IE UNIVERSIDAD

TESIS DOCTORAL/ DOCTORAL DISSERTATION

Filtraciones de Datos y Contabilidad: Tres Ensayos

Data Breaches and Accounting: Three Essays

CHRISTELLE ALKHOURY

SEGOVIA, 2025

IE UNIVERSIDAD

TESIS DOCTORAL/ DOCTORAL DISSERTATION

Filtraciones de Datos y Contabilidad: Tres Ensayos

Data Breaches and Accounting: Three Essays

CHRISTELLE ALKHOURY

Doctoral Thesis Advisors:

Nieves Carrera

Amanda Wilford

Abstract

This dissertation examines key challenges in firms' data security, focusing on board interlocks, scientific disclosures, and audit offices, through an analysis of secondary data from multiple sources.

In Chapter 1, I examine whether spillover effects of data breaches present themselves through board interlocks. In testing this, I empirically show that interlocking to previously breached firms or firms that will experience a breach, puts focal firms at a higher risk of experiencing a breach. The findings show a dark side of interlocking, and contribute to the corporate governance, spillover effects, data security, and social capital literatures.

In Chapter 2, I investigate how exposure to data breach risk affects firms' issuance of scientific publications. In doing so, I examine how peer breaches impact firms' publishing decisions. I argue that firms publish less to avoid signaling valuable intellectual property, revealing an indirect economic cost of data breaches. This study contributes to the literature on non-financial disclosures, and answers calls for more research on publishing incentives and non-traditional disclosures.

In Chapter 3, I move the focus to audit offices and assess the effect of audit offices' cybersecurity experience and range of industry experiences on their clients' breach likelihood. I find that clients are more likely to experience a breach when the audit office has had a previously breached client or will have a breached client. However, this breach likelihood is decreased when auditors have a broad industry range. This study contributes to the cybersecurity research within the auditing literature and to the literature on the effects of audit office experiences. Also, it is valuable to regulators in the public domain.

Taken together, these three chapters are relevant and timely given the increased impact and significance of data breaches. As such, these studies should have a contemporary appeal to both academics and practitioners.

Resumen

Esta tesis examina los principales desafíos en la seguridad de los datos de las empresas centrándose en las interconexiones (*interlocks*) de los consejos de administración, las divulgaciones científicas y los auditores externos, a través de un análisis de datos secundarios de múltiples fuentes.

En el Capítulo 1, analizo si los efectos secundarios de las filtraciones de datos se manifiestan a través de las interconexiones (*interlocks*) entre los miembros de los consejos de administración. Con este análisis se demuestra empíricamente que las conexiones con empresas que previamente han experimentado una filtración o con aquellas que experimentarán una filtración el futuro, pone a las empresas focales en un mayor riesgo de sufrir una filtración. Los resultados muestran un "lado oscuro" de las interconexiones vía consejos de administración y contribuyen a los estudios sobre gobierno corporativo, el contagio de prácticas vía miembros del consejo de administración, seguridad de los datos y capital social.

En el Capítulo 2, investigo cómo la exposición al riesgo de filtración de datos afecta a la publicación de resultados científicos por parte de las empresas. De este modo, examino cómo las filtraciones de datos entre pares impactan las decisiones de publicación de las empresas. Sostengo que las empresas publican menos para evitar revelar propiedad intelectual valiosa, lo que revela un costo económico indirecto de las filtraciones de datos. Este estudio contribuye a los estudios sobre divulgación no financiera y responde a la demanda de más investigación sobre incentivos de publicación y divulgación no tradicional.

En el Capítulo 3, redirijo el enfoque hacia las oficinas de las firmas de auditoría, evaluando el efecto de su experiencia en ciberseguridad y en la industria de sus clientes sobre la probabilidad de filtración de datos de sus clientes. Los resultados indican que los clientes son más propensos a experimentar una filtración cuando su auditor pertenece a una oficina que ha tenido, o tendrá, un cliente afectado por una filtración. Sin embargo, esta probabilidad de filtración disminuye cuando los auditores tienen una amplia experiencia en la industria. Este estudio contribuye a la investigación en ciberseguridad y su relación con los auditores externos y su experiencia a nivel de oficina. Además, los resultados son de relevancia para los reguladores.

En conjunto, estos tres capítulos son especialmente relevantes y oportunos debido al creciente impacto y la importancia de las filtraciones de datos. Por ello, sus resultados deberían resultar de interés tanto para la comunidad académica como para los profesionales.

Acknowledgments

A PhD is challenging yet incredibly rewarding! This journey has been marked by its share of late nights, early morning efforts, sacrifices, and lots of coffee. As I reflect on the past five years, I am deeply grateful for the unwavering support of those who helped me navigate this process.

First and foremost, my sincerest gratitude goes to my advisors, Dr. Nieves Carrera and Dr. Amanda Wilford. Their guidance and encouragement were invaluable throughout this journey. They consistently challenged me to strive for excellence and always went the extra mile, even on short notice.

They say family is everything, and indeed it is! I am so blessed and grateful for my parents and brother, to whom I owe everything. Ramy, from childhood, you've always had my back. You are, and always will be, my inspiration and role model. Thank you for your unwavering support from miles away, for lifting me up every time I stumbled, for pushing me toward my goals, and for assuring me that my efforts would be rewarded. Samir, your support these past five years has been invaluable. You always believed in me, encouraged me to pursue my dream, and consistently checked on my progress, reminding me of how far I had come. Mona, I don’t even know where to begin to describe your importance in my life. You are not just my mother, but my best friend, my rock, my haven, and so much more. You’ve always pushed me to be the best version of myself and to aim high. Thank you for tirelessly listening to me talk for hours about my research and my hectic schedule. To my grandmother, Nour, thank you for your constant prayers and for asking God to light my way. Antonella, you’re not just my brother’s wife, you’re my sister. Thank you for your continuous support, our long conversations, and the many cherished moments we've shared.

To my friends and cousins, both near and far, thank you for your encouragement and for listening to my endless venting about work. A special thanks to Marita, Jennifer, Rita, Pamela, and Cindy – my invaluable support system, with whom I can share anything.

I extend my sincere appreciation to my co-authors, Tim Martens, Robert Felix, Xiaochi Ge, and Christoph Sextroh, for their significant contributions, constructive feedback, and inspiring collaboration on various aspects of my thesis. Your expertise and insights were crucial to the successful completion of my research.

I would also like to express my gratitude to the committee members for dedicating their time to review and evaluate my thesis. A special thank you to Marco Trombetta and Lei Zhou for their valuable feedback, which was instrumental in shaping my research. Thank you as well to Pietro Bianchi for serving as the president of my PhD committee; I eagerly anticipate your feedback.

Finally, a heartfelt thank you to my colleagues, with whom I shared workspaces, ideas, notes, long calls, and countless venting sessions.

I couldn’t have done it without the support of each and every one of you. Thank you!

List of Contents

Introduction

In a world where technology evolves at the speed of light, the risk of data breaches is dramatically increasing. Data breaches have become so prevalent, reaching a record high of $4.88 million globally in 2024. The significance of these events warrants further investigation into the leading factors, and the effects these events have on firms. This dissertation aims to address questions related to data breaches and accounting from different perspectives.

In the first chapter "The Dark Side of Board Interlocks: Evidence from Data Breaches" I test whether non-breached firms are protected against breaches when interlocked, through board members, to firms that have or will experience a breach. Board interlocks have been associated with positive and negative spillover effects (e.g., Bizjak et al., 2009; Cheng et al., 2021; Chiu et al., 2013; Kang, 2008) and prior studies have shown that the effect of data breaches is not exclusive to breached firms (Ashraf, 2022; Islam et al., 2022). However, board interlocks have been underexplored in the context of cybersecurity risks. This study builds upon prior research and investigates the impact of interlocking on the risk of data breaches. The results indicate that interlocking leads to negative spillovers by exposing non-breached firms to higher likelihoods of experiencing a breach. They are more pronounced when the breach is internal rather than external and when firms have more interlocks. The presence of interlocked breached female directors or interlocked breached directors who also serve on the audit committee is associated with a lower breach likelihood, indicating their effective engagement in cybersecurity matters. Additional analysis indicates that the presence of low-quality or high-quality directors on boards increases the breach likelihood, indicating that our results are unlikely to be driven by director quality. Moreover, the presence of breached suppliers increases the likelihood of breaches for firms, highlighting the negative consequences of operating within the same supply chain.

In the second chapter "Becoming Invisible? Data Breach Risk and Scientific Publications" I examine the effect of peer breaches, in the same state and industry, on firms' issuance of scientific publications. Prior studies have investigated the effect of data breaches on firms' corporate innovation (Lattanio & Ma, 2023) and firms' investment in corporate governance (Ashraf, 2022), however, the effect on firms' scientific disclosure behavior remains largely unstudied. To understand the incentives that drive scientific disclosure decisions and the economic consequences of data breaches, I test whether firms' publishing frequency changes when exposed to higher breach risks. The use of peer breaches, which are exogenous to the non-breached firms, mitigates the endogeneity concern that breached firms might directly react to a breach. The results indicate that increased breach risks drive firms to issue less scientific publications. This could be interpreted as a risk-mitigation strategy adopted by firms to protect their intellectual property (IP). To understand whether breach risk affects other non-financial disclosures beyond scientific publications, I test the effect on trade secret disclosure. Consistent with the previous argument, I find a decrease in trade secret disclosure, which is in line with firms disclosing less IP-related content when faced with increased risks of data breaches.

In the third chapter "Audit Offices' Cybersecurity Experience and Industry Range" I investigate the effect of audit office cybersecurity experience on audit clients in pre and post cyberattack periods. Although auditors are not required to provide assurance on their clients' cybersecurity practices (Li et al., 2024), current auditing standards require auditors to play a significant role in cybersecurity (CAQ, 2019; Hamm, 2019). A breach to one client may be interpreted as auditors failing to properly evaluate a client's internal control over financial reporting, a failure that could spill over to other clients of the same audit office (Li et al., 2024). Alternatively, auditors of breached clients gain first-hand cybersecurity experience and may become more knowledgeable in areas that include information technology management

actions, internal controls, etc. (Li et al., 2024; Smith et al., 2019). By focusing on the informal spillover channel, through sharing the same audit office, I examine whether cybersecurity information may be shared. By doing so, I examine whether audit offices act as knowledge intermediaries or as risk intermediaries. Additionally, I test the effect of auditor industry range on client likelihood of experiencing a data breach, to determine whether a broader range of industry experiences plays a significant role in this setting of data breaches. Findings suggest that clients are more likely to experience a cyber incident when their audit office has had a previously breached client or will have a breached client. However, this breach likelihood is decreased when auditors are identified as having a broad range of industry experiences.

The methodological approach of this dissertation is quantitative, using regression analysis (i.e., logistic and OLS) and other advanced statistical techniques to test the hypotheses and examine the relationship between key variables across the three studies. In all chapters, I use U.S. secondary data extracted from different sources: Privacy Rights Clearinghouse (PRC), Institutional Shareholder Services (ISS), Audit Analytics, Compustat, CRSP, 10-K reports, Arora et al. (2024a, 2024b, 2021), and Kogan et al. (2017).

Overall, this dissertation contributes to the literatures on data security, corporate governance, spillover effects, social capital, non-financial disclosures, and auditing. It provides insights for boards, management, regulators, and shareholders, suggesting that interlocking boards and audit offices can act as risk spillover mechanisms for data breaches. Additionally, it provides insights for policy makers and potential investors by suggesting that data breaches affect firms' scientific disclosures, consequently imposing indirect economic costs. All three chapters provide evidence that data breaches do not only affect the breached firms, but spill over to other firms.

Introducción

En un mundo donde la tecnología evoluciona a la velocidad de la luz, el riesgo de filtraciones de datos está creciendo drásticamente. Las filtraciones de datos se han vuelto tan comunes, alcanzando un récord de 4,88 millones de dólares a nivel mundial en 2024. La magnitud e importancia de estos sucesos justifica una investigación más profunda sobre los factores principales y los efectos que estos sucesos provocan en las empresas. Esta tesis tiene como objetivo abordar cuestiones relacionadas con las filtraciones de datos y la contabilidad desde diferentes perspectivas.

En el primer capítulo, "El lado oscuro de las interconexiones (*interlocks*) de los consejos de administración: Evidencia de las filtraciones de datos", pruebo si las empresas no afectadas por filtraciones están protegidas contra las filtraciones cuando tienen conexiones, a través de miembros del consejo de administración, con empresas han tenido o tendrán una filtración. Las interconexiones o *interlocks* del consejo de administración se han asociado con "efecto contagio" tanto positivos como negativos (p. ej., Bizjak et al., 2009; Cheng et al., 2021; Chiu et al., 2013; Kang, 2008) y estudios previos han demostrado que los efectos de las filtraciones de datos no se limitan exclusivamente a las empresas afectadas por la filtración (Ashraf, 2022; Islam et al., 2022). Sin embargo, hasta la fecha los efectos de las interconexiones de los consejos de administración en el contexto de los riesgos de ciberseguridad han sido poco explorados. A partir de los resultados de investigaciones previas, este estudio analiza el impacto de los *interlocks* en el riesgo de filtraciones de datos. Los resultados indican que las interconexiones conducen a un efecto contagio negativo al exponer a las empresas no afectadas por la filtración a una mayor probabilidad de experimentar una filtración. Estos efectos son más importantes cuando la filtración es interna en lugar de externa y cuando las empresas tienen más vínculos entre los miembros del consejo de administración. La presencia de consejeras mujeres vinculadas a consejos de empresas afectadas por la filtración o de consejeros

vinculados a consejeros que forman parte del comité de auditoría se asocia con una menor probabilidad de filtración, lo que indica la participación efectiva de estos consejeros en temas de ciberseguridad. Un análisis adicional indica que la presencia de miembros del consejo de "baja" o "alta" calidad en los consejos de administración aumenta la probabilidad de filtración, lo que indica que nuestros resultados no están, probablemente, influenciados por la calidad de los directores. Además, la presencia de proveedores afectados por una filtración aumenta la probabilidad de filtraciones para las empresas relacionadas, resaltando las consecuencias negativas de operar dentro de la misma cadena de suministro.

En el segundo capítulo "¿Volverse Invisible? Riesgo de Filtraciones de Datos y Publicaciones Científicas", investigo el efecto de las filtraciones de datos de empresas entre pares, en el mismo estado e industria, sobre la publicación de resultados científicos por parte de las empresas. Estudios previos han analizado el efecto de las filtraciones de datos sobre la innovación de las empresas (Lattanio & Ma, 2023) y sobre la inversión de las empresas en gobierno corporativo (Ashraf, 2022). Sin embargo, el impacto de estas filtraciones sobre el comportamiento de divulgación científica de las empresas sigue siendo, en gran medida, un área poco explorada. Para comprender los incentivos que guían las decisiones de divulgación científica y las consecuencias económicas de las filtraciones de datos, este capítulo investiga si la frecuencia de publicación de las empresas de sus resultados científicos cambia al estar expuestas a mayores riesgos de filtraciones. El uso de filtraciones de empresas entre pares, las cuales son exógenas a las empresas no afectadas por la filtración, mitiga el problema de endogeneidad de que las empresas afectadas por la filtración puedan reaccionar directamente ante una filtración. Los resultados indican que un mayor riesgo de filtraciones induce a las empresas a reducir la decisión de publicar sus resultados científicos. Esto podría interpretarse como una estrategia de mitigación de riesgos que las empresas adoptan para proteger su propiedad intelectual (PI). Para entender si el riesgo de filtraciones influye en otras

divulgaciones no financieras más allá de las publicaciones científicas, también se estudia el efecto sobre la divulgación de secretos comerciales. En línea con el argumento anterior, los resultados indican una disminución en la divulgación de secretos comerciales, lo cual concuerda con la idea de que las empresas divulgan menos contenido relacionado con la PI cuando se enfrentan a un mayor riesgo de filtraciones de datos.

En el tercer capítulo "Experiencia en Ciberseguridad y en Múltiples Industrias de las Firmas de Auditoría", estudio el efecto de la experiencia en ciberseguridad de las firmas auditoras, a nivel de oficina, sobre los clientes de auditoría en los períodos previos y posteriores a un ciberataque. Aunque los auditores no están obligados a proporcionar aseguramiento sobre las prácticas de ciberseguridad de sus clientes (Li et al., 2024), los estándares actuales de auditoría exigen que los auditores jueguen un papel importante en materia de ciberseguridad (CAQ, 2019; Hamm, 2019). Una filtración en un cliente podría interpretarse como un "fallo" de los auditores al evaluar los controles internos sobre la información financiera de un cliente, lo cual podría repercutir en otros clientes de la misma firma de auditoría (Li et al., 2024). Alternativamente, los auditores de clientes afectados por una filtración adquieren experiencia directa en ciberseguridad y pueden volverse más conocedores en áreas relacionadas con la gestión de tecnología de la información y controles internos (Li et al., 2024; Smith et al., 2019). El análisis del canal informal de contagio de información que supone utilizar la misma firma auditora y la misma oficina, permite investigar si la información sobre ciberseguridad puede ser compartida vía auditor. De este modo, el capítulo examina si las oficinas de auditoría actúan como intermediarios de conocimiento o intermediarios de riesgos vinculados a los ciberataques.

Además, con el objetivo de determinar si la experiencia en múltiples industrias del auditor afecta a las filtraciones de datos, en el estudio se estudia el efecto que dicha experiencia tiene sobre la probabilidad de que el cliente experimente una filtración de datos. Los resultados

indican que los clientes tienen mayor probabilidad de experimentar un incidente cibernético cuando la oficina de su firma auditora ha tenido un cliente previamente afectado por una filtración o tendrá un cliente afectado en el futuro. Sin embargo, esta probabilidad de filtración disminuye cuando los auditores tienen experiencia en múltiples industrias.

El enfoque metodológico de esta tesis es de tipo cuantitativo, utilizando análisis de regresión (OLS y regresión logística) y otras técnicas estadísticas avanzadas para contrastar las hipótesis y estudiar la relación entre las variables clave de los tres estudios. En todos los capítulos, utilizo datos secundarios de EE. UU. extraídos de diferentes fuentes: Privacy Rights Clearinghouse (PRC), Institutional Shareholder Services (ISS), Audit Analytics, Compustat, CRSP, informes 10-K, Arora et al. (2024a, 2024b, 2021) y Kogan et al. (2017).

En términos generales, esta tesis contribuye a los estudios de seguridad de datos, gobierno corporativo, efecto contagio en el ámbito del gobierno corporativo y auditoría, capital social, divulgaciones no financieras, y auditoría a nivel de oficina. Los resultados son de potencial interés para los consejos de administración, la gerencia, los reguladores y los accionistas, sugiriendo que los miembros de los consejos de administración y los auditores externos, mediante sus interconexiones, pueden actuar como mecanismos de "transmisión" de riesgos ante filtraciones de seguridad de datos. Además, los resultados relativos al efecto que las filtraciones de seguridad de datos tienen sobre las divulgaciones científicas de las empresas, resultando en costes económicos secundarios, son relevantes para potenciales inversores, reguladores y responsables políticos. En conclusión, los tres capítulos ofrecen evidencia de que las filtraciones de datos no solo afectan a las empresas que las sufren, sino que también se transmiten a otras empresas.

Chapter 1

The Dark Side of Board Interlocks:

Evidence from Data Breaches

ABSTRACT

This paper examines whether spillover effects of data breaches present themselves through board interlocks. By definition, breached firms have had their confidential data intentionally or unintentionally exposed to unauthorized parties. Data breaches put affected firms at risk because of the proprietary nature of the breached data. However, the effect is not exclusive to the breached firm and can spill over to interlocked focal firms. On one hand, interlocking may mitigate the risk of data breaches by facilitating the transfer of information and knowledge that enhances data security management. On the other hand, interlocking might increase the likelihood of breaches if interlocked directors are not well-equipped to share information or are too occupied to effectively disseminate the acquired knowledge. We find that focal firms are more likely to experience a data breach if they are interlocked to a firm that has already experienced a breach or that will experience a breach in the future. The breach likelihood is more pronounced when the breach is internal rather than external. Further, the higher the number of interlocks, the greater the breach likelihood. Additional analysis indicates that the presence of low-quality directors on boards increases the likelihood of a focal firm experiencing a breach more than the presence of high-quality directors. However, the spillover effect is mitigated in the presence of interlocked female directors with prior or post breach experience, consistent with the notion that female directors are more risk-averse and have strong firm monitoring behaviors. Also, the presence of directors who are either executives or also serve on the audit committee of the focal firm decreases breach likelihood, which provides evidence that executives and audit committee members actively engage in data security matters. Moreover, the presence of breached suppliers increases the likelihood of breaches for firms, highlighting the negative consequences of operating within the same supply chain. Overall, our study sheds light on how data breaches spread through the interlock network.

Keywords: Data breaches, board interlocks, spillover, internal breach, external breach.

1. Introduction

According to the IBM Cost of Data Breach Report 2023, the average cost of a data breach in 2023 was $4.45 million, representing a 15% increase over the previous three years (IBM, 2023). Despite increased security investments, managerial efforts, and resource allocation to combat cybersecurity breaches, breaches continue to increase in number and severity (Garg, 2020; Rothrock et al., 2018). Breaches have become so prevalent that only the breaches with the largest impact (i.e., Equifax, Target, Sony, and Facebook) make headlines (CSIS, 2021; Héroux & Fortin, 2022). Reported breaches generally expose a large amount of sensitive personal information and cost companies millions of dollars in both direct and indirect costs (Li et al., 2018). Following a security breach, firms must manage public scrutiny and reputation damages (Li et al., 2018). More than 20% of breached firms experience substantial drops in revenues, customers, and business opportunities (CISCO, 2017). In a recent report, the U.S. Securities and Exchange Commission (hereafter, SEC) indicates that information security should be prioritized and is a perennial focus area (SEC, 2024). Even though firms are working to combat data breaches and research studies seek to disentangle the root causes of the breaches and their associated losses, these breaches continue.

To combat breaches, firms have established cybersecurity frameworks. Building a firm-wide cybersecurity framework is not simply a process of security software installation, but a complex endeavour that involves technological and organizational pieces (Yeoh et al., 2022). As such, cybersecurity risk is no longer the sole responsibility of the information technology department (Rothrock et al., 2018). Rather, cybersecurity should be strategically addressed from the top down, and boards of directors are expected to place cybersecurity as a top priority (Garg, 2020; Rothrock et al., 2018). While IT experts on the board could help with cybersecurity governance, firms have had difficulty finding these qualified individuals for their boards (Bonime-Blanc, 2017; Héroux & Fortin, 2022).

Accordingly, the board at-large must play an important role in influencing a firm's cybersecurity governance. Board members have exclusive access to monitoring and decision-making information (Cheng et al., 2019), oversee the management of cybersecurity, and ensure the protection of personal and other sensitive data (Bonime-Blanc, 2017; Héroux & Fortin, 2022). With this in mind, we seek to examine whether shared board memberships, or "interlocking boards" (Cai et al., 2014; Chiu et al., 2013), can have an impact on data security. Board members are regarded as a "trusted source", with whom information sharing is acceptable and the shared information is considered valid (Johansen & Pettersson, 2013). Also, board members manage cybersecurity risk (Bonime-Blanc, 2017; Héroux & Fortin, 2022) and contribute to the overall governance of the firm (Hauser, 2018).

Building on governance relationships, studies based on social capital theory provide invaluable insights into how connections impact firms (e.g. Kim & Cannella, 2008; Kor & Sundaramurthy, 2009) by creating social networks that shape governance practices (Bertrand et al., 2000). Social capital theory traditionally emphasizes the bright side of social connections (Bianchi, 2018; Intintoli et al., 2018; Lin, 2002) however, social ties can also have a dark side (Bruynseels & Cardinaels, 2014; Carrera et al., 2017; He et al., 2017; Portes, 1998; Van Deth & Zmerli, 2010). Board interlocks, among the most studied types of social networks, have been associated with positive and negative spillover effects (Bizjak et al., 2009; Kang, 2008).

Using the incidence of data breaches as our key variable, we investigate whether and how board interlocks affect the spread of data breaches. We choose interlocked firms' data breaches as our setting because, besides being an important economic question to address, data security is a significant risk faced by firms (Ashraf, 2022). Additionally, the relation between this risk and firm governance is not clear and is difficult to understand (Rajgopal & Srinivasan, 2016). As security breaches become increasingly unavoidable, firms may choose to respond to them only after their occurrence (Sonnemaker 2019), rather than proactively taking actions to

prevent their occurrence. This implies a reactive, rather than a preventive, approach when dealing with security breaches. Ettredge and Richardson (2003) show that focal firms' exposure to risk might be heightened when connected to breached firms. As such, breached firms may represent a plausible exogenous signal of security risk (or protection) to the connected focal firms. Accordingly, interlocked firms' data breaches represent a unique setting to evaluate whether connected board members take real actions to reduce exposure to such risks. To the best of our knowledge, no prior studies have investigated whether board interlocks impact the likelihood of data breaches. More specifically, this study aims to investigate whether board interlocks benefit connected focal firms by lowering their likelihood of experiencing a breach, or whether they put them at a higher risk of experiencing a breach.

While the evidence is clear that breaches can negatively impact breached firms, these effects are not exclusive to the breached firm and can impact, either positively or negatively, other firms (Garg, 2020). Research also indicates that focal firms may benefit from data breaches experienced by their breached peers. For example, Ashraf (2022) finds that following a breach, companies employ cybersecurity experts on their top management team to enhance their governance over cyber risk. Moreover, IT and cybersecurity processes are not transparent, and firms may rely on private channels to learn about other firms' practices. One way to indirectly facilitate this communication could be through board interlocks. As indicated by prior research, board interlocks can create knowledge spillovers between connected firms (Cheng et al., 2019). Furthermore, communication between the expert outside board members and interlocked members can lead to knowledge spillovers regarding management of and actions to minimize security risks (Cheng et al., 2021). Also, interlocked directors can transfer corporate practices, behavior, experiences, and knowledge within their network of connected firms (Chiu et al., 2013; Stuart & Yim, 2010).

Although positive effects have been noted in association with board interlocks, studies have also indicated the downside of interlocks. Interlocked directors might poorly monitor the firm, and their independence might be compromised, thus exposing the firm to certain risks (Fich & Shivdasani, 2006; Larcker et al., 2005). Further, interlocking to firms that commit financial reporting fraud might expose non-accused interlocked firms to reputational penalties (Kang, 2008). Prior research indicates that earning management practices are more pronounced for firms interlocked to others undertaking these practices (Chiu et al., 2013). Also, firms have a higher likelihood of backdating stock options if interlocked to firms that have previously done so (Bizjak et al., 2009). Overall, the evidence suggests that poor practices may spread through networks of connected firms, as interlocked directors can learn and transmit these practices.

While the literature indicates that interlocking can have both positive and negative effects, its specific impact in the context of data breaches remains unclear. On one hand, learning may accrue, and directors will be better able to address a firm's weaknesses and protect it from facing the same consequences of the breached firm. On the other hand, even after experiencing a breach, directors might not understand or engage enough and may not be able to benefit the focal firms. The economic and reputational consequences of data breaches are severe and understanding breach-related spillovers is important.

Our main results suggest that being connected through interlocked directors, to breached firms or firms that will experience a breach, exposes the focal firms to higher likelihoods of experiencing a breach. This suggests that interlocks are positively associated with the focal firm's risk of data breaches. Furthermore, we find that our results are more significant when the breaches are internal rather than external.[1] Previous studies have distinguished between two

[1] Following prior studies, breaches were identified as internal or external based on their description. Physical loss, unintended disclosure, payment card fraud, insiders, and unknown are classified as internal breaches. Portable device theft, hacking, and stationary theft are classified as external breaches (Higgs et al., 2016; Islam et al., 2022).

categories of data breaches: internal and external (e.g., Feng & Wang, 2019; Higgs et al., 2016; Islam et al., 2022) and have emphasized the need for further research on the various types of breaches firms encounter (Hartmann & Carmenate, 2021). This distinction is crucial for firms to understand the factors that attract or deter external attackers and those associated with internal data security issues. Even though we acknowledge that we cannot provide conclusive causal evidence on the effect of board interlocks on data breach likelihood, the extensive set of analyses that we conduct offers reassurance that the association between interlocking and data breach risk is unlikely to be spurious.

We examine whether interlocked board member attributes have an impact on the likelihood of a breach. We find that the presence of a female director with prior or future breach experience on the board of a focal firm, reduces the likelihood that the focal firm will experience a breach. This is consistent with prior literature that shows that women are more risk-averse than men (Croson & Gneezy, 2009) and enhance boards' monitoring (Adams & Ferreira, 2009) and effectiveness (Nielsen & Huse, 2010). Additionally, we find that when the interlocking director is a member of the audit committee of the focal firm, the focal firm's likelihood of a data breach decreases. While audit committee members are responsible for ensuring that the firm complies with public and regulatory expectations concerning the protection of confidential data (Lanz, 2014), our results provide evidence that audit committees have an impact on data security. Further, the presence of at least one interlocked breach-experienced executive decreases the breach likelihood, which aligns with SOX requirements for executives and boards' increased attention to governance over IT controls (Li et al., 2007). Overall, we document that director attributes are important in mitigating the spread of data breaches and provide preliminary evidence that these director attributes lead to a lower likelihood that the focal firm will experience a breach.

Moreover, our results indicate that firms with a greater number of board interlocks have a higher likelihood of experiencing a breach, consistent with the notion that spillovers are more likely as inter-firm connections increase (Kang, 2008). Furthermore, as data security breaches can impact an entire supply-chain (Rajagopal, 2019; Symantec, 2019), we test and find that a firm's breach likelihood rises in the presence of a breached supplier. To address alternative explanations for our findings, we assess whether interlocked directors' intrinsic quality—rather than their breach experience—drives the observed effects. We measure director quality based on tenure, age, and IT expertise. Prior IT expertise, and tenure and age greater than the median reflect a higher quality (Fairchild & Li, 2005; Hartmann & Carmenate, 2021; Poletti-Hughes & Martinez Garcia, 2022). Our results show that boards with low-quality interlocked directors face a higher breach likelihood than those with high-quality directors. Additionally, we employ a propensity score matching technique to address potential selection bias, confirming that our results hold and are not merely a consequence of a firm's deliberate choice to interlock with a breached firm. Overall, interlocking in the context of data breaches is detrimental; however, we provide a more nuanced view of when and how interlocks spread and mitigate breaches.

Our study makes several contributions to the literature. First, we contribute to the corporate governance literature. Prior research questions whether boards of directors take data security seriously (Rajgopal & Srinivasan, 2016). Certain studies provide unclear evidence (e.g., Lawrence et al., 2018; Richardson et al., 2019), while Ashraf (2022) proves that, in the context of peer breaches, boards take cyber risk seriously. In contrast, looking at connected firms in the *pre* and *post* breach periods, we provide evidence that board interlocks alone do not seem to protect firms from experiencing a breach. This is further supported by board busyness, which limits their engagement in data security matters. Second, we contribute to the empirical literature on spillover effects, and we examine risk spillovers across firms. We provide empirical evidence that data security risks negatively spill over across interlocked firms. Third,

we contribute to the data security literature by examining whether and how the effect of interlocks varies by breach type (internal vs. external). We find that in the setting of interlocked firms, an internal breach to one firm, increases the breach likelihood of a focal firm by a greater extent compared to an external breach. Finally, our study contributes to the social capital literature by showing another "dark side" of social connections. Overall, given that breaches represent an increasing economy-wide risk, our study is relevant and timely and should be of interest to both academics and practitioners.

Our study proceeds as follows. In section 2, we present background on the data breaches, spillovers, board interlocks, data security, and social capital-related literatures and, and we position our study through our hypotheses' development. In section 3, we present the sample selection and the empirical research design. We present our analysis in section 4. Finally, we provide concluding remarks in section 5.

2.1. Literature Review

2.1.1. Data Breaches

Several studies have investigated the impact of data breaches on breached firms. For example, negative abnormal returns have been noted following a breach disclosure (e.g., Amir et al., 2018; Goel & Shawky, 2009; Gordon et al., 2011). Further, while extreme data breaches can cause catastrophic losses, the average breach is associated with negative market reactions and limited economic losses (Richardson et al., 2019). On average, the total cost of a single breached record has been estimated at $225, which includes the costs of lost customers and forgone business opportunities resulting from negative reputation effects after reporting the breach to the victims (Ponemon Institute, 2017). The decrease in a breached firm's sales has been estimated at 33%, which is attributed to either a complete loss of customers or a decline in customer purchases (Janakiraman et al., 2018). Cyberattacks disrupt business operations,

expose firm records, data, and assets, and eventually lead to losses of customers, revenues, and business opportunities (CISCO, 2017; SEC, 2024).

Data protection initiatives have become a priority for almost every company and the increased use of information technology is accelerating the prevalence of data breach incidents (Walton et al., 2021). These attacks have attracted practitioner, academic, regulator, and media attention (Ettredge et al., 2018). As academics have examined this area, five key cybersecurity topics have dominated the research landscape: cybersecurity risk disclosure; information security governance; cybersecurity investment; customer, auditor, and manager responses to data security breaches; and the market's reaction to cyberattacks, and spillover effects (Walton et al., 2021). As the effects of a data breach are not isolated to the breached firm, we seek to determine whether they would spillover from breached firms to connected focal firms and in so doing add to the information security governance and spillover streams of literature.

2.1.2. Spillover Effects

Prior studies have documented different aspects of data breach spillover effects. Garg (2020) investigates the effect of cyberattacks on cash holdings of affected firms. The results indicate that not only breached firms increase cash holdings following a breach, but also focal peer firms, suppliers of the affected firms' big clients, and firms with geographical proximity to the attacked firms' headquarters (Garg, 2020). Islam et al. (2022) find that following the announcement of cybersecurity breaches at breached firms, the expectations of rival firms' investors are affected. The latter is reflected by higher abnormal trading volume in these firms. This spillover effect is mitigated by the presence of chief information officers at competing firms (Islam et al., 2022). Ashraf (2022) finds that following a data breach, focal peer firms employ a cybersecurity expert on their top management team. This enhances their governance over cyber risk and eventually reduces future internal control material weaknesses (Ashraf,

2022). Although spillover effects in the context of data breaches have been studied from different angles, they have not been addressed in the context of board interlocks.

2.1.3. Board Interlocks and Data Security

The main objective of our study is to determine whether interlocking affects the risk of data breaches spilling over from breached to focal firms. As board members oversee the management of cybersecurity (Bonime-Blanc, 2017; Héroux & Fortin, 2022), they are held accountable for cyber incidents and any related legal implications that might follow (Von Solms & Von Solms, 2018). Vincent et al. (2019) identify competent IT monitors as having either IT expertise or IT experience. IT expertise is gained through education and training, enabling the board member to better guide and supervise IT risks (Vincent et al., 2019), while IT experience is gained through exposure to IT-related operations and enhanced through exposure to IT-related failures such as data breaches (Benaroch & Chernobai, 2017). Of course, board members with first-hand experience are better suited for involvement in IT surveillance (Jewer & McKay, 2012; Yayla & Hu, 2014), better understand the extent and significance of data breaches (Ashraf et al., 2020), and are better able to advise on data-related risks. However, prior studies on information security governance indicate that board members generally lack necessary IT expertise and experience (Ashraf et al., 2020; Héroux & Fortin, 2022).

