IIQ		β_{21}		0.010***		0.004***
		(+)		(5.32)		(2.60)
GDPRRISK	β_{11}	β_{22}	0.132***	-0.023***	0.095***	-0.015***
	(+)	(−)	(4.67)	(-3.62)	(3.63)	(-3.42)
SIZE			0.058***	0.084***	0.012	0.055***
			(5.62)	(38.45)	(0.44)	(14.18)
LEV			-0.222***	0.014***	-0.076***	0.006**
			(-9.76)	(6.70)	(-3.33)	(2.55)
GROWTH			-0.106***	0.006**	-0.009	0.014***
			(-5.81)	(2.44)	(-0.50)	(6.40)
INVENTORY			0.204	0.225***	0.225	0.176***
			(1.62)	(8.86)	(0.87)	(5.57)
FOREIGN			0.020	-0.005	-0.012	-0.001
			(0.77)	(-0.68)	(-0.25)	(-0.20)
SEG			-0.089***	-0.005	-0.005	-0.004
			(-3.61)	(-0.81)	(-0.11)	(-0.60)
MA			-0.067***	-0.002	-0.022	0.010***
			(-3.77)	(-0.50)	(-1.41)	(4.17)
REST			0.008	-0.029***	0.007	-0.015***
			(0.48)	(-6.66)	(0.39)	(-5.56)
AGE			0.121***	-0.009**	0.101	-0.017
			(6.05)	(-2.33)	(1.08)	(-1.44)
AUDBIG4			0.128***	0.014*	-0.177***	-0.001
			(4.25)	(1.86)	(-2.84)	(-0.07)
ACCEL			0.032	0.008	-0.032	0.012**
			(0.82)	(0.98)	(-0.67)	(1.98)
INSTOWN			0.273***	0.039***	0.106**	0.032***
			(7.80)	(4.17)	(2.47)	(4.07)
ANALYSTS			0.083***	0.004	0.089***	0.004
			(4.89)	(1.06)	(2.96)	(0.98)
BREACHFIRM			-0.066*	-0.011	-0.015	-0.009
			(-1.75)	(-1.16)	(-0.47)	(-1.64)
BREACHIND			-0.026	0.009***	0.003	0.009***
			(-1.48)	(2.83)	(0.21)	(3.56)
CYBERRISK			0.026***	0.006***	0.016**	0.003**

		(3.44)	(3.32)	(2.32)	(2.37)
WRDCNTRISK		-0.025	-0.036***	-0.332***	-0.030***
		(-0.95)	(-6.72)	(-6.15)	(-4.35)
MKTSHARE			0.603***		2.448***
			(3.56)		(6.73)
CONCEN			0.001		0.002
			(0.09)		(0.09)
Observations		20,225	20,225	20,225	20,225
Adjusted R^2		0.216	0.588	0.575	0.858
Indirect and Net Effects					
Indirect Effect ($\beta_{11}*\beta_{21}$)	+		0.0013*** (3.51)		0.0004** (2.17)
Net Effect ($\beta_{22} + (\beta_{11}*\beta_{21})$)	−		-0.0216*** (-3.41)		-0.0149*** (-3.34)

Panel B: Change in Operating Efficiency for Firms with EU Operations

Fixed Effects			**Industry / Year**		**Firm / Year**	
Dependent Var.			***IIQ***	***OPEREFF***	***IIQ***	***OPEREFF***
Model [2] Eq.			**2A**	**2B**	**2A**	**2B**
	Pred.		**(1)**	**(2)**	**(3)**	**(4)**
	Eq. 2A	**Eq. 2B**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**
IIQ		β_{21}		0.011***		0.004**
		(+)		(5.15)		(2.36)
*EUOPS*POST*	β_{11}	β_{22}	0.093***	-0.026***	0.089***	-0.027***
	(+)	(−)	(2.84)	(-4.86)	(2.78)	(-5.56)
EUOPS			0.003	-0.004		
			(0.09)	(-0.38)		
Observations			16,883	16,883	16,883	16,883
Model [2] Controls			Included	Included	Included	Included
Adjusted R^2			0.180	0.583	0.512	0.849
Indirect and Net Effects						
Indirect Effect ($\beta_{11}*\beta_{21}$)	+		0.0010** (2.40)		0.0003* (1.80)	
Net Effect ($\beta_{22} + (\beta_{11}*\beta_{21})$)	−		-0.0252*** (-4.67)		-0.0271*** (-5.49)	

Table 5 reports the results of estimating model [2] studying the effects of GDPR on operating efficiency (a) indirectly via improvements in IIQ, (b) directly, and (c) in net. Panels A and B report the results of running models using *GDPRRISK* and

*EUOPS***POST*, respectively, as GDPR impact proxies. Model [2] consists of equations [2A] and [2B], which are reported side-by-side in each panel. The indirect and net effects estimated as combinations of coefficients across models are reported at the bottom of each panel. The sample sizes in Panels A and B differ since I require firms to have at least one observation with available data in each of the pre-adoption and post-adoption periods when using *EUOPS***POST* as my GDPR impact proxy. All variables are defined in Appendix D. z-statistics calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

CHAPTER 8

ADDITIONAL ANALYSES

8.1. Abnormal Returns to Early Governmental Support for GDPR

Next, considering the overall negative effect on operating efficiency that I observe as a consequence of GDPR, I explore whether the market perceived the costs that would be associated with the regulation when it first became apparent that the regulation would be enacted. Specifically, I explore whether firms that eventually made a GDPR-related risk factor disclosure (*GDPRFIRM*) or that had EU operations (*EUOPS*) exhibited significant negative abnormal returns compared to other firms around two events that represented early support for the regulation by the two primary EU legislative bodies, the EU Commission and the EU Parliament. As highlighted in Appendix A, the first event date (January 25, 2012) is that on which the initial proposal for GDPR was unveiled by the European Commission, and the second event date (March 12, 2014) is that on which the European Parliament signaled strong support for the regulation with an overwhelmingly affirmative vote (621 votes in favor to 10 votes against) on the draft text of the regulation.

For each of these two events, I regress the 5-day Fama-French 3-Factor cumulative abnormal returns against *GDPRFIRM* and *EUOPS* after controlling for differences in firm fundamentals (*SIZE*, *LEV*, and *ROA*). In Table 6, I report the results of running these models and find that both of my GDPR impact proxies identify firms with negative market reactions relative to other firms, particularly around the European Parliament's affirmative vote on the draft text of the regulation. These analyses provide

some evidence that the market perceived that GDPR would be costly for firms impacted by the regulation.

Table 6
Cumulative Abnormal Returns (CAR) for Early GDPR Support

Dependent Var.	**Fama-French 3-Factor CAR (-2,2)**			
Event Date	**Initial GDPR Proposal by European Commission (January 25, 2012)**		**European Parliament Strong Support for GDPR (March 12, 2014)**	
	(1) Coefficient (t-statistic)	**(2) Coefficient (t-statistic)**	**(3) Coefficient (t-statistic)**	**(4) Coefficient (t-statistic)**
GDPRFIRM	-0.004		-0.005***	
	(-1.59)		(-3.01)	
EUOPS		-0.003		-0.007***
		(-1.52)		(-3.81)
SIZE	-0.002***	-0.002***	0.003***	0.003***
	(-3.10)	(-2.92)	(5.07)	(5.29)
LEV	0.012**	0.011**	0.001	0.000
	(2.56)	(2.36)	(0.38)	(0.04)
ROA	-0.011*	-0.011*	0.005	0.005
	(-1.77)	(-1.79)	(0.85)	(0.81)
Observations	3,190	3,190	3,203	3,203
Adjusted R^2	0.013	0.013	0.017	0.019

Table 6 reports the results of regressing Fama-French 5-day cumulative abnormal returns against indicators for *GDPRFIRM* and *EUOPS*. Columns 1 and 2 report the results for cumulative abnormal returns around the date on which GDPR was initially proposed by the European Commission, and columns 3 and 4 report the results for cumulative abnormal returns around the date on which the European Parliament signaled strong support for GDPR with a plenary vote on the regulation. All variables are defined in Appendix D. t-statistics calculated based on robust standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

8.2. Cross-Sectional Tests

The evidence provided above in my main tests supports the view that GDPR led to improvements in IIQ, but overall, negatively affected operating efficiency. Next, I explore whether these effects are stronger for two sets of firms that should be more significantly affected by the regulation. Specifically, I expect the effects of GDPR to be more acute for high-tech firms, which collect and process significant amounts of personal data, compared to other firms, and smaller firms, which are less likely to have had the pre-existing infrastructure in place to handle the requirements of the regulation. Initial empirical support for using these firm attributes for cross-sectional tests is provided in Chapter 6, where the determinants of GDPR-related risk factor disclosures are reported. As reported in Table 2 Panel B, *HITECH* was positively (and *SIZE* was negatively) associated with the likelihood of a firm making a GDPR-related risk factor disclosure.

To study the differences in the effects of GDPR across these cross-sections, I run models [1] and [2] in subsamples of high-tech and low-tech firms as well as small and large firms. As above, I define high-tech firms as those in high-tech intangibles industries per Barron et al. (2002), and low-tech firms as all other firms. Small and large firm subsamples are defined based on whether a firm's average total assets in the period prior to the adoption of GDPR (2012 to 2015) was below or above the sample median.