If board members lack the necessary technical/cybersecurity experience (Aguilar, 2014), an important source of information for them could be the expertise and experience of other board members at their other directorships (Cai et al., 2014), especially those who have experienced breaches in prior periods. Communication with IT-experienced board members helps firms lower their IT-related risks (Cheng et al., 2021; Karpovich, 2002). Bizjak et al. (2009) indicate that one important mechanism that could facilitate the communication and flow of information across firms is shared directors (i.e., interlocking boards). We focus on board

interlocks for three reasons. First, board members have access to monitoring and decision-making information, beyond what is directly available, within their networks (Cheng et al., 2019). Second, board members oversee IT and cybersecurity risk management (Bonime-Blanc, 2017; Héroux & Fortin, 2022) and play a role in the firm's governance (Hauser, 2018). Third, and most importantly, it is not clear what impact, if any, board interlocks would have on firm data security.

2.1.4. Board Interlocks and Social Capital

Research based on social capital theory offers significant insights into how board interlocks affect firms (e.g. Kim & Cannella, 2008; Kor & Sundaramurthy, 2009). Shared directors form a social network among firms, where the governance practices of these firms influence each other, like how behaviors in other social networks are shaped by their members (Bertrand et al., 2000). A social network is the channel through which social capital is formed, maintained, and utilized (Javakhadze et al., 2016). Mainstream social capital theory emphasizes the positive effect of social connections (Lin, 2002), as supported by prior studies (e.g., Bianchi, 2018; Intintoli et al., 2018). However, 'sociability cuts both ways' (Portes, 1998, p.18), and while there is a bright side to it, a dark side also exists (Bruynseels & Cardinaels, 2014; Carrera et al., 2017; He et al., 2017; Van Deth & Zmerli, 2010). Among the most studied types of social networks, board interlocks are associated with positive and negative spillover effects (Bizjak et al., 2009; Kang, 2008). As our main objective is to determine whether interlocking affects interlocked firms' breach likelihood, we rely on social capital theory to assess whether the bright side or the dark side of social connections in the form of board interlocks would prevail.

A network perspective can be applied for teams, individuals, and firms, enabling the examination of almost any social system defined by connections between entities (Borgatti & Ofem, 2010). The network lens acknowledges the web of relationships that provides

opportunities and imposes constraints on actors (Borgatti & Ofem, 2010), affecting their economic outcomes (Granovetter, 2005). Network theorists examine the relationships that firms have with other firms, such as the relationship through common directors. This relation could be described as a "flow" which reflects the transmission of tangible and intangible elements through interactions (Borgatti & Ofem, 2010). Flows are not measured, instead they are inferred from relational and interactional data (Borgatti & Ofem, 2010).

Social capital is defined as the benefit derived from an individual's position within a network of relationships (Burt, 2005), and it is positively related to the diversity of information available within the network and the transfer of knowledge among the members (e.g. Anderson, 2008; Reagans & McEvily, 2003). An individual's social capital reflects the existing and potential resources embedded in and obtained through their network of relationships (Nahapiet & Ghoshal, 1998).

Board interlocks create valuable connections between firms, fostering the transfer of information, resources, knowledge, experiences, behaviors, and organizational practices within the network (Brown et al., 2019; Chiu et al., 2013; Omer et al., 2020; Stuart & Yim, 2010). This transfer reduces information acquisition costs (Caiazza et al., 2019). Haunschild (1993) found that corporate managers imitate acquisition activities performed by firms connected to them through shared board members. Likewise, Gulati and Westphal (1999) determined that the likelihood of forming strategic alliances is influenced by relationships between outside directors and CEOs. According to Srinivasan (2005), directors serving on multiple boards have incentives to work well in each assigned position, and this is rewarded in the labor market.

Firms with inside directors appointed as outside directors for other firms exhibit higher market-to-book ratios, better operating performance, more cash holdings, less earnings management, and improved acquisition decisions compared to peers without such board

appointments (Masulis & Mobbs, 2011). Multiple directorships may enhance a director's reputation as an indicator of the board member's abilities (Fama & Jensen, 1983). Additionally, busy boards[2] have been associated with superior firm value for listed firms and for firms requiring major advising (Field et al., 2013; Loderer & Peyer, 2002). Moreover, interlocked boards equipped with cybersecurity risk information can assess managers' remediation strategies, force resource allocation to address unresolved concerns, and, in cases of noncompliance, employ disciplining tools (Ashraf et al., 2020).

Although the mentioned studies largely indicate positive effects of board interlocks, social connections are not always advantageous and, as mentioned above, can have a dark side. Interlocking could compromise board members' keen attention to management, weakening their monitoring abilities and exposing the firm to avoidable risks (Fich & Shivdasani, 2006). For example, Dhaliwal et al. (2010) found that multiple social connections degrade directors' monitoring abilities, leading to lower accruals quality. Additionally, reputational penalties could spill over from firms that commit financial reporting fraud to non-accused interlocked firms (Kang, 2008).

Financial reporting behaviors appear to spread across board-connected firms, as demonstrated by Chiu et al. (2013), who found that earnings management is more pronounced for firms that share a director with another firm undertaking such practices. Board interlocking has also facilitated option backdating, whereby firms connected by a board member to another firm that has previously backdated stock options, have a higher probability of following the same path (Bizjak et al., 2009). Carrera et al. (2017) found a negative relationship between audit committee members' financial reporting quality and their social capital. Besides,

[2] "[B]usy boards [are] defined as those in which a majority of independent directors hold three or more directorships" (Field et al., 2013, p. 63).

malevolent cooperations, including criminal groups might stem from social connections (Harraka, 2002).

Studies investigating the impact of multiple directorships on firm performance indicate that "busy" directors lack sufficient time to dedicate to their responsibilities at each firm, consequently harming firm performance (Ahn et al., 2010; Brown et al., 2019; Cashman et al., 2012; Falato et al., 2014; Fich & Shivdasani, 2006; Hauser 2018). The Non-Executive Directors Association recommends allocating a day and a half per week to each board seat (Brown et al., 2019). Therefore, it comes as no surprise that over-boarded directors might be inattentive and ineffective (Brown et al., 2019). Busy boards have been associated with lower market-to-book ratios and profitability (Fich & Shivdasani, 2006), and weaker compensation monitoring (Core et al., 1999). Also, busy boards make poor acquisition decisions as evidenced by lower M&A announcement returns (Ahn et al., 2010). Over-committed boards may be too busy to effectively advise and monitor the firm (Cashman et al., 2012; Erel et al., 2021). Also, they are more likely to skip board meetings (Jiraporn et al., 2009). The presence of multiple interlocked directors on a board may therefore negatively affect firms, increasing their risk of data breaches. As identified through this discussion, whether board interlocks have a positive or a negative impact in the context of data breaches is yet to be determined.

2.2. Hypotheses Development

While prior research has examined spillover effects of interlocked boards from different angles, it has not addressed whether breached firms impact the connected focal firms' likelihood of experiencing a breach. This study aims to investigate whether board interlocking reveals a bright side benefitting connected focal firms by lowering their likelihood of experiencing a breach, or whether it reveals a dark side putting them at a higher risk of experiencing a breach.

As indicated in our literature review above, some studies show that interlocked boards lead to positive spillovers among firms through improved monitoring (Zhong et al., 2017), helping firms avoid data breaches or mitigate the effects when avoidance is not an option (Ashraf et al., 2020). However, other studies indicate negative spillover effects. More specifically, busy directors might not be able to pay enough attention, providing weak advising and monitoring (Brown et al., 2019; Cashman et al., 2012; Erel et al., 2021). This tension in the literature is interesting and we seek to investigate how board interlocks may impact data security risk, as measured through reported data breaches.

Prior literature points to two possible impacts to data security risk that may result from interlocked boards. For directors who have experienced breaches at their other firms, learning could accrue (Ahmad et al., 2020). These members' confidence in addressing such issues would rise to the extent that they would promote cyber-related debates within the boardroom (Schinagl & Shahim, 2020). As a next step, these directors could then disseminate information from their experience with breach events to their other firms within the network (Gale et al., 2022). More information has the potential to decrease uncertainty, and connected boards' first-hand experience may decrease ambiguity (Cai et al., 2014). The focal firms would then be better protected against breaches or better able to mitigate the consequences in case the breach could not be avoided (Ashraf et al., 2020). Hence, firms' risk of experiencing a breach would be lower when connected, through board members, to firms that experience breaches. In this case, social ties represent a medium through which knowledge is exchanged, enhancing individuals' skills and competencies, and in turn benefiting the firms (Burt, 1997).

Drawing on social capital theory, we determine a mechanism through which interlocked directors are linked to data security. Information flows within social networks may enhance interlocked directors' monitoring effectiveness. Social capital can function as a disciplinary

mechanism, incentivizing individuals operating within networks to maintain a good reputation by adhering to norms of appropriate behavior (Burt, 2005; Javakhadze et al., 2016).

Alternatively, a high likelihood of experiencing a breach may arise if the learning processes within the organization are ineffective or insufficient to safeguard against emerging breaches. This inadequacy could stem from a lack of updated knowledge or an inability to adapt to evolving data security threats. Furthermore, if board members are not strategically positioned to share critical information, or if they are too occupied with other responsibilities to effectively disseminate acquired knowledge, the potential negative consequences of inter-firm social ties, or the 'dark side,' could manifest. In such cases, even if information is available, its transmission would be hindered. Moreover, when board members lack a comprehensive understanding of the technical aspects of IT and data security, they may struggle to communicate this information accurately and efficiently, further impeding the firm's ability to mitigate security risks.

Boards often hesitate to engage in cybersecurity discussions. If directors' knowledge in the area is limited, they may remain silent to protect their ego; if their knowledge is not limited, they may also remain silent and rely on other knowledgeable colleagues to handle data security matters, protecting themselves from personal liability (Gale et al., 2022). In Gale et al. (2022), half of the interviewed directors admitted that because of its mysterious nature and difficulty of alleviation, cybersecurity is not placed on the board's agenda unless a cyberattack takes place.

Additionally, research indicates that even when placed on the board's agenda, cybersecurity remains on the periphery and is solely a topic "of interest" rather than an actionable one (Aguilar, 2014). More often, it is discussed among board members either during board meetings or via emails. The level of superficiality at which the information is shared reflects

that, at the board level, there is no awareness about what is really happening, and everyone assumes that it will be handled by the head of the IT department (Mishra, 2015). Another explanation for a potential increased breach likelihood is that busy directors might be distracted by numerous responsibilities (Cashman et al., 2012) to the extent they neglect data protection matters.

As we consider board interlocks, timing is also crucial in defining the terms of our investigation. We differentiate between interlocks that took place before the breach occurred (i.e., *Pre*), and after the breach occurred (i.e., *Post*). For firms that were interlocked to firms that experienced a breach at any point in the three years following the interlock, the breach and its consequences could be described as "private" information (Cheng et al., 2019). It is important to recognize that there are risks to interlocking to a firm that has data security weaknesses that are yet to be acknowledged by the public. A board member's busyness may shift the director toward neglecting matters of data security, because the occurrence of an upcoming breach, and its associated costs, are unknown. Conversely, networks diffuse knowledge between connected firms, where interlocked directors might better monitor and advise the firms. Testing interlocks prior to connected firms' breach experience, we will determine whether focal firms are "also" at risk of experiencing a breach. To examine this likelihood, we present the following hypothesis in its null form.

H1a: The likelihood that a focal firm will report a data breach is not associated with interlocking to a firm that will report a future breach.

For firms that are interlocked to breached firms in any of the three years after the breached firms experienced their breach, the breach could be described as "public" information (Cheng et al., 2019). Like the point made in our discussion above, interlocking to a firm that has data security weaknesses known to the public may have a different impact and should be tested

separately. On one hand, economically rational directors would be more attentive following a breach of another firm within the network, now that the breach's consequences are known. Further, with the breach being public, directors might be more attentive to maintain their reputation. On the other hand, even with the known consequences, busy directors might be constrained by time and overwhelmed with responsibilities, that they end up neglecting data matters at other connected firms. As such, we test whether focal connected firms take proactive actions to evade data breaches. To examine this likelihood, we present the following hypothesis in its null form.

H1b: The likelihood that a focal firm will report a data breach is not associated with interlocking to a firm that has reported a breach in the past.

By comparing the breach likelihood of firms connected before the breach *(Pre)* to firms connected after the breach *(Post)*, we will be able to identify whether board members take any action to shield against breaches when they are aware of them. If the breach likelihood is high when the information is private *(Pre)* and when it is public *(Post)*, we can say that even when the risk is known, board members are not adequately using the data security information to protect their firms, which can be attributed to inefficiencies within their networks or their over-boarded schedules. This can be interpreted as firms being at risk of experiencing a breach regardless of whether the interlock took place before or after the connected firm's breach.

Important distinctions between breaches exist and should be considered. Specifically, the literature has pointed to two categories, internal and external (Islam et al., 2022). Previous studies differentiating between these categories found that firms that have risk and compliance committees are more likely to experience internal breaches, whereas firms with board-level technology committees are more likely to experience external breaches (Higgs et al., 2016). The risk aversion level of Chief Information Officers (CIOs) has been negatively linked to the

probability of experiencing breach incidents, especially when the breach is internal (Feng & Wang, 2019). In an experiment conducted to understand management's responsibility acceptance, a higher degree of acceptance was a better strategy when an external breach occurred (Tan & Yu, 2018). Studying the effect of focal firms' breach announcements on rival firms' abnormal trading volume, Islam et al. (2022) find the same significant positive effect regardless of breach type.

Research also indicates that firms that disclose the presence of a CIO have a higher likelihood of experiencing both external and internal breaches (Smith et al., 2021). However, CIO structural capital attributes (i.e., recognized commitment to support IT and multitasking) and human characteristics (i.e., past technology experience, external board member, CIO tenure and firm tenure) are associated with internal breaches (Smith et al., 2021). Importantly, recent research by Hartmann and Carmenate (2021) has called for more research on the different types of breaches that firms might experience.

Given the significant implications behind breach type in this study, we also investigate whether the likelihood of experiencing a breach differs based on the breach type. Following prior studies (Higgs et al., 2016; Islam et al., 2022), we classify physical loss, unintended disclosure, payment card fraud, insiders, and unknown as internal breaches, and portable device theft, hacking, and stationary theft as external breaches. As with our prior examination, tension also exists here. On the one hand, it could be possible that focal firms connected via board members to externally breached firms will be more attractive to hackers, who might find it more convenient and easier to attack given that knowledge resources are shared with the focal firm. On the other hand, it could be possible that focal firms interlocked via board members with internally breached firms reflect weaknesses in the latter's internal processes, that can thus transfer to the connected firms through the social networks. We posit the below hypothesis in the null form:

H2: The likelihood of experiencing a data breach does not differ with the type of breach reported by the breached firm.

3. Research Methods

3.1. Data and Sample Selection

Our study sample is composed of breached and focal U.S. public firms over the period 2007 through 2021. We focus on the US due to its significant influence on corporate governance globally (Beattie et al., 2012). Following prior research, we rely on the Privacy Rights Clearinghouse (PRC)[3] database to extract breach-related data. PRC has been widely used in many reputable research studies (e.g., Ashraf, 2022; Higgs et al., 2016; Kamiya et al., 2018; Li et al., 2018; Rosati et al., 2017). PRC reports data on breaches from different sources that include regulatory bodies (e.g., the Federal Bureau of Investigation and attorney general offices), verifiable public media sources, customers, and state governments (Say & Vasudeva, 2020). Some breach incidents might not be included in the PRC database because firms are either unaware of their occurrence or are not required to disclose them based on different reporting laws (PRC, 2017).

We exclude breaches at government agencies and NGOs due to data limitations (Ashraf, 2022; Kamiya et al., 2018). To supplement our set of breach observations, we obtain board-related data from the Institutional Shareholder Services (ISS) database, and financial data from Compustat. Next, to ensure accuracy of the data, manual completion, adjustments and matching of breach incidents to firm-year observations from Compustat were conducted. Details of our final sample are included in Table 1. Our final sample includes 18,731 firm-year observations (Panel B) where 651 breaches (Panel A) and 2,079 firms are represented.

[3] We compared the data from PRC with that of Audit Analytics. Results showed an almost complete similarity.

[Insert Table 1 here]

3.2. Research Design

Based on the hypotheses presented above, we examine the effect a firm may experience when it is interlocked with a firm that will report a future breach or with a firm that has previously reported a breach. More specifically, we examine the connected focal firm's likelihood of experiencing a breach. To do so, we use a logit model, the most suitable for time-to-event settings (Cheng et al., 2019; Chiu et al., 2013; Higgs et al., 2016). The model we specify for our main analysis is presented below.

$$Logit\ (Breached_{i,t}=1) = \beta_0 + \beta_1\ Interlocked\ Breached\ Firm\ Pre_{i,t} + \beta_2\ Interlocked\ Breached\ Firm\ Post_{i,t} + \beta_3\ Firm\ Size_{i,t-1} + \beta_4\ Leverage_{i,t-1} + \beta_5\ Loss_{i,t-1} + \beta_6\ High\ Tech_{i,t} + \beta_7\ Board\ Size_{i,t} + \beta_8\ Foreign_{i,t} + \beta_9\ Merger_{i,t} + \beta_{10}\ Industry + \beta_{11}\ Year + \varepsilon$$

The dependent variable *(Breached)* is an indicator variable that takes the value of 1 if a firm has reported at least one breach in a given year t. The main independent variables are *Interlocked Breached Firm Pre* and *Interlocked Breached Firm Post*, both of which are dummy variables. *Interlocked Breached Firm Pre* is set equal to 1 if the focal firm is interlocked, through a shared director, with another firm that will report a breach at any point in the next three years. *Interlocked Breached Firm Post* is set equal to 1 if the focal firm is interlocked, through a shared director, with another firm that has already reported a breach at any point in the previous three years.

If the coefficient on *Interlocked Breached Firm Pre* is positive, it would indicate that focal interlocked firms are associated with a high risk of experiencing a breach through their connection to firms that are about to experience a breach. Based on hypothesis H1a, this would lead us to conclude that there is an increased likelihood that a focal firm will report a data breach when interlocked with a firm that will report a future breach. Alternatively, a negative

coefficient would indicate that focal interlocked firms have a lower likelihood of experiencing a breach in the future. If the coefficient on *Interlocked Breached Firm Post* is positive, this would indicate that interlocked firms are associated with a high risk of facing data breaches through their connection to firms that have already experienced a breach. Based on hypothesis H1b, we would then conclude that there is an increased breach likelihood for focal firms when interlocked with a firm that has reported a breach in the past. However, if the coefficient is negative, it would indicate that focal firms have a lower likelihood of experiencing a breach when connected to firms that have already experienced a breach. The *Interlocked Breached Firm Pre* and *Post* variables are coded 1 only for the focal interlocked firms to capture the spillover effect from breached to focal firms. In other words, breached firms are coded 0.

To test H2, we replace the dependent variable *Breached* with *Internally Breached* and *Externally Breached* to determine whether interlocked firms' likelihood of experiencing a data breach differs with the type of breach reported by the breached firm. If the coefficients on *Interlocked Breached Firm Pre* and *Interlocked Breached Firm Post* are different, we would conclude that breach likelihood differs based on breach type.

The two main independent variables are developed at the firm-level by checking whether there is an interlocked director on the board of directors (hereafter, BOD) of the focal firm that also served on the BOD of a breached firm in the three-year period leading up to the breach (*Pre*), or in the three-year period following the breach (*Post*). Focal firms with at least one interlocked director that has experienced a breach, in a given year, are coded 1. For example, assume that Firm X disclosed a breach in 2015. Also, assume that Firm Y has a board member that simultaneously served on the board of Firm X during at least one of the following years: 2012, 2013, 2014, 2016, 2017, and 2018. In this situation, Firm Y is the focal firm and the *Interlocked Breached Firm Post* variable for that firm is coded 1 in year(s) 2016, 2017, and

2018, and 0 for all other years. Further, *Interlocked Breached Firm Pre* variable for that focal firm is coded 1 in year(s) 2012, 2013, and 2014, and 0 for all other years.

Prior studies show that several firm and board-related characteristics correlate with the likelihood of firms experiencing a breach (Higgs et al., 2016; Hsu & Wang, 2014; Kobelsky et al., 2008; Li et al., 2018; Say & Vasudeva, 2020). We utilize the prior literature to identify relevant control variables for our model. First, given that large firms are more attractive to attackers than small firms, we control for firm size as reflected through total assets (Higgs et al., 2016; Li et al., 2018; Say & Vasudeva, 2020). It has also been identified that financially constrained firms have a higher risk of experiencing a breach due to the limited resources for adequate IT security investments. To control for financial constraints, we utilize variables that measure firm leverage and whether a firm has reported a loss (Higgs et al., 2016; Li et al., 2018; Say & Vasudeva, 2020). We also control for whether a firm is considered high-tech since these firms generally have higher IT budgets and this may be linked to lower breach likelihoods (Kobelsky et al., 2008). Since, larger boards have been associated with lower risks of information security breaches, we control for board size as reflected by the number of directors on the board (Hsu & Wang, 2014). Variables that represent foreign operations and merger activities are included to control for business complexity. Dispersed and complex operations may lead to inconsistent and ineffective controls resulting in higher breach likelihoods (Li et al., 2018).

To reduce potential endogeneity concerns (i.e., reverse causality), firm size, leverage and loss are lagged by one year (Bouwman, 2011). Finally, we control for year and industry effects. Thriving industries present good opportunities for attackers and thus are associated with higher risks of firms experiencing data breaches (Say & Vasudeva, 2020). Variables descriptions are presented in Appendix A.

4. Results

4.1. Summary Statistics

To begin our analysis, we examine the distribution and types of breaches reported within our sample. This univariate analysis is presented in Figure 1 for our sample period of 2007 through 2021. Breaches caused by hacking are on average the highest over the studied period. In our sample, the highest number of breaches (113) occurred in 2019[4]. The total number of internal breaches (382) is greater than that of external breaches (269).

[Insert Figure 1 here]

Table 2 presents information on our sample composition. As indicated in Panel A, the number of breached firms in our sample is 548. This number is less than the number of breaches (651) because some firms report more than one breach. In Panel B, we indicate that out of the 548 breached firms, 502 are interlocked. Finally, in Panel C we present the number of breach incidents per industry. As indicated in Panel C, firms in the Finance and Insurance industry appear to be the most vulnerable to data breaches, occupying 30.66% of the total breached firms. This industry is followed by manufacturing (16.61%) and information (11.13%).

[Insert Table 2 here]

Table 3 presents descriptive statistics for our main sample. These statistics indicate that 2.9% of firm-year observations have reported data breaches. Additionally, 24.3% of firm-year observations are interlocked to breached firms in the 3 years prior to the breach occurrence, while 25.7% are interlocked to breached firms in the 3 years after the breach occurrence. High tech firms make up a significant proportion of the total sample (33%). The statistics indicate that our sample firms have an average leverage ratio of 0.57 and 12.5% of firms exhibit reported

[4] In 2019, T-Mobile experienced a data breach that affected over 1 million of its customers.

losses. On average, the natural logarithm of firm size is 8.212 million dollars, which corresponds to an actual value of 3.685 million dollars in total assets, and the board includes 9 members. In terms of firm complexity, around 31% had foreign operations and 33% engaged in merger activities.

[Insert Table 3 here]

Table 4 displays the pairwise correlations among the dependent and main independent variables. Statistically significant (at the 1% level) and positive correlations are observed between *Breached firm* and *Interlocked breached firm pre* and *Interlocked breached firm post*. This provides a preliminary indication of the relationship between the dependent variable and the main independent variables. Consistent with prior studies, we observe a statistically significant and positive relationship between *Firm Size*, *Leverage,* and the dependent variable, and a negative relationship between *Foreign* and the dependent variable (Higgs et al., 2016; Li et al., 2018; Say & Vasudeva, 2020).

Our results for *Loss*, *High tech* (negative and significant), and *Board Size* (positive and significant) are inconsistent with those of prior studies (Higgs et al., 2016; Hsu & Wang, 2014; Kobelsky et al., 2008; Li et al., 2018; Say & Vasudeva, 2020). Firms that report losses might not be an attractive target for attackers, and firms in the high-tech industry might be seen as difficult targets, thus explaining the negative association with the dependent variable. As for board size, we can say that members in bigger boards may rely on other colleagues to handle data security matters to the extent no one ends up taking responsibility, leading to more breaches and hence the positive relationship with the dependent variable.

[Insert Table 4 here]

4.2. Main findings

Table 5 Panel A presents the results when we apply our model to the data. Consistent with our univariate tests, we estimate a statistically significant and positive coefficient (coefficient= 0.803, p-value < 0.01) on the *Interlocked Breached Firm Pre* variable, suggesting that interlocked firms have 2.23[5] times higher odds than non-interlocked firms of experiencing a breach when connected to a firm that will experience a breach in any of the upcoming three years. Specifically, the odds of a focal firm experiencing a data breach are approximately 123 percent[6] higher if the firm is interlocked to a firm that will experience a breach in the future. An average marginal effect (AME) of 0.021 (Panel B) shows that the predicted probability of a firm experiencing a breach is higher by 0.021 for firms interlocked to a firm that will experience a breach compared to non-interlocked firms. Further, interlocked firms have a statistically significant and positive coefficient (coefficient= 0.620, p-value < 0.01) on the *Interlocked Breached Firm Post* variable, suggesting that firms interlocked to a breached firm in any of the three years after the breach occurrence have 1.86[7] times higher odds of experiencing a breach than non-interlocked firms. This could be interpreted as the odds of a focal firm experiencing a data breach are approximately 86 percent[8] higher if the firm is interlocked post-breach to a firm that experienced a breach in the past. The AME in Panel B reflects that the predicted probability of a firm experiencing a breach is higher by 0.016 for firms interlocked to firms that experienced a breach, compared to non-interlocked firms. Taken together, these results provide evidence that leads us to reject H1a and provide preliminary evidence that interlocking is associated with a significant negative spillover effect on connected focal firms by putting them at a higher risk of experiencing data breaches in the future. Additionally, these results provide compelling evidence that focal firms connected to

[5] Calculated as: e^0.803.
[6] Calculated as: (e^0.803 - 1) x 100 ≈ 123.22%.
[7] Calculated as: e^0.620.
[8] Calculated as: (e^0.620 - 1) x 100 ≈ 85.89%.

previously breached firms are exposed to a negative spillover effect, and this evidence leads us to reject H1b.

The results in Table 5 also indicate that operating in a high-tech industry, reporting losses, board size, or merger involvement are not significantly linked to the likelihood of experiencing a breach within our sample. However, our results do indicate that bigger firms are more likely to experience a data breach (coefficient=0.334, p-value<0.01), whereas more leveraged firms (coefficient = -0.444, p-value<0.1) and firms that engaged in foreign operations (coefficient = -0.295, p-value<0.05) are less likely to experience a data breach. Overall, our results indicate that interlocking to breached firms does not reduce the likelihood that a focal firm will experience a breach. Rather, it seems that instead of experience and knowledge transferring from exposure to data breaches, the risk of breach is being spilled over among connected firms. This is evident when the breached firm's disclosure is still private (i.e., *Pre*) as well as when it is already public (i.e., *Post*). As such, we claim that there is a very strong association between the likelihood that a focal firm will report a data breach when interlocked to a firm that has already reported a data breach or that will report a breach in the future.

[Insert Table 5 here]

Table 6 presents the results for our test of H2. Internal breaches are defined as those stemming from within the firm and include instances of fraud involving debit and credit cards, unintended disclosures, insider malevolent actions by employees, contractors, or customers, and losses of physical documentation. External breaches originate from outside sources and include hacking, malware infections, loss of portable devices, and loss of stationary computers. As mentioned above, prior research has highlighted the differing consequences of internal versus external breaches, motivating our investigation into this crucial distinction. To compare the coefficients of our main independent variables (*Interlocked Breached Firm Pre* and

Interlocked Breached Firm Post) across the logit regression with two dependent variables, we run a joint logit regression regressing the dependent variables on the main independent variables only[9]. Table 6 presents the logit regression output for each of the two categories of dependent variables *(Internally Breached* and *Externally Breached)*. The coefficients of the independent variables differ significantly across the dependent variables (prob>chi2 = 0.0016). Differentiating between the coefficients, *Interlocked Breached Firm Pre* has a coefficient of 1.233 for Internally Breached and 0.928 for Externally Breached. While the coefficient is positive and statistically significant when interlocking to internally breached firms, it is not significant when interlocking to externally breached firms. This indicates that firms interlocked to a firm that will experience an internal breach in any of the upcoming three years have a high likelihood of experiencing a breach compared to firms that are interlocked to another firm that will experience an external breach. The coefficients for *Interlocked Breached Firm Post* for Internally Breached and Externally Breached are 1.229 and 0.760, respectively, and are both statistically significant. This means that firms interlocked to a firm that experienced an internal breach in any of the prior three years have a higher likelihood of experiencing a breach than firms connected to an externally breached firm. The results indicate that interlocking pre-breach to internally breached firms only is associated with a higher likelihood of experiencing a breach, while interlocking post-breach to either internally or externally breached firms is associated with a higher breach likelihood. Overall, the results show that the likelihood of experiencing a data breach differs with the type of breach reported by the breached firm, resulting in the rejection of H2.

[Insert Table 6 here]

4.3. Additional Analyses

[9] Control variables are not included because only dummy variables are allowed in this type of regression.

4.3.1. Female Directors

Prior studies have examined the role of women on the board of directors. Enhanced board effectiveness (Nielsen & Huse, 2010) and monitoring (Adams & Ferreira, 2009) have been identified for BODs with female directors. Moreover, research indicates that boards with gender diversity experience increased levels of financial performance and social and ethical compliance (Francoeur et al., 2008; Isidro & Sobral, 2015). Further explanation for these previous results indicates that women are more involved in corporate social responsibility activities than their male counterparts (Williams, 2003), and this has a positive influence on firm value (Campbell & Mínguez-Vera, 2008). Additionally, the presence of women on boards is associated with higher levels of transparency (Gul et al., 2011; Larkin et al., 2013) and risk-reporting (Bravo, 2018).

Women's role on boards has also been recognized in the cybersecurity context. The percentage of women on boards is positively linked with the presence and degree of cybersecurity disclosure in annual reports (Radu & Smaili, 2021), and with the disclosure level of specific cybersecurity aspects (Héroux & Fortin, 2022). Given the impact of women directors, as an additional analysis, we examine whether the presence of breached interlocked female directors on the board of a focal firm impacts the focal firm's likelihood of experiencing a breach. We combine the pre and post periods for the interlocked breached female variable since the main analysis showed that they both have the same effect on breach likelihood. We present the results of this analysis in Table 7. We find that firms, with at least one interlocked breached female director are less likely to experience a breach, as evidenced by the statistically significant and negative coefficient (coefficient= -3.592, p-value<0.01). Even though board interlocks in general increase the breach likelihood, the presence of an interlocked breach-experienced female director on board decreases the breach likelihood.

[Insert Table 7 here]

4.3.2. Audit Committee Members

Another key area of interest in the interlock literature relates to the audit committee. Prior research indicates that firms connected through directors or members of the audit committee have a lower likelihood of reporting internal control material weaknesses than non-connected firms, firms operating in the same industry, or firms that share the same auditor (Cheng et al., 2019). Cheng et al. (2019) also show that prior directors' experiences outside a firm affect the work of the firm's audit committee. Audit committee members with cybersecurity knowledge and experience carry more weight than their colleagues and are better able to understand the scope of data breaches through their interactions with management (Ashraf et al., 2020). As such, audit committee members are well-positioned to provide advice on matters related to cybersecurity risks (Ashraf et al., 2020). Audit committee membership provides the committee members with the means to perform their corporate governance responsibilities (Peterson & Philpot, 2007). For a long time, audit committee members have been navigating data security risks that firms face (Lanz, 2014). Potential losses push audit committee members to be more involved in breach-related matters and to identify ways to better manage these responsibilities (Lanz, 2014). As such, we aim to see whether audit committee members with breach experience affect the firm's breach likelihood.

We test whether the presence of at least one director that is on the audit committee (AC) of the focal firm and connected to a breached firm affects the likelihood of the focal firm experiencing a breach. We test our main model with this additional independent variable where we combine the interlocked breached AC pre and interlocked breached AC post variables into one variable, since the main analysis showed that they both have the same effect on the breach likelihood. Our results in Table 8 indicate that firms with at least one interlocked breached director on the audit committee are less likely to experience a breach. This is evident in the

negative and statistically significant coefficient on the *Interlocked Breached AC variable* (coefficient= -3.920, p-value<0.01) compared to firms with no breach-experienced interlocked AC member. Our results provide preliminary evidence that indicates that aside from cybersecurity management responsibilities, audit committee members with breach experience aid in diffusing insights to help firms avoid pitfalls that could lead to data breaches.

[Insert Table 8 here]

4.3.3. Executive Directors

The Sarbanes-Oxley Act (SOX) holds executives of public companies directly accountable for creating, assessing, and overseeing the effectiveness of internal controls related to financial reporting and disclosure (Li et al., 2007). With the vital role IT-based systems play in the success of many firms and due to growing regulatory demands, senior management is becoming more and more responsible for ensuring IT control effectiveness (Li et al., 2007). Executive directors include the Chief Executive Officer (CEO), the Chief Information Officer (CIO), the Chief Operating Officer (COO), the Chief Financial Officer (CFO) and other senior business executives, such as Executive Vice Presidents, responsible for key business areas. As SOX requirements have made a significant change with respect to executives and boards' attention to governance over IT controls, we investigate whether the presence of interlocked breached executives is associated with a lower breach likelihood for focal connected firms. As evidenced in Table 9, the coefficient on *Interlocked Breached Executive* is negative and statistically significant (coefficient = -3.184, p-value<0.01), indicating a lower breach likelihood in the presence of breach-experienced executives.

[Insert Table 9 here]

The above tests indicate that female directors, directors on the audit committee, and executive directors are more incentivized than their social network peers to better monitor

firms. They appear to use their networks to transfer information and ensure appropriate behavior and adherence to the norms, showcasing a bright side to social capital.

4.3.4. Number of Interlocks

The number of connections through board interlocks is reflective of information sources accessible to the connected firm (Haunschild & Beckman, 1998). Akbas et al. (2016) find that more connected boards provide firms with a competitive advantage by providing insights that are otherwise inaccessible to less connected boards. Alternatively, more board links to firms undertaking controversial practices (e.g., option backdating) result in connected firms initiating the same practices (Bizjak et al., 2009). As spillovers from breached to associated firms may be more likely to occur with more ties between the accused and non-accused firms (Kang, 2008), we look for whether the contagion effect is positive or negative in our setting (Chiu et al., 2013). In Table 10, we test the effect of the number of interlocks on the breach likelihood. We observe that the coefficient on the *Number of Interlocks* is positive (coefficient = 0.047, p-value<0.1) and statistically significant at the 10% level. This result supports our main findings that interlocking seems bad for firms in the context of data breaches, as shown in the coefficients for *Interlocked Breached Firm Pre* and *Interlocked Breached Firm Post* variables which are positive and statistically significant (coefficient = 0.776 and coefficient = 0.584, respectively). Together, our results provide evidence that the contagion effect is negative when it comes to breach likelihoods, and the more interlocks the higher the likelihood that a firm will experience a breach.

[Insert Table 10 here]

4.3.5. Breaches in the Supply Chain

Cybersecurity breaches have a detrimental impact not only on the affected firm's market value, reputation, and financial stability (Gwebu et al., 2018), but also on the broader supply chain in which the firm operates (Rajagopal, 2019; Symantec, 2019). He et al. (2022) observe that firms within the same supply chain, referred to as targeted firms, respond strategically to supply-chain breaches by lowering transaction costs. They achieve this through increased use of real activities, including managing discretionary cash flows, expenses, and production costs. He et al. (2023) identify a negative relationship between customers' reported cyberattacks and suppliers' innovative investments. Moreover, they find that cybersecurity breaches experienced by customers heighten the likelihood of disruptions in supplier-customer relationships. Given the interconnected nature of firms within a supply chain, where breaches can negatively affect not only the breached firm but also impose spillover costs on stakeholders, including customers and suppliers (Hovav & Gray, 2014), we are interested in examining the impact of breaches across the supply chain. We examine whether the number of breached customers and suppliers influences the likelihood of breaches among firms within the same supply chain. Our findings (Table 11) reveal a positive and significant spillover effect from breached suppliers (coefficient = 2.361, p-value < 0.05), but not from breached customers, indicating that a higher number of breached suppliers within the supply chain is associated with a higher likelihood of firms experiencing a breach. Additionally, we investigate whether the number of suppliers and customers interlocked with breached firms affects the breach likelihood of these connected firms. We do not observe any significant results. Overall, our study provides evidence of negative spillovers from breached suppliers to other firms across the supply chain.

[Insert Table 11 here]

4.4. Robustness Checks

4.4.1. Additional Control Variables

To provide additional robustness to our results, we follow prior studies and control for the number of geographic and business segments in a fiscal year (*Segments*) as a measure of a firm's complexity (Li et al., 2018). Firms with more dispersed and complex operations are expected to have more ineffective controls, thus higher likelihoods of experiencing breaches. We also use book value to market value (*BTM*) as an additional measure of a firm's size (Ettredge et al., 2018). Finally, we control for *Intangibles* as firms with more intangible assets have higher chances of experiencing a breach (Kamiya et al., 2021). We include these additional controls as a robustness check because the main sample is reduced significantly, due to data limitations, when these control variables are included. In applying this modified dataset to our model, we find that *Intangibles* is significant and positive, in alignment with previous research. However, our results indicate no significance with respect to *BTM* and *Segments*. More importantly, inclusion of these additional control variables did not have an impact on our main results. That is, higher likelihoods of experiencing a breach if interlocked pre or post to the connected firm's breach remained consistent[10].

[Insert Table 12 here]

4.4.2. Director Quality

We run additional tests to rule out alternative explanations and illustrate the robustness of our results. Interlocked directors might be intrinsically of higher quality compared to non-interlocked directors, and thus our results might be affected by director quality rather than their experiences (Cheng et al., 2019). For this reason, we test whether the presence of high-quality and low-quality directors on the board affect breach likelihood. First, we create a Director Quality Index (DQI) that includes director's tenure, age, and IT expertise, all dummy variables.

[10] We repeated the main tests without the *High-tech* variable to address any correlation with the industry fixed effect. The results remained the same.

Previous studies determined that these three attributes are characteristics that determine a director's quality (Fairchild & Li, 2005; Hartmann & Carmenate, 2021; Poletti-Hughes & Martinez Garcia, 2022). *IT Expertise* is coded 1 if a director has prior IT expertise. *Tenure* is coded 1 if the director served on the board for a period greater than or equal to the sample median of 8 years. *Age* is coded 1 if the director is 63 years old or older (the sample median is 63).

The average score of DQI at the firm level is computed by summing the total DQI scores of all directors on the board and dividing by the number of directors. The DQI is coded 1 if a firm has an average equal to or greater than the median DQI-firm level of all firms (that being 1.154). The values for which DQI is above the median are coded as high-quality (HQ) directors, while those below the median are coded as low-quality (LQ) directors. As indicated by the results in Table 13, interlocked breached HQ and LQ directors both pre (column 1) and post (column 2) breach increase the breach likelihood of the connected firm. These results imply that regardless of whether the board is made up of high or low-quality directors, the effect of interlocking on the breach remains the same. Thus, we interpret this to mean that our results are unlikely to be driven by director quality.

To determine which of low-quality (LQ) directors or high-quality (HQ) directors is associated with the highest breach likelihood, we run simultaneous tests to differentiate between the magnitudes of these variables. The probability (Prob > chi2) of 0.00 indicates that the p-value, which is 0.00, is less than the common significance levels suggesting that the null hypothesis, that the coefficients of HQ and LQ directors are equal, can be rejected at the 0.01 significance level. The coefficients of *Interlocked Breached Pre LQ* and *Interlocked Breached Post LQ* are larger than the coefficients of *Interlocked Breached Pre HQ* and *Interlocked Breached Post HQ*, respectively. These results suggest that while both low-quality and high-

quality interlocked directors increase the connected firms' breach likelihood, directors of a lower quality increase the likelihood by a greater extent.

[Insert Table 13 here]

4.4.3. Propensity Score Matching Technique

As our outcome variable has a relatively low percentage (2.9%) compared to the whole sample, we also employ a propensity score matching (PSM) technique. We use the nearest neighbor PSM method without replacement whereby we require that each interlocked breached observation is matched to one non-interlocked breach observation within a 3% calliper distance. This method has been used in prior studies (e.g., Ettredge et al., 2018; Smith et al., 2021). This match is performed to exclude non-interlocked breached firm observations that are different from the interlocked breached firm observations. This technique allows us to have a balanced sample of firm-year observations where half of the sample is interlocked to a breached firm and the other half is not. This results in a sample of 4,313 (4,234) interlocked breached firms pre (post) matched to 4,313 (4,234) non-interlocked breached firms pre (post). The mean differences between the treated interlocked group and the matched non-interlocked control group are balanced for all variables.

Table 14 indicates that our coefficients for both pre and post are statistically significant and positive. The area under the ROC curve is 0.746 and 0.707 for the pre and post models, respectively suggesting that the models we used for matching are adequately fit[11]. This technique enhances confidence in our main findings and serves as a method to reduce any bias that may have been present in the earlier samples that were tested. Collectively, our findings

[11] "The empirical receiver operating characteristic (ROC) curve is a standard statistical tool used to evaluate the performance of a binary classifier" (Lieli & Hsu, 2019, p.100). The area under the ROC curve ranges from 0 to 1, and values closer to the upper boundary reflect fit tests (Qin & Zhang, 2003).

suggest that the increased likelihood of the focal firm experiencing a breach is due to the breached firm interlock, pre or post breach occurrence, and not due to other differences between the treatment and control samples.

[Insert Table 14 here]

5. Conclusion

Our study finds that firms connected to breached firms through shared directors face a higher likelihood of experiencing a breach, especially when the breach is internal. Notably, the presence of female directors, directors with audit roles, or executive directors with breach experience mitigates this risk. Firms with a greater number of interlocks and breached suppliers are more susceptible to breaches. We conducted additional analyses to rule out alternative explanations, such as differences in director quality or firm-specific factors unrelated to breach experience. These analyses support our findings and underscore the need to reinforce data protection regulations and clarify board responsibilities in data security. While we cannot definitively establish causality between board interlocks and data breach likelihood, our rigorous analyses provide compelling evidence that the observed association is not a spurious correlation.

We provide insights for boards, management, regulators, and shareholders, suggesting that interlocking board members can act as a risk spillover mechanism for data breaches. Specifically, when firms share directors with breached firms, a negative spillover effect is observed, revealing a previously unrecognized drawback of interlocking boards. This effect is more pronounced when the breach is internal, revealing that security weaknesses within the firms are mainly responsible for the inter-firm increased breach likelihood. In the pre-period, this may arise from directors not fully addressing data breach risks, potentially due to a lack of knowledge in data security matters or information asymmetry. In the post-period, it may result

from directors not being involved much in security matters or failing to leverage their breach experience to improve risk management in interlocked firms, possibly due to time constraints. Consequently, firms remain vulnerable to breaches despite board members' prior exposure to them. Additionally, there appears to be a gap between board actions for managing breach risks and the actual level of risk exposure, leaving firms unprepared to handle data-related attacks (Higgs et al., 2016). This finding aligns with Ashraf (2022), highlighting the risk that firms might focus on managing breach consequences rather than strengthening data protection proactively.

To address these issues, firms could convert data breach risks into learning opportunities by leveraging board members' experiences and addressing the root causes of breaches. Ashraf (2022) suggests that peer breaches should prompt focal firms to mitigate cyber risk. Firms can use these findings to review and enhance their security practices and risk management discussions by prioritizing data protection (Garg, 2020; Rothrock et al., 2018; SEC, 2024). Further, consideration may also be given to Héroux and Fortin's (2022) findings, which link certain board characteristics to improved cybersecurity. Specifically, having more female directors, audit committee members, and executive directors with breach experience could lower breach risk. Regulatory authorities might consider mandating specific board composition criteria, such as including at least one female director, one audit committee member, and one executive director with breach experience, to protect shareholder investments and safeguard sensitive data.

While contributing to corporate governance literature by examining determinants of data breaches and interlocking boards, our findings should be evaluated in view of their limitations. We aimed to capture as many breaches as possible; however, undisclosed breaches may have led to sample misrepresentation, potentially affecting observed spillover effects. Future research could extend this study as more data becomes available and consider breach

magnitude and the cyber vulnerability and defense capabilities of focal firms. Additionally, examining the barriers to knowledge spillover, such as time constraints, information asymmetry, or difficulty in applying learned lessons across diverse settings, could yield further insights into improving data security.

References

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, *94*(2), 291-309.

Aguilar, L. A. (2014, June). Boards of directors, corporate governance and cyber-risks: Sharpening the focus. In *Cyber Risks and the Boardroom conference, New York Stock Exchange*. Retrieved December 9, 2022, from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://insidecybersecurity.com/sites/insidecybersecurity.com/files/documents/jun2014/cs06102014_BOD_Corporate_Governance_Cyber_Risks.pdf

Ahmad, A., Desouza, K. C., Maynard, S. B., Naseer, H., & Baskerville, R. L. (2020). How integration of cyber security management and incident response enables organizational learning. *Journal of the Association for Information Science and Technology*, *71*(8), 939-953.

Ahn, S., Jiraporn, P., & Kim, Y. S. (2010). Multiple directorships and acquirer returns. *Journal of Banking & Finance*, *34*(9), 2011-2026.

Akbas, F., Meschke, F., & Wintoki, M. B. (2016). Director networks and informed traders. *Journal of Accounting and Economics*, *62*(1), 1-23.

Anderson, M. H. (2008). Social networks and the cognitive motivation to realize network opportunities: A study of managers' information gathering behaviors. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *29*(1), 51-78.

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

Ashraf, M. (2022). The role of peer events in corporate governance: Evidence from data breaches. *The Accounting Review*, *97*(2), 1-24.

Ashraf, M., Michas, P. N., & Russomanno, D. (2020). The impact of audit committee information technology expertise on the reliability and timeliness of financial reporting. *The Accounting Review*, *95*(5), 23–56.

Beattie, V., Fearnley, S., & Hines, T. (2012). Do UK audit committees really engage with auditors on audit planning and performance? *Accounting and Business Research*, *42*(3), 349-375.

Benaroch, M., & Chernobai, A. (2017). Operational IT failures, IT value destruction, and board-level IT governance changes. *MIS Quarterly*, *41*(3), 729-A6.

Bertrand, M., Luttmer, E. F., & Mullainathan, S. (2000). Network effects and welfare cultures. *The Quarterly Journal of Economics*, *115*(3), 1019-1055.

Bianchi, P. A. (2018). Auditors' joint engagements and audit quality: Evidence from Italian private companies. *Contemporary Accounting Research*, *35*(3), 1533–1577.

Bizjak, J., Lemmon, M., & Whitby, R. (2009). Option backdating and board interlocks. *The Review of Financial Studies*, *22*(11), 4821–4847.

Bonime-Blanc, A. (2017). A strategic cyber roadmap for the board. Retrieved December 2, 2022, from https://corpgov.law.harvard.edu/2017/01/12/a-strategic-cyber-roadmap-for-the-board/

Borgatti, S. P., & Ofem, B. (2010). Social network theory and analysis. *Social Network Theory and Educational Change*, 17-29.

Bouwman, C. H. (2011). Corporate governance propagation through overlapping directors. *The Review of Financial Studies*, *24*(7), 2358-2394.

Bravo, F. (2018). Does board diversity matter in the disclosure process? an analysis of the association between diversity and the disclosure of information on risks. *International Journal of Disclosure and Governance*, *15*(2), 104-114.

Brown, A. B., Dai, J., & Zur, E. (2019). Too busy or well-connected? Evidence from a shock to multiple directorships. *The Accounting Review*, *94*(2), 83-104.

Bruynseels, L., & Cardinaels, E. (2014). The audit committee: Management watchdog or personal friend of the CEO? *The Accounting Review*, *89*(1), 113-145.

Burt, R. S. (2005). *Brokerage and closure: An introduction to social capital*. Oxford University Press, Incorporated.

Burt, R. S. (1997). The contingent value of social capital. *Administrative Science Quarterly*, *42*(2), 339–365.

Cai, Y., Dhaliwal, D. S., Kim, Y., & Pan, C. (2014). Board interlocks and the diffusion of disclosure policy. *Review of Accounting Studies*, *19*, 1086-1119.

Caiazza, R., Cannella Jr, A. A., Phan, P. H., & Simoni, M. (2019). An institutional contingency perspective of interlocking directorates. *International Journal of Management Reviews*, *21*(3), 277–293.

Campbell, K., & Mínguez-Vera, A. (2008). Gender diversity in the boardroom and firm financial performance. *Journal of Business Ethics*, *83*, 435-451.

Carrera, N., Sohail, T., & Carmona, S. (2017). Audit committees' social capital and financial reporting quality. *Accounting and Business Research*, *47*(6), 633-672.

Cashman, G. D., Gillan, S. L., & Jun, C. (2012). Going overboard? On busy directors and firm value. *Journal of Banking & Finance*, *36*(12), 3248-3259.

Center for Strategic and International Studies (CSIS) – Washington, D. C. (2021). Significant cyber incidents. Retrieved December 1, 2022, from https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents

Cheng, S., Felix, R., & Indjejikian, R. (2019). Spillover effects of internal control weakness disclosures: the role of audit committees and board connections. *Contemporary Accounting Research*, *36*(2), 934–957.

Cheng, Z., Rai, A., Tian, F., & Xu, S. X. (2021). Social learning in information technology investment: the role of board interlocks. *Management Science*, *67*(1), 547-576.

Chiu, P. C., Teoh, S. H., & Tian, F. (2013). Board interlocks and earnings management contagion. *The Accounting Review*, *88*(3), 915-944.

CISCO (2017). Annual Cybersecurity Report. Retrieved December 7, 2022, from https://learningnetwork.cisco.com/s/article/cisco-2017-annual-cybersecurity-report-pdf

Core, J. E., Holthausen, R. W., & Larcker, D. F. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics*, *51*(3), 371-406.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448-474.

Dhaliwal, D. A. N., Naiker, V. I. C., & Navissi, F. (2010). The association between accruals quality and the characteristics of accounting experts and mix of expertise on audit committees. *Contemporary Accounting Research*, *27*(3), 787-827.

Erel, I., Stern, L. H., Tan, C., & Weisbach, M. S. (2021). Selecting directors using machine learning. *The Review of Financial Studies*, *34*(7), 3226-3264.

Ettredge, M., Guo, F., & Li, Y. (2018). Trade secrets and cyber security breaches. *Journal of Accounting and Public Policy*, *37*(6), 564-585.

Ettredge, M. L., & Richardson, V. J. (2003). Information transfer among internet firms: the case of hacker attacks. *Journal of Information Systems*, *17*(2), 71-82.

Fairchild, L., & Li, J. (2005). Director quality and firm performance. *Financial Review*, *40*(2), 257-279.

Falato, A., Kadyrzhanova, D., & Lel, U. (2014). Distracted directors: Does board busyness hurt shareholder value? *Journal of Financial Economics*, *113*(3), 404-426.

Fama, E.F., & Jensen, M.C. (1983). Separation of ownership and control. *Journal of Law and Economics 26*, 301–325.

Feng, C. Q., & Wang, T. (2019). Does CIO risk appetite matter? Evidence from information security breach incidents. *International Journal of Accounting Information Systems*, *32*, 59-75.

Fich, E. M., & Shivdasani, A. (2006). Are busy boards effective monitors? *The Journal of Finance*, *61*(2), 689-724.

Field, L., Lowry, M., & Mkrtchyan, A. (2013). Are busy boards detrimental? *Journal of Financial Economics*, *109*(1), 63-82.

Francoeur, C., Labelle, R., & Sinclair-Desgagné, B. (2008). Gender diversity in corporate governance and top management. *Journal of Business Ethics*, *81*, 83-95.

Gale, M., Bongiovanni, I., & Slapnicar, S. (2022). Governing cybersecurity from the boardroom: challenges, drivers, and ways ahead. *Computers & Security*, *121*, 102840.

Garg, P. (2020). Cybersecurity breaches and cash holdings: Spillover effect. *Financial Management*, *49*(2), 503-519.

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

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

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives*, *19*(1), 33–50.

Gul, F. A., Srinidhi, B., & Ng, A. C. (2011). Does board gender diversity improve the informativeness of stock prices? *Journal of Accounting and Economics*, *51*(3), 314-338.

Gulati, R., & Westphal, J. D. (1999). Cooperative or controlling? The effects of CEO-board relations and the content of interlocks on the formation of joint ventures. *Administrative Science Quarterly*, *44*(3), 473-506.

Gwebu, K. L., Wang, J., & Wang, L. (2018). The role of corporate reputation and crisis response strategies in data breach management. *Journal of Management Information Systems*, *35*(2), 683-714.

Harraka, M. (2002). Bowling alone: The collapse and revival of American community, by Robert D. Putnam. *Journal of Catholic Education, 6* (2).

Hartmann, C. C., & Carmenate, J. (2021). Academic research on the role of corporate governance and IT expertise in addressing cybersecurity breaches: Implications for practice, policy, and research. *Current Issues in Auditing*, *15*(2), A9-A23.

Haunschild, P. R. (1993). Interorganizational imitation: The impact of interlocks on corporate acquisition activity. *Administrative Science Quarterly*, *38*(4), 564–592.

Haunschild, P. R., & Beckman, C. M. (1998). When do interlocks matter?: Alternate sources of information and interlock influence. *Administrative Science Quarterly*, 815-844.

Hauser, R. (2018). Busy directors and firm performance: Evidence from mergers. *Journal of Financial Economics, 128*(1), 16-37.

He, C. Z., HuangFu, J., Kohlbeck, M., & Wang, L. (2023). The impact of customer-reported cybersecurity breaches on key supplier innovations and relationship disruption. *Journal of Information Systems, 37*(2), 21-49.

He, X., Pittman, J. A., Rui, O. M., & Wu, D. (2017). Do social ties between external auditors and audit committee members affect audit quality? *The Accounting Review*, *92*(5), 61-87.

He, Z., HuangFu, J., & Walton, S. (2022). Cybersecurity breaches in the supply chain and earnings management. *Journal of Information Systems*, *36*(3), 83-113.

Héroux, S., & Fortin, A. (2022). Board of directors' attributes and aspects of cybersecurity disclosure. *Journal of Management and Governance*, 1-46.

Higgs, J. L., Pinsker, R. E., Smith, T. J., & Young, G. R. (2016). The relationship between board-level technology committees and reported security breaches. *Journal of Information Systems*, *30*(3), 79-98.

Hovav, A., & Gray, P. (2014). The ripple effect of an information security breach event: A stakeholder analysis. *Communications of the Association for Information Systems*, *34*(1), 50.

Hsu, C., & Wang, T. (2014). Exploring the association between board structure and information security breaches. *Asia Pacific Journal of Information Systems*, *24*(4), 531-557.

IBM (2023). *IBM Report: Half of Breached Organizations Unwilling to Increase Security Spend Despite Soaring Breach Costs*. Retrieved January 9, 2024, from https://newsroom.ibm.com/2023-07-24-IBM-Report-Half-of-Breached-Organizations-Unwilling-to-Increase-Security-Spend-Despite-Soaring-Breach-Costs?mhsrc=ibmsearch_a&mhq=cost%20of%20a%20data%20breach%20report%202023

Intintoli, V. J., Kahle, K. M., & Zhao, W. (2018). Director connectedness: Monitoring efficacy and career prospects. *Journal of Financial and Quantitative Analysis*, *53*(1), 65-108.

Isidro, H., & Sobral, M. (2015). The effects of women on corporate boards on firm value, financial performance, and ethical and social compliance. *Journal of Business Ethics*, *132*, 1-19.

Islam, M. S., Wang, T., Farah, N., & Stafford, T. (2022). The spillover effect of focal firms' cybersecurity breaches on rivals and the role of the CIO: Evidence from stock trading volume. *Journal of Accounting and Public Policy*, *41*(2), 106916.

Janakiraman, R., Lim, J. H., & Rishika, R. (2018). The effect of a data breach announcement on customer behavior: Evidence from a multichannel retailer. *Journal of Marketing*, *82*(2), 85-105.

Javakhadze, D., Ferris, S. P., & French, D. W. (2016). Social capital, investments, and external financing. *Journal of Corporate Finance*, *37*, 38-55.

Jewer, J., & McKay, K. N. (2012). Antecedents and consequences of board IT governance: Institutional and strategic choice perspectives. *Journal of the Association for Information Systems*, *13*(7), 1.

Jiraporn, P., Davidson III, W. N., DaDalt, P., & Ning, Y. (2009). Too busy to show up? An analysis of directors' absences. *The Quarterly Review of Economics and Finance*, *49*(3), 1159-1171.

Johansen, T. R., & Pettersson, K. (2013). The impact of board interlocks on auditor choice and audit fees. *Corporate Governance: An International Review*, *21*(3), 287-310.

Kamiya, S., Kang, J. K., Kim, J., Milidonis, A., & Stulz, R. M. (2018). What is the impact of successful cyberattacks on target firms? (No. w24409). *National Bureau of Economic Research*.

Kamiya, S., Kang, J. K., Kim, J., Milidonis, A., & Stulz, R. M. (2021). Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics*, *139*(3), 719-749.

Kang, E. (2008). Director interlocks and spillover effects of reputational penalties from financial reporting fraud. *Academy of Management Journal*, *51*(3), 537-555.

Karpovich, T. (2002). Baltimore-based eChapman Inc. abandons Arthur Andersen. *Daily Record*. Retrieved December 9, 2022, from https://thedailyrecord.com/2002/03/18/city-firm-abandons-arthur-andersen/

Kim, Y., & Cannella Jr, A. A. (2008). Toward a social capital theory of director selection. *Corporate Governance: An International Review*, *16*(4), 282-293.

Kobelsky, K. W., Richardson, V. J., Smith, R. E., & Zmud, R. W. (2008). Determinants and consequences of firm information technology budgets. *The Accounting Review*, *83*(4), 957-995.

Kor, Y. Y., & Sundaramurthy, C. (2009). Experience-based human capital and social capital of outside directors. *Journal of Management*, *35*(4), 981-1006.

Lanz, J. (2014). Cybersecurity governance: The role of the audit committee and the CPA. *The CPA Journal*, *84*(11), 6.

Larcker, D. F., Richardson, S. A., Seary, A., & Tuna, A. (2005). Back door links between directors and executive compensation. *Available at SSRN 671063*.

Larkin, M. B., Bernardi, R. A., & Bosco, S. M. (2013). Does female representation on boards of directors associate with increased transparency and ethical behavior? *Accounting and the Public Interest*, *13*(1), 132-150.

Lawrence, A., Minutti-Meza, M., & Vyas, D. (2018). Is operational control risk informative of financial reporting deficiencies? *Auditing: A Journal of Practice & Theory*, *37*(1), 139-165.

Li, C., Lim, J. H., & Wang, Q. (2007). Internal and external influences on IT control governance. *international Journal of Accounting information Systems*, *8*(4), 225-239.

Li, H., No, W. G., & Wang, T. (2018). SEC's cybersecurity disclosure guidance and disclosed cybersecurity risk factors. *International Journal of Accounting Information Systems*, *30*, 40-55.

Lieli, R. P., & Hsu, Y. C. (2019). Using the area under an estimated ROC curve to test the adequacy of binary predictors. *Journal of Nonparametric Statistics*, *31*(1), 100-130.

Lin, N. (2002). *Social capital: A theory of social structure and action* (Vol. 19). Cambridge University Press.

Loderer, C., & Peyer, U. (2002). Board overlap, seat accumulation and share prices. *European Financial Management*, *8*(2), 165-192.

Masulis, R. W., & Mobbs, S. (2011). Are all inside directors the same? Evidence from the external directorship market. *The Journal of Finance*, *66*(3), 823-872.

Mishra, S. (2015). Organizational objectives for information security governance: a value focused assessment. *Information & Computer Security*, *23*(2), 122-144.

Nahapiet, J., & Ghoshal, S. (1998). Social Capital, Intellectual Capital, and the Organizational Advantage. *The Academy of Management Review*, *23*(2), 242.

Nielsen, S., & Huse, M. (2010). The contribution of women on boards of directors: Going beyond the surface. *Corporate Governance: An International Review*, *18*(2), 136-148.

Omer, T. C., Shelley, M. K., & Tice, F. M. (2020). Do director networks matter for financial reporting quality? Evidence from audit committee connectedness and restatements. *Management Science*, *66*(8), 3361-3388.

Peterson, C. A., & Philpot, J. (2007). Women's Roles on U.S. Fortune 500 Boards: Director Expertise and Committee Memberships. *Journal of Business Ethics*, *72*(2), 177–196.

Poletti-Hughes, J., & Martinez Garcia, B. (2022). Leverage in family firms: The moderating role of female directors and board quality. *International Journal of Finance & Economics*, *27*(1), 207-223.

Ponemon Institute. (2017). *2017 Cost of data breach study*. Retrieved December 10, 2022, from https://documents.ncsl.org/wwwncsl/Task-Forces/Cybersecurity-Privacy/IBM_Ponemon2017CostofDataBreachStudy.pdf

Portes, A. (1998). SOCIAL CAPITAL: Its origins and applications in modern sociology. *Annual Review of Sociology*, *24*, 1–24.

Privacy Rights Clearinghouse. (2017). *Chronology of data breaches: FAQ*. Retrieved December 11, 2022, from https://web.archive.org/web/20170617004210/https://www.privacyrights.org/chronology-data-breaches-faq

Qin, J., & Zhang, B. (2003). Using logistic regression procedures for estimating receiver operating characteristic curves. *Biometrika*, *90*(3), 585-596.

Radu, C., & Smaili, N. (2021). Board gender diversity and corporate response to cyber risk: evidence from cybersecurity related disclosure. *Journal of Business Ethics*, *177*, 351-374.

Rajagopal, A. (2019). Cyber attacks top list of risks impacting supply chain. *Cyber Security Hub*. Retrieved August 5, 2024, from https://www.cshub.com/attacks/articles/cyber-attacks-top-list-of-risks-impacting-supply-chain

Rajgopal, S., & Srinivasan, S. (2016). Why the market yawned when Yahoo was hacked. Wall Street Journal. Retrieved December 7, 2022, from https://www.wsj.com/articles/why-the-market-yawned-when-yahoo-was-hacked-1475537076

Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, *48*(2), 240-267.

Richardson, V. J., Smith, R. E., & Watson, M. W. (2019). Much ado about nothing: The (lack of) economic impact of data privacy breaches. *Journal of Information Systems*, *33*(3), 227-265.

Rosati, P., Cummins, M., Deeney, P., Gogolin, F., Van der Werff, L., & Lynn, T. (2017). The effect of data breach announcements beyond the stock price: Empirical evidence on market activity. *International Review of Financial Analysis*, *49*, 146-154.

Rothrock, R. A., Kaplan, J., & Van Der Oord, F. (2018). The board's role in managing cybersecurity risks. *MIT Sloan Management Review*, *59*(2), 12-15.

Say, G., & Vasudeva, G. (2020). Learning from digital failures? The effectiveness of firms' divestiture and management turnover responses to data breaches. *Strategy Science*, *5*(2), 117-142.

Schinagl, S., & Shahim, A. (2020). What do we know about information security governance? "From the basement to the boardroom": towards digital security governance. *Information & Computer Security*, *28*(2), 261-292.

Sonnemaker, T. (2019). Facing inevitable data breaches and new privacy laws, companies shift focus to response. Retrieved December 7, 2022, from https://news.medill.northwestern.edu/chicago/facing-inevitable-data-breaches-and-new-privacy-laws-companies-shift-focus-to-response/

Smith, T., Tadesse, A. F., & Vincent, N. E. (2021). The impact of CIO characteristics on data breaches. *International Journal of Accounting Information Systems*, *43*, 100532.

Srinivasan, S. (2005). Consequences of financial reporting failure for outside directors: Evidence from accounting restatements and audit committee members. *Journal of Accounting Research*, *43*(2), 291-334.

Stuart, T. E., & Yim, S. (2010). Board interlocks and the propensity to be targeted in private equity transactions. *Journal of Financial Economics*, *97*(1), 174-189.

Symantec, C. (2019). Internet security threat report: Volume 24. *Symantee Enterprise Security*. Retrieved August 5, 2024, from https://docs.broadcom.com/doc/istr-24-2019-en

Tan, H. T., & Yu, Y. (2018). Management's responsibility acceptance, locus of breach, and investors' reactions to internal control reports. *The Accounting Review*, *93*(6), 331-355.

U.S. Securities and Exchange Commission (SEC) (2024). *Examination Priorities – Division of Examinations*. Retrieved December 15, 2024, from www.sec.gov/files/2024-exam-priorities.pdf

Van Deth, J. W., & Zmerli, S. (2010). Introduction: Civicness, equality, and democracy- A "dark side" of social capital? *American Behavioral Scientist*, *53*(5), 631-639.

Vincent, N. E., Higgs, J. L., & Pinsker, R. E. (2019). Board and management-level factors affecting the maturity of IT risk management practices. *Journal of Information Systems*, *33*(3), 117-135.

Von Solms, B., & Von Solms, R. (2018). Cybersecurity and information security–what goes where? *Information and Computer Security*, *26* (1), 2–9.

Walton, S., Wheeler, P. R., Zhang, Y. I., & Zhao, X. R. (2021). An integrative review and analysis of cybersecurity research: Current state and future directions. *Journal of Information Systems, 35*(1), 155-186.

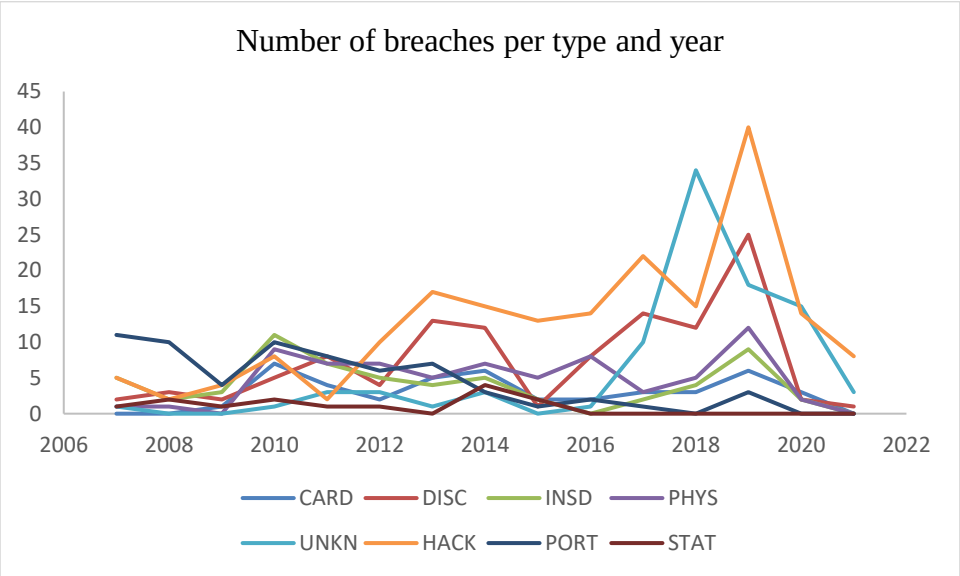
Williams, R. J. (2003). Women on corporate boards of directors and their influence on corporate philanthropy. *Journal of Business Ethics, 42,* 1-10.

Yayla, A. A., & Hu, Q. (2014). The effect of board of directors' IT awareness on CIO compensation and firm performance. *Decision Sciences, 45*(3), 401-436.

Yeoh, W., Wang, S., Popovič, A., & Chowdhury, N. H. (2022). A systematic synthesis of critical success factors for cybersecurity. *Computers & Security, 118,* 102724.

Zhong, Q., Liu, Y., & Yuan, C. (2017). Director interlocks and spillover effects of board monitoring: Evidence from regulatory sanctions. *Accounting & Finance, 57*(5), 1605-1633.

Figure 1. Time trend of breach types.



Note: Figure 1 displays the number of breaches of each of the 8 types over the period 2007 through 2021. Payment card fraud (CARD), unintended disclosure (DISC), insiders (INSD), physical loss (PHYS), and unknown (UNKN) reflect internal breaches. Hacking (HACK), portable device theft (PORT), and stationary theft (STAT) reflect external breaches (breach definitions are available in Appendix B).

Table 1
Sample Selection

Panel A: Data Breach Incidents

	N
Total data breach incidents from 2007 to 2021 (Privacy Rights Clearinghouse)	8,047
Less: Not for profit, education, and government organizations	-1,441
Less: Duplicate observations and firms not included in Compustat	-5,955
Final sample of data breach incidents	651

Panel B: Final Sample

	n
Firm-year observations from Compustat	168,623
Director observations from Institutional Shareholder Services	210,682
Final matched firm-level dataset	**18,731**
Breached firm-year observations	548
Focal firm-year observations	18,183

Notes: Panel A presents the number of breach incidents from 2007 to 2021, and the final sample of breaches used. Panel B shows the construction of our final sample. The total number of data breach incidents (651) represents all breaches experienced by firms over the sample period, with some firms experiencing more than one breach in certain years. The number of breached firm-year observations (548) represents the number of unique firm-year instances where at least one breach occurred.

Table 2. Frequencies and percentages within different categories

Panel A. Number of breached and focal firm-year

Year	Focal Firms	Breached Firms	Total
2007	765	25	790
2008	1,088	20	1,108
2009	1,153	15	1,168
2010	1,140	43	1,183
2011	1,173	36	1,209
2012	1,217	38	1,255
2013	1,234	43	1,277
2014	1,245	49	1,294
2015	1,294	26	1,320
2016	1,323	31	1,354
2017	1,341	48	1,389
2018	1,329	62	1,391
2019	1,371	65	1,436
2020	1,328	35	1,363
2021	1,182	12	1,194
Total	18,183	548	18,731
Percent	97.07	2.93	100

Panel B. Breached and Interlocked firm-year

Breached Firm	Interlocked Firm		
	0	1	Total
0	3,360	14,823	18,183
1	46	502	548
Total	3,406	15,325	18,731

Panel C. Breaches per industry.

Industry	Breached Firm	
	Number	Percentage
Accommodation and Food Services	25	4.56%
Administrative and Support and Waste Management and Remediation Services	20	3.65%
Construction	2	0.36%
Educational Services	1	0.18%
Finance and Insurance	168	30.66%
Health Care and Social Assistance	38	6.93%
Information	61	11.13%
Manufacturing	91	16.61%
Mining, Quarrying, and Oil and Gas Extraction	5	0.91%
Not Classified	20	3.65%
Other Services (except Public Administration)	1	0.18%
Professional, Scientific, and Technical Services	15	2.74%
Real Estate and Rental and Leasing	12	2.19%
Retail Trade	54	9.85%
Transportation and Warehousing	13	2.37%
Utilities	7	1.28%
Wholesale Trade	15	2.74%
Total	548	100%

Notes: Panel A presents the number of breached and focal firm-year observations over the studied period. Panel B includes information about breached and interlocked firm-year observations. Panel C includes the breaches by industry.