In Table 7, I report the results of studying the effects of GDPR in high- and low-tech subsamples. In Panel A, I find that both GDPR impact proxies load positively and significantly in high-tech firm (columns 1 and 3) subsamples. Further, as reported at the bottom of the panel, the magnitude of the coefficients in the high-tech firm subsamples are larger than those in the low-tech firm subsamples (p-values 0.11 and 0.05). In other

words, the improvements in IIQ around the adoption of GDPR are significantly stronger for high-tech firms than for low-tech firms. In Panel B, I report similar inferences for GDPR's effects on operating efficiency although the difference in the coefficients on *GDPRRISK* across the two subsamples is not significant at traditional levels.

Table 7
Cross-Sectional Test – High-Tech and Low-Tech Firms

Panel A: GDPR and IIQ

Dependent Variable	***IIQ***			
GDPRIMPACT =	***GDPRRISK***		***EUOPS*POST***	
Sample	**High-Tech**	**Low-Tech**	**High-Tech**	**Low-Tech**
	(1)	**(2)**	**(3)**	**(4)**
	Coefficient (t-statistic)	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
GDPRIMPACT (β_1)	0.144***	0.064**	0.216***	0.090***
	(3.59)	(2.26)	(3.62)	(3.62)
Observations	8,244	23,386	6,630	20,148
Model [1] Controls	Included	Included	Included	Included
Fixed Effects	Firm / Year	Firm / Year	Firm / Year	Firm / Year
Adjusted R^2	0.563	0.551	0.520	0.478
Diff in β_1	0.079		0.126*	
(P-Value)	(0.105)		(0.050)	

Panel B: GDPR and Operating Efficiency

Dependent Variable	***OPEREFF***			
GDPRIMPACT =	***GDPRRISK***		***EUOPS*POST***	
Sample	**High-Tech**	**Low-Tech**	**High-Tech**	**Low-Tech**
	(1)	**(2)**	**(3)**	**(4)**
	Coefficient (t-statistic)	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
GDPRIMPACT (β_1)	-0.022***	-0.014**	-0.044***	-0.024***
	(-3.14)	(-2.31)	(-4.53)	(-4.18)
Observations	6,742	13,483	5,425	11,458
Model [2] Controls	Included	Included	Included	Included
Fixed Effects	Firm / Year	Firm / Year	Firm / Year	Firm / Year
Adjusted R^2	0.882	0.847	0.876	0.835
Diff in β_1	-0.008		-0.019*	
(P-Value)	(0.379)		(0.090)	

Table 7 reports the results of estimating models studying GDPR's effects on IIQ (in Panel A) and operating efficiency (in Panel B) in subsamples of high-tech and low-tech firms using both *GDPRRISK* and *EUOPS*POST* as GDPR impact proxies. The bottom of each panel reports the difference in the coefficients on the respective GDPR impact proxies between the two subsamples. All variables are defined in Appendix D. t-statistics calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

In Table 8, I conduct similar analyses using subsamples of small and large firms. Across GDPR impact proxies, I find that GDPR is associated with larger improvements in IIQ (as shown in Panel A) and larger declines in operating efficiency (as shown in Panel B) for small firms as opposed to larger firms. Collectively, these results provide cross-sectional support for the findings presented in the main analyses. The regulation's effects on IIQ and operating efficiency vary in an expected fashion across dimensions of industry and firm size.

Table 8
Cross-Sectional Test – Small and Large Firms

Panel A: GDPR and IIQ

Dependent Variable	***IIQ***			
GDPRIMPACT =	***GDPRRISK***		***EUOPS*POST***	
Sample	**Small**	**Large**	**Small**	**Large**
	(1)	**(2)**	**(3)**	**(4)**
	Coefficient (t- statistic)	**Coefficient (t- statistic)**	**Coefficient (t- statistic)**	**Coefficient (t- statistic)**
GDPRIMPACT (β_1)	0.130***	0.023	0.159***	0.016
	(3.70)	(0.89)	(4.28)	(0.70)
Observations	13,423	13,355	13,423	13,355
Model [1] Controls	Included	Included	Included	Included
Fixed Effects	Firm / Year	Firm / Year	Firm / Year	Firm / Year
Adjusted R^2	0.514	0.385	0.515	0.384
Diff in β_1	0.106**		0.143***	
(P-Value)	(0.015)		(0.001)	

Panel B: GDPR and Operating Efficiency

Dependent Variable	***OPEREFF***			
GDPRIMPACT =	***GDPRRISK***		***EUOPS*POST***	
Sample	**Small**	**Large**	**Small**	**Large**
	(1)	**(2)**	**(3)**	**(4)**
	Coefficient (t- statistic)	**Coefficient (t- statistic)**	**Coefficient (t- statistic)**	**Coefficient (t- statistic)**
GDPRIMPACT (β_1)	-0.034***	0.004	-0.038***	-0.014*
	(-5.01)	(0.71)	(-6.01)	(-1.70)
Observations	9,700	7,183	9,700	7,183
Model [2] Controls	Included	Included	Included	Included
Fixed Effects	Firm / Year	Firm / Year	Firm / Year	Firm / Year
Adjusted R^2	0.809	0.780	0.810	0.780
Diff in β_1	-0.039***		-0.025**	
(P-Value)	(0.000)		(0.019)	

Table 8 reports the results of estimating models studying GDPR's effects on IIQ (in Panel A) and operating efficiency (in Panel B) in subsamples of small and large firms using both *GDPRRISK* and *EUOPS*POST* as GDPR impact proxies. The bottom of each panel reports the difference in the coefficient on the respective GDPR impact proxy between the two subsamples. All variables are defined in Appendix D. t-statistics calculated based on firm-clustered standard errors are presented in parentheses below coefficient

estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

8.3. Hedonic Adaptation Models

In the main analyses above, my difference-in-differences specification uses the year of GDPR adoption (2016) as the first year of the post period. However, it is conceivable that the effects of GDPR might strengthen after the regulation is enforced beginning in 2018. To explore this notion further, I decompose the *POST* indicator into two indicators reflecting (a) the years between GDPR adoption and enforcement (2016 and 2017) (*POSTADOPT*) and (b) the years after enforcement (2018 and 2019) (*POSTENFORCE*) and interact these indicators with *EUOPS*. In Table 9 Panel A, I report the results of running these models to understand changes in IIQ. I find that my results strengthen slightly (although not in a statistically significant manner) after GDPR enforcement. Although weak, such a trend is consistent with information management processes aimed at GDPR compliance continuing to mature after GDPR enforcement, further improving IIQ.

Next, I conduct similar analyses for operating efficiency to understand whether the negative effects of GDPR on operating efficiency have persisted for the entirety of the post-adoption period. In Table 9 Panel B, I find that the regulation's negative effect on operating efficiency has persisted into the post-enforcement period (2018 and 2019). However, the magnitudes of the coefficients on the interactions with *POSTENFORCE* are smaller than (although not statistically different from) the interactions with *POSTADOPT*, which would be consistent with the negative effect on efficiency attenuating over time as firms adjust to the regulation.

Table 9
Hedonic Adaptation Analysis

Panel A: GDPR and IIQ

Dependent Var.		***IIQ***			
Fixed Effects		**Industry / Year**		**Firm / Year**	
		(1)		**(2)**	
	Pred.	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
EUOPS POSTADOPT*	+	0.087***	*(3.65)*	0.093***	*(3.71)*
EUOPS POSTENFORCE*	+	0.108***	*(3.89)*	0.115***	*(3.83)*
EUOPS		-0.019	*(-0.64)*		
Observations		26,778		26,778	
Model [1] Controls		Included		Included	
Adjusted R^2		0.147		0.491	

Panel B: GDPR and Operating Efficiency

Fixed Effects			**Industry / Year**		**Firm / Year**	
Dependent Var.			***IIQ***	***OPEREFF***	***IIQ***	***OPEREFF***
Model [2] Eq.			**2A**	**2B**	**2A**	**2B**
	Pred.		**(1)**	**(2)**	**(3)**	**(4)**
	Eq. 2A	**Eq. 2B**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**
IIQ		β_{21}		0.011***		0.004**
		(+)		(5.15)		(2.36)
*EUOPS**	β_{11A}	β_{22A}	0.081**	-0.030***	0.089***	-0.032***
POSTADOPT	(+)	(−)	(2.34)	(-5.74)	(2.65)	(-6.60)
*EUOPS**	β_{11E}	β_{22E}	0.106***	-0.022***	0.089**	-0.022***
POSTENFORCE	(+)	(−)	(2.61)	(-3.12)	(2.25)	(-3.45)
EUOPS			0.003	-0.004		
			(0.08)	(-0.38)		
Observations			16,883	16,883	16,883	16,883
Model [2] Controls			Included	Included	Included	Included
Adjusted R^2			0.180	0.583	0.514	0.849

Indirect and Net Effects – POSTADOPT			
Indirect Effect (β_{11A}*β_{21})	+	0.0009** (2.07)	0.0003* (1.75)
Net Effect (β_{22A} + (β_{11A}*β_{21}))	–	-0.0291*** (-5.58)	-0.0313*** (-6.52)
Indirect and Net Effects – POSTENFORCE			
Indirect Effect (β_{11E}*β_{21})	+	0.0012** (2.27)	0.0003 (1.64)
Net Effect (β_{22E} + (β_{11E}*β_{21}))	–	-0.0206*** (-2.95)	-0.0217*** (-3.40)