Table 3. Descriptive Statistics

Variable	N	Mean	Std. Dev.	Min	Max
Breached Firm $_{i,t}$	18,731	0.029	0.169	0	1
Interlocked Breached Firm Pre $_{i,t}$	18,731	0.243	0.429	0	1
Interlocked Breached Firm Post $_{i,t}$	18,731	0.257	0.437	0	1
Firm Size $_{i,t-1}$	18,731	8.212	1.703	3.968	15.035
Leverage $_{i,t-1}$	18,731	0.570	0.243	0.032	4.350
Loss $_{i,t-1}$	18,731	0.125	0.331	0	1
High Tech $_{i,t}$	18,731	0.331	0.470	0	1
Board Size $_{i,t}$	18,731	9.310	2.314	1	34
Foreign $_{i,t}$	18,731	0.309	0.462	0	1
Merger $_{i,t}$	18,731	0.330	0.470	0	1
Interlocked Breached Female $_{i,t}$	18,731	0.037	0.189	0	1
Interlocked Breached AC $_{i,t}$	18,731	0.051	0.220	0	1
Number of Interlocks $_{i,t}$	18,731	2.254	1.838	0	11
Interlocked Breached Executive $_{i,t}$	18,731	0.029	0.167	0	1

Notes: This table provides descriptive statistics for the main analysis' variables. The variables are defined in Appendix A.

Table 4. Pairwise Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Breached Firm $_{i,t}$	1.000													
(2) Interlocked Breached Firm Pre $_{i,t}$	**0.135**	1.000												
(3) Interlocked Breached Firm Post $_{i,t}$	**0.129**	**0.331**	1.000											
(4) Firm Size $_{i,t-1}$	**0.152**	**0.291**	**0.371**	1.000										
(5) High Tech $_{i,t}$	**-0.038**	-0.013*	-0.006	**-0.072**	1.000									
(6) Leverage $_{i,t-1}$	**0.063**	**0.114**	**0.171**	**0.472**	**-0.206**	1.000								
(7) Loss $_{i,t-1}$	**-0.027**	**-0.065**	**-0.046**	**-0.104**	**0.065**	0.013*	1.000							
(8) Board Size $_{i,t}$	**0.088**	**0.255**	**0.257**	**0.576**	**-0.079**	**0.329**	**-0.106**	1.000						
(9) Foreign $_{i,t}$	**-0.034**	0.002	*0.019*	**-0.040**	**0.212**	**-0.118**	**0.028**	*-0.018*	1.000					
(10) Merger $_{i,t}$	0.011	0.002	**0.067**	**0.043**	**0.135**	-0.015*	**-0.021**	**0.048**	**0.136**	1.000				
(11) Interlocked Breached Female $_{i,t}$	**-0.031**	**0.176**	**0.190**	**0.143**	-0.006	**0.070**	*-0.018*	**0.108**	*0.015*	0.005	1.000			
(12) Interlocked Breached AC $_{i,t}$	**-0.037**	**0.197**	**0.208**	**0.156**	0.007	**0.061**	**-0.020**	**0.114**	*0.014*	*0.015*	**0.407**	1.000		
(13) Number of Interlocks $_{i,t}$	**0.101**	**0.345**	**0.403**	**0.479**	**0.047**	**0.195**	**-0.043**	**0.428**	**0.092**	**0.073**	**0.182**	**0.197**	1.000	
(14) Interlocked Breached Executive $_{i,t}$	**-0.026**	**0.154**	**0.145**	**0.129**	0.013*	**0.050**	**-0.024**	**0.106**	**0.033**	0.003	**0.174**	**0.231**	**0.152**	1.000

Notes: This table presents the correlations among the variables. Values in bold represent significance at the 1% level, values in italic at the 5% level, and * denotes significance at the 10% level. The variables are defined in Appendix A.

Table 5: Effect of interlocks on breach likelihood and the average marginal effects.
Panel A: Logit Regression output

Breached firm	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.803***	0.109	0.589	1.018
Interlocked Breached Firm Post $_{i,t}$	0.620***	0.111	0.403	0.837
Firm Size $_{i,t-1}$	0.334***	0.035	0.266	0.403
Leverage $_{i,t-1}$	-0.444*	0.253	-0.940	0.051
Loss $_{i,t-1}$	-0.094	0.176	-0.439	0.251
High Tech $_{i,t}$	0.123	0.157	-0.184	0.430
Board Size $_{i,t}$	-0.021	0.024	-0.068	0.026
Foreign $_{i,t}$	-0.295**	0.115	-0.520	-0.071
Merger $_{i,t}$	0.051	0.101	-0.147	0.249
Industry	Yes			
Year	Yes			
Constant	-7.103***	0.584	-8.248	-5.958
Pseudo R2	0.174			
Observations	18,731			

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

Panel B: Average Marginal Effects output

	AME	Std.Err.	P-value	[95%Conf.	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.021	0.003	0.000	0.015	0.027
Interlocked Breached Firm Post $_{i,t}$	0.016	0.003	0.000	0.010	0.022
Firm Size $_{i,t-1}$	0.009	0.001	0.000	0.007	0.011
Leverage $_{i,t-1}$	-0.012	0.007	0.079	-0.025	0.001
Loss $_{i,t-1}$	-0.002	0.005	0.592	-0.011	0.007
High Tech $_{i,t}$	0.003	0.004	0.431	-0.005	0.011
Board Size $_{i,t}$	-0.001	0.001	0.380	-0.002	0.001
Foreign $_{i,t}$	-0.008	0.003	0.010	-0.014	-0.002
Merger $_{i,t}$	0.001	0.003	0.614	-0.004	0.006
Industry	YES				
Year	YES				

Notes: Panel A presents the logit regression output for our model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$ and *Interlocked Breached Firm Post* $_{i,t}$, which include whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach, or in any of the 3 years following the breach, respectively. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively. Panel B presents the average marginal effects output where AME is the discrete change from the base level.

Table 6: Effect of interlocking to internally or externally breached firm on breach likelihood.

Dependent Variables:	**Internally Breached**				**Externally Breached**			
	Coef.	St.Err.	[95% Conf	Interval]	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	1.233***	0.126	0.986	1.481	0.928	0.193	-0.683	0.073
Interlocked Breached Firm Post $_{i,t}$	1.229***	0.127	0.979	1.479	0.760**	0.194	-0.849	-0.090
Pseudo R2	0.074							
Observations	37,462							
Prob>chi2	0.0016							

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

Notes: This table presents the joint logit regressions' coefficients of the main independent variables *(Interlocked Breached Firm Pre* $_{i,t}$ and *Interlocked Breached Firm Post* $_{i,t}$*)* for each of the two categories of dependent variables *(Internally Breached* and *Externally Breached)*. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

The number of observations is doubled because the system runs the model twice, once for Internally Breached as a DV and once for Externally Breached as a DV.

Table 7: Effect of the presence of female director with breach experience on breach likelihood

Logit regression

Breached firm	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.891***	0.110	0.676	1.106
Interlocked Breached Firm Post$_{i,t}$	0.699***	0.110	0.483	0.915
Interlocked Breached Female $_{i,t}$	-3.592***	0.714	-4.991	-2.194
Firm Size $_{i,t-1}$	0.350*	0.035	0.281	0.418
Leverage $_{i,t-1}$	-0.459	0.264	-0.977	0.058
Loss $_{i,t-1}$	-0.057	0.176	-0.403	0.289
High Tech $_{i,t}$	0.098	0.158	-0.211	0.407
Board Size $_{i,t}$	-0.019	0.024	-0.065	0.028
Foreign $_{i,t}$	-0.286**	0.116	-0.514	-0.059
Merger $_{i,t}$	0.035	0.102	-0.165	0.234
Industry	YES			
Year	YES			
Constant	-7.323***	0.588	-8.474	-6.171
Pseudo R2	0.194			
Observations	18,731			

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

Notes: This table presents the logit regression output for our model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach, *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach, and *Interlocked Breached Female* $_{i,t}$ variable which is coded 1 when there is at least 1 female on the board of a focal firm that is interlocked to a firm that has experienced/will experience a breach. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 8: Effect of the presence of a director on the Audit Committee of a focal firm connected to a breached firm

Logit regression

Breached firm	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.906***	0.110	0.691	1.120
Interlocked Breached Firm Post $_{i,t}$	0.723***	0.110	0.507	0.939
Interlocked Breached AC $_{i,t}$	-3.920***	0.713	-5.317	-2.522
Firm Size $_{i,t-1}$	0.358***	0.035	0.289	0.427
Leverage $_{i,t-1}$	-0.370	0.255	-0.871	0.131
Loss $_{i,t-1}$	-0.107	0.177	-0.455	0.240
High Tech $_{i,t}$	0.122	0.157	-0.186	0.430
Board Size $_{i,t}$	-0.016	0.024	-0.063	0.031
Foreign $_{i,t}$	-0.331***	0.116	-0.558	-0.103
Merger $_{i,t}$	0.038	0.102	-0.161	0.238
Industry	YES			
Year	YES			
Constant	-7.334***	0.586	-8.483	-6.185
Pseudo R2	0.202			
Observations	18,731			

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

Notes: This table presents the logit regression output for our main model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach and *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach. The *Interlocked Breached AC* $_{i,t}$ variable is coded 1 when there is at least 1 director on the audit committee of a focal firm that is interlocked to a firm that has experienced/will experience a breach. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 9: Effect of the presence of executive directors on breach likelihood

Logit regression

Breached firm	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.858***	0.109	0.644	1.072
Interlocked Breached Firm Post $_{i,t}$	0.657***	0.111	0.440	0.874
Interlocked Breached Executive $_{i,t}$	-3.184***	0.714	-4.583	-1.785
Firm Size $_{i,t-1}$	0.348***	0.035	0.279	0.417
Leverage $_{i,t-1}$	-0.434*	0.256	-0.935	0.068
Loss $_{i,t-1}$	-0.120	0.177	-0.466	0.227
High Tech $_{i,t}$	0.114	0.157	-0.193	0.422
Board Size $_{i,t}$	-0.017	0.024	-0.063	0.030
Foreign $_{i,t}$	-0.268**	0.115	-0.494	-0.042
Merger $_{i,t}$	0.032	0.102	-0.167	0.231
Industry	YES			
Year	YES			
Constant	-7.256***	0.586	-8.405	-6.107
Pseudo R2	0.187			
Observations	18,731			

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

Notes: This table presents the logit regression output for our model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach, *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach, and *Interlocked Breached Executive* $_{i,t}$ variable which is coded 1 when there is at least 1 executive director on the board of a focal firm that is interlocked to a firm that has experienced/will experience a breach. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 10: Effect of the number of interlocks on breach likelihood

Logit regression

Breached firm	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.776***	0.111	0.559	0.993
Interlocked Breached Firm Post $_{i,t}$	0.584***	0.113	0.363	0.805
Firm Size $_{i,t-1}$	0.317***	0.036	0.246	0.389
Leverage $_{i,t-1}$	-0.447*	0.252	-0.941	0.048
Loss $_{i,t-1}$	-0.092	0.176	-0.437	0.252
High Tech $_{i,t}$	0.124	0.157	-0.184	0.431
Board Size $_{i,t}$	-0.030	0.025	-0.079	0.019
Foreign $_{i,t}$	-0.303***	0.115	-0.528	-0.078
Merger $_{i,t}$	0.052	0.101	-0.146	0.250
Number of Interlocks $_{i,t}$	0.047*	0.028	-0.009	0.102
Industry	YES			
Year	YES			
Constant	-6.958***	0.591	-8.117	-5.800
Pseudo R2	0.175			
Observations	18,731			

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

Notes: This table presents the logit regression output for our model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach and *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach. The *Number of Interlocks* $_{i,t}$ variable reflects the number of firms to which a firm is connected. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 11: Effect of breaches across the supply chain

Logit regression

Breached firm	Coef.	St.Err.	p-value	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	2.214***	0.244	0.000	1.784	2.748
Interlocked Breached Firm Post $_{i,t}$	1.798***	0.201	0.000	1.445	2.238
Firm Size $_{i,t-1}$	1.388***	0.050	0.000	1.292	1.490
Leverage $_{i,t-1}$	0.626*	0.160	0.067	0.380	1.033
Loss $_{i,t-1}$	0.915	0.161	0.615	0.648	1.293
High Tech $_{i,t}$	1.137	0.179	0.414	0.835	1.548
Board Size $_{i,t}$	0.978	0.024	0.368	0.933	1.026
Foreign $_{i,t}$	0.754**	0.087	0.014	0.601	0.945
Merger $_{i,t}$	1.070	0.109	0.505	0.877	1.307
Nb Breached Customers $_{i,t}$	1.316	0.393	0.357	0.733	2.363
Nb Breached Suppliers $_{i,t}$	2.361**	0.804	0.012	1.212	4.602
Nb Interlocked Breached Customers Pre $_{i,t}$	0.990	0.226	0.964	0.633	1.548
Nb Interlocked Breached Customers Post $_{i,t}$	0.777	0.161	0.224	0.518	1.167
Nb Interlocked Breached Suppliers Pre $_{i,t}$	0.920	0.085	0.368	0.768	1.103
Nb Interlocked Breached Suppliers Post $_{i,t}$	1.117	0.098	0.206	0.941	1.326
Industry	YES				
Year	YES				
Constant	0.001***	0.000	0.000	0.000	0.003
Pseudo R2	0.175				
Observations	18,467				

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

Notes: This table presents the logit regression output for our model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach, *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach. *Nb Breached Customers (Suppliers)* $_{i,t}$ reflect the number of breached customers (suppliers) of the firm. *Nb Interlocked Breached Customers Pre* and *Post* variables reflect the number of customers that are interlocked to firms that will experience a breach in any of the coming 3 years or that have experienced a breach in any of the past 3 years, respectively. *Nb Interlocked Breached Suppliers Pre* $_{i,t}$ and *Post* $_{i,t}$ variables reflect the number of suppliers that are interlocked to firms that will experience a breach in any of the coming 3 years or that have experienced a breach in any of the past 3 years, respectively. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 12. Effect of interlock on breach likelihood with additional control variables

Logit regression

Breached Firm	Coef.	St.Err.	[95% Conf	Interval]
Interlocked Breached Firm Pre $_{i,t}$	0.702***	0.166	0.376	1.027
Interlocked breached Firm Post $_{i,t}$	0.811***	0.168	0.482	1.14
Firm Size $_{i,t-1}$	0.186***	0.070	0.049	0.323
Leverage $_{i,t-1}$	-0.782**	0.379	-1.526	-0.039
Loss $_{i,t-1}$	-0.092	0.277	-0.634	0.451
High Tech $_{i,t}$	-0.349	0.231	-0.801	0.103
Board Size $_{i,t}$	-0.010	0.039	-0.085	0.066
Foreign $_{i,t}$	-0.278*	0.167	-0.605	0.050
Merger $_{i,t}$	0.052	0.150	-0.242	0.346
Segments $_{i,t}$	0.035	0.106	-0.174	0.243
BTM $_{i,t}$	0.017	0.142	-0.261	0.296
Intangibles $_{i,t}$	0.076*	0.045	-0.012	0.163
Industry	YES			
Year	YES			
Constant	-7.677***	1.149	-9.929	-5.424
Pseudo R2	0.178			
Observations	7,807			

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

Notes: This table presents the logit regression output for our model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach and *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach. All the variables are defined in Appendix A. Year and Industry fixed effects are included. The number of observations is lower due to missing segments data before 2014. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 13: Effect of the presence of high and low-quality directors on breach likelihood

Dependent Variable: Breached firm (dummy)		
	1	**2**
Variables	Interlocked in the 3 years prior to the breach occurrence	Interlocked in the 3 years following the breach occurrence
Interlocked Breached Pre HQ $_{i,t}$	1.136***	
Interlocked Breached Pre LQ $_{i,t}$	2.393***	
Interlocked Breached Post HQ $_{i,t}$		1.206***
Interlocked Breached Post LQ $_{i,t}$		2.292***
Firm Size $_{i,t-1}$	0.187***	0.163***
Leverage $_{i,t-1}$	-0.167	-0.363
Loss $_{i,t-1}$	-0.031	-0.096
High Tech $_{i,t}$	0.180	0.161
Board Size $_{i,t}$	-0.037	-0.017
Foreign $_{i,t}$	-0.200	-0.266**
Merger $_{i,t}$	0.067	0.032
Constant	-6.035***	-5.195***
Pseudo R2	0.306	0.302
Observations	18,731	18,731

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

Notes: This table presents the logit regression output for the main model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables in column 1 are *Interlocked Breached Pre HQ* $_{i,t}$ and *Interlocked Breached Pre LQ* $_{i,t}$, which include whether a firm has been interlocked to a breached firm with high- and low-quality directors, respectively in any of the 3 years leading to the breach. The main independent variables in column 2 are *Interlocked Breached Post HQ* $_{i,t}$ and *Interlocked Breached Post LQ* $_{i,t}$, which include whether a firm is interlocked to a breached firm with high- and low-quality directors, respectively in any of the 3 years following the breach. The *HQ* (*LQ)* reflects whether the average of directors serving on a firm's board are considered of high (low) quality based on tenure, age, and IT expertise. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 14: Propensity score matching

Dependent Variable: Breached firm (dummy)		
	1	**2**
Variables	Interlocked in the 3 years prior to the breach occurrence	Interlocked in the 3 years following the breach occurrence
Interlocked Breached Firm Pre $_{i,t}$	3.348***	
Interlocked Breached Firm Post $_{i,t}$		2.270***
Firm Size $_{i,t-1}$	1.538***	1.426***
Leverage $_{i,t-1}$	0.563*	0.664
Loss $_{i,t-1}$	0.651*	0.685
High Tech $_{i,t}$	0.582***	0.511***
Board Size $_{i,t}$	0.989	1.034
Foreign $_{i,t}$	0.685***	0.700***
Merger $_{i,t}$	1.195	1.034
Constant	0.001***	0.001***
Pseudo R2	0.098	0.063
Observations	8,626	8,468
Area under ROC curve	0.746	0.707

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

Notes: This table presents the logit regression output of the propensity score matching technique. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variable in column 1 is *Interlocked Breached Firm Pre* $_{i,t}$, which includes whether a firm has been interlocked to a breached firm in any of the 3 years leading to the breach. The main independent variable in column 2 is *Interlocked Breached Firm Post* $_{i,t}$, which includes whether a firm is interlocked to a breached firm in any of the 3 years following the breach. All the variables are defined in Appendix A.***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Appendix A. Variables description and source

Variable	Description
Breached Firm $_{i,t}$	Coded 1 if the firm is breached, 0 otherwise *(PRC)*
Internally Breached $_{i,t}$	Coded 1 if the firm is internally breached, 0 otherwise *(PRC)*
Externally Breached $_{i,t}$	Coded 1 if the firm is externally breached, 0 otherwise *(PRC)*
Interlocked Breached Firm Pre $_{i,t}$	Coded 1 if a firm shares at least 1 director in a given year with another firm that is about to experience a breach in any of the next 3 years. The breached firms are removed to only account for the spillover effect from breached to focal firms *(PRC, ISS)*
Interlocked Breached Firm Post $_{i,t}$	Coded 1 if a firm shares at least 1 director in a given year with another firm that has experienced a breach in any of the previous 3 years. The breached firms are removed to only account for the spillover effect from breached to focal firms *(PRC, ISS)*
Firm Size $_{i,t-1}$	= natural logarithm of Total Assets *(Compustat)*
Leverage $_{i,t-1}$	=Total Liabilities/ Total Assets *(Compustat)*
Loss $_{i,t-1}$	Coded 1 if Net Income is negative, 0 otherwise *(Compustat)*
High Tech $_{i,t}$	Coded 1 if NAICS is classified as High tech, 0 otherwise *(Compustat, Heckler, D. E. (2005). High-technology employment: a NAICS-based update. Monthly Lab. Rev., 128, 57)*
Board Size $_{i,t}$	Count of the number of directors on board *(ISS)*
Foreign $_{i,t}$	Coded 1 if the firm has foreign operations (based on FCA), 0 otherwise *(Compustat)*
Merger $_{i,t}$	Coded 1 if the firm is involved in a merger activity (based on AQP), 0 otherwise *(Compustat)*
BTM $_{i,t}$	= book value of equity (CEQ) / market value (PRCC_F×CSHO) *(Compustat)*
Intangibles $_{i,t}$	= natural logarithm of (1+ total intangible assets) *(Compustat)*
Segments $_{i,t}$	= natural logarithm of the number of geographic and business segments *(Compustat)*
Interlocked Breached Female $_{i,t}$	Coded 1 if the firm has at least one female director that also served on the BOD of another firm in a common year and has experienced a breach. The breached firms are removed *(PRC, ISS)*

Interlocked Breached AC $_{i,t}$	=1 if there's at least 1 director that is on the Audit Committee of a focal firm and connected to a breached firm *(PRC, ISS)*
Number of Interlocks $_{i,t}$	= number of other firms a firm is linked to through a shared director(s) for at least 1 year *(Directors' id, name, year service began and ended from ISS)*
Interlocked Breached Executive $_{i,t}$	Coded 1 if the firm has at least one executive director that also served on the BOD of another firm in a common year and has experienced a breach. The breached firms are removed *(PRC, ISS)*
Directors' Quality $_{i,t}$	*High-quality (HQ)* if the average directors' score on board is greater than or equal to 1.154 (median score of the sample). *Low-quality (LQ)* if the average directors' score on board is less than 1.154 (median score of the sample). The sum of the score for each firm is = sum of 3 dummy variables (IT expertise, Tenure, and Age). The average is equal to the sum for each firm divided by the number of directors on board *(ISS).*
Nb Breached Customers $_{i,t}$	= number of breached customers within the supply chain *(PRC, Compustat)*
Nb Breached Suppliers $_{i,t}$	= number of breached suppliers within the supply chain *(PRC, Compustat)*
Nb Interlocked Breached Customers Pre $_{i,t}$	= number of customers within the supply chain interlocked to firms that will experience a breach *(PRC, Compustat, ISS)*
Nb Interlocked Breached Customers Post $_{i,t}$	= number of customers within the supply chain interlocked to firms that experienced a breach *(PRC, Compustat, ISS)*
Nb Interlocked Breached Suppliers Pre $_{i,t}$	= number of suppliers within the supply chain interlocked to firms that will experience a breach *(PRC, Compustat, ISS)*
Nb Interlocked Breached Suppliers Post $_{i,t}$	= number of suppliers within the supply chain interlocked to firms that experienced a breach *(PRC, Compustat, ISS)*
Director Quality Index (DQI) $_{i,t}$	=1 if a firm has an average equal to or greater than the median of 1.154 DQI-firm level of all firms, calculated using Tenure, Age and IT Expertise
Tenure $_{i,t}$	=1 if the director served on the board for a period greater than or equal to the sample median of 8 years *(ISS)*
Age $_{i,t}$	=1 if the director is as old or older as the same median of 63 years *(ISS)*
IT Expertise $_{i,t}$	=1 if a director has prior IT expertise *(ISS)*

Notes: The sources are indicated within parentheses. PRC stands for Privacy Rights Clearinghouse and ISS stands for Institutional Shareholder Services.

Appendix B. Explanation of breach types

Internal Breach	External Breach
CARD- Fraud Involving Debit and Credit Cards Not Via Hacking (skimming devices at point-of-service terminals, etc.)	**HACK**- Hacked by an Outside Party or Infected by Malware
DISC- Unintended Disclosure Not Involving Hacking, Intentional Breach or Physical Loss (sensitive information posted publicly, mishandled or sent to the wrong party via publishing online, sending in an email, sending in a mailing or sending via fax)	**PORT**- Portable Device (lost, discarded or stolen laptop, PDA, smartphone, memory stick, CDs, hard drive, data tape, etc.)
INSD- Insider (employee, contractor or customer)	**STAT**- Stationary Computer Loss (lost, inappropriately accessed, discarded or stolen computer or server not designed for mobility)
PHYS- Physical (paper documents that are lost, discarded or stolen)	
UNKN- Unknown (not enough information about breach to know how exactly the information was exposed)	

Chapter 2

Becoming Invisible?

Data Breach Risk and Scientific Publications

ABSTRACT

We examine how exposure to data breach threats shapes firms' disclosure decisions regarding corporate innovation. Exploiting data breaches of peer firms as exogenous variation in firms' data security risk, we document that following peer firms' data breaches, firms reduce the number of scientific publications, a measure of firms' disclosure of fundamental research and innovation. This suggests that firms avoid signaling valuable intellectual property to potential attackers when the risk of data breaches increases. Besides the *direct* consequences of data breaches on corporate innovation and investment, security risks thus also impose substantial *indirect* economic costs if firms become more secretive and avoid publications that otherwise would facilitate corporate innovation and economic growth.

Keywords: Data breaches; Scientific publications; Corporate innovation; Trade secrets.

"*What is the real cost of a breach?"*

The Wall Street Journal (2020)

1. Introduction

Does data breach risk affect firms' scientific disclosure behavior? While prior literature examines the effect of data breach risk on firms investments in corporate governance (Ashraf, 2022) and corporate innovation (Lattanzio & Ma, 2023), the effect on firms' scientific disclosure behavior remains largely unstudied. Whether firms adjust their scientific disclosure behavior in response to data breach risk disclosures matters for regulators who aim to understand the incentives that drive disclosure decisions. It also matters for policy makers who want to understand the economic consequences of data breaches, and for potential investors who seek to understand whether firms react to data breaches. This study uses the number of scientific publications, as a proxy for disclosure, and the exogenous timing of competitors' data breaches, as variation in the data breach risk, to examine how firms react to an increase in data breach risk.

According to IBM, costs of a data breach reached a record high of $4.88 million globally in 2024, reflecting a 10% increase over last year, where 43% of these breaches involved intellectual property (IP) (IBM, 2024). Despite the economic significance of data breaches, how firms manage this risk, particularly in terms of scientific disclosure behavior, remains unclear. Disclosing IP through scientific publications might increase the risk by signaling valuable assets to potential attackers (Ettredge et al., 2018). Additionally, the risk of a data breach could disrupt firms' internal information-sharing practices. Internal processes often require multiple employees to review and approve material before publication. Consequently, if employees fear industrial espionage within the firm, they might simply decide not to publish scientific findings to avoid sharing the material

with other staff members. However, making research public could also reduce the appeal of data breaches by diminishing the value of stolen information, thereby potentially lowering the risk of an attack. Furthermore, such disclosure could also serve as a first step in protecting the firm's IP, since it would effectively shield the firm from patent infringement lawsuits and enable it to file a patent for the invention up to twelve months after publication. As a consequence, it is unclear whether firms adjust their scientific disclosure behavior[12].

To examine whether the risk of becoming the target of a data breach affects a firm's disclosure behavior, we exploit the data breaches of firms in the same industry and state. We concentrate on variation within industry and state because it enables us to control for developments, such as publishing trends within the microchips industry, within an industry and within a state. Hence, this research design allows us to minimize the impact of firm characteristics on the results.

We find that, all else equal, the occurrence of a data breach leads to a decrease in the number of scientific publications by around 0.0413 per million dollars of R&D expenses. Using our sample where the average R&D expenses is $149.321 million and the average number of firms per state-industry-year is 3.9, all else equal this translates to an average reduction of 24.05[13] scientific publications across all firms in a state-industry-year. Given that the average number of publications per state-industry-year is sixty-three, this represents a sixty-two percent decrease. This effect primarily emerges when peer firms experience data breaches that expose employee account information (i.e., emails and passwords) granting attackers access credentials to company systems and allowing them to steal proprietary information that could include R&D data, trade secrets, and

[12] It is important to note that while our focus in this study is the response firms have to potential attackers, by limiting their scientific disclosures, these firms will also impact all other potential users of this information. While this impact is important and should be addressed in future research, our focus and discussion will be around potential attackers.
[13] Calculated as 0.0413*149.321*3.9.

patent applications. Our results further indicate that data breach risk is also associated with fewer disclosed trade secrets on an R&D expense basis. Overall, these results suggest that conditional on their investment into innovation, firms decide to avoid disclosures that could potentially signal the value of internal IP to potential outside attackers.

These findings provide additional insights into the overall costs of data breach risks. Previous research on cybersecurity risks and data breaches has explored various potential consequences, including peer firms' future internal control material weaknesses (Ashraf, 2022), peer firms' abnormal trading volume (Islam et al., 2022), firms' reputation (Asthana et al., 2021), withholding of bad news (Obaydin et al., 2024), and the disclosure of cybersecurity risk factors (Chen et al., 2023). Our paper contributes to the current literature in three distinct ways.

First, this study advances this stream of literature by demonstrating that data breaches lead to a reduction in the level of non-financial disclosures. Besides the direct impact of data breaches on corporate innovation and investment, our results suggest that data breach risks also impose substantial indirect economic costs if firms become more secretive and avoid scientific disclosures that otherwise would facilitate corporate innovation and economic growth.

Second, we respond to the call by Glaeser and Lang (2024) for more research examining the incentives behind disclosing scientific publications. Previous studies have primarily focused on factors that include the role of prestige (Almeida et al., 2011; Arora et al., 2018), the advantages of a firm's commercialization strategy (Arora et al., 2021; Huang, 2016; Polidoro Jr & Theeke, 2012), increasing awareness among investors (Baruffaldi et al., 2024), establishing property rights through prior art (Godin, 1996; Nelson, 1990), and signaling a firm's leading position in the patent

race (Baker & Mezzetti, 2005)[14]. Our research extends this body of literature by demonstrating that the risk of data breaches is also a significant determinant in the decision to publish scientific findings.

Third, by focusing on scientific publications, this study also answers the call of Leuz and Wysocki (2016) for more research on nontraditional disclosure. In this stream of literature, several studies have examined the role of patents as corporate disclosures (e.g., Glaeser et al., 2020; Glaeser & Landsman, 2021; Kim & Valentine, 2021; Martens, 2023); however, the role of scientific publications is less clear. This study extends this stream of literature by showing that, similar to other corporate disclosures, firms carefully consider the distinct risks associated with scientific publications and adjust their disclosure behavior accordingly.

Our study proceeds as follows. In section 2, we offer background information on scientific publications, patents, data breaches and spillovers, and we develop the hypothesis. In section 3, we present the sample construction process, the outcome and treatment variables, and the empirical research design. In section 4, we present our analysis. In section 5, we provide a discussion, and we conclude in section 6.

2. Prior Literature and Hypothesis Development

2.1. Scientific Publications

Research and development is an important source of competitive advantage for companies (Kamiya et al., 2021). As many companies have been actively disclosing their findings in scientific journals, research focused on firms' publishing strategy has grown (Rotolo et al., 2022).

[14] Please refer to Rotolo et al. (2022) for a literature review that covers these topics.

Scientific publishing is the outcome of basic research and development and is a disclosure method where firms selectively reveal technical solutions (Alexy et al., 2013; Blind et al., 2022; Rotolo et al., 2022). Scientific publishing requires significant contributions to theory, empirical evidence, or research methodology (Blind et al., 2022). Publications are available within scientific journals as open access or fee-based, making them non-exclusive (Blind et al., 2022). Compared to prior research on trade secrets and patenting, the literature on firms' engagement in scientific publishing is still premature (Blind et al., 2022). Prior studies have shown that publishing requires certain capabilities that few firms have, such as R&D personnel with scientific education and an extensive knowledge base (Simeth & Lhuillery, 2015). As such, firms might be incentivized or hesitant to engage in scientific publications.

The incentives for firms to publish their results in a non-proprietary fashion can be grouped into five general categories. First, publishing firms gain access to external knowledge (Zucker et al., 2002), important technologies for new product development (Hayter & Link, 2018), and engage with other researchers in the production of new knowledge (Rotolo et al., 2022). Further, publishing stimulates suppliers and competitors in the same industry to reciprocate in knowledge sharing, nurturing the culture in the industry (Alexy et al., 2013; Pénin, 2007). Second, firms publish to attract and retain knowledge by building connections with other research communities (Zucker et al., 2002). Third, firms publish to establish property rights and to manage and safeguard their intellectual property portfolio (Godin, 1996; Nelson, 1990). By doing so, firms alter prior art, restricting competing firms from patenting their disclosed innovations (Rotolo et al., 2022). Prior studies have referred to this phenomenon as strategic disclosure/publication (Lichtman et al., 2000), defensive publishing (Rotolo et al., 2022), and defensive disclosure (Bar-Gill & Parchomovsky, 2003). By pursuing a defensive publishing strategy, the leader in this patent race

can drive out competitors (Baker & Mezzetti, 2005; Lichtman et al., 2000). Fourth, disclosing scientific knowledge helps the firm elevate its prestige and credibility and build its reputation as a knowledge producer (Almeida et al., 2011; Arora et al., 2018; Penders & Nelis, 2011). Fifth, and finally, publishing supports a firm's commercialization strategy by signaling its products and services and their respective quality (Arora et al., 2021; Huang, 2016; Polidoro Jr & Theeke, 2012). In addition to these benefits and compared to patents, publishing is a cheaper and faster way to build awareness among investors (Baruffaldi et al., 2024). A positive relation between publishing and firm valuation has been noted in technology-intensive industries (Arora et al., 2018; Pellens & Della Malva, 2018; Simeth & Cincera, 2016). Overall, publishing positively affects firms' innovation performance in the long run (Blind et al., 2022).

While the literature has shown the positive side of publishing, a few studies have raised concerns with scientific publication. Open disclosures represent an indirect cost of research and lead to knowledge spillovers from the publishing firms to competing firms (Arora et al., 2021). As knowledge becomes non-exclusive, firms will not be able to use it internally in innovation, resulting in reductions in the publishing firm's private returns (Arora et al., 2021). Further, the use of a firm's science by its rivals has been associated with a lower firm market value (Arora et al., 2018). Scientific publications can limit the publishing firms' competitive advantage as the results of their research become publicly available (Gans et al., 2017). Also, publishing limits firms' ability to generate profits on their commercialized innovations (Gans et al., 2017). For these reasons, firms may be less willing to invest in upstream research and more in downstream development, leading to a drop in innovation (Arora et al., 2021). In summary, the spillovers generated by firms' scientific publications may limit the respective benefits and success of their innovations in the short run (Blind et al., 2022).

2.2. Patents

Patenting is another strategy of selective knowledge disclosure (Alexy et al., 2013). It is the most prominent and formal instrument, and it allows firms, through its exclusivity characteristic, to appropriate the returns of the disclosed inventions (Blind et al., 2022). Patent disclosure should reveal the novelty and usefulness of an invention, the necessary inventive steps, and the ability of industrial application (Blind et al., 2022). A patent is the output of research and development (Blind et al., 2022). It gives the patentee property rights on the intangible assets and a legal monopoly under intellectual property and patent laws (Dass et al., 2015; Ettredge et al., 2018; Glaeser, 2018; Klasa et al., 2018). Prior studies have revealed a positive relationship between innovation and patenting (e.g., Hagedoorn & Cloodt, 2003; Nerkar & Roberts, 2004). Even for small firms that face patenting challenges, patenting assists in commercializing their innovations by providing legal protection for their inventions, granting them exclusive rights to make, use, or sell their innovations (Andries & Faems, 2013; Hall et al., 2013). Besides, patents are used for strategic purposes (Blind et al., 2006; Torrisi et al., 2016). These purposes include signaling to the capital market and blocking competitors (Blind et al., 2022). Although patenting positively affects a firm's innovation performance, the strategic aspect of patenting blurs this effect (Blind et al., 2022). Using patenting as a strategy to improve their positions relative to competitors, firms might sacrifice their returns from innovation activities (Blind et al., 2022). Compared to research disclosed in scientific publications, innovations protected by patents are less likely to spill over (Arora et al., 2021).

2.3. Data Breaches and Spillovers

For many years, malicious actors have been conducting cyber intrusions into firms' commercial networks, targeting firms' intellectual assets and other confidential business information (Basuchoudhary & Searle, 2019). The technologies that have stimulated business and economic growth by making it easier to access, store, and publish confidential information, have put vital assets such as intellectual property at risk (Government of the United States, 2013). A successful cyberattack reduces firms' competitiveness due to data losses and IP becoming publicly available (Anderson et al., 2013; Gordon & Loeb, 2002; Lagazio et al., 2014). IP theft may have insidious and longer-term negative effects on affected firms compared to other types of cyberattacks (Andrijcic & Horowitz, 2006). Therefore, IP theft can be considered a crucial strategic concern for firms (Basuchoudhary & Searle, 2019).

Corporate strategies, such as innovation, are important for firms to gain and retain a competitive advantage (Wang et al., 2024), and as such, identifying the drivers of corporate strategies has been an area of interest for researchers (e.g., Aghion et al., 2013; Bernstein, 2015; Cornaggia et al., 2015; He & Tian, 2013). Wang et al. (2024) found that firms innovation inputs (i.e., R&D expenses) and outputs (i.e., patents and patent citations) are negatively associated with cybersecurity risks. In a recent study, Lattanzio and Ma (2023) found that following increased cybersecurity threats, firms redesign their corporate innovation and appropriation strategies to safeguard their intellectual capital. For example, an increase in cybersecurity disclosure in 10-K reports is associated with firms filing for "simpler" patents to speed up their innovation cycle and rely less on trade secrets (Lattanzio & Ma, 2023). Investigating firm innovation levels after a cyberattack, He et al. (2020) found a 10% decrease in R&D spending one year after a cybersecurity breach. Further, patents decreased two years following the breach (He et al., 2020).

While research on the effect of data breaches on corporate decisions, such as appropriation strategies, is becoming more prevalent (Wang et al., 2024), the effect of peer breaches on these decisions is yet to be examined. Data breaches clearly affect the breached firms, yet their effects are not isolated to these firms (Ashraf, 2022; Garg, 2020; Islam et al., 2022). Following a breach, not only breached firms, but also non-breached peer firms and suppliers of breached firms increase their cash holdings (Garg, 2020). Further, non-breached firms seem to experience an increase in their abnormal trading volumes compared to their breached peers (Islam et al., 2022). A decrease in future internal control material weaknesses was observed for non-breached firms following a peer firm's breach (Ashraf, 2022). Further, non-breached firms take real actions and employ cybersecurity experts on their management team to improve their governance over cyber risk (Ashraf, 2022).

2.4. Hypothesis Development

Prior research has examined the effect of breaches on firms' innovation strategies, however, and to the best of our knowledge, no study has looked at the effect of peer breaches on firm innovation strategies, and specifically scientific disclosures. How firms respond to increased data breach risks through scientific disclosures has two different perspectives to consider.

IP thieves primarily target corporate secrets, rather than IP already in the public domain, such as patents and trademarks. When peer firms get breached, a firm's assessment of the presence of possible undetected IP leakages within its premises may also increase. Since R&D activities are very costly, perpetrators are mostly interested in the proprietary intellectual assets that can be monetized quickly. For example, they may be interested in trade secrets that could include drug trial data, a manufacturing process, a paint formula, or a unique design (Gelinne et al., 2016). To

prevent the attackers from patenting the breached IP, a firm can strategically disclose pre-patenting proprietary information in scientific publications. Further, by disclosing the research publicly, firms can decrease the benefits that attackers can get from their data breaches, thus decreasing their possibility of experiencing a data breach. Therefore, when the threats of data breaches increase, firms may increase their scientific disclosure to avoid being attacked.

As an alternative perspective, scientific publications are a good signal of firms' R&D competencies and internal value of trade secrets (Baruffaldi et al., 2024; Rotolo et al., 2022). Although these signals improve the firms' reputation and enable potential academic and industrial partners, they can also attract IP thieves who might interpret publications as a signal of more valuable trade secrets safeguarded within the firm (Ettredge et al., 2018). As these reflect a higher firm value, the firm will become an interesting target for hackers and will then be more vulnerable to all types of data breaches. In this case, managers may decide to decrease scientific publications to avoid being attacked (Ettredge et al., 2018). Any type of breach experienced by a peer firm may signal material weaknesses in the firm's security infrastructure, may lead to IP theft, and may influence managerial decisions. As no definitive direction can be predicted ex-ante, we present our hypothesis as H1 in the null form.

H1: Changes in data breach risks do not affect firms' scientific publications.

3. Data and Research Design

3.1. Sample Construction

To build our sample, we start with Compustat data from 2004 to 2020 and merge this data with Gao et al. (2021) historical location data, Audit Analytics data breach data, Arora et al.

(2024a, 2024b, 2021) scientific publication data and Kogan et al. (2017) patent data. Furthermore, we merge this data with CRSP daily stock data and I/B/E/S summary data. The sample period is determined by the start (2004) and the end of the data breach data (2020).

To ensure a clean identification we exclude all firms that have experienced a data breach firsthand over our sample period since these firms might directly react to the data breach. Furthermore, to ensure that the firms in our sample can adjust the number of scientific publications we limit the sample to firms that are research active. We define firms as research active if they have positive R&D expenses in year t and publish at least one scientific publication in year t. Finally, we exclude singletons to ensure unbiased standard errors (Breuer & deHaan, 2024), and end up with 6,933 observations. Table 1 presents the description of the sample construction.

[Insert Table 1]

3.2. Outcome Variable: Scientific Publications

We use firms' scientific publications as our outcome variable because they convey important R&D information and firms have a substantial degree of control over these publications. We use Arora et al. (2024a, 2024b, 2021) data to construct this measure. *Scientific publications* is the number of scientific publications of firm i in year t+1 scaled by the R&D expenses in year t. By scaling the variable by R&D expenses we can interpret any change in the number of scientific publications as a change in scientific disclosure since we hold the input constant. Figure 1 shows an example of a scientific publication by Facebook.

[Insert Figure 1]

3.3. Treatment Variable: Data Breach Risk

To measure the risk of becoming the target of a data breach we use the observed breaches in the same industry[15] in the same state as firm i. We develop the *Data breach risk* measure which reflects the number of data breaches in the same state and industry as firm i in year t, excluding breached firms. Successful data breaches affect not only breached firms, but also their rivals (Garg, 2020; Kelton & Pennington, 2020). As such, we use peer breaches, which are staggered events exogenous to the non-breached firms; this way we mitigate endogeneity concerns (Lattanzio & Ma, 2023).

Table 2 shows the number of data breaches reported across the years. The increasing number of data breaches over time highlights their growing prevalence. In addition, the growing number of reported breaches is also consistent with the introduction of Data Breach Disclosure Laws that require firms to disclose data breaches (Ashraf, 2022; Obaydin et al., 2024).

[Insert Table 2]

Tables 3 and 4 present the number of data breaches reported across states and industries, respectively. Overall, our sample includes 704 data breaches. The most affected states are California (133, 18.89%), New York (79, 11.22%), and Texas (55, 7.81%). The most affected industries are Business services (135, 19.18%), Communications (57, 8.10%), and Chemical & allied products (37, 5.26%).

[Insert Tables 3 and 4]

[15] Industry is defined using the 2-digit Standard Industrial Classification (SIC2) code.

3.4. Research Design

To examine the effect of data breach risk on scientific publications, we estimate the following OLS regression:

$$R\&D\ Output_{i,t+1} = \beta_0 + \beta_1\ Data\ breach\ risk_{i,t} + \Sigma\ Controls_{i,t} + \Sigma\ Firm + \Sigma\ Industry\ x\ state + \Sigma\ State\ x\ year + \Sigma\ Industry\ x\ year + \varepsilon \quad (1)$$

In the main analysis, *R&D Output* refers to *Scientific publications.* In an additional analysis, *R&D Output* refers to *Patents* or *Trade secrets.* Following previous studies, we include several control variables to control for firm characteristics that might be correlated with our outcome variable and treatment variable. *Analyst following* has been associated with both a decrease (He & Tian, 2013) and an increase in firms' innovation (Guo et al., 2019), which we measure through *Scientific publications*. Firm size (i.e., *Size*) affects firms' innovative activities (Vaona & Pianta, 2008); and firm profitability (i.e., *ROA*) has been positively associated with innovation (Love et al., 2009). Additionally, larger firms and firms with higher profitability are more likely to experience breaches (Higgs et al., 2016; Li et al., 2018; Li et al., 2024; Say & Vasudeva, 2020). The length of 10-K filings (*10-K filesize)* has been identified as a measure for company innovation (e.g., Nousiainen et al., 2024) and has been associated with subsequent breaches (Ettredge et al., 2018). Firms' *Cyber defense* and *Cyber vulnerability* capture 10-K disclosures related to firms' cyber defense capabilities and vulnerability to cyberattacks (Ettredge et al., 2018; Gordon et al., 2006, 2010; Lawrence et al., 2018) which could affect the extent of firms' scientific publications.

We include firm fixed effects to control for time-invariant firm characteristics (e.g., IBM might disclose more than GM). We include industry x state fixed effects to control for all

fundamentals within the industry and state. We include state x year fixed effects to control for all fundamentals within the state in year t. We include industry x year fixed effects to control for all fundamentals within the industry in year t. Consequently, this specification constitutes a generalized difference-in-difference approach allowing us to use changes in the treatment variable to identify its effect (Bertrand et al., 2004). Since the location of a firm can vary over time, the industry x state fixed effects do not subsume the firm fixed effects. By clustering residuals by industry-state (i.e., level of treatment), we compute standard errors that allow each industry-state to have its own unobserved effect on scientific publications (Abadie et al., 2023; Petersen, 2008).

4. Results

4.1. Descriptive Statistics

Table 5 presents our descriptive statistics for the period 2004-2020. The firms in our sample issue on average 16.152 *Scientific publications* per year, file an average of 24.405 *Patents*, and mention 2.507 *Trade secrets* terms per year in the 10-K. *R&D* expenses are on average $149.321 million per year. The mean of *Data breach risk* is 0.494[16], which implies a moderate variation in the number of security breaches that are observed (this can be translated to every 2 firms having a breached peer firm). The variation in *Data breach risk* ranges from 0 to 13 observed security breaches in the same industry and state as firm i.

Our sample includes 704 breaches, a value similar to that of other studies (e.g., Asthana et al., 2021; Li et al., 2024). The means of the number of data breaches that involve access to

[16] The number of treated observations is 1,127, which represents 16.26% of our sample. No effect can be identified if we conduct the analysis on the treated observations only. In other words, we need both the treated (*Data breach risk*>0) and the non-treated observations (*Data breach risk*=0).

personal information *(Data breach risk - Personal information)*, financial information *(Data breach risk - Financial information)*, other information *(Data breach risk - Other information)*, and non-disclosed information *(Data breach risk – Information not disclosed)* are 0.285, 0.031, 0.113, and 0.065, respectively. The means of the number of data breaches that involve malware *(Data breach risk - Malware)*, unauthorized access *(Data breach risk – Unauthorized access)*, phishing *(Data breach risk - Phishing)*, ransomware *(Data breach risk - Ransomware)*, misconfiguration *(Data breach risk - Misconfiguration)*, and non-disclosed cyberattack *(Data breach risk – Cyberattack not disclosed)* are 0.054, 0.078, 0.099, 0.062, 0.067, and 0.137, respectively. The mean for *Analyst Following* is 3.727, firm *Size* is 2,356.160, return on assets -0.148 *(ROA)*, and *10-K file size* is 9,647,090.439. The mean for *Cyber defense* terms is 1.237 and for *Cyber vulnerability* terms is 0.230.

Table 6 presents the correlation between our variables. Scientific publications are positively associated ROA and cyber defense, and negatively associated with all categories of data breach risk and accessed information, analyst following, firm size, 10-K file size, and cyber vulnerability. A full description of all variables is provided in Appendix A.

[Insert Tables 5 and 6]

4.2. Main Analysis: Data Breach Risk and Scientific Publications

In Table 7, we examine the association between data breach risk and scientific publications. In column (1), we show the estimate of Equation 1 with fixed effects. The result of column (1) suggests that firms that experience an increase in *Data breach risk* of one unit decrease the number of scientific publications by 0.0408 per million dollars of R&D expenses in contrast to firms that do not experience a change in data breach risk.

To ensure that these results are not driven by other characteristics, we include a vector of control variables in column (2). These variables account for characteristics that might be correlated with both the data breach risk and the decision to disclose scientific publications. Similar to column (1), the result of column (2) indicates that firms that experience an increase in *Data breach risk* of one unit decrease the number of scientific publications by 0.0413 per million dollars in R&D expenses, respectively, in contrast to firms that do not experience a change in security data breach risk. The coefficient estimates for *Data breach risk* are negatively and statistically significant at the $p \leq 0.05$ level across all specifications. The findings suggest that firms' publishing of less scientific research when exposed to higher breach risks acts as a risk-mitigation strategy to protect their IP. Along the lines of Ettredge et al. (2018), we argue that as publishing reveals valuable IP to potential attackers, firms tend to issue less scientific publications in risky environments to avoid experiencing a breach. In sum, the results of our analysis underline that firms decrease the number of scientific publications in response to the increase in data breach risk, leading to the rejection of the hypothesis.

[Insert Table 7]

4.3. Verification of the Data Breach Risk Measure: Data Breach Risk and Patents

To ensure that our data breach risk measure behaves in the same way as the data breach risk measure used in other studies (e.g., Florackis et al., 2023; Lattanzio & Ma, 2023), we test its association with the number of patents. Lattanzio and Ma (2023) show that if firms' exposure to data breach risk increases, firms seek to protect their firms' intellectual capital through patents. Consequently, in Table 8, we examine the association between data breach risk and the number of patents by estimating Equation 1 with the number of patents in year t+1

scaled by the R&D expenses in year t as outcome variable. Column (1) shows the coefficient estimate just including fixed effects. We add control variables in column (2). The coefficient estimates for *Data breach risk* are positive and statistically significant at the $p \leq 0.05$ level across all specifications. The results suggest that firms that experience an increase in data breach risk of one unit, increase the number of patents by 0.017 per million dollars in R&D expenses, in contrast to firms that do not experience a change in data breach risk. These results are consistent with Lattanzio and Ma (2023) and underline the reliability of our data breach risk measure.

[Insert Table 8]

4.4. Robustness Tests

To rule out the possibility that our results are driven by our regression design we use a randomization inference to calculate Fisher's exact p-values (Bind & Rubin, 2020). This approach is advantageous as it is non-parametric and hence does not require a correctly specified regression model (Imbens & Rubin, 2015). For example, this approach generates p-values that are correct even if the correct model is non-linear.

To calculate Fisher's exact p-values, we start by randomly reassigning our treatment variable across our sample to estimate a pseudo-coefficient. We then repeat this process 100,000 times and count how often these pseudo-coefficients exceed the coefficient of our main specification. In Figure 2, we show the distribution of pseudo-coefficients. We find that less than 1% of these pseudo-coefficients exceed the original coefficient which rules out the possibility that our results are driven by the design of our regression model.

[Insert Figure 2]

To rule out the alternative explanation that our results are driven by correlated omitted variables we follow Call et al. (2018) and gauge the bias introduced by correlated omitted variables using the theoretical framework developed by Altonji et al. (2005) and Oster (2019). This analysis leverages a proportional selection relationship that considers both changes in coefficients (between models with and without controls) and shifts in R-squared values to detect omitted variable bias. Oster (2019) introduces a coefficient of proportionality, δ, which incorporates the movement of the coefficient of interest and the explanatory power (R-squared) of linear regression models with and without controls.

Oster (2019) emphasizes the importance of R-squared changes in evaluating unobservable selection. To estimate this selection, researchers rely on R_{max}, which represents the R-squared from a hypothetical regression that includes the outcome variable, the treatment variable, observed controls, and unobserved controls. By comparing R_{max} and the differences in coefficients and R-squared values between OLS regressions with and without controls, one can calculate δ to assess the robustness of the treatment effect against unobservable selection.

Oster (2019) advises researchers to report the value of δ at which the coefficient of interest, β, becomes zero. At $R_{max} = 0.8$, we calculate a δ of 1.60 for our main analysis. This positive delta suggests that including the unobservable factors would increase our negative coefficient estimate and hence make our results stronger. This finding strengthens our confidence that endogeneity due to unobserved heterogeneity is unlikely to account for our results.

To test the parallel trend assumption, we examine whether our treatment variable *Data breach risk* is associated with our outcome variable *Scientific publications* in t-1 and t-2 (i.e., before the data breaches occurred). A significant coefficient would indicate that firms in the treatment group differ systematically from those in the control group, suggesting a violation of the parallel trend assumption. However, the insignificant coefficients in columns (1) and (2) of Table 9 do not point to any such violation, thereby supporting the parallel trend assumption.

Furthermore, to better understand the duration of the treatment effect, we examine the effect of *Data breach risk* on *Scientific publications* in t+2 and t+3. We again find insignificant coefficients in columns (3) and (4) of Table 9, suggesting that any effect is short term, likely an immediate reaction to the observed data breach within the industry and location, rather than a long-term effect.

[Insert Table 9]

4.5. Heterogenous Treatment Effects: Data Breach Characteristics

To understand whether our results are driven by specific categories (i.e., are stronger or weaker for specific categories), we examine if the effect of data breach risk varies across different types of accessed information during the breach. Columns (1) to (4) of Table 10 include the coefficient estimates for data breaches that involve private information (*Data breach risk - Personal information*), financial information (*Data breach risk - Financial information*), other information (*Data breach risk - Other information*), and data breaches that do not disclose the information accessed (*Data breach risk - Information not disclosed)*.

Additionally, to understand whether a specific breach type drives our results, we conduct our analysis for each type. Columns (5) to (10) of Table 10 include the coefficient estimates for data breaches that involve malware (*Data breach risk - Malware*), unauthorized access (*Data breach risk - Unauthorized access*), phishing (*Data breach risk - Phishing*), ransomware (*Data breach risk - Ransomware*), misconfiguration (*Data breach risk - Misconfiguration*), and data breaches that do not disclose the type of cyberattack (*Data breach risk - Cyberattack not disclosed*)[17].

Results are significant for the personal information category but not for the others. This category includes sensitive data such as emails and passwords that can be used to monitor employees' activities, access companies' systems, and eventually access IP information. However, any data breach—regardless of whether it involves IP—signals a security deficiency and may directly influence managerial judgment. Although the results are not significant for other categories, the insignificant coefficient estimates should be interpreted with caution, as they might be due to insufficient statistical power rather than the absence of an effect. The statistically significant findings are associated with the more frequently occurring categories.

Additionally, the results seem to be driven by certain types of data breaches, namely *Phishing* and *Misconfiguration*, and *Cybersecurity not disclosed*. Phishing involves malicious actors, disguised as trustworthy entities in electronic communication, attempting to trick individuals to obtain sensitive information, such as account credentials. Misconfiguration refers to a flawed setup of network devices, computer systems, or software applications which create security vulnerabilities. Both phishing and misconfiguration facilitate attackers' access

[17] Please refer to Appendix B for a full description of these data breach characteristics.

to a company's system, whereby proprietary information such as IP could be stolen. The non-disclosed cybersecurity attack could refer to any breach category that is not yet publicized. The insignificant coefficients on the other categories should be interpreted cautiously, because similar to the type of information accessed, they might be due to insufficient statistical power, and not due to the absence of an effect.

[Insert Table 10]

4.6. Alternative Outcome Variable: Trade Secrets

In Table 11, we examine if the effect of data breach risk extends to trade secret term disclosure in the 10-Ks. We repeat our main analysis using the number of trade secret terms disclosed in year t+1 scaled by the R&D expenses in year t as an alternative outcome variable. We conduct this analysis to understand whether data breach risk affects firms' disclosures, besides those related to scientific publications. Ettredge et al. (2018) find that companies that disclose more content related to trade secrets in their 10-K, subsequently have a higher probability of experiencing a breach relative to firms that do not disclose as much. To understand the overall consequences of data breach risk, we test whether firms that are at a higher risk of experiencing a breach would disclose less trade secret-related content in their 10-K to protect themselves from falling victims to cyberattacks. The results of columns (1) and (2) of Table 11 suggest that firms that experience an increase in *Data breach risk* of one unit decrease the number of trade secret terms in their 10-K by 0.0141 to 0.0146 per million dollars in R&D expenses in contrast to firms that do not experience a change in data breach risk. These results are consistent with the argument that the disclosure of trade secrets increases the risk of becoming the target of a data breach (Ettredge et al., 2018).

[Insert Table 11]

5. Discussion of Socio-Economic Consequences

To understand the indirect welfare effects of data breaches, we conduct back-of-the-envelope calculations of the socio-economic value of the foregone scientific publications. Socio-economic consequences of a reduction in scientific publications encompass slower technological advancements, innovation, and productivity, diminished public knowledge and scientific output, reduced competitiveness and economic growth, and limited reputational and prestige enhancement (Morretta et al., 2022). We start with our estimated reduction in scientific publications of 0.0413 per million dollars in R&D expenses. We multiply this reduction with the average R&D expenses in our sample of $149.321 million and the average number of firms per state and industry of 3.9 (untabulated). Finally, we use the estimates of Morretta et al. (2022)[18] to approximate the socio-economic value of the foregone scientific publications in our sample. They estimate an average value of $31,833[19] per scientific publication. To compute the social value of a publication, they conduct a cost-benefit analysis where they claim that the marginal social value is equal to at least the cost of producing the publication. They estimate the marginal cost of a scientific publication by taking the first derivative of the total cost, which

[18] We use Morretta et al. (2022) estimates because computing the value of a scientific publication is out of the scope of this paper.

[19] Calculated as: (cost of all publications in euros/number of publications) * exchange rate from euro to dollar on 24/3/2025 = 36,317,985 euros/ 1,235 publications * 1.0825. Publications pertain to several subject areas: Earth and Planetary Sciences, Computer Science, Engineering, Physics and Astronomy, Mathematics, Materials Science, Environmental Science, Social Sciences, Agricultural and Biological Sciences, and Other. Costs are taken from a sample of heterogeneous researchers and countries (Morretta et al., 2022).

includes research and publication costs. Taken together, each data breach creates on average 24.05[20] foregone scientific publications and a loss of socio-economic value of $765,584[21].

Yet, one should be careful with the interpretation of this estimate because there are several caveats. First, it is unclear whether the information contained in the scientific publications is disclosed at a later point in time. Second, it is ambiguous whether the socio-economic value foregone by the lower number of scientific publications is offset by the socio-economic value of the higher number of patents. Third, it is unclear whether the data breach risk also affects firms in other industries and states.

6. Conclusion

We investigate whether data breach risk affects a firm's decision to issue scientific publications. We take advantage of the data breaches of peer firms as exogenous variation in data breach risk. We find that firms decrease their number of scientific publications in response to the increase in data breach risk. Our findings suggest that to avoid becoming an interesting target for cyber attackers, firms might issue less scientific publications to indirectly signal a lack of intellectual property. As such, reduced publication can be viewed as a potential risk-mitigation strategy when operating in a cyber-risky environment. In an additional analysis, we find that firms also decrease their trade secret disclosures and increase the number of patents filed to protect themselves from potential attackers. Taken together, our results provide evidence that firms adapt their IP strategies when faced with high data breach risks to protect

[20] Calculated as the coefficient of the decrease in scientific publications per million $ R&D expenses from our main results * average R&D expenses in our sample * average number of firms per state-industry-year = 0.0413 * 149.321 * 3.9.

[21] Calculated as 24.05 * 31,833.

their proprietary information. It is worth noting that firms' withholding of information, as a response to data breach risks, may have negative externalities. These findings are important in light of efforts to better understand the incentives to disclose information, especially in the innovation context. Furthermore, these findings are crucial to understanding the welfare implications of data breaches since the decrease in scientific publications is an additional indirect welfare loss beyond the direct negative consequences.

Our results should be evaluated in light of several limitations. First, we do not include scientific publications from all publication outlets. Second, we cannot capture whether the decrease in scientific publications leads to decreased knowledge in the economy; we look at knowledge *output* not at knowledge *per se*. Third, our breach risk measure has a mean of 0.494, however this is justifiable as the percentage of breached firms in previous studies is around 3% (e.g., Asthana et al., 2021; Li et al., 2024). In our study the percentage of treated firms is 16.26%. Fourth, due to data limitations, we cannot verify whether IP was stolen in every breach. The SEC Commissioner Robert J. Jackson Jr. claims that policy makers and firms face numerous challenges due to the lack of a representative dataset for cybersecurity breaches. Future research could test whether the results would differ in the absence of data breach disclosure regulations. That is, whether data breach disclosure regulations affect the number of scientific publications through more disclosed data breaches. Additionally, future studies might use alternative measures for data breach risks, when more details on breaches become available. Lastly, it could be interesting to explore the direct consequences of diminished scientific publications.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1):1–35.

Aghion, P., Van Reenen, J., and Zingales, L. (2013). Innovation and institutional ownership. *American Economic Review*, 103(1):277–304.

Alexy, O., George, G., and Salter, A. J. (2013). Cui bono? The selective revealing of knowledge and its implications for innovative activity. *Academy of Management Review*, 38(2):270–291.

Almeida, P., Hohberger, J., and Parada, P. (2011). Individual scientific collaborations and firm-level innovation. *Industrial and Corporate Change*, 20(6):1571–1599.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184.

Anderson, R., Barton, C., Böhme, R., Clayton, R., Van Eeten, M. J., Levi, M., Moore, T., and Savage, S. (2013). Measuring the cost of cybercrime. *The Economics of Information Security and Privacy*, pages 265–300.

Andries, P. and Faems, D. (2013). Patenting activities and firm performance: does firm size matter? *Journal of Product Innovation Management*, 30(6):1089–1098.

Andrijcic, E. and Horowitz, B. (2006). A macro-economic framework for evaluation of cyber security risks related to protection of intellectual property. *Risk Analysis*, 26(4):907–923.

Arora, A., Belenzon, S., Cioaca, L., Sheer, L., Shin, H. M., and Shvadron, D. (2024a). Discern 2.0: Duke innovation & scientific enterprises research network dataset.

Arora, A., Belenzon, S., Cioaca, L., Sheer, L., and Shvadron, D. (2024b). Back to the future: Are big firms regaining their scientific and technological dominance? Evidence from discern 2.0.

Arora, A., Belenzon, S., and Patacconi, A. (2018). The decline of science in corporate R&D. *Strategic Management Journal*, 39(1):3–32.

Arora, A., Belenzon, S., and Sheer, L. (2021). Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, 111(3):871–898.

Ashraf, M. (2022). The role of peer events in corporate governance: Evidence from data breaches. *The Accounting Review*, 97(2):1–24.

Asthana, S. C., Kalelkar, R., and Raman, K. (2021). Does client cyber-breach have reputational consequences for the local audit office? *Accounting Horizons*, 35(4):1–22.

Baker, S. and Mezzetti, C. (2005). Disclosure as a strategy in the patent race. *Journal of Law and Economics*, 48(1):173–194.

Bar-Gill, O. and Parchomovsky, G. (2003). The value of giving away secrets. *Vanderbilt Law Review*, 89:1857.

Baruffaldi, S., Simeth, M., and Wehrheim, D. (2024). Asymmetric information and R&D disclosure: evidence from scientific publications. *Management Science*, 70(2):1052–1069.

Basuchoudhary, A. and Searle, N. (2019). Snatched secrets: Cybercrime and trade secrets modelling a firm's decision to report a theft of trade secrets. *Computers & Security*, 87:101591.

Bernstein, S. (2015). Does going public affect innovation? *The Journal of Finance*, 70(4):1365–1403.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Bind, M.-A. and Rubin, D. (2020). When possible, report a fisher-exact p value and display its underlying null randomization distribution. *Proceedings of the National Academy of Sciences*, 117(32):19151–19158.

Blind, K., Edler, J., Frietsch, R., and Schmoch, U. (2006). Motives to patent: Empirical evidence from Germany. *Research Policy*, 35(5):655–672.

Blind, K., Krieger, B., and Pellens, M. (2022). The interplay between product innovation, publishing, patenting and developing standards. *Research Policy*, 51(7):104556.

Breuer, M. and deHaan, E. (2024). Using and interpreting fixed effects models. *Journal of Accounting Research*, 62(4):1183–1226.

Call, A. C., Martin, G. S., Sharp, N. Y., and Wilde, J. H. (2018). Whistleblowers and outcomes of financial misrepresentation enforcement actions. *Journal of Accounting Research*, 56(1):123–171.

Chen, J., Henry, E., and Jiang, X. (2023). Is cybersecurity risk factor disclosure informative? Evidence from disclosures following a data breach. *Journal of Business Ethics*, 187(1):199–224.

Cornaggia, J., Mao, Y., Tian, X., and Wolfe, B. (2015). Does banking competition affect innovation? *Journal of Financial Economics*, 115(1):189–209.

Dass, N., Nanda, V., and Xiao, S. C. (2015). Intellectual property protection and financial markets: Patenting vs. secrecy. *Working Paper*.

Ettredge, M., Guo, F., and Li, Y. (2018). Trade secrets and cyber security breaches. *Journal of Accounting and Public Policy*, 37(6):564–585.

Florackis, C., Louca, C., Michaely, R., & Weber, M. (2023). Cybersecurity risk. *The Review of Financial Studies*, *36*(1), 351-407.

Gans, J. S., Murray, F. E., and Stern, S. (2017). Contracting over the disclosure of scientific knowledge: Intellectual property and academic publication. *Research Policy*, 46(4):820–835.

Gao, M., Leung, H., and Qiu, B. (2021). Organization capital and executive performance incentives. *Journal of Banking & Finance*, 123:106017.

Garg, P. (2020). Cybersecurity breaches and cash holdings: Spillover effect. *Financial Management*, 49(2):503–519.

Gelinne, J., Fancher, J., and Mossburg, E. (2016). The hidden costs of an IP breach: Cyber theft and the loss of intellectual property. *Deloitte Review*, 19.

Glaeser, S. (2018). The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. *Journal of Accounting and Economics*, 66(1):163–193.

Glaeser, S. and Lang, M. (2024). Measuring innovation and navigating its unique information issues: A review of the accounting literature on innovation. *Journal of Accounting and Economics*, 101720.

Glaeser, S., Michels, J., and Verrecchia, R. E. (2020). Discretionary disclosure and manager horizon: Evidence from patenting. *Review of Accounting Studies*, 25:597–635.

Glaeser, S. A. and Landsman, W. R. (2021). Deterrent disclosure. *The Accounting Review*, 96(5):291–315.

Godin, B. (1996). Research and the practice of publication in industries. *Research Policy*, 25(4):587– 606.

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

Gordon, L. A., Loeb, M. P., & Sohail, T. (2010). Market value of voluntary disclosures concerning information security. *MIS Quarterly*, 567-594.

Gordon, L. A., Loeb, M. P., Lucyshyn, W., & Sohail, T. (2006). The impact of the Sarbanes-Oxley Act on the corporate disclosures of information security activities. *Journal of Accounting and Public Policy*, *25*(5), 503-530.

Guo, B., Pérez-Castrillo, D., & Toldrà-Simats, A. (2019). Firms' innovation strategy under the shadow of analyst coverage. *Journal of Financial Economics*, *131*(2), 456-483.

Hagedoorn, J. and Cloodt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators? *Research Policy*, 32(8):1365–1379.

Hall, B. H., Helmers, C., Rogers, M., and Sena, V. (2013). The importance (or not) of patents to UK firms. *Oxford Economic Papers*, 65(3):603–629.

Hayter, C. S. and Link, A. N. (2018). Why do knowledge-intensive entrepreneurial firms publish their innovative ideas? *Academy of Management Perspectives*, 32(1):141–155.

He, C. Z., Frost, T., and Pinsker, R. E. (2020). The impact of reported cybersecurity breaches on firm innovation. *Journal of Information Systems*, 34(2):187–209.

He, J. J. and Tian, X. (2013). The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics*, 109(3):856–878.

Higgs, J. L., Pinsker, R. E., Smith, T. J., & Young, G. R. (2016). The relationship between board-level technology committees and reported security breaches. *Journal of Information Systems*, *30*(3), 79-98.

Huang, K. G.-L. (2016). Uncertain intellectual property conditions and knowledge appropriation strategies: Evidence from the genomics industry. *Industrial and Corporate Change*, page dtw015.

IBM Security (2024). Cost of a data breach report 2024.

Imbens, G. W. and Rubin, D. B. (2015). *Fisher's Exact P-Values for Completely Randomized Experiments*, pages 57–82. Cambridge University Press.

Islam, M. S., Wang, T., Farah, N., and Stafford, T. (2022). The spillover effect of focal firms' cybersecurity breaches on rivals and the role of the cio: Evidence from stock trading volume. *Journal of Accounting and Public Policy*, 41(2):106916.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., and Stulz, R. M. (2021). Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics*, 139(3):719–749.

Kelton, A. S., & Pennington, R. R. (2020). Do voluntary disclosures mitigate the cybersecurity breach contagion effect? *Journal of Information Systems*, *34*(3), 133-157.

Kim, J. and Valentine, K. (2021). The innovation consequences of mandatory patent disclosures. *Journal of Accounting and Economics*, 71(2-3):101381.

Klasa, S., Ortiz-Molina, H., Serfling, M., and Srinivasan, S. (2018). Protection of trade secrets and capital structure decisions. *Journal of Financial Economics*, 128(2):266–286.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, 132(2):665–712.

Lagazio, M., Sherif, N., and Cushman, M. (2014). A multi-level approach to understanding the impact of cyber crime on the financial sector. *Computers & Security*, 45:58–74.

Lattanzio, G. and Ma, Y. (2023). Cybersecurity risk and corporate innovation. *Journal of Corporate Finance*, 82:102445.

Lawrence, A., Minutti-Meza, M., & Vyas, D. (2018). Is operational control risk informative of financial reporting deficiencies? *Auditing: A Journal of Practice & Theory*, *37*(1), 139-165.

Leuz, C. and Wysocki, P. D. (2016). The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research*, 54(2):525–622.

Li, H., No, W. G., & Wang, T. (2018). SEC's cybersecurity disclosure guidance and disclosed cybersecurity risk factors. *International Journal of Accounting Information Systems, 30*, 40-55.

Li, H., Sun, Z., & Huang, F. (2024). The impact of audit office cybersecurity experience on nonbreach client's audit fees and cybersecurity risks. *Journal of Information Systems, 38*(1), 177-206.

Lichtman, D., Baker, S., and Kraus, K. (2000). Strategic disclosure in the patent system. *Vanderbilt Law Review*, 53:2175.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

Love, J. H., Roper, S., & Du, J. (2009). Innovation, ownership and profitability. *International Journal of Industrial Organization*, *27*(3), 424-434.

Martens, T. (2023). The disclosure function of the us patent system: Evidence from the ptdl program and extreme snowfall. *Review of Accounting Studies*, 28(1):237–264.