Table 9 presents the results of estimating models [1] (in Panel A) and [2] (in Panel B) after decomposing the *POST* indicator interacted with *EUOPS* into two indicators reflecting the years between GDPR adoption and enforcement (*POSTADOPT*; 2016 and 2017) and years after GDPR enforcement (*POSTENFORCE*; 2018 and 2019). Model [2] consists of equations [2A] and [2B], which are reported side-by-side in each panel. The indirect and net effects estimated as combinations of coefficients across models are reported at the bottom of each panel. t-statistics (in Panel A) and z-statistics (in Panel B) calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

8.4. Movement Away from the EU Market

Lastly, recognizing the costs of operating in EU that accompany the adoption of GDPR, it is conceivable that firms adjust their geographic footprint away from the region. To understand whether this phenomenon has manifested, I run models where I explore whether the presence and count of EU-based segments and subsidiaries declines after the adoption of GDPR. Specifically, I compare the change in reported geographic footprints for high-tech firms compared to low-tech firms since high-tech firms appear to

have been more acutely affected by GDPR based on my above analyses.[29] The model described above takes the following form:

$$EUVAR_{i,t} = \alpha + \beta_1 HITECH_i * POST_t + \beta_2 HITECH_i + \gamma CONTROLS_{i,t} + Year\ F.E. + e_{i,t} \quad [3]$$

EUVAR reflects either *EUSEG (EUSEGCNT)* or *EUSUB (EUSUBCNT)*, indicators for the presence of a (the logged counts of) EU segment(s) or subsidiary(ies) reported by the firm in a given year. In Table 10, I report the results of running logit models with both *EUSEG* and *EUSUB* as dependent variables and linear models with *EUSEGCNT* and *EUSUBCNT* as dependent variables after controlling for differences in firm fundamentals (*SIZE*, *LEV*, and *ROA*). Across models, it appears that high-tech firms reduce their EU footprint after the adoption of GDPR more than other firms although the difference in the change around the adoption of GDPR is significant only in relation to the segments-based proxies (and not subsidiary-based proxies). This could be because the data necessary to construct subsidiary-based proxies only extends through 2018, weakening the power of these tests. However, the tenor of the evidence presented is consistent with the notion that, high-tech companies, which appear to experience greater

[29] The use of measures based on EU segments and subsidiaries as my dependent variable precludes the use of my other GDPR impact proxies as test variables since these proxies are either based directly upon or determined by a firms' geographic footprint.

costs associated with GDPR, reduce their geographic footprint in the EU after the adoption of the regulation.[30]

Table 10
Changes in EU-Based Segments and Subsidiaries

Dependent Var.	***Pr (EUSEG = 1)***	***EUSEGCNT***	***Pr (EUSUB = 1)***	***EUSUBCNT***
	(1) **Coefficient (z-statistic)**	**(2)** **Coefficient (t-statistic)**	**(3)** **Coefficient (z-statistic)**	**(4)** **Coefficient (t-statistic)**
*HITECH*POST*	-0.159***	-0.035***	-0.033	-0.036
	(-2.91)	(-3.75)	(-0.59)	(-1.52)
HITECH	1.659***	0.252***	1.526***	0.676***
	(19.43)	(18.07)	(19.89)	(19.24)
SIZE	0.150***	0.024***	0.329***	0.212***
	(8.64)	(10.91)	(19.65)	(28.65)
LEV	-0.244**	-0.005	-0.384***	0.027**
	(-2.16)	(-1.02)	(-3.24)	(2.10)
ROA	0.740***	0.021***	0.462***	-0.044***
	(5.34)	(3.95)	(4.01)	(-3.08)
Observations	31,630	31,630	27,977	27,977
Fixed Effects	Year	Year	Year	Year
Pseudo / Adjusted R^2	0.091	0.084	0.120	0.168

Table 10 reports the results of running difference-in-differences models to explore the change in the extent of firms' EU operations using high-tech firms as treatment firms and other firms as control firms around the adoption of GDPR. Models reported in columns 1 and 3 (2 and 4) are logit (linear regression) models. Models using subsidiary-based measures as dependent variables (columns 3 and 4) are limited to fiscal years 2012 through 2018 due to data availability to construct *EUSUB* and *EUSUBCNT*. All variables are defined in Appendix D. t-statistics (and z-statistics) calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

[30] These broad findings are consistent with recent anecdotal evidence that smaller technology firms have begun to shift focus away from the European market after the adoption of GDPR (e.g., Davies 2018; Kottasová 2018).

CHAPTER 9

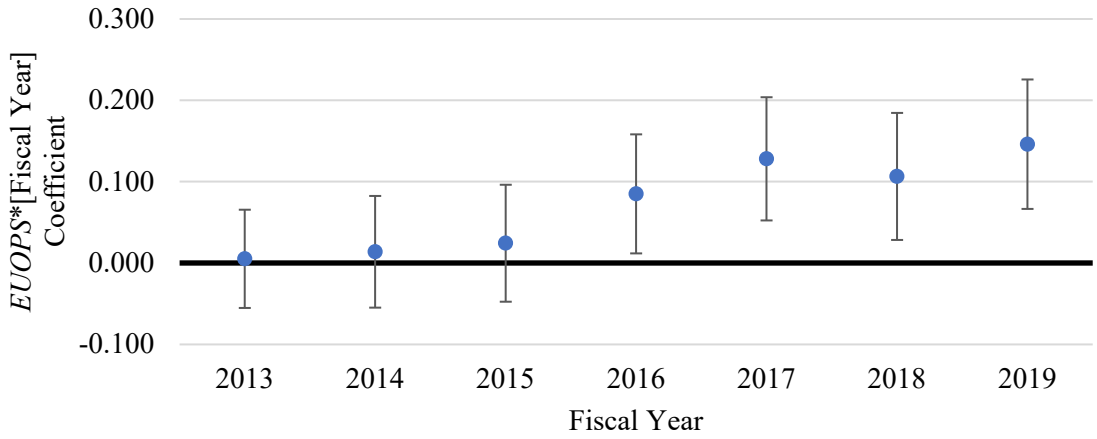
ROBUSTNESS ANALYSES

9.1. Analysis of Parallel Trends Assumption

A key assumption in running difference-in-differences models, like those employing *EUOPS*POST* as my GDPR-impact proxy, is that the trend between treatment and control observations in the pre period is similar. To validate this assumption, I re-run models used in main tests (including firm-fixed effects) after replacing *POST* with indicators for each year in my sample (2013 to 2019; 2012 is dropped due to collinearity). Figure 5 plots the coefficients (with 95% confidence intervals) on the interaction terms between the year indicators and *EUOPS*. For both models studying *IIQ* and *OPEREFF*, the trend in the pre period does not appear to differ significantly between my treatment and control firms.

Figure 5
Parallel Trends Assumption

Panel A: IIQ Model



Panel B: Operating Efficiency Model

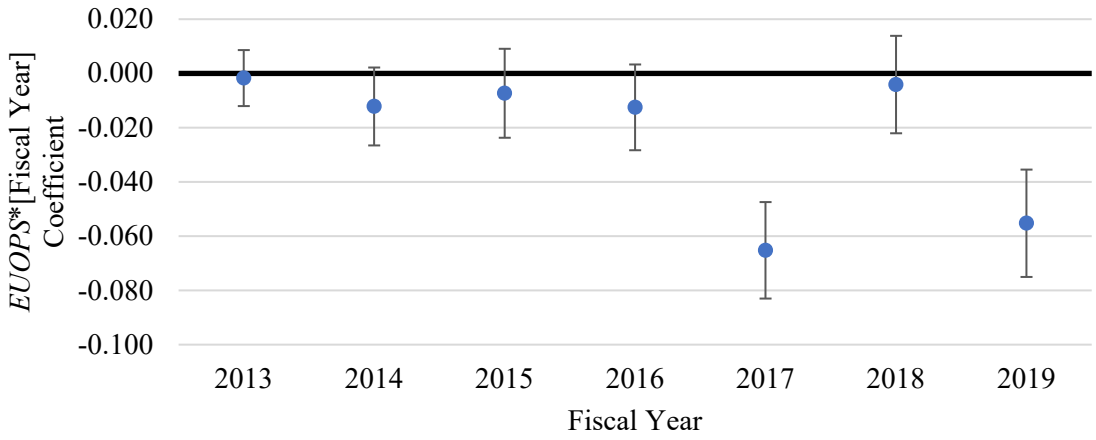


Figure 5 plots the coefficients (along with 95% confidence intervals) on interactions between *EUOPS* and indicators for each year in my sample period (2012 is dropped due to collinearity).

9.2. GDPR Effects on Individual Components of *IIQ*

In the main analyses, *IIQ* is an index composed of six variables reflecting various dimensions of IIQ. As discussed above, this index is beneficial since no one individual publicly available measure is capable of capturing IIQ in the broad sense that I am attempting to proxy. However, for illustrative purposes, I next explore the effects of GDPR on each of the components of *IIQ*. Specifically, I re-run model [1] with the dependent variable set to be each of the individual components of *IIQ* (*NOICMW*, *NOITMW*, *NORESTATE*, *NOLATEFILER, GUIDANCE*, and *GUIDANCEACC*). Table 11 reports the coefficients obtained from regressing each of the GDPR impact proxies against these index components. The coefficients on the GDPR impact proxies are positive and significant at traditional levels in 9 of the 12 models. Further, each component of *IIQ*, except for *GUIDANCEACC*, is explained significantly by either

GDPRRISK and/or *EUOPS*POST* providing assurance that a single component is not exclusively driving the results that I observe in main analyses.

Table 11
Individual Components of *IIQ*

Dependent Var.	GDPR Impact Proxy	Coefficient	t-statistic
Internal Control Quality			
NOICMW	*GDPRRISK*	0.015*	*(1.87)*
	*EUOPS*POST*	0.016**	*(2.23)*
NOITMW	*GDPRRISK*	0.016**	*(2.45)*
	*EUOPS*POST*	0.012**	*(2.01)*
Financial Reporting Accuracy and Speed			
NORESTATE	*GDPRRISK*	0.017*	*(1.75)*
	*EUOPS*POST*	0.026***	*(2.92)*
NOLATEFILER	*GDPRRISK*	0.009	*(1.52)*
	*EUOPS*POST*	0.011*	*(1.92)*
Management Guidance			
GUIDANCE	*GDPRRISK*	0.024*	*(1.92)*
	*EUOPS*POST*	0.026**	*(2.34)*
GUIDANCEACC	*GDPRRISK*	-0.001	*(-0.43)*
	*EUOPS*POST*	-0.001	*(-0.11)*

Table 11 reports the coefficients and t-statistics on the GDPR impact proxies obtained after estimating model [1] when replacing *IIQ* with each of the components used to construct it. All variables are defined in Appendix D. t-statistics calculated based on firm-clustered standard errors are presented in parentheses next to coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

9.3. Alternative Proxies for Operating Efficiency

In the main analyses, I use an industry-year percentile-ranked measure of operating efficiency based on Demerjian et al. (2012). In Table 12, I report the results of re-running model [2] using alternative proxies for operating efficiency. First, I use the

continuous output from the DEA estimation routine (prior to percentile ranking to arrive at *OPEREFF*), and my results are consistent with those in my main tests. I also use *ROA* as a coarser proxy for operating efficiency. Here again, my results regarding the indirect, direct, and net effects of GDPR on operating efficiency are consistent with those that I report in the main analyses.

Table 12
Alternative Measurements of Operating Efficiency

Fixed Effects	GDPR Impact Proxy	Indirect Effect via IIQ	Direct Effect	Net Effect
		(1) Coefficient (z-statistic)	(2) Coefficient (z-statistic)	(3) Coefficient (z-statistic)
Dependent Variable: Continuous Version of DEA-Calculated Firm Efficiency				
Industry / Year	*GDPRRISK*	0.0013***	-0.0387***	-0.0375***
		(3.79)	(-7.01)	(-6.77)
	*EUOPS*POST*	0.0009**	-0.0358***	-0.0349***
		(2.50)	(-7.35)	(-7.16)
Firm / Year	*GDPRRISK*	0.0003*	-0.0245***	-0.0242***
		(1.88)	(-5.76)	(-5.69)
	*EUOPS*POST*	0.0003	-0.0391***	-0.0388***
		(1.60)	(-8.65)	(-8.58)
Dependent Variable: Return on Assets				
Industry / Year	*GDPRRISK*	0.0073***	-0.0289**	-0.0216*
		(3.81)	(-2.41)	(-1.80)
	*EUOPS*POST*	0.0052**	-0.0378**	-0.0325**
		(2.56)	(-2.28)	(-1.98)
Firm / Year	*GDPRRISK*	0.0013**	-0.0407***	-0.0394***
		(2.08)	(-4.02)	(-3.88)
	*EUOPS*POST*	0.0011*	-0.0391***	-0.0379***
		(1.79)	(-2.71)	(-2.62)

Table 12 reports the results of estimating model [2] after replacing *OPEREFF* with each of the operating efficiency measures described in the table. z-statistics calculated based on firm-clustered standard errors are presented in parentheses below coefficient

estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

9.4. Controlling for Changes in Board Cyber Focus

Klein et al. (2022) find that firms' boards became more focused on cybersecurity around the adoption of GDPR. Considering past literature connecting board-level information technology expertise to financial reporting quality (Ashraf et al. 2020), it is conceivable that the improvements to IIQ I observe could be a consequence of these changes in board-level focus rather than GDPR itself. To rule out this possibility, I follow Klein et al. (2022) and use firm's proxy statements to identify the extent to which a firm's board is focused on cybersecurity in the given year.[31] From these statements, I identify the scaled count of the number of keywords related to cybersecurity (*BOARDCYBER*) using a cybersecurity keyword list provided by Li, No, and Boritz (2020). I include *BOARDCYBER* along with three other board-level variables from BoardEx (*BOARDSIZE*, or the logged count of the number of board members; *BOARDIND*, or the percentage of independent board members; and *ACFINEXP*, or the percentage of financial experts on the audit committee of the board (Krishnan 2005)) as additional controls in model [1].

In Table 13 Panel A, I report descriptive statistics for variables used in these analyses. Due to requiring both an available proxy statement and BoardEx data for these analyses, my sample size shrinks (from 31,630 to 26,755). While there is not a clear interpretation for the mean value of *BOARDCYBER* (due to scaling), untabulated raw

[31] Proxy statements are the primary SEC filings providing shareholders the information (expertise, background, etc.) needed to evaluate nominated directors and should reflect areas of board-level focus for the firm.

counts of cybersecurity keywords reflect that each proxy statement in my sample has, on average, 1.2 words / phrases related to cybersecurity. On average, firms have 8 members on their board and 78% of these members are independent. 50% of firms' audit committees are financial experts.

Next, before running my IIQ models, I first replicate the results documented in Klein et al. (2022) using my GDPR impact proxies. Specifically, I regress *BOARDCYBER* against my GDPR impact proxies and a series of controls including firm fundamentals (*SIZE*, *LEV*, and *ROA*), *AUDBIG4*, *ACCEL*, *INSTOWN*, *ANALYSTS*, *BOARDSIZE*, *BOARDIND*, *ACFINEXP,* the cyber risk proxies (*BREACHFIRM*, *BREACHIND*, and *CYBERRISK*), and *WRDCNTRISK*. In Table 13 Panel B, I report similar results to those documented in Klein et al. (2022) with significant increases in boards' attention to cybersecurity after the adoption of GDPR.

Lastly, in Table 13 Panel C, I report the results of running my IIQ models (model [1]) after controlling for *BOARDCYBER*, *BOARDSIZE*, *BOARDIND,* and *ACFINEXP*. Results from my primary analyses hold as my GDPR impact proxies continue to be positively associated with IIQ. *BOARDCYBER* is positively, but not significantly (p-value = 0.12), associated with *IIQ*.

Table 13
Controlling for Board-Level Changes

Panel A: Descriptive Statistics

Firm-Year Sample = 26,755					
	Mean	**SD**	**Q1**	**Median**	**Q3**
GDPR-Impact Proxies					
GDPRRISK	0.094	0.292	0.000	0.000	0.000
EUOPS	0.509	0.500	0.000	1.000	1.000
Board Cyber Focus					
BOARDCYBER	0.045	0.122	0.000	0.000	0.039
IIQ Proxy					
IIQ	0.086	0.851	0.119	0.119	0.571
Controls					
Total Assets (millions)	6,848.951	19,298.937	231.095	1,125.400	4,237.187
SIZE	6.846	2.223	5.443	7.026	8.352
LEV	0.626	0.497	0.391	0.589	0.808
ROA	-0.079	0.460	-0.027	0.015	0.057
AUDBIG4	0.671	0.470	0.000	1.000	1.000
ACCEL	0.813	0.390	1.000	1.000	1.000
INSTOWN	0.582	0.339	0.284	0.680	0.879
Analysts (#)	8.008	7.961	2.000	6.000	12.000
ANALYSTS	1.762	1.011	1.099	1.946	2.565
Board Members (#)	8.487	2.452	7.000	8.000	10.000
BOARDSIZE	2.096	0.291	1.946	2.079	2.303
BOARDIND	0.775	0.131	0.714	0.800	0.875
ACFINEXP	0.504	0.281	0.250	0.400	0.667
BREACHFIRM	0.025	0.155	0.000	0.000	0.000
BREACHIND	2.437	1.002	1.609	2.565	3.258
CYBERRISK	1.734	1.648	0.503	1.340	2.492
WRDCNTRISK	8.316	0.620	7.923	8.334	8.748

Panel B: GDPR and Board Cyber Focus

Dependent Var.	*BOARDCYBER*			
	(1)		(2)	
	Coefficient	t-statistic	Coefficient	t-statistic
GDPRRISK	0.024***	*(3.86)*		
*EUOPS*POST*			0.010**	*(2.14)*
SIZE	0.010***	*(2.96)*	0.010***	*(3.09)*
LEV	-0.002	*(-0.70)*	-0.002	*(-0.67)*
ROA	-0.003	*(-1.32)*	-0.003	*(-1.25)*
AUDBIG4	0.006	*(0.87)*	0.006	*(0.96)*
ACCEL	-0.006	*(-1.35)*	-0.006	*(-1.20)*
INSTOWN	-0.025***	*(-3.51)*	-0.025***	*(-3.38)*
ANALYSTS	-0.002	*(-0.62)*	-0.002	*(-0.52)*
BOARDSIZE	0.019**	*(2.49)*	0.020**	*(2.51)*
BOARDIND	0.015	*(1.14)*	0.016	*(1.14)*
ACFINEXP	0.019**	*(2.09)*	0.020**	*(2.08)*
BREACHFIRM	0.025**	*(2.41)*	0.026**	*(2.49)*
BREACHIND	0.002	*(0.90)*	0.003	*(1.35)*
CYBERRISK	0.006***	*(4.05)*	0.006***	*(4.09)*
WRDCNTRISK	-0.021***	*(-3.02)*	-0.019***	*(-2.66)*
Observations	26,755		22,699	
Fixed Effects	Firm / Year		Firm / Year	
Adjusted R^2	0.463		0.452	