Morretta, V., Vurchio, D., and Carrazza, S. (2022). The socio-economic value of scientific publications: The case of earth observation satellites. *Technological Forecasting and Social Change*, 180:121730.

Nelson, R. R. (1990). Capitalism as an engine of progress. *Research Policy*, 19(3):193–214.

Nerkar, A. and Roberts, P. W. (2004). Technological and product-market experience and the success of new product introductions in the pharmaceutical industry. *Strategic Management Journal*, 25(8-9):779–799.

Nousiainen, E., Ranta, M., Ylinen, M., & Järvenpää, M. (2024). Using machine learning and 10-K filings to measure innovation. *Accounting & Finance, 64*(4), 3211-3239.

Obaydin, I., Xu, L., and Zurbruegg, R. (2024). The unintended cost of data breach notification laws: Evidence from managerial bad news hoarding. *Journal of Business Finance & Accounting*.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Pellens, M. and Della Malva, A. (2018). Corporate science, firm value, and vertical specialization: Evidence from the semiconductor industry. *Industrial and Corporate Change*, 27(3):489–505.

Penders, B. and Nelis, A. P. (2011). Credibility engineering in the food industry: linking science, regulation, and marketing in a corporate context. *Science in Context*, 24(4):487–515.

Pénin, J. (2007). Open knowledge disclosure: An overview of the evidence and economic motivations. *Journal of Economic Surveys*, 21(2):326–347.

Petersen, M. A. (2008). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1):435–480.

Polidoro Jr, F. and Theeke, M. (2012). Getting competition down to a science: The effects of technological competition on firms' scientific publications. *Organization Science*, 23(4):1135–1153.

Robert J. Jackson Jr. (2018). Statement on Commission Statement and Guidance on Public Company Cybersecurity Disclosures. U.S. Securities and Exchange Commission. Retrieved March 23, 2025, from https://www.sec.gov/newsroom/speeches-statements/statement-jackson-2018-02-21

Rotolo, D., Camerani, R., Grassano, N., and Martin, B. R. (2022). Why do firms publish? A systematic literature review and a conceptual framework. *Research Policy*, 51(10):104606.

Say, G., & Vasudeva, G. (2020). Learning from digital failures? The effectiveness of firms' divestiture and management turnover responses to data breaches. *Strategy Science*, *5*(2), 117-142.

Simeth, M. and Cincera, M. (2016). Corporate science, innovation, and firm value. *Management Science*, 62(7):1970–1981.

Simeth, M. and Lhuillery, S. (2015). How do firms develop capabilities for scientific disclosure? *Research Policy*, 44(7):1283–1295.

The Wall Street Journal (2020). What is the real cost of a breach? https://www.wsj.com/video/events/what-is-the-real-cost-of-a-breach/31C314D2-1F52-473A-81CB-C867758EC110?mod=WSJvidctr_wsjprocybersecurityexecutiveforum2020_pos2. Accessed: 2025-03-10.

Torrisi, S., Gambardella, A., Giuri, P., Harhoff, D., Hoisl, K., and Mariani, M. (2016). Used, blocking and sleeping patents: Empirical evidence from a large-scale inventor survey. *Research Policy*, 45(7):1374–1385.

Vaona, A., & Pianta, M. (2008). Firm Size and Innovation in European Manufacturing. *Small Business Economics : An Entrepreneurship Journal*, *30*(3), 283–299.

Wang, J., Ho, C. Y. C., and Shan, Y. G. (2024). Does cybersecurity risk stifle corporate innovation activities? *International Review of Financial Analysis*, 91:103028.

Zucker, L. G., Darby, M. R., and Armstrong, J. S. (2002). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. *Management Science*, 48(1):138–153.

Figure 1

Example of a scientific publication

DrawFromDrawings: 2D Drawing Assistance via Stroke Interpolation with a Sketch Database

Yusuke Matsui, *Member, IEEE*, Takaaki Shiratori, *Member, IEEE*, and Kiyoharu Aizawa, *Fellow, IEEE*

Abstract—We present *DrawFromDrawings*, an interactive drawing system that provides users with visual feedback for assistance in 2D drawing using a database of sketch images. Following the traditional imitation and emulation training from art education, DrawFromDrawings enables users to retrieve and refer to a sketch image stored in a database and provides them with various novel strokes as suggestive or deformation feedback. Given regions of interest (ROIs) in the user and reference sketches, DrawFromDrawings detects as-long-as-possible (ALAP) stroke segments and the correspondences between user and reference sketches that are the key to computing seamless interpolations. The stroke-level interpolations are parametrized with the user strokes, the reference strokes, and new strokes created by warping the reference strokes based on the user and reference ROI shapes, and the user study indicated that the interpolation could produce various reasonable strokes varying in shapes and complexity. DrawFromDrawings allows users to either replace their strokes with interpolated strokes (deformation feedback) or overlays interpolated strokes onto their strokes (suggestive feedback). The other user studies on the feedback modes indicated that the suggestive feedback enabled drawers to develop and render their ideas using their own stroke style, whereas the deformation feedback enabled them to finish the sketch composition quickly.

Index Terms—interactive drawing, 2D shape interpolation

1 Introduction

Imitation and emulation of high-quality artistic work represent essential training components in art education [20]. By observing how established artists draw different strokes, students can learn stroke-based interpretations of real scenes, which provides an entry point to establishing a unique interpretation. Consequently, many art teachers encourage students to emulate existing work and stroke patterns, a process facilitated by the wide availability of both professional and hobbyist artistic content on the Internet.

The aim of this paper is to push the traditional imitation and emulation paradigm a step further. In this context, we present *DrawFromDrawings*, an interactive drawing system that provides users with visual feedback for assistance in 2D drawing using a database of sketch images. Whenever the user feels unsatisfied with a sketch, an unsatisfactory region can be marked alongside an associated region of interest (ROI) in a reference sketch from the database (Figs. 1a and 1b). Using a marker position in an *assistance palette*, the user can explore new strokes that are interpolated with their original user strokes and the selected reference strokes (Fig. 1c). The user can integrate interpolated strokes as either *deformation feedback*, which replaces user strokes directly with interpolated strokes, or *suggestive feedback*, which overlays the interpolated strokes as a suggestive template.[1] Using a set of simple user interactions, the sketch evolves naturally (Fig. 1d) towards some satisfactory convergence (Fig. 1e).

The major technical challenge is interactively providing intuitive interpolation for drawing assistance, involving several issues. First, we need to properly handle significant structural differences between user strokes and reference strokes, including differences in length, curvature, and stroke count, as illustrated in Figs. 1a and 1b. Second, user interactions should be simple and intuitive for novice drawers. We need to avoid complicated user interactions such as specifying or refining point/stroke correspondences one-by-one, which might be acceptable for professional artists [4], [32]. Third, we want to be able to refer to rasterized sketch images where much of meta stroke information such as stroke order is lost. A system with this capability will benefit significantly from the vast quantity of sketches available on the Internet. Last but not least, all processes must be executable at an interactive speed to enable instant feedback with the user.

Our technical contribution lies in solving these issues by automatically detecting correspondences of as-long-as-possible (ALAP) stroke segments from strokes within the ROI perimeters. Sketch images in the database are preprocessed and represented as a collection of single curves without branches. ALAP segments and their many-to-many correspondences are then detected based on the geodesic distance along strokes. We also parameterize the ALAP segment interpolation based on the original user strokes,

- *Y. Matsui and K. Aizawa are with the Department of Information and Communication Engineering, The University of Tokyo, Tokyo 113-8654, Japan. E-mail: {matsui, aizawa}@hal.t.u-tokyo.ac.jp.*
- *T. Shiratori is with Oculus Research Pittsburgh, Facebook Inc, Pittsburgh, PA. E-mail: takaaki.shiratori@oculus.com.*

Manuscript received 24 Nov. 2014; revised 3 Feb. 2016; accepted 11 Feb. 2016. Date of publication 14 Apr. 2016; date of current version 26 May 2017. Recommended for acceptance by J. Heer. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TVCG.2016.2554113

1. In this paper, we refer to computing stroke shape changes as *interpolation*, and refer to replacing user strokes with computed stroke changes in the GUI as *deformation*.

Figure 2

Randomization inference

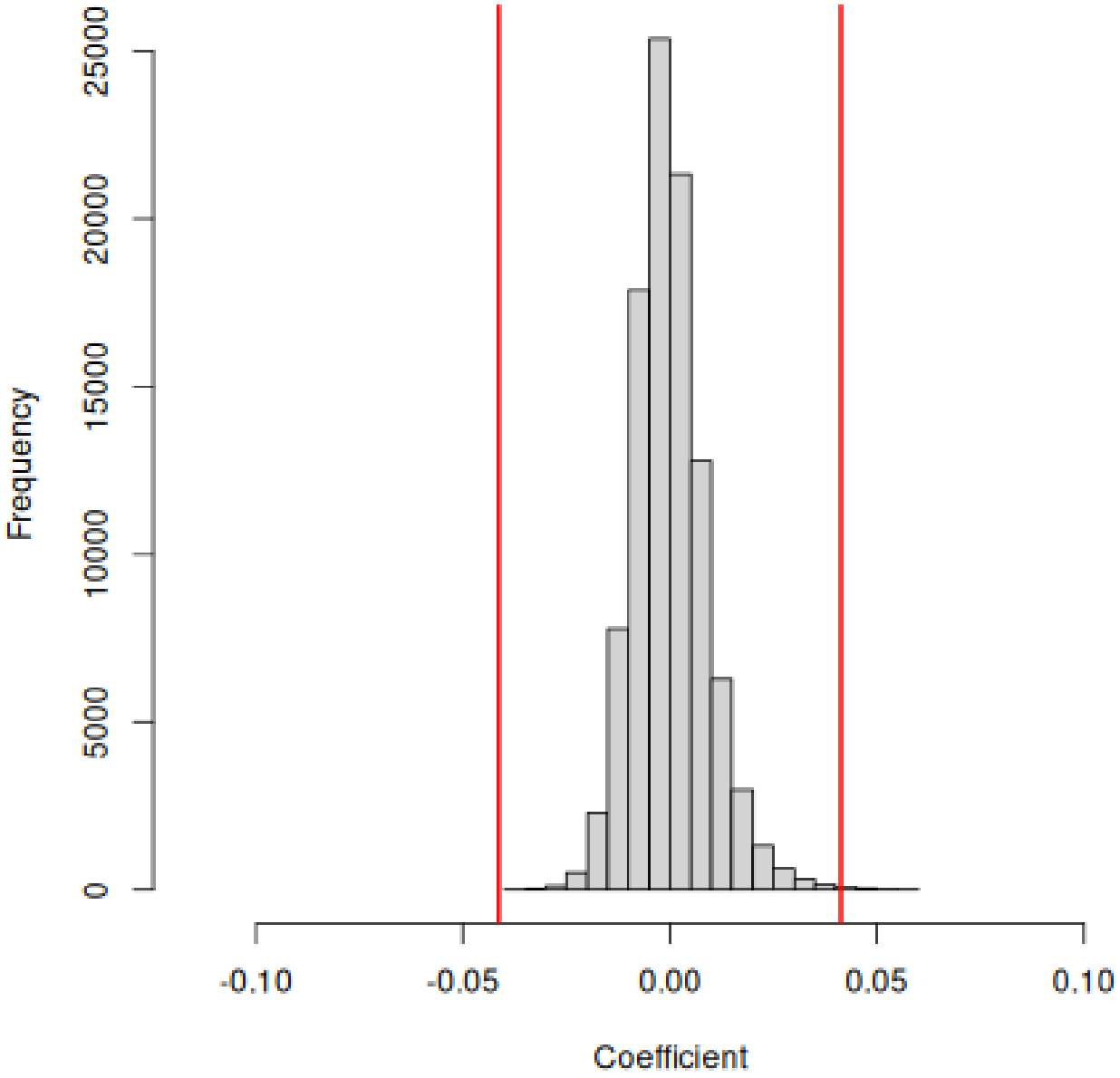


Table 1

Sample construction

Sample construction criteria	Observations
Compustat data from 2004 to 2020	96,643
Merged with additional data	64,340
Less firms that experienced a breach	59,371
Less observations with missing data	56,704
Less firms that are not research active	7,626
Less singletons	6,933

Table 2

Distribution of data breaches across years

Year	#Breaches	Percent
2004	1	0.14
2005	2	0.28
2006	1	0.14
2007	1	0.14
2008	2	0.28
2009	5	0.71
2010	15	2.13
2011	24	3.41
2012	29	4.12
2013	39	5.54
2014	55	7.81
2015	41	5.82
2016	49	6.96
2017	83	11.79
2018	109	15.48
2019	134	19.03
2020	114	16.19
	704	100.00

Table 3

Distribution of data breaches across states from 2004 to 2020

State	#Breaches	Percent
Alaska	1	0.14
Arizona	16	2.27
Arkansas	6	0.85
California	133	18.89
Colorado	10	1.42
Connecticut	10	1.42
Delaware	1	0.14
Florida	40	5.68
Georgia	21	2.98
Hawaii	1	0.14
Idaho	3	0.43
Illinois	41	5.82
Indiana	7	0.99
Iowa	4	0.57
Kansas	5	0.71
Kentucky	12	1.70
Louisiana	3	0.43
Maryland	17	2.41
Massachusetts	30	4.26
Michigan	8	1.14
Minnesota	13	1.85
Missouri	7	0.99
Nebraska	6	0.85
Nevada	11	1.56
New Hampshire	1	0.14
New Jersey	20	2.84
New York	79	11.22
North Carolina	7	0.99
Ohio	15	2.13
Oklahoma	2	0.28
Oregon	4	0.57
Pennsylvania	27	3.84
Puerto Rico	1	0.14
Rhode Island	4	0.57
South Carolina	2	0.28
Tennessee	18	2.56
Texas	55	7.81
Utah	4	0.57
Virginia	31	4.40
Washington	24	3.41
Wisconsin	4	0.57
	704	100.00

Table 4

Distribution of data breaches across industries from 2004 to 2020

Industry	#Breaches	Percent
Metal, mining	1	0.14
Coal mining	1	0.14
Nonmetallic minerals, except fuels	1	0.14
General building contractors	1	0.14
Special trade contractors	2	0.28
Food & kindred products	13	1.85
Apparel & other textile products	5	0.71
Printing & publishing	14	1.99
Chemical & allied products	37	5.26
Petroleum & coal products	1	0.14
Rubber & miscellaneous plastics products	4	0.57
Leather & leather products	1	0.14
Stone, clay, & glass products	2	0.28
Primary metal industries	4	0.57
Fabricated metal products	3	0.43
Industrial machinery & equipment	21	2.98
Electronic & other electric equipment	23	3.27
Transportation equipment	8	1.14
Instruments & related products	23	3.27
Miscellaneous manufacturing industries	6	0.85
Trucking & warehousing	9	1.28
Water transportation	7	0.99
Transportation by air	10	1.42
Transportation services	6	0.85
Communications	57	8.10
Electric, gas, & sanitary services	7	0.99
Wholesale trade & durable goods	10	1.42
Wholesale trade & nondurable goods	9	1.28
Building materials & gardening supplies	6	0.85
General merchandise stores	21	2.98
Food stores	8	1.14
Automative dealers & service stations	6	0.85
Apparel & accessory stores	8	1.14
Furniture & home furnishings stores	3	0.43
Eating & drinking places	25	3.55
Miscellaneous retail	29	4.12
Depository institutions	26	3.69
Non depository institutions	6	0.85
Security & commodity brokers	17	2.41
Insurance carriers	35	4.97
Insurance agents, brokers, & service	5	0.71
Real estate	2	0.28
Holding & other investment offices	15	2.13
Hotels & other lodging places	27	3.84
Personal services	2	0.28
Business services	135	19.18
Auto repair, services, & parking	2	0.28
Motion pictures	2	0.28
Amusement & recreation services	8	1.14
Health services	15	2.13
Educational services	6	0.85
Engineering & management services	9	1.28
	704	100.00

Table 5

Descriptive statistics

	N	Mean	Std.	Min	Max
Scientific publications $_{i,t+1}$	6,933	0.329	1.297	0.000	39.855
Scientific publications (unscaled) $_{i,t+1}$	6,933	16.152	78.122	0.000	2,539.167
Patents $_{i,t+1}$	6,933	0.164	0.477	0.000	12.353
Patents (unscaled) $_{i,t+1}$	6,933	24.405	218.905	0.000	8,883.000
Trade secrets $_{i,t+1}$	6,933	0.159	0.490	0.000	15.000
Trade secrets (unscaled) $_{i,t+1}$	6,933	2.507	2.676	0.000	68.000
R&D $_{i,t}$	6,933	149.321	479.126	0.031	10,895.000
Data breach risk $_{i,t}$	6,933	0.494	1.481	0.000	13.000
Data breach risk - Personal information $_{i,t}$	6,933	0.285	1.106	0.000	12.000
Data breach risk - Financial information $_{i,t}$	6,933	0.031	0.217	0.000	3.000
Data breach risk - Other information $_{i,t}$	6,933	0.113	0.377	0.000	2.000
Data breach risk - Information not disclosed $_{i,t}$	6,933	0.065	0.338	0.000	3.000
Data breach risk - Malware $_{i,t}$	6,933	0.054	0.308	0.000	4.000
Data breach risk - Unauthorized access $_{i,t}$	6,933	0.078	0.378	0.000	3.000
Data breach risk - Phishing $_{i,t}$	6,933	0.099	0.421	0.000	3.000
Data breach risk - Ransomware $_{i,t}$	6,933	0.062	0.322	0.000	2.000
Data breach risk - Misconfiguration $_{i,t}$	6,933	0.067	0.539	0.000	6.000
Data breach risk - Cyberattack not disclosed $_{i,t}$	6,933	0.137	0.513	0.000	5.000
Analyst following $_{i,t}$	6,933	3.727	2.805	0.000	17.638
Size $_{i,t}$	6,933	2,356.160	9,557.628	1.684	155,971.000
ROA $_{i,t}$	6,933	-0.148	0.491	-9.842	3.253
10-K filesize $_{i,t}$	6,933	9,647,090.439	8,408,480.359	171,382.000	106,200,750.000
Cyber defense $_{i,t}$	6,933	1.237	2.949	0.000	51.000
Cyber vulnerability $_{i,t}$	6,933	0.230	1.136	0.000	36.000

Notes: All variables are reported without logarithms. Please refer to Appendix A for a full description of all variables.

Table 6

Correlation analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Scientific publications $_{i,t+1}$		**0.619**	0.030*	-0.028*	**0.340**	**0.097**	**-0.296**	**-0.127**	**-0.120**	**-0.072**	**-0.056**	-0.031*
(2) Scientific publications (unscaled) $_{i,t+1}$	*0.037*		**0.194**	**0.302**	**-0.255**	**0.116**	**0.457**	0.015	0.010	-0.010	0.013	0.030*
(3) Patents $_{i,t+1}$	**0.144**	**0.043**		**0.944**	**-0.135**	0.008	**0.200**	**-0.138**	**-0.070**	**-0.074**	**-0.164**	**-0.158**
(4) Patents (unscaled) $_{i,t+1}$	-0.013	**0.756**	**0.130**		**-0.294**	0.007	**0.402**	**-0.098**	-0.031*	**-0.057**	**-0.152**	**-0.152**
(5) Trade secrets $_{i,t+1}$	**0.362**	**-0.046**	**0.043**	*-0.033*		**0.519**	**-0.746**	**-0.095**	**-0.096**	**-0.047**	*-0.036*	-0.024*
(6) Trade secrets (unscaled) $_{i,t+1}$	-0.005	-0.006	-0.021	-0.023	**0.201**		**0.040**	**0.093**	**0.069**	**0.041**	**0.071**	**0.086**
(7) R&D $_{i,t}$	**-0.053**	**0.636**	-0.000	**0.492**	**-0.091**	-0.008		**0.178**	**0.162**	**0.081**	**0.094**	**0.084**
(8) Data breach risk $_{i,t}$	-0.017	0.020	*-0.033*	-0.009	-0.030*	**0.056**	**0.105**		**0.816**	**0.360**	**0.684**	**0.446**
(9) Data breach risk - Personal information $_{i,t}$	-0.005	0.015	-0.000	0.004	-0.027*	*0.037*	**0.083**	**0.884**		**0.278**	**0.397**	**0.214**
(10) Data breach risk - Financial information $_{i,t}$	-0.008	**0.058**	-0.019	-0.012	-0.019	0.012	**0.100**	**0.482**	**0.253**		**0.221**	**0.269**
(11) Data breach risk - Other information $_{i,t}$	-0.024*	-0.000	**-0.066**	-0.024*	-0.020	**0.056**	**0.047**	**0.630**	**0.322**	**0.179**		**0.326**
(12) Data breach risk - Information not disclosed $_{i,t}$	-0.024*	0.003	**-0.057**	-0.020	-0.008	**0.056**	**0.070**	**0.476**	**0.083**	**0.444**	**0.475**	
(13) Data breach risk - Malware $_{i,t}$	-0.003	0.008	0.021	0.003	-0.021	0.020	**0.080**	**0.591**	**0.531**	**0.495**	**0.280**	**0.222**
(14) Data breach risk - Unauthorized access $_{i,t}$	-0.012	0.000	**-0.057**	-0.021	-0.010	**0.043**	**0.054**	**0.783**	**0.696**	**0.155**	**0.585**	**0.399**
(15) Data breach risk - Phishing $_{i,t}$	-0.017	0.015	**-0.042**	-0.007	-0.013	*0.034*	**0.058**	**0.511**	**0.393**	**0.271**	**0.443**	**0.289**
(16) Data breach risk - Cyberattack not disclosed $_{i,t}$	-0.007	0.029*	**0.046**	0.018	*-0.032*	*0.031*	**0.071**	**0.544**	**0.599**	**0.226**	**0.168**	**0.089**
(17) Data breach risk - Ransomware $_{i,t}$	-0.021	0.006	**-0.056**	-0.020	-0.005	**0.043**	**0.048**	**0.408**	**0.059**	**0.158**	**0.723**	**0.686**
(18) Data breach risk - Misconfiguration $_{i,t}$	-0.001	0.007	*-0.035*	-0.012	-0.020	*0.033*	**0.065**	**0.728**	**0.760**	**0.437**	**0.221**	**0.180**
(19) Analyst following $_{i,t}$	**-0.111**	**0.139**	0.008	**0.132**	**-0.199**	**0.055**	**0.334**	**0.152**	**0.139**	**0.112**	**0.058**	**0.075**
(20) Size $_{i,t}$	-0.030*	**0.625**	0.015	**0.478**	**-0.068**	-0.012	**0.740**	0.021	0.019	**0.053**	-0.007	0.005
(21) ROA $_{i,t}$	0.011	**0.072**	**0.089**	**0.059**	**-0.046**	**-0.046**	**0.123**	0.003	0.024*	**0.041**	-0.051	*-0.033*
(22) 10-K filesize $_{i,t}$	**-0.052**	**0.123**	**-0.065**	**0.065**	**-0.080**	**0.072**	**0.199**	**0.173**	**0.138**	**0.092**	**0.135**	**0.097**
(23) Cyber defense $_{i,t}$	0.008	**0.200**	**-0.040**	**0.130**	**-0.047**	0.015	**0.209**	-0.007	-0.010	0.008	0.003	-0.014
(24) Cyber vulnerability $_{i,t}$	-0.000	-0.006	*-0.037*	-0.014	-0.019	**0.042**	0.011	**0.074**	**0.064**	**0.054**	*0.039*	*0.037*

	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
(1) Scientific publications $_{i,t+1}$	**-0.107**	*-0.038*	**-0.063**	**-0.097**	-0.024*	**-0.084**	**-0.152**	**-0.201**	**-0.118**	**-0.094**	0.004	-0.016
(1) Scientific publications (unscaled) $_{i,t+1}$	-0.014	0.024	0.011	0.001	0.031*	-0.001	**0.290**	**0.362**	**0.099**	**0.164**	**0.160**	**0.048**
(3) Patents $_{i,t+1}$	*-0.034*	**-0.146**	**-0.123**	-0.021	**-0.150**	**-0.090**	**0.170**	**0.233**	**0.251**	**-0.145**	**-0.047**	**-0.114**
(4) Patents (unscaled) $_{i,t+1}$	-0.013	**-0.135**	**-0.103**	0.008	**-0.143**	**-0.075**	**0.280**	**0.381**	**0.312**	**-0.071**	0.021	**-0.098**
(5) Trade secrets $_{i,t+1}$	**-0.077**	-0.018	**-0.048**	**-0.078**	-0.018	**-0.056**	**-0.409**	**-0.588**	**-0.271**	**-0.210**	**-0.160**	*-0.033*
(6) Trade secrets (unscaled) $_{i,t+1}$	**0.043**	**0.070**	**0.059**	**0.053**	**0.060**	**0.065**	0.031*	**-0.065**	**-0.146**	**0.056**	-0.024*	**0.096**
(7) R&D $_{i,t}$	**0.119**	**0.071**	**0.101**	**0.129**	**0.064**	**0.104**	**0.557**	**0.722**	**0.265**	**0.316**	**0.199**	**0.087**
(8) Data breach risk $_{i,t}$	**0.451**	**0.525**	**0.579**	**0.660**	**0.445**	**0.344**	**0.140**	**0.103**	-0.022	**0.287**	0.019	**0.124**
(9) Data breach risk - Personal information $_{i,t}$	**0.368**	**0.594**	**0.621**	**0.496**	**0.226**	**0.377**	**0.133**	**0.109**	0.009	**0.248**	-0.010	**0.096**
(10) Data breach risk - Financial information $_{i,t}$	**0.515**	**0.206**	**0.229**	**0.245**	**0.145**	**0.666**	**0.086**	**0.086**	*0.035*	**0.117**	-0.006	**0.071**
(11) Data breach risk - Other information $_{i,t}$	**0.357**	**0.553**	**0.478**	**0.276**	**0.523**	**0.228**	**0.071**	*0.037*	**-0.061**	**0.175**	0.030*	**0.102**
(12) Data breach risk - Information not disclosed $_{i,t}$	**0.230**	**0.317**	**0.332**	**0.208**	**0.625**	**0.197**	**0.071**	0.017	**-0.089**	**0.126**	*0.034*	**0.081**
(13) Data breach risk - Malware $_{i,t}$		**0.139**	**0.136**	**0.292**	**0.139**	**0.499**	**0.104**	**0.111**	**0.059**	**0.143**	-0.003	**0.065**
(14) Data breach risk - Unauthorized access $_{i,t}$	**0.261**		**0.458**	**0.087**	**0.326**	**0.361**	**0.060**	0.021	**-0.066**	**0.155**	0.003	**0.100**
(15) Data breach risk - Phishing $_{i,t}$	**0.099**	**0.366**		**0.047**	**0.333**	**0.175**	**0.084**	**0.045**	**-0.063**	**0.185**	0.002	**0.086**
(16) Data breach risk - Cyberattack not disclosed $_{i,t}$	**0.300**	**0.146**	**0.123**		0.023	**0.213**	**0.094**	**0.078**	0.023	**0.171**	0.007	**0.051**
(17) Data breach risk - Ransomware $_{i,t}$	**0.109**	**0.401**	**0.207**	-0.004		**0.065**	**0.066**	0.025*	**-0.059**	**0.133**	0.029*	**0.063**
(18) Data breach risk - Misconfiguration $_{i,t}$	**0.458**	**0.635**	**0.096**	**0.227**	0.021		**0.093**	**0.089**	0.008	**0.109**	-0.001	**0.104**
(19) Analyst following $_{i,t}$	**0.119**	**0.074**	**0.057**	**0.129**	**0.051**	**0.107**		**0.538**	**0.253**	**0.280**	**0.128**	**0.069**
(20) Size $_{i,t}$	0.024*	-0.008	0.015	0.027*	0.007	0.008	**0.236**		**0.674**	**0.389**	**0.289**	**0.101**
(21) ROA $_{i,t}$	**0.057**	-0.016	**-0.053**	*0.033*	**-0.046**	0.024*	**0.211**	**0.132**		**0.123**	**0.135**	-0.000
(22) 10-K filesize $_{i,t}$	**0.107**	**0.118**	**0.116**	**0.113**	**0.102**	**0.073**	**0.253**	**0.255**	**0.127**		**0.248**	**0.172**
(23) Cyber defense $_{i,t}$	0.003	-0.014	-0.003	-0.000	0.000	-0.010	**0.105**	**0.263**	**0.082**	**0.210**		**0.070**
(24) Cyber vulnerability $_{i,t}$	0.030*	**0.055**	*0.037*	**0.041**	0.021	**0.067**	**0.047**	0.029*	*0.037*	**0.120**	**0.097**	

Notes: Numbers in bold represent significance at the 1% level, numbers in italic at the 5% level, and * denotes significance at the 10% level. The upper right triangle represents the Spearman correlations. The lower left triangle represents the Pearson correlations. All variables are reported without logarithms. Please refer to Appendix A for a full description of all variables.

Table 7

Main analysis: Data breach risk and scientific publications

	Scientific publications $_{i,t+1}$	
	(1)	(2)
Data breach risk $_{i,t}$	-0.0408***	-0.0413***
	(0.0155)	(0.0158)
Control variables		
Analyst following $_{i,t}$		-0.0197
		(0.1162)
Size $_{i,t}$		-0.1215***
		(0.0340)
ROA $_{i,t}$		0.1636***
		(0.0524)
10-K filesize $_{i,t}$		0.0306
		(0.0382)
Cyber vulnerability $_{i,t}$		0.0652
		(0.0634)
Cyber defense $_{i,t}$		-0.0833
		(0.0685)
Firm fixed effects	Yes	Yes
State x industry fixed effects	Yes	Yes
State x year fixed effects	Yes	Yes
Industry x year fixed effects	Yes	Yes
Observations	6,933	6,933
Adjusted R^2	0.43833	0.44101

Notes: Table 7 reports the regression results for the relationship between *Data breach risk* $_{i,t}$ and *Scientific publications* $_{i,t+1}$. The outcome variable *Scientific publications* $_{i,t+1}$ is the number of scientific publications in year t+1 scaled by the R&D expenses in year t. Please refer to Appendix A for a full description of all variables. Standard errors clustered by industry-state in parentheses. ∗, ∗∗, and ∗ ∗ ∗ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Table 8

Verification of data breach risk measure: Data breach risk and patents

	Patents $_{i,t+1}$	
	(1)	(2)
Data breach risk $_{i,t}$	0.0172*	0.0169*
	(0.0087)	(0.0087)
Control variables		
Analyst following $_{i,t}$		0.0079
		(0.0257)
Size $_{i,t}$		-0.0633***
		(0.0215)
ROA $_{i,t}$		0.0366**
		(0.0170)
10-K filesize $_{i,t}$		0.0051
		(0.0145)
Cyber vulnerability $_{i,t}$		-0.0012
		(0.0242)
Cyber defense $_{i,t}$		-0.0100
		(0.0164)
Firm fixed effects	Yes	Yes
State x industry fixed effects	Yes	Yes
State x year fixed effects	Yes	Yes
Industry x year fixed effects	Yes	Yes
Observations	6,933	6,933
Adjusted R^2	0.52858	0.53173

Notes: Table 8 reports the regression results for the relationship between *Data breach risk* $_{i,t}$ and *Patents* $_{i,t+1}$. The outcome variable *Patents* $_{i,t+1}$ is the number of patents in year t+1 scaled by the R&D expenses in year t. Please refer to Appendix A for a full description of all variables. Standard errors clustered by industry-state in parentheses. ∗, ∗∗, and ∗ ∗ ∗ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Table 9

Parallel trend and duration of treatment effect

	Scientific publications i			
	t-1	*t-2*	*t+2*	*t+3*
	(1)	(2)	(3)	(4)
Data breach risk $_{i,t}$	0.0055	0.0212	0.0088	0.0223
	(0.0491)	(0.0523)	(0.0208)	(0.0324)
Control variables				
Analyst following $_{i,t}$	-0.0706	0.0024	-0.0368	0.0671
	(0.1798)	(0.0810)	(0.0725)	(0.0828)
Size $_{i,t}$	0.0104	0.0265	-0.1375***	-0.0842**
	(0.0755)	(0.0524)	(0.0398)	(0.0413)
ROA $_{i,t}$	0.0458	-0.3224	0.0978**	-0.0021
	(0.0456)	(0.3741)	(0.0435)	(0.0350)
10-K filesize $_{i,t}$	0.1310	0.0014	0.0427	0.0016
	(0.1124)	(0.0476)	(0.0457)	(0.0232)
Cyber vulnerability $_{i,t}$	-0.0450	0.0440	0.1081*	0.0131
	(0.0781)	(0.0734)	(0.0625)	(0.0830)
Cyber defense $_{i,t}$	-0.0540	-0.0360	-0.0775	-0.0014
	(0.0880)	(0.0669)	(0.0509)	(0.0290)
Firm fixed effects	Yes	Yes	Yes	Yes
State x industry fixed effects	Yes	Yes	Yes	Yes
State x year fixed effects	Yes	Yes	Yes	Yes
Industry x year fixed effects	Yes	Yes	Yes	Yes
Observations	5,942	5,459	5,925	4,966
Adjusted R^2	0.25211	0.23350	0.47075	0.42369

Notes: Table 9 reports the regression results for the relationship between *Data breach risk* $_{i,t}$ and *Scientific publications* $_i$. The outcome variables *Scientific publications* $_i$ is the number of scientific publications in year t-1 scaled by the R&D expenses in year t-2 (column (1)), the number of scientific publications in year t-2 scaled by the R&D expenses in year t-3 (column (2)), the number of scientific publications in year t+2 scaled by the R&D expenses in year t+1 (column (3)), and the number of scientific publications in year t+3 scaled by the R&D expenses in year t+2 (column (4)). Please refer to Appendix A for a full description of all variables. Standard errors clustered by industry-state in parentheses. *, **, and *** represent significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Table 10

Heterogenous treatment effects: Data breach characteristics

	Scientific publications i,t+1									
	Type of information accessed				*Type of cyberattack*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Data breach risk - Personal information $_{i,t}$	-0.0543**									
	(0.0232)									
Data breach risk - Financial information $_{i,t}$		-0.0057								
		(0.0288)								
Data breach risk - Other information $_{i,t}$			0.0285							
			(0.0355)							
Data breach risk - Information not disclosed $_{i,t}$				-0.0310						
				(0.0528)						
Data breach risk - Malware $_{i,t}$					0.0033					
					(0.0908)					
Data breach risk - Unauthorized access $_{i,t}$						-0.0269				
						(0.0386)				
Data breach risk - Phishing $_{i,t}$							-0.0648**			
							(0.0306)			
Data breach risk - Ransomware $_{i,t}$								0.0358		
								(0.0455)		
Data breach risk - Misconfiguration $_{i,t}$									-0.0375*	
									(0.0195)	
Data breach risk - Cyberattack not disclosed $_{i,t}$										-0.1043**
										(0.0508)
Control variables										
Analyst following $_{i,t}$	-0.0195	-0.0212	-0.0213	-0.0211	-0.0212	-0.0208	-0.0212	-0.0212	-0.0211	-0.0188
	(0.1161)	(0.1165)	(0.1164)	(0.1165)	(0.1165)	(0.1166)	(0.1164)	(0.1164)	(0.1163)	(0.1163)
Size $_{i,t}$	-0.1216***	-0.1210***	-0.1211***	-0.1211***	-0.1210***	-0.1212***	-0.1208***	-0.1213***	-0.1214***	-0.1213***
	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0341)
ROA $_{i,t}$	0.1640***	0.1633***	0.1635***	0.1632***	0.1633***	0.1635***	0.1633***	0.1637***	0.1637***	0.1628***
	(0.0523)	(0.0523)	(0.0523)	(0.0523)	(0.0523)	(0.0523)	(0.0523)	(0.0522)	(0.0523)	(0.0522)
10-K filesize $_{i,t}$	0.0305	0.0298	0.0296	0.0301	0.0299	0.0301	0.0303	0.0295	0.0294	0.0308
	(0.0382)	(0.0382)	(0.0382)	(0.0382)	(0.0382)	(0.0382)	(0.0382)	(0.0382)	(0.0382)	(0.0383)
Cyber vulnerability $_{i,t}$	0.0659	0.0667	0.0671	0.0665	0.0667	0.0666	0.0659	0.0674	0.0669	0.0664
	(0.0639)	(0.0642)	(0.0644)	(0.0640)	(0.0642)	(0.0642)	(0.0638)	(0.0645)	(0.0643)	(0.0635)
Cyber defense $_{i,t}$	-0.0822	-0.0824	-0.0820	-0.0826	-0.0824	-0.0827	-0.0830	-0.0822	-0.0823	-0.0826
	(0.0684)	(0.0690)	(0.0688)	(0.0690)	(0.0691)	(0.0688)	(0.0691)	(0.0689)	(0.0688)	(0.0684)
Firm fixed effects (#1,088)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x industry fixed effects (#243)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State x year fixed effects (#340)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x year fixed effects (#225)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,933	6,933	6,933	6,933	6,933	6,933	6,933	6,933	6,933	6,933
Adjusted R^2	0.44114	0.44063	0.44065	0.44064	0.44063	0.44065	0.44077	0.44065	0.44070	0.44121

Notes: Table 10 reports the regression results for the relationship between different types of *Data breach risk* based on the *Type of information accessed* and the *Type of cyberattack* and *Scientific publications*. The outcome variable *Scientific publications* $_{i,t+1}$ is the number of scientific publications in year t+1 scaled by the R&D expenses in year t. Please refer to Appendix A for a full description of all variables. Standard errors clustered by industry-state in parentheses. *, **, and * * * represent significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Table 11

Alternative outcome variable: Data breach risk and trade secrets

	Trade secrets $_{i,t+1}$	
	(1)	(2)
Data breach risk $_{i,t}$	-0.0146***	-0.0141***
	(0.0040)	(0.0042)
Control variables		
Analyst following $_{i,t}$		-0.0691***
		(0.0201)
Size $_{i,t}$		-0.0658***
		(0.0176)
ROA $_{i,t}$		0.0690**
		(0.0219)
10-K filesize $_{i,t}$		-0.0113
		(0.0107)
Cyber vulnerability $_{i,t}$		0.0202
		(0.0256)
Cyber defense $_{i,t}$		-0.0283
		(0.0200)
Firm fixed effects	Yes	Yes
State x industry fixed effects	Yes	Yes
State x year fixed effects	Yes	Yes
Industry x year fixed effects	Yes	Yes
Observations	6,933	6,933
Adjusted R^2	0.51607	0.52442

Notes: Table 11 reports the regression results for the relationship between *Data breach risk* $_{i,t}$ and *Trade secrets* $_{i,t+1}$. The outcome variable *Trade secrets* $_{i,t+1}$ is the number of trade secret terms in year t+1 scaled by the R&D expenses in year t. Please refer to Appendix A for a full description of all variables. Standard errors clustered by industry-state in parentheses. ∗, ∗∗, and ∗ ∗ ∗ represent significance at the 10 percent, 5 percent, and 1 percent level, respectively.