Panel C: Change in IIQ Controlling for Board Cyber Focus

Dependent Var.	*IIQ*			
	(1)		(2)	
	Coefficient	t-statistic	Coefficient	t-statistic
GDPRRISK	0.072***	*(3.02)*		
*EUOPS*POST*			0.096***	*(3.86)*
BOARDCYBER	0.092	*(1.54)*	0.093	*(1.55)*
BOARDSIZE	0.125**	*(2.05)*	0.130**	*(2.11)*
BOARDIND	0.292***	*(2.67)*	0.301***	*(2.69)*
ACFINEXP	0.118**	*(2.54)*	0.112**	*(2.41)*
Observations	26,755		22,699	
Model [1] Controls	Included		Included	
Fixed Effects	Firm / Year		Firm / Year	
Adjusted R^2	0.447		0.398	

Table 13 reports descriptive statistics (Panel A) and analyses (Panels B and C) related to studying the effects of GDPR on boards' attention to cybersecurity and on IIQ after

controlling for any shift in board attention. t-statistics calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

9.5. Controlling for the Confounding Effects of CCPA

As discussed in Chapter 2, California's CCPA, which became effective in 2020, has similar features to GDPR. However, there are multiple reasons why the passage of this regulation is unlikely to confound my results. First, the regulation did not become effective until after my sample period, which ends in 2019. Second, although CCPA is similar to GDPR in many respects, the size of fines that can be levied under CCPA are far smaller than those under GDPR, and therefore, the reaction of firms is likely to be less severe.

Nonetheless, to rule out concerns that CCPA is the primary driver of the results identified, I consider whether firms headquartered in the state of California (*CAFIRM*), who are likely to be more affected by CCPA than other firms, experience significant changes in IIQ and operating efficiency compared to other firms around the years of GDPR implementation. Specifically, into models [1] and [2], I introduce an interaction (*CAFIRM*POST*) to control for changes in IIQ and operating efficiency for California firms (compared to other firms) after the adoption of GDPR. These analyses, which are reported in Table 14, reveal that there is not a significant differential effect of GDPR on California firms compared to other firms. Further, my results from main analyses are robust to including this interaction. As such, my results do not appear to be confounded by the effects of CCPA.

Table 14
Controlling for Confounding Effects of CCPA

Panel A: GDPR and IIQ

Dependent Var.		***IIQ***			
Fixed Effects		**Industry / Year**		**Firm / Year**	
		(1)		**(2)**	
	Pred.	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
*EUOPS*POST*	+	0.098***	*(4.38)*	0.102***	*(4.25)*
*CAFIRM*POST*		0.002	*(0.07)*	0.016	*(0.47)*
EUOPS		-0.016	*(-0.53)*		
CAFIRM		0.105***	*(3.00)*		
Observations		26,778		26,778	
Model [1] Controls		Included		Included	
Adjusted R^2		0.148		0.491	

Panel B: GDPR and Operating Efficiency

Fixed Effects			**Industry / Year**		**Firm / Year**	
Dependent Var.			***IIQ***	***OPEREFF***	***IIQ***	***OPEREFF***
Model [2] Eq.			**2A**	**2B**	**2A**	**2B**
	Pred.		**(1)**	**(2)**	**(3)**	**(4)**
	Eq. 2A	**Eq. 2B**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**
IIQ		β_{21}		0.011***		0.004**
		(+)		(5.13)		(2.36)
*EUOPS*POST*	β_{11}	β_{22}	0.093***	-0.026***	0.087***	-0.028***
	(+)	(−)	(2.83)	(-4.86)	(2.72)	(-5.60)
*CAFIRM*POST*			0.018	0.003	0.025	0.005
			(0.44)	(0.46)	(0.64)	(0.72)
EUOPS			0.005	-0.003		
			(0.13)	(-0.36)		
CAFIRM			0.070	0.001		
			(1.57)	(0.07)		
Observations			16,883	16,883	16,883	16,883
Model [2] Controls			Included	Included	Included	Included
Adjusted R^2			0.180	0.583	0.514	0.849
Indirect and Net Effects						
Indirect Effect (β_{11}*β_{21})	+		0.0010** (2.40)		0.0003* (1.79)	
Net Effect (β_{22} + (β_{11}*β_{21}))	−		-0.0254*** (-4.67)		-0.0274*** (-5.53)	

Table 14 presents the results of estimating models [1] (in Panel A) and [2] (in Panel B) after controlling for changes in IIQ and operating efficiency for California firms vs. other firms around the adoption of GDPR. Model [2] consists of equations [2A] and [2B], which are reported side-by-side in Panel B. The indirect and net effects estimated as combinations of coefficients across models are reported at the bottom of Panel B. t-statistics (in Panel A) and z-statistics (in Panel B) calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

9.6. Limiting Sample to Firms with I/B/E/S Coverage

The construction of *IIQ* relies partially on data from I/B/E/S (for *GUIDANCE* and *GUIDANCEACC*), for which there are fewer available observations than my full sample. In cases where a firm is not covered by I/B/E/S, I base my IIQ measure on the remaining four IIQ components (*NOICMW*, *NOITMW*, *NORESTATE*, and *NOLATEFILER*). While this approach allows me to retain a larger sample for my analyses, measuring IIQ differently for firms covered and not covered by I/B/E/S could confound my results. To address this concern, I re-run my main analyses after limiting my sample to firms covered by I/B/E/S. I report these results in Table 15, and they are consistent with those reported in the main analyses.

Table 15
Limiting Sample to Firms with I/B/E/S Coverage

Panel A: GDPR-Related Risk Factor Disclosures and IIQ

Dependent Var.		***IIQ***			
Fixed Effects		**Industry / Year**		**Firm / Year**	
		(1)		**(2)**	
	Pred.	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
GDPRRISK	+	0.083***	*(3.81)*	0.070***	*(2.82)*
Observations		22,437		22,437	
Model [1] Controls		Included		Included	
Adjusted R^2		0.118		0.405	

Panel B: Change in IIQ for Firms with EU Operations

Dependent Var.		***IIQ***			
Fixed Effects		**Industry / Year**		**Firm / Year**	
		(1)		**(2)**	
	Pred.	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
*EUOPS*POST*	+	0.074***	*(3.21)*	0.088***	*(3.46)*
EUOPS		-0.020	*(-0.73)*		
Observations		19,798		19,798	
Model [1] Controls		Included		Included	
Adjusted R^2		0.118		0.394	

Panel C: GDPR-Related Risk Factor Disclosures and Operating Efficiency

Fixed Effects			**Industry / Year**		**Firm / Year**	
Dependent Var.			***IIQ***	***OPEREFF***	***IIQ***	***OPEREFF***
Model [2] Eq.			**2A**	**2B**	**2A**	**2B**
	Pred.		**(1)**	**(2)**	**(3)**	**(4)**
	Eq. 2A	**Eq. 2B**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**
IIQ		β_{21}		0.012***		0.005***
		(+)		(4.89)		(2.78)
GDPRRISK	β_{11}	β_{22}	0.093***	-0.020***	0.084***	-0.008
	(+)	(−)	(3.71)	(-2.91)	(3.23)	(-1.61)
Observations			16,411	16,411	16,411	16,411
Model [2] Controls			Included	Included	Included	Included
Adjusted R^2			0.125	0.518	0.413	0.832

Indirect and Net Effects			
Indirect Effect (β_{11}*β_{21})	+	0.0011*** (3.00)	0.0004** (2.17)
Net Effect (β_{22} + (β_{11}*β_{21}))	–	-0.0184*** (-2.74)	-0.0071 (-1.52)

Panel D: Change in Operating Efficiency for Firms with EU Operations

Fixed Effects			**Industry / Year**		**Firm / Year**	
Dependent Var.			***IIQ***	***OPEREFF***	***IIQ***	***OPEREFF***
Model [2] Eq.			**2A**	**2B**	**2A**	**2B**
	Pred.		**(1)**	**(2)**	**(3)**	**(4)**
	Eq. 2A	**Eq. 2B**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**	**Coefficient (z-statistic)**
IIQ		β_{21} (+)		0.011*** (4.06)		0.005** (2.52)
*EUOPS*POST*	β_{11} (+)	β_{22} (–)	0.081** (2.57)	-0.021*** (-3.54)	0.084*** (2.66)	-0.023*** (-4.35)
EUOPS			-0.011 (-0.33)	-0.002 (-0.23)		
Observations			14,284	14,284	14,284	14,284
Model [2] Controls			Included	Included	Included	Included
Adjusted R^2			0.129	0.526	0.402	0.827
Indirect and Net Effects						
Indirect Effect (β_{11}*β_{21})	+			0.0009** (2.10)		0.0004* (1.84)
Net Effect (β_{22} + (β_{11}*β_{21}))	–			-0.0200*** (-3.38)		-0.0229*** (-4.28)

Table 15 presents the results of estimating models [1] (in Panels A and B) and [2] (in Panels C and D) after removing observations for firms not covered by I/B/E/S. Model [2] consists of equations [2A] and [2B], which are reported side-by-side in Panels C and D. The indirect and net effects estimated as combinations of coefficients across models are reported at the bottom of Panels C and D. t-statistics (in Panels A and B) and z-statistics (in Panels C and D) calculated based on firm-clustered standard errors are presented in parentheses below coefficient estimates. Significance levels are presented as follows based on a two-tailed test: *** p<0.01, ** p<0.05, * p<0.10.