Appendix A. Variables Description

Outcome variables	
Scientific publications $_{i,t+1}$	The number of scientific publications in year t+1 scaled by the R&D expenses in year t. Sources: Arora et al. (2024a, 2024b, 2021) scientific publication data and CRSP Compustat Merged data.
Patents $_{i,t+1}$	The number of patents in year t+1 scaled by the R&D expenses in year t. Sources: Kogan et al. (2017) patent data and CRSP Compustat Merged data.
Trade secrets $_{i,t+1}$	The number of trade secret terms[22] in year t+1 scaled by the R&D expenses in year t. Sources: Loughran and McDonald (2016) cleaned 10-K data and Glaeser (2018) trade secret word list.
Treatment variables	
Data breach risk $_{i,t}$	The number of security breaches in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Personal information $_{i,t}$	The number of security breaches that involve personal information in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Financial information $_{i,t}$	The number of security breaches that involve financial information in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Other information $_{i,t}$	The number of security breaches that involve other information in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Information not disclosed $_{i,t}$	The number of security breaches that do not disclose the type of information accessed in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Malware $_{i,t}$	The number of security breaches that involve malware in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Unauthorized access $_{i,t}$	The number of security breaches that involve unauthorized access in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Phishing $_{i,t}$	The number of security breaches that involve phishing in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Ransomware $_{i,t}$	The number of security breaches that involve ransomware in the same state and industry as firm i in year t. Source: Audit Analytics.

[22] Trade secret and trade secrecy.

Data breach risk - Misconfiguration $_{i,t}$	The number of security breaches that involve misconfiguration in the same state and industry as firm i in year t. Source: Audit Analytics.
Data breach risk - Cyberattack not disclosed $_{i,t}$	The number of security breaches that do not disclose the type of cyberattack in the same state and industry as firm i in year t. Source: Audit Analytics.
Control variables	
Analyst following $_{i,t}$	The arithmetic mean of the 12 monthly numbers of earnings forecasts for firm i over fiscal year t. Source: I/B/E/S summary data.
Size $_{i,t}$	The natural logarithm of firm i's book value of total assets in year t. Source: CRSP Compustat Merged data.
ROA $_{i,t}$	Firm i's operating income before depreciation (OIBDP) divided by book value of total assets (TA) in year t. Source: CRSP Compustat Merged data.
10-K filesize $_{i,t}$	The natural logarithm of firm i's 10-K filesize in kilobytes in year t. Source: Loughran and McDonald (2016) cleaned 10-K data.
Cyber defense $_{i,t}$	The natural logarithm of one plus the number of cyber defense terms[23] in the 10-K in year t. Source: Loughran and McDonald (2016) cleaned 10-K data and Ettredge et al. (2018) cyber defense word list.
Cyber vulnerability $_{i,t}$	The natural logarithm of one plus the number of cyber vulnerability terms[24] in the 10-K in year t. Source: Loughran and McDonald (2016) cleaned 10-K data and Ettredge et al. (2018) cyber vulnerability word list.

[23] Risk governance, risk model, risk control, risk policy, risk framework, risk document, risk system, risk technology, risk training, risk committee, risk management, risk board, risk review, risk oversight, risk governance, chief risk officer, CRO, enterprise risk management, ERM, risk compensation, risk incentive, risk method, risk compliance, risk system, risk data integration, risk limit, risk control.

[24] Data integrity, data risk, risk report, risk dashboard, operation failure, operational failure, operation risk, operational risk, IT risk, information technology risk, privacy breach, identity theft, computer virus, security, breach, hacker, cyber-attack, cyber risk, cybersecurity, security incident, computer breach, computer intrusion.

Appendix B. Data breach descriptions

Type of information accessed	
Personal information $_{i,t}$	Name, address, phone number, e-mail, username, password, and SSN.
Financial information $_{i,t}$	Bank accounts, debit and credit cards.
Other information $_{i,t}$	Intellectual property, proprietary business information, and all other types of information.

Type of cyberattack	
Unauthorized access $_{i,t}$	Unauthorized party gains access to protected systems and data.
Phishing $_{i,t}$	Fraudulent attempt to obtain sensitive information under the guise of trustworthy electronic communication.
Misconfiguration $_{i,t}$	Exploitation of incorrectly assembled or poorly written code and web applications.
Malware $_{i,t}$	Malicious software intentionally designed to cause damage.
Ransomware $_{i,t}$	Specific type of malware designed to hold systems hostage in exchange for demands being met.

Chapter 3

Audit Offices' Cybersecurity Experience and Industry Range

ABSTRACT

In this study, I aim to determine whether the risk of cyberattacks spills over across clients of the same audit office. I conduct my tests by assessing the effect of audit office cybersecurity experience on audit clients in pre and post cyberattack periods. Additionally, I investigate the effect of audit office industry range on their clients' breach likelihood. Experiences in diverse industries may lead to the development of flexibility and skepticism, enhancing auditors' abilities to identify client vulnerabilities. I use a sample of U.S. public firms over the period 2011 to 2024, and I run logistic regression models to test my hypotheses. I find that clients are more likely to experience a cyber incident when their audit office has had a previously breached client or will have a breached client in the future. However, this breach likelihood is decreased when auditors are identified as having a broad range of industry experiences. My results also indicate that industry specialization does not seem to play a significant role in the context of data breaches.

Keywords: cybersecurity experience, cyber incidents, breach, industry range.

1. Introduction

External auditors have exclusive access to top executives, engage in audit committee meetings, and review board meeting minutes and other general information about their clients during the execution of their audits (Dhaliwal et al., 2016). This allows auditors to discuss strategic initiatives, such as asset acquisition or disposition, with their clients (Dhaliwal et al., 2016). These strategic practices may then be transferred to other clients, either intentionally or unintentionally (Dhaliwal et al., 2016). Intentional information leakage could happen through discussions that justify auditing procedures, while unintentional leakage could occur through casual conversations about best practices (McAllister & Cripe, 2008). As such, auditors are positioned as information intermediaries, who can either harm or help their clients (Dhaliwal et al., 2016; Kang et al., 2022).

Prior research has highlighted that even if the actual information transfer incident is unknown, perceiving it as a costly risk is enough to influence firms' decisions on whether to share the same auditor with a competing firm (Bills et al., 2020). The risk of information leakage is intensified when rival firms can exploit the information (Cahan et al., 2008), and the cost is even higher for rivals with similar product offerings (Bills et al., 2020). Sharing the same auditor has financial reporting quality implications, and firms with the same external auditor tend to have financial statements that are set up in the same manner (Francis et al., 2014). Also, firms that share the same auditor are subject to the same auditing style (Cao & Pham, 2021). Investors believe that firms that share an auditor with other firms that have received a comment letter[25], will eventually be scrutinized by the SEC (Cao & Pham, 2021).

[25] A comment letter contains information related, but not limited to, the recognition of revenues, gains, or losses in firm 10k reports, that arises due to the SEC's regular review of firm reports (Cao & Pham, 2021).

While auditors are not required to provide assurance on their clients' cybersecurity practices under the current standards, they may face reputational damages when their clients get breached (Li et al., 2024). Auditors are responsible for cybersecurity risks that are related to internal control over financial reporting (ICFR), and a breach to one client may be interpreted as auditors failing to properly evaluate a client's ICFR, a failure that could then spill over to non-breached clients of the same audit office (Li et al., 2024). Further, a client's breach could signal deficiencies in auditing procedures systemic to the audit office or the entire audit firm (Li et al., 2024). For example, when Equifax, Ernst & Young's (EY) client, got breached, EY was accused of not detecting major internal control deficiencies, and this issue spilled over to the whole audit (Mckenna, 2018; Mintz, 2017).

The abovementioned risks may drive firms to avoid sharing an external auditor with another firm, however, some benefits may motivate them to employ the same auditor. Peer firms that offer similar products may have similarities in their accounting methods, business processes, and internal controls (Bills et al., 2020). With respect to cybersecurity, current auditing standards require auditors to play a significant role (CAQ, 2019; Hamm, 2019), and cybersecurity was placed on the PCAOB's 2020-2024 strategic plan (PCAOB, 2020). In evaluating the association between cybersecurity and financial statements' misstatement risks, auditors should focus on understanding their clients' use of automated internal controls over financial reporting and the use of information technology (Asthana et al., 2021). In the event of a cyberattack, auditors should assess whether it arose due to ICFR deficiencies and whether to revise and plan supplementary audit procedures (Asthana et al., 2021). Being at the forefront, auditors of breached clients gain first-hand experience through engagement in evaluating the breach incident, implementation of corrective actions, and performance of additional cybersecurity risk assessments (Li et al., 2024). Further, auditors of breached clients may become more knowledgeable in areas that include internal controls, information technology

(IT) management actions, the IT environment, etc. Experience in all of these areas may equip auditors with better risk (e.g. IT risks) identification and assessment capabilities (Smith et al., 2019). Also, auditors providing audit services to comparable firms will develop subspecialty knowledge that improves the efficiency and effectiveness of both audit engagements (Bills et al., 2020; Kang et al., 2022).

Many previous studies in auditing argue that auditors who have a specialization in an industry, provide better-quality audits to clients in that industry due to their strong competencies and reputational incentives (e.g., Balsam et al., 2003; Reichelt & Wang, 2010). As such, firms may be willing to employ the same auditor employed by a peer firm. Even though industry specialization can improve an auditor's specific industry knowledge, it often leads to constraints in the auditor's industry range, which is defined as the degree to which an auditor has experience auditing clients from various industries (Dekeyser et al., 2024). Narrowly focused experiences can help auditors develop deep knowledge in a certain domain, however, they might induce flawed intuition development (Dane, 2010). The "Einstellung" effect happens when a certain idea first comes to mind, triggered by similar previous experiences, which prevents alternative options from surfacing (Bilalić et al., 2008). Industry specialization is characterized by this effect, leading the specialist to reduce their skepticism and flexibility when performing auditing tasks (Dekeyser et al., 2024). Pursuing a broad range of experiences can alleviate this rigidity as individuals become more likely to engage in exceptional scenarios or inconsistencies in various areas thereby developing mental flexibility and skepticism (Srikantia & Pasmore, 1996). In line with this argument, experts with a wide range of experiences outside their specialty tend to perform better than experts with a limited range of experiences (Dekeyser et al., 2024). Prior research shows that auditors' broad experiences in various industries play a significant role in developing their way of thinking and

knowledge (Libby & Luft, 1993; Nelson & Tan, 2005), which in turn can enhance audit quality (Dekeyser et al., 2024).

Prior studies have revealed that although knowledge accrues at the individual-level when the same auditor is assigned to two engagements, knowledge is transferred to the audit firm through inter-office communication and process discussions across engagement teams (Nelson et al., 2016; Seavey et al., 2018). Auditors from Big Four firms help clients decrease cybersecurity risks through formal recommendations or through informal talks with the board of directors and management (Li et al., 2024). Audit offices' cybersecurity knowledge and experiences may be valuable for other clients who might be exposed to the same or similar hacking methods (PCAOB, 2019). Further, investors in non-breached firms tend to reassess the credibility of the audit office after a client of the same office was breached (Perols & Murthy, 2021). For this reason, auditors may be motivated to help non-breached clients shield against cyberattacks, avoiding potential reputational harm (Li et al., 2024).

Audit firms are organized into offices where the auditor network is strongest, client information is concentrated, and the chance of sharing client information is the highest (Dhaliwal et al., 2016). Informal spillover channels within the offices of shared auditors have been a focus in prior literature (Cai et al., 2016; Dhaliwal et al., 2016) that has indicated that auditor offices significantly affect client outcomes (Dhaliwal et al., 2016).

In this paper, I focus on informal spillover channels as reflected through sharing the same auditor office to examine whether cybersecurity information may be shared. I examine whether this channel puts clients at risk of experiencing a breach when other clients of the shared auditor have experienced a breach. Specifically, I test whether auditors act as knowledge intermediaries or as risk intermediaries by using the setting of cybersecurity breaches as a proxy for risk. Additionally, I test the effect of auditor industry range on client likelihood of

experiencing a data breach. By doing so, I determine whether diverse auditor industry experiences play a significant role within this setting of data breaches.

This study contributes to different streams of literature. First, it contributes to the growing stream of cybersecurity research within the auditing literature. Prior studies find that auditors increase audit fees following cyber incidents (Li et al., 2020; Smith et al., 2019), and this increase, conditional on auditors' prior cybersecurity experience, is negatively associated with future breach incidents (Li et al., 2024). This study extends Li et al. (2024) by showing that clients at the audit office level are subject to a higher breach likelihood if other clients of the same office have experienced a breach in the past *(i.e., breaches are known)* or will experience a breach in the future *(i.e., breaches are unknown)*. However, the breach likelihood decreases when the audit office has a broad industry range.

The second contribution is tied in with the literature on the effects of experience (Dekeyser et al., 2024; Hanlon et al., 2022), specifically audit office experience. Dekeyser et al. (2024) developed the concept of auditor industry range, applied it to a Chinese dataset, and found a positive impact on audit quality. I use the same variable for my U.S dataset, in the context of cybersecurity. My findings suggest that audit office experiences in diverse industries appear to assist the audit office in developing professional skepticism, therefore helping their clients by decreasing their likelihood of experiencing cyberattacks. While Li et al. (2024) find that increased audit fees are associated with a lower breach likelihood due to additional audit office efforts, I find that industry range, which reflects audit offices' improved knowledge and overall awareness, is associated with a decreased breach likelihood. Differentiating between audit offices at the city-level and at the state-level, I find significant results at both levels.

The third and final main contribution of my study is its value to regulators in the public domain. This study supports the current emphasis on the importance of prioritizing

cybersecurity (PCAOB, 2020) because cyber incidents put clients of the same audit office at risk (PCAOB, 2019).

This study is organized as follows. In section 2, I offer background information on prior studies and on cyber incidents, external auditors, and auditor's industry range and I develop the hypotheses. In section 3, I describe the sample selection process and the empirical research design. In section 4, I present the results. In section 5, I present the conclusions.

2. Background and Hypotheses Development

2.1. Background

Prior audit research has explored auditors' responses to cyber incidents, with a particular focus on the impact of such events on audit fees. Li et al. (2020) find that external auditors increase their audit fees following the occurrence of cyber incidents, however, this increase is lower following the 2011 Cybersecurity Disclosure Guidance issued by the SEC. Additionally, they find that auditors predict and price cybersecurity risks before the breach occurrence. Smith et al. (2019) also find a positive association between cyberattacks and audit fees; however, the result is largely determined by external breaches. In a more recent study, Li et al. (2024) find that audit offices with breached clients, charge higher fees to their non-breached clients. This increase in audit fees, conditional on audit offices' cybersecurity experience, is negatively associated with future cyber incidents.

While these studies focus primarily on the direct association between clients' breaches and audit fees, my study examines the indirect contagion effect of known (i.e., past) and unknown (i.e., future) breaches among clients of the same audit office. This study also examines the role that audit office industry range plays in the context of cybersecurity; a factor not investigated by previous studies. The breach data used in previous studies extends to 2014 (e.g., Li et al., 2020; Smith et al., 2019) or 2017 (e.g., Li et al., 2024), while the data in this study incorporates

breaches through 2024. Overall, this study complements and extends previous work by offering insights into intra-office cybersecurity contagion effects and the importance of industry diversification.

2.2. Cyber Incidents and External Auditors

Cybercrimes are defined as illegal activities conducted using the internet or computers to steal or destroy data, or to control or disrupt a target's infrastructure or computing environment (CSIS, 2018). Cyberattacks represent one of the biggest systemic risks faced by organizations and "clean financial reporting control assessments under SOX do not protect organizations from such risks." (Lawrence et al., 2018, p.143). They have been identified among the top five serious international threats (WEF, 2019), imposing considerable costs on the global economy (IBM, 2023). Cyberattacks are indicators of operational control risk deficiencies. If a firm's computer systems cannot be protected from cyberattacks, then its financial reporting systems are also vulnerable to attacks (Lawrence et al., 2018). The increase in data breaches reflects flaws in the protective mechanisms of operating and financial reporting activities (Lawrence et al., 2018). In 2020, the PCAOB issued a report that focused on the importance of understanding the risk-assessment strategies undertaken by auditors and their responses to identified risks (PCAOB, 2020). The costs of the breaches are not only borne by the affected firms, but spill over to their auditors in the form of reputational harm (Perols & Murthy, 2021).

Auditors are not responsible for assessing clients' cybersecurity risk, or the controls implemented to mitigate that risk (PCAOB, 2019). Instead, auditors focus on their clients' information technology used to prepare the financial statements, and on the automated controls around financial reporting (PCAOB, 2019). If a client firm experiences a cyberattack, the auditor must evaluate the nature and extent of that attack (i.e., identify what was stolen, destroyed, or modified), and should assess the implications on the company's operations

(PCAOB, 2019). Besides that, auditors should assess whether the breach was due to a client's ICFR deficiencies, and whether the client has implemented procedures to prevent future cyberattacks (PCAOB, 2019). As such, auditors are not held responsible nor do they face lawsuits for their clients' cyberattacks (Asthana et al., 2021; Valdetero et al., 2019). Considering this, there is likely little to no basis for expecting auditors to pay close attention to their clients' cybersecurity protocols and actions, especially since existing standards do not mandate that auditors provide assurance on their clients' cybersecurity (Li et al., 2024).

Although auditors are not required to detect and address cybersecurity risks, they could be subject to reputational damage (Perols & Murthy, 2021). Cyberattacks expose vulnerabilities in the internal controls related to the company's operations and to its financial reporting (Lawrence et al., 2018; PCAOB, 2019; Smith et al., 2019). As external auditors are responsible for ICFR, a cyberattack to one client may be interpreted as an audit failure to adequately assess ICFR; a failure that could then spill over to other clients of the same audit office (Li et al., 2024). Additionally, cyberattacks could reveal deficiencies in audit practices inherent to the audit firm or audit office (Li et al., 2024). Prior research indicates that investors of non-breached firms reconsider the reliability of the audit provided by the audit office following a breach to the office's clients (Li et al., 2024). In other words, cyberattacks reverse investors' favorable perception of auditor competence (Perols & Murthy, 2021). In this respect, a client cyberattack may taint the auditor by association (Asthana et al., 2021). Even though the client's negative event is beyond the auditor's responsibility and accountability, the auditor's audit engagement office may be viewed negatively (Asthana et al., 2021). There could also be a gap between the steps that the public expects auditors to undertake with respect to cybersecurity and the requirements that auditors need to comply with their professional standards (Franzel, 2016). To this extent, auditors may be motivated to provide additional support to their non-

breached clients to address cybersecurity risks, preventing potential reputational harm (Li et al., 2024).

A cyberattack does not necessarily reflect client misconduct, instead it shows that the client has fallen victim to a fraudulent act (SEC, 2018). Auditors' experiences with their cyberattacked clients may help them develop knowledge about internal controls, IT management, and the IT environment, which may enhance auditors' detection and evaluation of risks, including information technology risks (Smith et al., 2019). Perceived as rational learners, auditors can modify their behavior based on their experiences and can better engage, either formally through recommendations or informally through conversations, with their clients to help minimize cyber risks (Li et al., 2024). As auditors have access to senior executives and their strategic processes, they can transfer cybersecurity knowledge acquired from breached to non-breached clients (Li et al., 2024).

The effort and time dedicated by the audit office to investigate the determining factors and implications of cyberattacks may be deemed beneficial to other clients because cybercriminals often use similar techniques to attack firms (PCAOB, 2019). Nonetheless, auditors usually get involved after the occurrence of a client's cyberattack, where they evaluate the incident and implement corrective actions (Li et al., 2024). Although a client's cybersecurity incident helps auditors gain first-hand experience and puts them in a position to help other non-breached clients (Li et al., 2024; Smith et al., 2019), their unaccountability for their clients' breaches may not induce them to take additional measures. Incidentally, following the occurrence of a client's cybersecurity breach, the audit office may need to focus its resources and personnel on the affected client to assess the event and employ corrective measures. This diversion might compromise the attention to cybersecurity processes of other clients, exposing them to higher risks of breaches.

Audit offices that have already had clients experience a cybersecurity incident in the past, may become more knowledgeable about IT-related matters (Smith et al., 2019). Their increased awareness can be transferred to their clients, helping them decrease their risk of cyberattacks (Li et al., 2024; PCAOB, 2019). Alternatively, even with increased knowledge, auditors might not discuss IT issues with their clients because they are not responsible for assuring cybersecurity (Li et al., 2024; Perols & Murthy, 2021), or they might direct their resources to the breached clients, consequently failing to assist their other clients. Based on this discussion, I state the first hypothesis in the null form:

H1a: Sharing an auditor, at the audit office level, with a breached client is not associated with the likelihood that other clients will experience a breach.

Future cybersecurity weaknesses could be thought of as "private" information, as opposed to the previous argument where the weaknesses could be described as "public" information (Cheng et al., 2019). To the best of my knowledge, no previous studies have looked at the effect of having the same audit office as a client that will fall victim to a cyberattack. Cyberattacks may reveal failures in audit practices, such as inadequate cybersecurity assessments, insufficient communication of identified risks or recommendations to clients, and lack of regular monitoring, systemic to the audit office (Lawrence et al., 2018; Li et al., 2024; Smith et al., 2019). If these weaknesses are not yet revealed, clients of the same audit office might be subject to a high cybersecurity risk without knowing it because they may be using similar information systems and audit methodologies as others, and cybercriminals usually employ similar tactics to attack firms (PCAOB, 2019). The similarities in the systems used by client firms make it easier and more convenient for cybercriminals, who can employ similar intrusion strategies, to attack firms. Therefore, I argue that sharing the same audit office as clients that

will report breaches in the future is positively associated with the likelihood of experiencing a breach, which leads to the below hypothesis:

H1b: Sharing an auditor, at the audit office level, with a client that will report a breach is positively associated with the likelihood that other clients will experience a breach.

2.3. Auditor Industry Range

An important element that I believe might affect clients' breach likelihood is auditors' industry range because it influences auditors' familiarity with various industry risks and best practices. This concept reflects the degree to which auditors have experiences in auditing clients from various industries (Dekeyser et al., 2024). Prior studies focused on the concept of industry specialization which, although it can enhance an auditor's specific industry knowledge, usually results in limited auditor industry range; a notion associated with cognitive inflexibility (Dekeyser et al., 2024). Experts with a wide range of experiences outside their specialty usually outperform others with a limited range of experiences within their specialty (Epstein, 2019). Drawing on previous studies in cognitive science and psychology (Dane, 2010; Hargadon, 2006), auditors with wide industry experiences tend to better overcome flawed intuitions compared with auditors without such experiences (Dekeyser et al., 2024). Diversified experiences expose individuals to inconsistencies and exceptions in different domains leading to the development of flexibility and mental skepticism (Srikantia & Pasmore, 1996), whereby auditors can identify "could be" scenarios (Dane, 2010). Working with clients from diverse industries, audit offices may have encountered several client breaches, which might enhance their ability to proactively identify potential vulnerabilities and threats in their clients' current systems. Subsequently, they can transfer their acquired knowledge by providing better security recommendations to their clients, who can either implement new or modify existing security measures, resulting in a decreased exposure to cybersecurity risks. Dekeyser et al. (2024) tested

this concept using a sample of Chinese firms and auditors and found that auditor industry range can enhance audit quality. I apply the same concept to U.S. auditors in the context of cybersecurity to determine whether auditors' industry range plays an important role in a different setting. As auditor experiences in various industries help in developing their knowledge and mindset, I argue that auditor industry range is associated with a lower breach likelihood. I posit the below hypothesis:

H2: Audit offices with a broad range of industry experiences are associated with lower breach likelihood for their clients.

3. Research Methods

3.1. Data and Sample Selection

The sample used to test the hypotheses is composed of breached and non-breached U.S. public firms covering the period 2011-2024. Following prior studies (e.g., Asthana et al., 2021; Li et al., 2020; Li et al., 2024), I use the Audit Analytics Cybersecurity database to obtain data on cybersecurity incidents. To ensure the comprehensiveness of my breach-related data, I verified my dataset using the Privacy Rights Clearinghouse database. I use Compustat to obtain the comparative sample of non-breached firms. I obtain audit-related data from Audit Analytics and I focus on audit offices as they are responsible for many strategic functions such as assigning personnel, administering audit engagements, and providing audit opinions (Li et al., 2024). Additionally, inter-office communications and discussions across engagement teams facilitate the transfer of knowledge (Nelson et al., 2016; Seavey et al., 2018). After merging the samples and eliminating observations with missing values, I ended up with a final sample of 27,242 firm-year observations and 420 data breach incidents. Details of my final sample are included in Table 1.

[Insert Table 1 here]

3.2. Research Design

In this study, I examine the effect on a client when their audit office has clients that have already experienced a cyberattack or that will experience a cyberattack. More specifically, I study the client's likelihood of experiencing a cyber incident. In the main analyses, the audit office is selected at the city-level. This means that all clients that share the same audit firm in the same city are considered clients of the same audit office at the city-level[26]. To test the hypotheses, I use the below logit model:

$$Logit\ (Breached\ Firm_{i,t}=1) = \beta_0 + \beta_1\ AudCyber1\ Pre_{i,t} + \beta_2\ AudCyber1\ Post_{i,t} + \beta_3\ Range\ N3C_{i,t} + \beta_4\ Big4_{i,t} + \beta_5\ Firm\ Size_{i,t-1} + \beta_6\ ROA_{i,t-1} + \beta_7\ Loss_{i,t-1} + \beta_8\ Leverage_{i,t-1} + \beta_8\ Year + \beta_9\ Industry + \varepsilon$$

The dependent variable *(Breached Firm)* is an indicator variable that takes the value of 1 if a firm has reported a breach in a given year t. The main independent variables are *AudCyber1 Pre* and *AudCyber1 Post*, both of which are dummy variables, and *RangeN3C* which is a continuous variable; all variables are developed at the firm-level. In determining whether clients' cybersecurity risks affect others with the same audit office, we differentiate between audit offices that had breached clients in any of the previous three years and those that will have breached clients in any of the upcoming three years. *AudCyber1 Pre* is set equal to 1 if the audit office has clients that experienced a cyber incident at any point in the previous three years. I follow the methodology from previous studies (e.g., Haislip et al., 2016; Lennox & Li, 2014) and use a three-year window because auditors require a sufficient time to respond to client cyberattacks, and because cyberattacks are not frequent in a certain auditor-year, which limits the treatment sample size. To test whether there might be some cybersecurity weakness

[26] In an additional analysis, I repeat the main test at the state-level. For that, all clients that share the same audit firm in the same state, are considered clients of the same auditor at the state-level.

spillover inherent at the audit office but not known yet, I look at the future audit office cybersecurity experience. *AudCyber1 Post* is set equal to 1 if the audit office has clients that experience a cyber incident at any point in the following three years. A positive coefficient on *AudCyber1 Pre* would indicate that clients of an audit office with prior cybersecurity incidents experience is associated with a higher risk of experiencing a cyberattack, whereas a negative coefficient would indicate a lower risk. A positive coefficient on *AudCyber1 Post* would indicate that clients of an audit office that will have clients experiencing a cyberattack in any of the upcoming three years, is associated with a higher risk of experiencing a cyberattack. Conversely, a negative coefficient would indicate a lower cybersecurity risk. *RangeN3C* is the natural logarithm of the number of unique industries audited by an audit office in a specific city over the past 3 years. This variable is generated using the industry range measure developed by Dekeyser et al. (2024)[27] and reflects the extent of the audit office's experience with various industries. A negative coefficient on *RangeN3C* would indicate that the broader the audit office industry range, the lower the clients' breach likelihood. However, a positive coefficient would indicate that an auditor's industry range is associated with a higher breach likelihood for their clients.

I include several control variables that might affect the breach likelihood. Big Four firms (i.e., *Big4*) may have more expertise than non-Big Four firms in matters related to cybersecurity (Yen et al., 2018). Large firms and firms with high profitability (i.e., *Firm Size* and *ROA*) are attractive targets to hackers (Higgs et al., 2016; Li et al., 2018; Li et al., 2024; Say & Vasudeva, 2020). Additionally, financially constrained firms (i.e., *Loss* and *Leverage*) are more likely to experience a cyberattack due to inadequate IT security investment (Higgs et al., 2016; Li et al.,

[27] Dekeyser et al (2024) developed the auditor industry range variable which is "defined as the natural logarithm of the number of industries that an auditor's client portfolio covers in the past three years, t–2 to t" (pp. 21). The three-year period is chosen to allow enough time for auditors to work with clients from different industries.

2018; Say & Vasudeva, 2020). All control variables, except for Big4, are lagged by one year to lower potential endogeneity concerns (Bouwman, 2011). I also control for year and industry effects. In additional analyses, I test for the effect of audit office industry specialization (i.e., *Industry Specialist1* and *Industry Specialist2*) on clients' breach likelihood. For *Industry Specialist1*, I use an absolute measure where I follow Bills et al. (2015) and classify an audit office in a city-year as an industry specialist if that office holds a market share of audit fees, within a specific industry and year, greater than 50%, 0 otherwise. For *Industry Specialist2*, I use a relative measure from Audousset-Coulier et al. (2016) and classify an audit office in a city-year as an industry specialist if that office possesses the largest market share of audit fees in a given industry, 0 otherwise. Audit fees better capture auditor effort than other variables (i.e., client sales or assets), because audit fees are related to the client riskiness, complexity, and size (Audousset-Coulier et al., 2016). Detailed variable descriptions are provided in Appendix A.

4. Results

4.1. Summary Statistics

Table 2 presents the descriptive statistics of my sample. In the pre three-year window, 31.6% of audit offices at the city-level have at least one cyberattacked client *(AudCyber1 Pre)*, whereas in the post three-year window, 24.9% of audit offices have at least one client that experiences a cyberattack in the subsequent three years *(AudCyber1 Post)*. On average, the natural logarithm of audit office industry range at the city-level *(Range N3C)* during the prior three-year window (*t-2* to *t*) is 3.228. This translates to a value of approximately 25, indicating that auditors in this sample have experience across a wide range of industries. My sample, similar to other studies (e.g., Asthana et al., 2021; Li et al., 2024), includes 420 breached firms, which makes up 1.5% of the total observations *(Breached Firm)*. Around 63.7% of my sample firms are audited by big four firms *(Big4)* and 43.6% of the firms report losses in the prior year

(Loss). Using the absolute measure of industry specialist, 17.2% of client firms are audited by industry specialists *(Industry Specialist1)*, whereas using the relative measure, 7.3% of client firms are audited by industry specialists *(Industry Specialist2)*. The mean for the *Firm Size* variable is 6.127, return on assets has a mean of -2.229 *(ROA)*, and *Leverage* has a mean of 5.155. At the audit office state-level, 41.9% of offices have at least one cyberattacked client *(AudCyber2 Pre)*, and 35.9% have at least one client that experiences a cyberattack in the subsequent three years *(AudCyber2 Post)*. Auditor industry range at the state-level *(RangeN3S*[28]*)* has a mean of 3.718 which translates to approximately 40 industries, and the alternative industry range measure ($Range_{Herf}$) has a mean of 0.884.

[Insert Table 2 here]

Table 3 presents the pairwise correlations among the variables[29]. Breached firms are positively associated with auditor cybersecurity experience *Pre* and *Post*, auditor industry range (at the city and state-level), Big4, firm size, ROA, and the absolute measure of industry specialist. Breached firms are negatively associated with loss, leverage, and the relative measure of industry specialist.

[Insert Table 3 here]

4.2. Main Findings

Table 4 presents the results of my main analyses. I find a positive and statistically significant coefficient (coefficient= 0.573, p-value < 0.01) on the *AudCyber1 Pre* variable, suggesting that clients of city-based audit offices with at least one breached client in the

[28] *RangeN3S* is the natural logarithm of the number of unique industries audited by an audit office in a specific state over the past 3 years.