CHAPTER 10

CONCLUSION

In closing, GDPR represents a seismic shift in privacy regulation on a global level, and numerous jurisdictions (including the US) have passed, or are currently contemplating passing, similar privacy regulations.[32] Considering this global trend, understanding the effects of GDPR on firms' information management and operational processes is an important and timely research objective. Using a variety of measures capturing the differential effects of GDPR on US firms, I find that impacted firms exhibit improved IIQ in the face of the regulation. Additionally, while the overall effect of the regulation on firm operations is negative, these improvements in IIQ are themselves beneficial to firm operations. These preliminary findings should be of interest to regulators, who are currently looking to understand the broad costs and benefits of modern privacy regulations. Further, while this dissertation focuses on the relationship between IIQ and operating efficiency, in future work, I plan to explore whether there are other areas that stand to benefit from the improvements in information quality that manifest because of GDPR. These other unexplored benefits may further offset and counterbalance the significant operational costs imposed by the regulation.

[32] See the "US State Privacy Legislation Tracker" maintained by the International Association of Privacy Professionals at https://iapp.org/resources/article/us-state-privacy-legislation-tracker/ for more details on the current status of similar regulations in the US.

BIBLIOGRAPHY

Acquisti, A., C. Taylor, and L. Wagman. 2016. The Economics of Privacy. *Journal of Economic Literature* 54 (2): 442–492.

Aridor, G., Y. K. Che, and T. Salz. 2020. *The Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3522845 (last accessed 3/25/2022).

ARMA. 2017. *The Principles.* https://www.arma.org/page/principles# (last accessed 3/25/2022).

Ashbaugh-Skaife, H., D. W. Collins, and W. R. Kinney, Jr. 2007. The Discovery and Reporting of Internal Control Deficiencies Prior to SOX-Mandated Audits. *Journal of Accounting and Economics* 44 (1–22): 166–192.

Ashraf, M., P. N. Michas, and D. Russomanno. 2020. The Impact of Audit Committee Information Technology Expertise on the Reliability and Timeliness of Financial Reporting. *The Accounting Review* 95 (5): 23–56.

Ashraf, M. 2022. The Role of Peer Events in Corporate Governance: Evidence from Data Breaches. *The Accounting Review* 97 (2): 1–24.

Ashraf, M. and J. Sunder. 2021. *Does Consumer Protection Regulation Benefit Shareholders? Evidence from Data Breach Disclosure Laws and the Cost of Equity*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3308551 (last accessed 3/25/2022).

Baker McKenzie. 2020. *GDPR Survey: Benefits Beyond Compliance* (April 21). https://www.bakermckenzie.com/-/media/files/insight/publications/2020/04/gdpr_survey.pdf?la=en (last accessed 3/25/2022).

Banker, R., Y. Liang, and N. Ramasubbu. 2021. Technical Debt and Firm Performance. Management Science 67 (5): 3174–3194.

Barron, O. E., D. Byard, C. Kile, and E. J. Reidl. 2002. High-Technology Intangibles and Analysts' Forecasts. *Journal of Accounting Research* 40 (2): 289–312.

BDO. 2018. *Countdown to GDPR: How Will the GDPR Affect Information Management?* https://www.bdo.com/insights/business-financial-advisory/information-governance-privacy/countdown-to-gdpr-how-will-the-gdpr-affect-inform (last accessed 3/25/2022).

Burtch, G., A. Ghose, and S. Wattal. 2015. The Hidden Cost of Accommodating Crowdfunder Privacy Preferences: A Randomized Field Experiment. *Management Science* 61 (5): 949–962.

Campbell, J. L., H. Chen, D. S. Dhaliwal, H. Lu, and L. B. Steele. 2014. The Information Content of Mandatory Risk Factor Disclosures in Corporate Filings. *Review of Accounting Studies* 19: 396–455.

Cheng, Q., B. W. Goh, and J. B. Kim. 2018. Internal Control and Operating Efficiency. *Contemporary Accounting Research* 35 (2): 1102–1139.

Chin, C. 2019. *Highlights: The GDPR and CCPA as Benchmarks for Federal Privacy Legislation*. https://www.brookings.edu/blog/techtank/2019/12/19/highlights-the-gdpr-and-ccpa-as-benchmarks-for-federal-privacy-legislation (last accessed 3/25/2022).

Christensen, D., D. Lynch, and C. Partridge. 2021. *You Don't Know What You Don't Know: Improvements in Investment Efficiency Prior to a Mandated Accounting Change*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3825083 (last accessed 3/25/2022).

Davies, J. 2018. *Ad Tech Firms Are Quitting Europe, Blaming the GDPR (Often as a Scapegoat)*. https://digiday.com/media/ad-tech-firms-quitting-europe-blaming-gdpr-often-scapegoat/ (last accessed 6/27/2022).

Demerjian, P., B. Lev, and S. McVay. 2012. Quantifying Managerial Ability: A New Measure and Validity Tests. *Management Science* 58 (7): 1229–1248.

Dorantes, A., C. Li, G. F. Peters, and V. J. Richardson. 2013. The Effect of Enterprise Systems Implementation on the Firm Information Environment. *Contemporary Accounting Research* 30 (4): 1427–1461.

Dyreng, S. D., J. L. Hoopes, P. Langtieg, and J. H. Wilde. 2020. Strategic Subsidiary Disclosure. *Journal of Accounting Research* 58 (3): 643–692.

European Commission (EC). 2010. *A Comprehensive Approach on Personal Data Protection in the European Union*. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0609:FIN:EN:PDF (last accessed 3/25/2022).

European Data Protection Supervisor (EDPS). 2021. *The History of the General Data Protection Regulation.* https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en (last accessed 3/25/2022).

Feng, M., C. Li, and S. McVay. 2009. Internal Control and Management Guidance. *Journal of Accounting and Economics* 48 (2): 190–209.

Forbes. 2018a. *Data Privacy Vs. Data Protection: Understanding the Distinction in Defending Your Data*. https://www.forbes.com/sites/forbestechcouncil/2018/12/19/data-privacy-vs-data-

protection-understanding-the-distinction-in-defending-your-data/ (last accessed 3/25/2022).

Forbes. 2018b. *The GDPR Racket: Who's Making Money from This $9bn Business Shakedown.* https://www.forbes.com/sites/oliversmith/2018/05/02/the-gdpr-racket-whos-making-money-from-this-9bn-business-shakedown/?sh=32e55f3c34a2 (last accessed 3/25/2022).

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder Litigation and Corporate Disclosures. *Journal of Accounting Research* 32 (2): 137–164.

Gal, M. S., and O. Aviv. 2020. The Competitive Effects of the GDPR. *Journal of Competition Law and Economics* 16 (3): 349–391.

Gallemore, J., and E. Labro. 2015. The Importance of the Internal Information Environment for Tax Avoidance. *Journal of Accounting and Economics* 60 (1): 149–67.

Ge, W., A. Koester, and S. McVay. 2017. Benefits and Costs of Sarbanes-Oxley Section 404(b) Exemption: Evidence from Small Firms' Internal Control Disclosures. *Journal of Accounting and Economics* 63 (2–3): 358–384.

Goldberg, S., G. Johnson, and S. Shriver. 2021. *Regulating Privacy Online: An Economic Evaluation of GDPR*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3421731 (last accessed 3/25/2022).

Gordon, L. A., M. P. Loeb, and T. Sohail. 2010. Market Value of Voluntary Disclosures Concerning Information Security. *MIS Quarterly* 34 (3): 567–594.

Heitzman, S. and M. Huang. 2019. Internal Information Quality and the Sensitivity of Investment to Market Prices and Accounting Profits. *Contemporary Accounting Research* 36 (3): 1699–1723.

Hilary, G., B. Segal, and M. H. Zhang. 2016. *Cyber-Risk Disclosure: Who Cares?* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2852519 (last accessed 3/25/2022).

Hope, OK., D. Hu, and H. Lu. 2016. The Benefits of Specific Risk-Factor Disclosures. *Review of Accounting Studies* 21: 1005–1045.

Hope, OK., D. Wang, H. Yue, and J. Zhao. 2022. Information Quality and Workplace Safety. *Journal of Management Accounting Research* 34 (1): 133–162.

Horngren, C. T., S. M. Datar, G. Foster, M. Rajan, and C. D. Ittner. 2012. *Cost Accounting: A Managerial Emphasis, 14th Edition*. Prentice Hall.

Hui, K., H. H. Teo, and S. T. Lee. 2007. The Value of Privacy Assurance: An Exploratory Field Experiment. *MIS Quarterly* 31 (1): 19–33.

International Association of Privacy Professionals (IAPP). 2017. *Global 500 Companies to Spend $7.8B on GDPR Compliance.* https://iapp.org/news/a/survey-fortune-500-companies-to-spend-7-8b-on-gdpr-compliance/ (last accessed 8/4/2021).

Iron Mountain. 2014. *A Practical Guide to Information Governance.* https://www.ironmountain.ca/en/resources/whitepapers/a/a-practical-guide-to-information-governance (last accessed 3/25/2022).

Jamal, K., M. Maier, and S. Sunder. 2005. Enforced Standards versus Evolution by General Acceptance: A Comparative Study of E-Commerce Privacy Disclosure and Practice in the United States and the United Kingdom. *Journal of Accounting Research* 43 (1): 73–96.

Janger, E. J. and P. M. Schwartz. 2002. The Gramm-Leach-Bliley Act, Information Privacy, and the Limits of Default Rules. *Minnesota Law Review* 86: 1219–1262.

Jia, J., G. Z. Jin, and L. Wagman. 2021. The Short-Run Effects of the General Data Protection Regulation on Technology Venture Investment. *Marketing Science* 40 (4): 661–684.

Kottasová, I. 2018. *These Companies Are Getting Killed by GDPR*. https://money.cnn.com/2018/05/11/technology/gdpr-tech-companies-losers/index.html (last accessed 6/27/2022).

Krishnan, J. 2005. Audit Committee Quality and Internal Control: An Empirical Analysis. *The Accounting Review* 80 (2): 649–675.

Krishnan, J., D. Rama., and Y. Zhang. 2008. Costs to Comply with SOX Section 404. *AUDITING: A Journal of Practice & Theory* 27 (1): 169–186.

Kim, G., V. J. Richardson, and M. W. Watson. 2018. IT Does Matter: The Folly of Ignoring IT Material Weaknesses. *Accounting Horizons* 32 (2): 37–55.

Kinney, W. 1999. *Information Quality Assurance and Internal Control for Management Decision Making.* New York: McGraw-Hill Higher Education.

Klein, A., R. Manini, and Y. Shi. 2022. Across the Pond: How US Firms' Boards of Directors Adapted to the Passage of the GDPR. *Contemporary Accounting Research* 39 (1): 199–233.

Laplante, S. K., D. P. Lynch, and M. E. Vernon. 2021. Internal Information Quality and State Tax Planning. *Contemporary Accounting Research* 38 (4): 2589–2621.

Li, H., W. G. No, and T. Wang. 2018. SEC's Cybersecurity Disclosure Guidance and Disclosed Cybersecurity Risk Factors. *International Journal of Accounting Information Systems* 30 (September): 40–55.

Li, H., W. G. No, and J. E. Boritz. 2020. Are External Auditors Concerned about Cyber Incidents? Evidence from Audit Fees. *AUDITING: A Journal of Practice & Theory* 39 (1): 151–171.

Lomas, N. 2020. France Fines Google $120M and Amazon $42M for Dropping Tracking Cookies without Consent. *TechCrunch.com (December 10)*. https://techcrunch.com/2020/12/10/france-fines-google-120m-and-amazon-42m-for-dropping-tracking-cookies-without-consent/ (last accessed 3/25/2022).

Microsoft. 2019. 10-K Filing. *SEC EDGAR*. https://www.sec.gov/Archives/edgar/data/0000789019/000156459019027952/msft-10k_20190630.htm (last accessed 3/25/2022).

Mithas, S., N. Ramasubbu, and V. Sambamurthy. 2011. How Information Management Capability Influences Firm Performance. *Management Information Systems Quarterly* 35 (1): 237–256.

Moore, S. 2020. *Gartner Predicts for the Future of Privacy 2020*. https://www.gartner.com/smarterwithgartner/gartner-predicts-for-the-future-of-privacy-2020/ (last accessed 3/25/2022).

National Archives, The. 2017. *Information Assets Factsheet.* https://www.nationalarchives.gov.uk/documents/information-management/information-assets-factsheet.pdf (last accessed 3/25/2022).

Nelson, R. R. 2007. IT Project Management: Infamous Failures, Classic Mistakes, and Best Practices. *MIS Quarterly Executive* 6 (2): 67–78.

PWC. 2017. *Pulse Survey: US Companies Ramping Up General Data Protection Regulation (GDPR) Budgets*. https://www.pwc.com/us/en/increasing-it-effectiveness/publications/assets/pwc-gdpr-series-pulse-survey.pdf (last accessed 3/25/2022).

Rice, S. C., and D. P. Weber. 2012. How Effective Is Internal Control Reporting under SOX 404? Determinants of the (Non-)Disclosure of Existing Material Weaknesses. *Journal of Accounting Research* 50 (3): 811–843.

RSM. 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 (last accessed 3/25/2022).

Securities and Exchange Commission (SEC). 2022. *Press Release: SEC Proposes Rules on Cybersecurity Risk Management, Strategy, Governance, and Incident*

Disclosure by Public Companies. https://www.sec.gov/news/press-release/2022-39 (last accessed 5/27/2022).

Sedona Conference (Sedona). 2019. The Sedona Conference Commentary on Information Governance, Second Edition. *Sedona Conference Journal* 20: 96–178.

Sheneman, A. G. 2022. *Cybersecurity Risk and the Cost of Debt*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3406217 (last accessed 3/25/2022).

Smallwood, R. F. 2019. *Information Governance: Concepts, Strategies, and Best Practices (2nd Edition)*. John Wiley & Sons, Inc.

Tsai, J. Y., S. Egelman, L. Cranor, and A. Acquisti. 2011. The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research* 22 (2): 254–268.

Yang, Y., B. Pan, and H. Song. 2013. Predicting Hotel Demand Using Destination Marketing Organization's Web Traffic Data. *Journal of Travel Research* 53 (4): 433–447.

APPENDIX A

TIMELINE OF GDPR DEVELOPMENT AND PASSAGE

Date[33]	Event
January 25, 2012	The European Commission proposes a comprehensive reform of the EU's 1995 data protection rules to strengthen online privacy rights.
March 7, 2012	The EDPS adopts an Opinion on the Commission's data protection reform package.
March 23, 2012	The Article 29 Working Party adopts an Opinion on the data protection reform proposal.
October 5, 2012	The Article 29 Working Party provides further input on the data protection reform discussions.
March 12, 2014	The European Parliament demonstrates strong support for the GDPR by voting in plenary with 621 votes in favor, 10 against, and 22 abstentions.
June 15, 2015	The European Council reaches a general approach on the GDPR.
December 15, 2015	The European Parliament, the Council, and the Commission reach an agreement on the GDPR.
February 2, 2016	The Article 29 Working Party issues an action plan for the implementation of the GDPR.
April 27, 2016	GDPR passed by the European Parliament and European Council repealing Directive 95/46/EC (the 1995 Data Protection Directive).
May 25, 2018	GDPR applies from this day forward.

Key:	
	Initial signals of political support used in CAR analyses (Table 6)
	Date of regulation adoption

[33] Adapted from "The History of the General Data Protection Regulation" published by the European Data Protection Supervisor (EDPS) available at: https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en.

APPENDIX B

GDPR PRINCIPLES REGARDING PROCESSING OF PERSONAL DATA

Principle[34]	**Description**
Lawfulness, Fairness, and Transparency	Personal data is to be processed lawfully, fairly, and in a transparent manner in relation to the data subject.
Purpose Limitation	Personal data is to be collected for specified, explicit, and legitimate purposes and not further processed in a manner that is incompatible with those purposes.
Data Minimization	Personal data processed is to be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed.
Accuracy	Personal data is to be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.
Storage Limitation	Personal data is to be kept in a form that permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed.
Integrity and Confidentiality	Personal data is to be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction, or damage, using appropriate technical or organizational measures.

[34] Adapted from GDPR Article 5 available at: https://gdpr-info.eu/art-5-gdpr/

APPENDIX C

GDPR VERSUS OTHER US IMPACTING PRIVACY REGULATIONS

Regulation	The General Data Protection Regulation (GDPR)	California Consumer Protection Act (CCPA)	Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule
	Date Effective: May 25, 2018 **Government Body:** European Union **Source:** https://gdpr-info.eu/	**Date Effective:** January 1, 2020 **Government Body:** State of California **Source:** https://oag.ca.gov/privacy/ccpa	**Date Effective:** April 14, 2003 **Government Body:** US Government **Source:** https://www.hhs.gov/hipaa/index.html
Summary	Data protection law aimed at protecting the privacy rights of European Union citizens based on six principles of processing personal data (see Appendix B)	Secures the following privacy rights for California consumers including: • Right to know what information is collected • Right to delete personal information collected • Right to opt-out of the sale of information • Right to non-discrimination (in CCPA rights)	Establishes national standards to protect individuals' medical records and other individually identifiable health information
Scope of Impacted Organizations	Any person or entity that processes (collects, stores, uses, etc.) personal data, unless such processing is for purely personal or household purposes	Similar to GDPR with the limitation that it is focused on businesses that operate for profit	More targeted than GDPR in that the rule only applies to health plans, health care clearinghouses, and those health care providers that conduct certain health care transactions electronically

Geographic Scope	Extraterritorial scope applying to all organizations collecting personal data of EU subjects	Extraterritorial in scope but applies to businesses that offer goods and services to persons or entities in CA	Unlike GDPR, HIPAA lacks extraterritorial scope and applies only to individuals treated within the US
Scope of Protected Data	Any information related to an identified or identifiable natural person including identifiers (e.g., name), locations (e.g., IP addresses), or any other data traceable to individual	Similar to GDPR and also includes data connected to a household even if not tied to an individual	Unlike GDPR, HIPAA is focused only on protected health information (PHI) which includes individuals' medical records and other individually identifiable health information
Size of Potential Fines	Up to €20 million or 4% of annual worldwide turnover of the preceding financial year, whichever is greater	Smaller than GDPR at $2,500 per violation or $7,500 per intentional violation	Smaller than GDPR at a maximum level of $25,000 per violation category per calendar year

APPENDIX D

VARIABLE DEFINITIONS

Variable	Definition
Test Variables (in Order of Appearance)	
GDPRRISK	An indicator coded 1 if the firm references either of the terms GENERAL DATA PROTECTION REGULATION or GDPR in Item 1A (risk factors section) of the firm's 10-K filing, and 0 otherwise (Source: SEC EDGAR).
EUOPS	An indicator coded 1 if the firm reports an EU-based segment in Compustat or an EU-based subsidiary in WRDS Company Subsidiary database over the sample period, and 0 otherwise. EU-based segments were identified by manually reviewing the list of Compustat geographic segment names (SNMS) for my in-scope firms to determine those reflecting EU or European Economic Area (EEA) regions (EU, EMEA, etc.) or member states as of 2019. EU-based subsidiaries were identified based on whether the subsidiary was listed with the COUNTRY_CODE reflecting an EU or EEA member state as of 2019 including (AT, BE, BG, HR, CY, CZ, DK, EE, FI, FR, DE, GR, HU, IS, IE, IT, LV, LI, LT, LU, MT, NL, NO, PL, PT, RO, SK, SI, ES, SE, and GB). (Source: Compustat / WRDS Subsidiary Database).
POST	An indicator coded 1 for fiscal years 2016 and after, and 0 otherwise.
GDPRFIRM	An indicator coded for the maximum value of *GDPRRISK* for the firm over the sample period (Source: SEC EDGAR).
POSTADOPT	An indicator coded 1 for fiscal years 2016 and 2017, and 0 otherwise.
POSTENFORCE	An indicator coded 1 for fiscal years 2018 and 2019, and 0 otherwise.
CAFIRM	An indicator coded 1 for firms headquartered in the state of California, and 0 otherwise (Source: Audit Analytics).
Dependent Variables (in Order of Appearance)	
IIQ	The average of standardized values of the following internal information quality proxies: *NOICMW*, *NOITMW*, *NORESTATE*, *NOLATEFILER*, *GUIDANCE*, and *GUIDANCEACC*.
NOICMW	An indicator coded 1 if the firm does not report an internal control material weakness in its annual SOX 404(a) or 404(b) report, and 0 otherwise (Source: Audit Analytics).

NOITMW	An indicator coded 1 if the firm does not report an internal control material weakness of the following type in its annual SOX 404(a) or 404(b) report based on Audit Analytics coding following Ashraf et al. (2020): 22 (information technology, software, security and access), 42 (segregations of duties/design of controls), 76 (journal entry control issues), 12 (untimely or inadequate account reconciliations), 77 (non-routine transaction control issues), and 57 (treasury control issues) (Source: Audit Analytics).
NORESTATE	An indicator coded 1 for firms that do not subsequently restate their year-end financial statements, and 0 otherwise (Source: Audit Analytics).
NOLATEFILER	An indicator coded 1 for firms that file their annual 10-K timely (i.e., do not file a form NT 10-K), and 0 otherwise (Source: Audit Analytics).
GUIDANCE	An indicator coded 1 for firms that issue sales guidance (MEASURE = "SAL") during the 365 days leading up to the fiscal year end, and 0 otherwise (Source: I/B/E/S).
GUIDANCEACC	The average of the negative absolute value of the difference between sales guidance issuances and actual sales reported (scaled by actual sales reported) for sales guidance issuances made during the 365 days leading up to the fiscal year end, and 0 otherwise (Source: I/B/E/S).
OPEREFF	Operating efficiency for the given firm based on the DEA model outlined in Demerjian et al. (2012). This model estimates firm operational efficiency by using one output of revenue and seven inputs: net PP&E; cost of goods sold; selling, general, and administrative expenses; capitalized operating leases; capitalized R&D costs; purchased goodwill; and other intangibles. This model is run by year across non-financial services (SIC codes 6000–6999) and non-utilities (SIC codes 4900–4999) firms. The variable used in analyses is the percentile ranking of this efficiency score by year and Fama-French 48 industry classification (taking the value from 0.01 to 1.00 (Source: Compustat).
EUSEG	An indicator coded 1 if the firm reports an EU-based segment in Compustat for the given fiscal year, and 0 otherwise (Source: Compustat).
EUSEGCNT	Natural log of 1 plus the number of EU-based segments in Compustat for the given fiscal year (Source: Compustat).
EUSUB	An indicator coded 1 if the firm reports an EU-based subsidiary in WRDS Company Subsidiary database for the given fiscal year, and 0 otherwise (Source: WRDS Subsidiary Database).
EUSUBCNT	Natural log of 1 plus the number of EU-based subsidiaries in WRDS Company Subsidiary database for the given fiscal year (Source: WRDS Subsidiary Database).

Control Variables (in Alphabetical Order)

ACCEL	An indicator coded 1 for firms reporting as accelerated (or large accelerated) filers, and 0 otherwise (Source: Audit Analytics).
ACFINEXP	The percentage of financial experts on the firm's audit committee (Source: BoardEx).
AGE	The natural log of the number of years that the firm has been present in Compustat (Source: Compustat).
ANALYSTS	The natural log of 1 plus the number of analysts issuing earnings forecasts for a firm in the given year (Source: I/B/E/S).
AUDBIG4	An indicator coded 1 for firms employing a Big 4 auditor, and 0 otherwise (Source: Audit Analytics).
BOARDCYBER	The count of references to information security/cybersecurity as represented by the list of terms provided in Appendix B of Li et al. (2020) multiplied by 1000 and scaled by the total number of the words in the firm's proxy filing (Source: SEC EDGAR).
BOARDIND	The percentage of independent directors on the firm's board of directors (Source: BoardEx).
BOARDSIZE	The natural log of the number of members of the firm's board of directors (Source: BoardEx).
BREACHFIRM	An indicator coded 1 for firms reporting a cyber breach in the two years prior to the fiscal year end, and 0 otherwise (Source: Audit Analytics).
BREACHIND	Natural log of the number of reported cyber breaches by firms in the same Fama-French 12 industry classification over the two years leading up to the firm's fiscal year end date (Source: Audit Analytics).
CONCEN	The ratio of individual business segment sales to total sales, summed across all business segments. If the firm is not in the Compustat segments file, it is assigned a concentration of 1 (Source: Compustat).
CYBERRISK	The count of references to the list of terms provided in Appendix B of Li et al. (2020) multiplied by 1000 and scaled by the total number of the words in Item 1A (risk factors section) of the 10-K filing (Source: SEC EDGAR).
FOREIGN	An indicator coded 1 if the firm reports foreign income (PIFO non-blank / non-zero) in Compustat, and 0 otherwise (Source: Compustat).
GROWTH	The percentage change in sales (SALE) from year *t-1* to year *t* (Source: Compustat).

HITECH	An indicator coded 1 for high technology intangibles industries per Barron et al. (2002), and 0 otherwise; high technology intangibles industries include the following 3-digit SIC codes: 283, 284, 357, 366-367, 371, 382, 384, 737 (Source: Audit Analytics).
INSTOWN	The percentage of shares owned by institutions at the end of year *t* (Source: Thomson Reuters).
INVENTORY	The ratio of inventory (INVT) to total assets (AT) (Source: Compustat).
LEV	The ratio of total liabilities (LT) to total assets (AT) (Source: Compustat).
LITIGATION	An indicator coded 1 for high litigation risk industries per Francis, Philbrick, and Schipper (1994), and 0 otherwise; high litigation risk industries include pharmaceuticals and biotechnology (SIC Codes 2833-2836 and 8731-8734), computers (SIC Codes 3570-3577 and 7370-7374), electronics (SIC Codes 3600-3674), and retail (SIC Codes 5200-5691) (Source: Audit Analytics).
MA	An indicator coded 1 for firms with acquisition expenses (AQC > 0), and 0 otherwise (Source: Compustat).
MKTSHARE	The percentage of revenue (SALE) in the Fama-French 48 industry earned by the given firm in the given year (Source: Compustat).
REST	An indicator coded 1 for firms with restructuring expenses (RCP > 0), and 0 otherwise (Source: Compustat).
ROA	The ratio of earnings (IB) to total assets (AT) (Source: Compustat).
SEG	The natural log of 1 plus the number of firm geographic and operating segments (Source: Compustat).
SIZE	Natural log of total assets (AT) (Source: Compustat).
WRDCNTRISK	Natural log of the word count of the contents of Item 1A (risk factors section) of the firm's 10-K filing (Source: SEC EDGAR).