[29] To assess the degree of multicollinearity among the variables, I conducted a Variance Inflation Factor (VIF) test. All the VIF values turned out to be below the cutoff point of 5 (Marcoulides & Raykov, 2019), suggesting that there is not an issue with multicollinearity.

previous three years, are associated with a higher likelihood of experiencing a breach. Specifically, having an audit office that has cybersecurity experience at another client, is associated with approximately 1.774 times[30] higher odds of experiencing a breach compared to clients of audit offices without such experience. Stated differently, clients of these audit offices have 77 percent[31] higher odds of breach occurrence. These results indicate that audit offices may not be communicating cybersecurity protective measures with their clients, either due to their unaccountability or due to their limited resources, leading to the rejection of H1a.

My findings provide additional evidence that auditors seem to evaluate a cyber incident and implement corrective actions after the occurrence of a breach (Li et al., 2024). Additionally, I find a positive and statistically significant coefficient (coefficient= 0.421, p-value < 0.01) on the *AudCyber1 Post* variable, suggesting that clients of audit offices that have at least one breached client in the following three years, are associated with a higher likelihood of experiencing a breach. More specifically, having an audit office that has at least one breached client at any point in the next three years, is associated with approximately 1.523 times[32] higher odds of experiencing a breach compared to clients of other audit offices. In other words, clients of these audit offices have approximately 52 percent[33] higher odds of breach occurrence. These results indicate that audit offices auditing breached clients tend to have more clients who experience breaches, supporting H1b. Taken together, my findings suggest that there is a positive association between the breach likelihoods of clients audited in the same audit office, regardless of whether the breach is known or not.

Additionally, Table 4 presents the results for my test of H2. Industry range is the degree to which an auditor has experience auditing clients from various industries (Dekeyser et al.,

[30] Calculated as: e^0.573.
[31] Calculated as: (e^0.573 – 1) * 100 ≈ 77.36% ≈ 77%.
[32] Calculated as: e^0.421.
[33] Calculated as: (e^0.421 – 1) * 100 ≈ 52.35% ≈ 52%.

2024). The variable for industry range reflects the natural logarithm of the number of unique industries audited by the audit office in a city over the prior three-year period. I find a negative and statistically significant coefficient (coefficient= -0.177, p-value<0.05) on the *RangeN3C* variable, suggesting that clients of audit offices with a wider range of industry experience over the past year, are associated with a lower likelihood of experiencing a breach. Specifically, this coefficient can be translated to signify that a one unit increase in the log of unique industries in an audit office industry range is associated with a decrease in the predicted odds of being breached by a multiplicative factor[34] of approximately 0.84[35]. In simpler terms, a one-unit increase in *RangeN3C* (i.e., an increase of three[36] unique industries) in an audit office industry range is associated with a 16.2 percent[37] decrease in the odds of a client's breach occurrence.

This finding supports H2 and indicates that audit offices with a more diversified portfolio of industries are associated with lower odd that a client will be breached. This can be attributed to several factors. A broad industry range might reduce an audit office's exposure to high-risk industries, as opposed to having a narrow industry range. Alternatively, working with clients from different industries might expose the audit office to a wide array of cybersecurity challenges and measures, enabling it to better advise its clients. Additionally, engaging with clients from multiple industries might foster increased awareness of cyber threats, which, if communicated to clients effectively, can help them shield against potential cyberattacks.

Table 4 also shows that higher leverage is not significantly associated with the likelihood of experiencing a breach. However, my results, consistent with Li et al. (2024) and Yen et al.

[34] For example, if the odds of experiencing a breach for a client firm with an audit office that has *RangeN3C*=10 is 0.3, then the predicted odds of another client firm with an audit office with *RangeN3C*=11 is = 0.3 x 0.84 = 0.252.

[35] Calculated as: e^-0.177. It shows the ratio of odds for those one-unit higher to those one-unit lower on the independent variable *RangeN3C*.

[36] A one unit of natural log is equivalent to about 2.718, rounded up to 3.

[37] Percentage change = (e^-0.177 – 1) * 100 = -16.2%.

(2018), show that firms that experience a loss (coefficient= -0.256, p-value<0.10) and firms audited by a big four audit firm are less likely to experience a breach (coefficient= -0.508, p-value<0.05), whereas bigger firms (coefficient= 0.490, p-value<0.01) and firms with a higher return on assets (coefficient= 0.899, p-value<0.01) are more likely to experience a breach.

[Insert Table 4 here]

4.3. Additional Analyses

4.3.1. Auditor Industry Specialization

Auditor industry specialization has been associated with higher auditing quality due to its increased and focused knowledge in a certain domain (e.g., Balsam et al., 2003; Reichelt & Wang, 2010). Additionally, specialists have been found to provide more effective audits and to better detect errors (Owhoso et al., 2002). Compared to non-specialists, specialists give the impression that they can offer a higher level of assurance (Beasley & Petroni, 2001), and their clients' earnings forecast future cash flows more precisely (Gramling et al., 2001). In contrast, enhanced industry specialization leads to a narrower industry range (Dekeyser et al., 2024), limiting auditor flexibility and skepticism (Bilalić et al., 2008; Dekeyser et al., 2024). To determine the effect of auditor industry specialization in the context of cybersecurity, I replace the industry range variable with the industry specialization variable in my main model. Tables 5 and 6 indicate that the coefficients on the *Industry Specialist1* and *Industry Specialist2* variables are not statistically significant (coefficient= 0.060, p-value >0.1; coefficient= -0.204, p-value >0.1, respectively). This suggests that auditors' industry specialization does not play a role in the breach likelihood of their clients.

[Insert Tables 5 and 6 here]

4.3.2. Auditor Office – State Level

The main tests are conducted by looking at the effect of having the same city-based audit office as other clients. Such offices are considered semi-autonomous within an audit firm, and they have their own client base (Choi et al., 2010). Previous studies in auditing have examined the effect of auditor city-level on audit quality, audit fees, and client cybersecurity risk (e.g., Ferguson et al., 2003; Francis et al., 2005; Li et al., 2024). However, and to the best of my knowledge, no study has examined the auditor effect at the state-level on a client's likelihood of experiencing a breach. State-level tests allow for the assessment of a more diverse clientele across various sectors and industries, which can lead to different concentrations of risk, compared to city-level tests. By broadening the geographic scope of analysis, I can capture a cumulative impact of multiple audit offices across various cities. I use an alternative measure for audit office cybersecurity experience, through breached clients *Pre* and *Post*, whereby I consider common auditors at the state-level instead of at the city-level. Using the state-level, I test the effect of belonging to an audit firm in a specific state, operating at a larger scale, auditing a bigger number of client firms, and having a wider range of experiences across various industries. The results presented in Table 7 reveal similar findings to those identified at the city-level. Having previously breached clients or clients that will experience a breach, is associated with a higher breach likelihood at client firms. However, auditor industry range is associated with a lower breach likelihood for clients.

[Insert Table 7 here]

4.3.3. Alternative Industry Range Measure

In the main model, the audit office industry range variable represents the number of unique industries audited by the audit office in the past 3 years. To further validate my main results, I use an alternative industry range measure, $Range_{Herf}$, which reflects the relative importance of

each industry in the audit office's portfolio. This variable is adopted from Dekeyser et al. (2024) and is based on the Herfindahl index for an audit office client's portfolio over the previous three years (*t-2* to *t*). The variable is generated by first calculating each industry's relative importance (i.e., share) in the audit office portfolio in terms of client-years (*Pi*). The Herfindahl index is then calculated by using the portfolio shares as $[\sum_{i=1}^{N} P_i^2]$, where *N* is the number of industries in the audit office's portfolio. The $Range_{Herf}$ variable is constructed by subtracting the Herfindahl index from one; that way a larger $Range_{Herf}$ value reflects a broader industry range. I replace *RangeN3C* with $Range_{Herf}$ in the main model. The results in Table 8 show that a broader industry range is associated with lower breach odds (coefficient= -0.498, p-value<0.1), providing further support to the main results.

[Insert Table 8 here]

5. Conclusion

My study reveals a potential positive association between the breach likelihoods of clients audited by the same audit office. I argue that cybersecurity risks are being transferred between clients of an audit office. This suggests that auditors are either not placing cybersecurity as a top priority or are not well-equipped to provide support for their clients on matters related to cybersecurity, regardless of whether they have experienced breaches at other clients. However, client breach likelihood decreases when auditors have a broad range of industry experiences. This can be attributed to the enhanced cognitive flexibility and skepticism that auditors develop through their engagements with various industry experiences, helping their clients against cyberattacks. By testing for auditors' industry specialization, I find no significant effect on the breach likelihood, which further confirms my results that auditors' varied experiences, rather than focused ones, seem to help their clients in the setting of data breaches. My findings might

be of interest to academics and regulators due to the timeliness and urgency of cybersecurity-related research.

The results are limited to audit offices which are seen as the channel through which knowledge and risk are transferred. Future studies may shift the focus from the office-level to the audit engagement partner level. Another caveat to my findings is that I do not differentiate between the magnitude of the cyber incidents. Future studies could extend my study when data on the breach magnitude becomes available. Additionally, conducting interviews with external auditors may provide a thorough understanding of the measures taken (if any) by auditors to help their clients shield against cyber incidents. Lastly, I do not differentiate between the types of industries in an audit office clients' portfolio, rather I control for unobserved industry factors that can affect the odds of experiencing a breach. Future studies could deepen the analysis by determining whether the breach likelihood differs based on the client's industry.

References

Asthana, S. C., Kalelkar, R., & Raman, K. K. (2021). Does client cyber-breach have reputational consequences for the local audit office? *Accounting Horizons*, *35*(4), 1-22.

Audousset-Coulier, S., Jeny, A., & Jiang, L. (2016). The validity of auditor industry specialization measures. *Auditing: A Journal of Practice & Theory*, *35*(1), 139-161.

Balsam, S., Krishnan, J., & Yang, J. S. (2003). Auditor industry specialization and earnings quality. *Auditing: A Journal of Practice & Theory*, *22*(2), 71-97.

Beasley, M. S., & Petroni, K. R. (2001). Board independence and audit-firm type. *Auditing: A Journal of Practice & Theory*, *20*(1), 97-114.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*(3), 652-661.

Bills, K. L., Cobabe, M., Pittman, J., & Stein, S. E. (2020). To share or not to share: The importance of peer firm similarity to auditor choice. *Accounting, Organizations and Society*, *83*, 101115.

Bills, K. L., Jeter, D. C., & Stein, S. E. (2015). Auditor industry specialization and evidence of cost efficiencies in homogenous industries. *The Accounting Review*, *90*(5), 1721-1754.

Bouwman, C. H. (2011). Corporate governance propagation through overlapping directors. *The Review of Financial Studies*, *24*(7), 2358-2394.

Cahan, S. F., Godfrey, J. M., Hamilton, J., & Jeter, D. C. (2008). Auditor specialization, auditor dominance, and audit fees: The role of investment opportunities. *The Accounting Review*, *83*(6), 1393-1423.

Cai, Y., Kim, Y., Park, J. C., & White, H. D. (2016). Common auditors in M&A transactions. *Journal of Accounting and Economics*, *61*(1), 77-99.

Cao, V. N., & Pham, A. V. (2021). Behavioral spillover between firms with shared auditors: The monitoring role of capital market investors. *Journal of Corporate Finance*, *68*, 101914.

Center for Audit Quality (CAQ). (2019). Understanding Cybersecurity and the External Audit. Washington, DC: CAQ. Retrieved January 15, 2025 from www.thecaq.org/wp-content/uploads/2019/03/cybersecurity_and_external_audit_final.pdf

Center for Strategic and International Studies (CSIS). (2018). The Economic Impact of Cybercrime—No Slowing Down. McAfee. Retrieved January 10, 2025, from www.marylandnonprofits.org/wp-content/uploads/mcafee.pdf

Cheng, S., Felix, R., & Indjejikian, R. (2019). Spillover effects of internal control weakness disclosures: the role of audit committees and board connections. *Contemporary Accounting Research*, *36*(2), 934–957.

Choi, J. H., Kim, C., Kim, J. B., & Zang, Y. (2010). Audit office size, audit quality, and audit pricing. *Auditing: A Journal of Practice & Theory*, *29*(1), 73-97.

Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *Academy of Management Review*, *35*(4), 579-603.

Dekeyser, S., He, X., Xiao, T., & Zuo, L. (2024). Auditor industry range and audit quality. *Journal of Accounting and Economics*, *77*(2-3), 101669.

Dhaliwal, D. S., Lamoreaux, P. T., Litov, L. P., & Neyland, J. B. (2016). Shared auditors in mergers and acquisitions. *Journal of Accounting and Economics*, *61*(1), 49-76.

Epstein, D. (2019). Range: Why Generalists Triumph in a Specialized World. Riverhead Books, New York.

Ferguson, A., Francis, J. R., & Stokes, D. J. (2003). The effects of firm-wide and office-level industry expertise on audit pricing. *The Accounting Review*, *78*(2), 429-448.

Francis, J. R., Pinnuck, M. L., & Watanabe, O. (2014). Auditor style and financial statement comparability. *The Accounting Review*, *89*(2), 605-633.

Francis, J. R., Reichelt, K., & Wang, D. (2005). The pricing of national and city-specific reputations for industry expertise in the US audit market. *The Accounting Review*, *80*(1), 113-136.

Franzel, J. (2016). Audit Expectations Gap: A Framework for Regulatory Analysis. PCAOB Board member speech at International Institute on Audit Regulations. (December 13). Retrieved January 16, 2025, from https://pcaobus.org/news-events/speeches/speech-detail/audit-expectations-gap-a-framework-for-regulatory-analysis_640

Gramling, A. A., & Stone, D. N. (2001). Audit firm industry expertise: A review and synthesis of the archival literature. *Journal of Accounting Literature*, *20*, 1.

Haislip, J. Z., Peters, G. F., & Richardson, V. J. (2016). The effect of auditor IT expertise on internal controls. *International Journal of Accounting Information Systems*, *20*, 1-15.

Hamm, K. (2019). Cybersecurity: Where we are; what more can be done? A call for auditors to lean in. Speech by PCAOB board member. Retrieved January 16, 2025, from https://pcaobus.org/news-events/speeches/speech-detail/cybersecurity-where-we-are-what-more-can-be-done-a-call-for-auditors-to-lean-in_700

Hanlon, M., Yeung, K., & Zuo, L. (2022). Behavioral economics of accounting: A review of archival research on individual decision makers. *Contemporary Accounting Research*, *39*(2), 1150-1214.

Hargadon, A. B. (2006). Bridging old worlds and building new ones: Toward a microsociology of creativity. In *Creativity and Innovation in Organizational Teams* (pp. 219-236). Psychology Press.

Higgs, J. L., Pinsker, R. E., Smith, T. J., & Young, G. R. (2016). The relationship between board-level technology committees and reported security breaches. *Journal of Information Systems*, *30*(3), 79-98.

IBM (2023). Cost of a Data Breach Report 2023. Retrieved January 9, 2024, from https://www.ibm.com/reports/data-breach

Kang, J. K., Lennox, C., & Pandey, V. (2022). Client concerns about information spillovers from sharing audit partners. *Journal of Accounting and Economics*, *73*(1), 101434.

Lawrence, A., Minutti-Meza, M., & Vyas, D. (2018). Is operational control risk informative of financial reporting deficiencies? *Auditing: A Journal of Practice & Theory*, *37*(1), 139-165.

Lennox, C., & Li, B. (2014). Accounting misstatements following lawsuits against auditors. *Journal of Accounting and Economics*, *57*(1), 58-75.

Li, H., No, W. G., & Boritz, J. E. (2020). Are external auditors concerned about cyber incidents? Evidence from audit fees. *Auditing: A Journal of Practice & Theory*, *39*(1), 151-171.

Li, H., No, W. G., & Wang, T. (2018). SEC's cybersecurity disclosure guidance and disclosed cybersecurity risk factors. *International Journal of Accounting Information Systems*, *30*, 40-55.

Li, H., Sun, Z., & Huang, F. (2024). The impact of audit office cybersecurity experience on nonbreach client's audit fees and cybersecurity risks. *Journal of Information Systems, 38*(1), 177-206.

Libby, R., & Luft, J. (1993). Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society, 18*(5), 425-450.

Marcoulides, K. M., & Raykov, T. (2019). Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educational and Psychological Measurement, 79*(5), 874-882.McAllister, B., & Cripe, B. (2008). Improper release of proprietary information. *The CPA Journal*, *78*(3), 52.

Mckenna, F. (2018). Unit of Equifax's auditor EY certified the information security that was later breached. *MarketWatch.* Retrieved December 28, 2024, from www.marketwatch.com/story/unit-of-equifaxs-auditor-ey-certified-the-information-security-that-was-later-breached-2018-12-20

Mintz, S. (2017). Ernst & Young auditors should be held responsible for fraud at Equifax. Retrieved December 28, 2024, from www.stevenmintzethics.com/single-post/2017/10/11/ernst-young-auditors-should-be-held-responsible-for-fraud-at-equifax

Nelson, M. W., Proell, C. A., & Randel, A. E. (2016). Team-oriented leadership and auditors' willingness to raise audit issues. *The Accounting Review*, *91*(6), 1781-1805.

Nelson, M., & Tan, H. T. (2005). Judgment and decision making research in auditing: A task, person, and interpersonal interaction perspective. *Auditing: A Journal of Practice & Theory, 24*(s-1), 41-71.

Owhoso, V. E., Messier, Jr, W. F., & Lynch, Jr, J. G. (2002). Error detection by industry-specialized teams during sequential audit review. *Journal of Accounting Research*, *40*(3), 883-900.

Perols, R. R., & Murthy, U. S. (2021). The impact of cybersecurity risk management examinations and cybersecurity incidents on investor perceptions and decisions. *Auditing: A Journal of Practice & Theory*, *40*(1), 73-89.

Public Company Accounting Oversight Board (PCAOB). (2019). "Keep Calm and Carry on": The Role of Regulators in Cybersecurity and Resiliency. Washington, DC: PCAOB. Retrieved January 10, 2025, from https://pcaobus.org/news-events/speeches/speech-detail/-keep-calm-and-carry-on-the-role-of-regulators-in-cybersecurity-and-resiliency_705

Public Company Accounting Oversight Board (PCAOB). (2019). Cybersecurity: Where We Are; What More Can be Done? A Call for Auditors to Lean In. New York, NY: PCAOB. Retrieved March 8, 2025, from https://pcaobus.org/news-events/speeches/speech-detail/cybersecurity-where-we-are-what-more-can-be-done-a-call-for-auditors-to-lean-in_700

Public Company Accounting Oversight Board (PCAOB). (2020). Strategic Plan 2020-2024. Washington, DC: PCAOB. Retrieved January 10, 2025, from https://pcaob-assets.azureedge.net/pcaob-dev/docs/default-source/about/administration/documents/strategic_plans/strategic-plan-2020-2024.pdf?sfvrsn=776073d3_4

Reichelt, K. J., & Wang, D. (2010). National and office-specific measures of auditor industry expertise and effects on audit quality. *Journal of Accounting Research*, *48*(3), 647-686.

Say, G., & Vasudeva, G. (2020). Learning from digital failures? The effectiveness of firms' divestiture and management turnover responses to data breaches. *Strategy Science*, *5*(2), 117-142.

Seavey, S. E., Imhof, M. J., & Westfall, T. J. (2018). Audit firms as networks of offices. *Auditing: A Journal of Practice & Theory*, *37*(3), 211-242.

Securities and Exchange Commission (SEC). (2018). Commission Statement and Guidance on Public Company Cybersecurity Disclosures. Washington, DC: SEC. Retrieved January 16, 2025, from https://www.sec.gov/rules/interp/2018/33-10459.pdf

Smith, T. J., Higgs, J. L., & Pinsker, R. E. (2019). Do auditors price breach risk in their audit fees? *Journal of Information Systems*, *33*(2), 177-204.

Srikantia, P., & Pasmore, W. (1996). Conviction and doubt in organizational learning. *Journal of Organizational Change Management*, *9*(1), 42-53.

Valdetero, J., Zetoony, D., & Maciejewski, A. (2019). Data breach litigation report. 2019 Edition. Retrieved January 15, 2025, from https://www.bclplaw.com/images/content/1/6/v6/163774/2019-Litigation-Report.pdf

World Economic Forum (WEF) (2019). The Global Risks Report 2019-14th Edition. Geneva. Retrieved February 16, 2025, from https://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf

Yen, J. C., Lim, J. H., Wang, T., & Hsu, C. (2018). The impact of audit firms' characteristics on audit fees following information security breaches. *Journal of Accounting and Public Policy*, *37*(6), 489-507.

Table 1. Sample Selection

Panel A: Data Breach Incidents

	N
Total data breach incidents from 2011 to 2024 (Audit Analytics Cybersecurity)	1,145
Less: Duplicate observations and firms not included in Compustat	-725
Final sample of data breach incidents	420

Panel B: Final Sample

	N
Firm-year observations from Compustat	152,732
Auditor-related observations from Audit Analytics	112,765
Final matched firm-level dataset	**27,242**
Number of unique firm observations	3,352

Panel C: Year Breakdown

Year	Number of Observations	Percent
2011	1,403	5.15
2012	1,464	5.37
2013	1,584	5.81
2014	1,726	6.34
2015	1,791	6.57
2016	1,865	6.85
2017	1,995	7.32
2018	2,096	7.69
2019	2,248	8.25
2020	2,353	8.64
2021	2538	9.32
2022	2,845	10.44
2023	2,767	10.16
2024	567	2.08
Total	27,242	100.00

Notes: Panel A presents the number of breach incidents from 2011 to 2024, and the final sample of breaches used. Panel B presents the construction of my final sample covering the period 2011-2024. Panel C presents the sample breakdown by year.

Table 2. Descriptive Statistics

Variables	N	Mean	Std. Dev.	Min	Max
Breached Firm $_{i,t}$	27,242	0.015	0.123	0	1
AudCyber1 Pre $_{i,t}$	27,242	0.316	0.465	0	1
AudCyber1 Post $_{i,t}$	27,242	0.249	0.432	0	1
RangeN3C $_{i,t}$	27,242	3.228	0.915	1.099	5.501
Big4 $_{i,t}$	27,242	0.637	0.481	0	1
Firm Size $_{i,t-1}$	27,242	6.127	2.729	0.001	13.221
ROA $_{i,t-1}$	27,242	-2.229	57.074	-5,325.5	7.253
Loss $_{i,t-1}$	27,242	0.436	0.496	0	1
Leverage $_{i,t-1}$	27,242	5.155	150.273	0	15,567
AudCyber2 Pre $_{i,t}$	27,242	0.419	0.493	0	1
AudCyber2 Post $_{i,t}$	27,242	0.359	0.480	0	1
RangeN3S $_{i,t}$	27,242	3.718	0.980	1.099	5.497
Industry Specialist1 $_{i,t}$	27,242	0.172	0.377	0	1
Industry Specialist2 $_{i,t}$	27,242	0.073	0.260	0	1
Range$_{Herf\ i,t}$	27,242	0.884	0.242	-1.250	0.999

Notes: This table provides descriptive statistics for the analyses' variables. The variables are defined in Appendix A.

Table 3. Pairwise correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Breached Firm $_{i,t}$	1.000													
(2) AudCyber1 Pre $_{i,t}$	**0.057**	1.000												
(3) AudCyber1Post $_{i,t}$	**0.094**	**0.390**	1.000											
(4) RangeN3C $_{i,t}$	**0.028**	**0.416**	**0.390**	1.000										
(5) Big4 $_{i,t}$	**0.071**	**0.378**	**0.394**	**0.309**	1.000									
(6) Firm Size $_{i,t-1}$	**0.134**	**0.306**	**0.328**	**0.156**	**0.702**	1.000								
(7) ROA $_{i,t-1}$	0.005	**0.025**	**0.022**	**-0.019**	**0.052**	**0.087**	1.000							
(8) Loss $_{i,t-1}$	**-0.070**	**-0.103**	**-0.083**	**0.031**	**-0.321**	**-0.517**	**-0.044**	1.000						
(9) Leverage $_{i,t-1}$	-0.004	**-0.018**	**-0.017**	**0.017**	**-0.040**	**-0.068**	**-0.291**	**0.027**	1.000					
(10) AudCyber2 Pre $_{i,t}$	**0.040**	**0.800**	**0.296**	**0.344**	**0.462**	**0.355**	**0.031**	**-0.114**	**-0.024**	1.000				
(11) AudCyber2 Post $_{i,t}$	**0.081**	**0.349**	**0.769**	**0.343**	**0.503**	**0.383**	**0.029**	**-0.083**	**-0.023**	**0.426**	1.000			
(12) RangeN3S $_{i,t}$	**0.030**	**0.362**	**0.335**	**0.763**	**0.429**	**0.267**	0.001	0.002	0.002	**0.494**	**0.512**	1.000		
(13) Industry Specialist1 $_{i,t}$	**0.030**	**0.040**	**0.042**	0.004	**0.152**	**0.227**	**0.017**	**-0.183**	*-0.012*	**0.039**	**0.037**	0.002	1.000	
(14) Industry Specialist2 $_{i,t}$	-0.010	**-0.024**	**-0.016**	**-0.022**	**0.036**	**0.045**	0.009	**-0.091**	-0.007	**-0.028**	**-0.016**	**-0.036**	**0.616**	
(15) $Range_{Herf\ i,t}$	*0.014*	**0.205**	**0.177**	**0.602**	**0.192**	**0.124**	0.004	-0.008	0.008	**0.171**	**0.165**	**0.465**	**0.018**	1.000

Notes: This table presents the correlations among the variables. Values in bold represent significance at the 1% level, values in italic at the 5% level, and * denotes significance at the 10% level. The variables are defined in Appendix A.

Table 4: Effect of Auditor Cybersecurity Experience and Industry Range on Breach Likelihood

Logit Regression – Dependent Variable: *Breached Firm* $_{i,t}$				
Independent Variables	Coef.	St.Err.	[95% Conf	Interval]
AudCyber1 Pre $_{i,t}$	0.573***	0.141	0.297	0.849
AudCyber1 Post $_{i,t}$	0.421***	0.123	0.179	0.662
RangeN3C $_{i,t}$	-0.177**	0.071	-0.317	-0.037
Big4 $_{i,t}$	-0.508**	0.210	-0.919	-0.097
Firm Size $_{i,t-1}$	0.490***	0.032	0.428	0.552
ROA $_{i,t-1}$	0.899***	0.279	0.352	1.447
Loss $_{i,t-1}$	-0.256*	0.156	-0.561	0.049
Leverage $_{i,t-1}$	0.008	0.014	-0.019	0.035
Year FE	YES			
Industry FE	YES			
Constant	-9.844***	0.744	-11.303	-8.385
Pseudo R2	0.198			
Observations	27,242			

Notes: This table presents the logit regression output for my model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *AudCyber1 Pre* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the prior 3 years, *AudCyber1 Post* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the following 3 years, and *RangeN3C* $_{i,t}$ which reflects the number of unique industries in the auditor's client base in the previous 3 years; all at the city-level. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 5: Effect of Auditor Industry Specialist (*Industry Specialist1*) on Breach Likelihood

Logistic regression – Dependent Variable: *Breached Firm* $_{i,t}$				
Independent Variables	Coef.	St.Err.	[95% Conf	Interval]
AudCyber1 Pre $_{i,t}$	0.471***	0.135	0.207	0.735
AudCyber1 Post $_{i,t}$	0.329***	0.117	0.099	0.560
Industry Specialist1 $_{i,t}$	0.060	0.123	-0.181	0.301
Big4 $_{i,t}$	-0.595***	0.207	-1.001	-0.189
Firm Size $_{i,t-1}$	0.494***	0.032	0.431	0.556
ROA $_{i,t-1}$	0.890***	0.272	0.357	1.423
Loss $_{i,t-1}$	-0.269*	0.155	-0.573	0.036
Leverage $_{i,t-1}$	0.007	0.016	-0.024	0.039
Year FE	YES			
Industry FE	YES			
Constant	-10.322***	0.720	-11.734	-8.910
Pseudo R2	0.196			
Observations	27,242			

Notes: This table presents the logit regression output for my model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *AudCyber1 Pre* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the prior 3 years, and *AudCyber1 Post* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the following 3 years, and *Industry Specialist1* $_{i,t}$ which reflects whether the audit office is an industry specialist or not; all at the city-level. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 6: Effect of Auditor Industry Specialist (*Industry Specialist2*) on Breach Likelihood

Logistic regression – Dependent Variable: *Breached Firm* $_{i,t}$				
Independent Variables	Coef.	St.Err.	[95% Conf	Interval]
AudCyber1 Pre $_{i,t}$	0.467***	0.135	0.203	0.731
AudCyber1 Post $_{i,t}$	0.328***	0.118	0.098	0.559
Industry Specialist2 $_{i,t}$	-0.204	0.230	-0.655	0.247
Big4 $_{i,t}$	-0.590***	0.207	-0.997	-0.184
Firm Size $_{i,t-1}$	0.495***	0.032	0.433	0.557
ROA $_{i,t-1}$	0.889***	0.272	0.357	1.422
Loss $_{i,t-1}$	-0.276*	0.156	-0.581	0.029
Leverage $_{i,t-1}$	0.007	0.016	-0.025	0.040
Year FE	YES			
Industry FE	YES			
Constant	-10.299***	0.721	-11.712	-8.886
Pseudo R2	0.196			
Observations	27,242			

Notes: This table presents the logit regression output for my model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *AudCyber1 Pre* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the prior 3 years, and *AudCyber1 Post* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the following 3 years, and *Industry Specialist2* $_{i,t}$ which reflects whether the audit office is an industry specialist or not; all at the city-level. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 7: Effect of Auditor Cybersecurity Experience and Industry Range on Breach Likelihood – State-level

Logistic regression – Dependent Variable: *Breached Firm* $_{i,t}$

Independent Variables	Coef.	St.Err.	[95% Conf	Interval]
AudCyber2 Pre $_{i,t}$	0.298**	0.147	0.009	0.586
AudCyber2 Post $_{i,t}$	0.367***	0.132	0.109	0.625
RangeN3S $_{i,t}$	-0.149**	0.070	-0.285	-0.012
Big4 $_{i,t}$	-0.477**	0.213	-0.895	-0.058
Firm Size $_{i,t-1}$	0.506***	0.032	0.444	0.568
ROA $_{i,t-1}$	0.896***	0.275	0.357	1.435
Loss $_{i,t-1}$	-0.261*	0.155	-0.565	0.043
Leverage $_{i,t-1}$	0.008	0.014	-0.019	0.035
Year FE	YES			
Industry FE	YES			
Constant	-10.074***	0.750	-11.545	-8.604
Pseudo R2	0.193			
Observations	27,242			

Notes: This table presents the logit regression output for my model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *AudCyber2 Pre* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the prior 3 years, and *AudCyber2 Post* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the following 3 years, and *RangeN3S* $_{i,t}$ which reflects the number of unique industries in the auditor's client base in the previous 3 years; all at the state-level. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 8: Effect of Auditor Cybersecurity Experience and Industry Range on Breach Likelihood

Logit Regression – Dependent Variable: *Breached Firm* $_{i,t}$				
Independent Variables	Coef.	St.Err.	[95% Conf	Interval]
AudCyber1 Pre $_{i,t}$	0.514***	0.137	0.245	0.784
AudCyber1 Post $_{i,t}$	0.355***	0.118	0.123	0.587
$Range_{Herf\ i,t}$	-0.498*	0.258	-1.003	0.008
Big4 $_{i,t}$	-0.564***	0.208	-0.972	-0.156
Firm Size $_{i,t-1}$	0.495***	0.032	0.433	0.556
ROA $_{i,t-1}$	0.897***	0.273	0.361	1.433
Loss $_{i,t-1}$	-0.268*	0.155	-0.572	0.037
Leverage $_{i,t-1}$	0.008	0.015	-0.022	0.037
Year FE	YES			
Industry FE	YES			
Constant	-9.958***	0.742	-11.412	-8.504
Pseudo R2	0.197			
Observations	27,242			

Notes: This table presents the logit regression output for my model. The dependent variable is *Breached Firm* $_{i,t}$, a dummy for whether the firm experienced a breach in year t. The main independent variables are *AudCyber1 Pre* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the prior 3 years, *AudCyber1 Post* $_{i,t}$, which reflects whether the audit office has clients that experienced a breach in the following 3 years, and $Range_{Herf\ i,t}$ which reflects the relative importance of an industry in the audit office portfolio in the previous 3 years; all at the city-level. All the variables are defined in Appendix A. Year and Industry fixed effects are included. ***, **, and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Appendix A. Variables description and source

Variable	Description
Breached Firm $_{i,t}$	Coded 1 if the firm is breached in the period 2011-2024, 0 otherwise *(Audit Analytics)*
AudCyber1 Pre $_{i,t}$	Coded 1 if the audit office (city-level) has clients that experienced a cyber incident at any point in the previous three years, *t-3* to *t-1 (Audit Analytics)*
AudCyber1 Post $_{i,t}$	Coded 1 if the audit office (city-level) has clients that will experience a cyber incident at any point in the following three years, *t+1* to *t+3 (Audit Analytics)*
RangeN3C $_{i,t}$	= natural logarithm of the number of unique industries audited by the audit office in the past 3 years, *t-2* to *t*, at the city-level *(Audit Analytics)*
Big4 $_{i,t}$	Coded 1 if the firm is audited by a Big4 audit firm in year *t*, 0 otherwise *(Audit Analytics)*
Firm Size $_{i,t-1}$	= natural logarithm of Total Assets in year *t-1 (Compustat)*
ROA $_{i,t-1}$	= Operating income scaled by total assets, in year *t-1(Compustat)*
Loss $_{i,t-1}$	Coded 1 if Net Income is negative in year *t-1*, 0 otherwise *(Compustat)*
Leverage $_{i,t-1}$	=Total Liabilities/Total Assets, in year *t-1 (Compustat)*
Industry Specialist1 $_{i,t}$	Coded 1 if the audit office has market share of audit fees > 50% in year *t*, 0 otherwise *(Audit Analytics)*
Industry Specialist2 $_{i,t}$	Coded 1 if the audit office has the largest market share of audit fees in a given industry in year *t*, 0 otherwise *(Audit Analytics)*
AudCyber2 Pre $_{i,t}$	Coded 1 if the audit offices (state-level) have clients that experienced a cyber incident at any point in the previous three years, *t-3* to *t-1 (Audit Analytics)*
AudCyber2 Post $_{i,t}$	Coded 1 if the audit offices (state-level) have clients that will experience a cyber incident at any point in the following three years, *t+1* to *t+3 (Audit Analytics)*
RangeN3S $_{i,t}$	= natural logarithm of the number of unique industries audited by the auditor in the past 3 years, *t-2* to *t*, at the state-level *(Audit Analytics)*
$Range_{Herf\ i,t}$	= 1 - $\sum_{i=1}^{N} P_i^2$ in the past 3 years (*t-2* to *t*) at the city-level. *N* is the number of industries in an audit office's portfolio. *Pi* is the share of industry *i* in that portfolio *(Audit Analytics)*

Conclusion

In conclusion, this dissertation investigates a timely topic in today's business world and is relevant in light of the significance of data breaches. The three chapters demonstrate that the effect of breaches is not exclusive to the breached firms, but rather it negatively spills over to other firms. This dissertation identifies indirect effects of breaches and suggests that interlocked boards and audit offices act as networking channels through which the risk of data breaches is transferred. However, certain attributes help decrease the likelihood of breach occurrence. For instance, the presence of female directors, directors with audit roles, or executive directors, all with breach experience, and audit offices with a broad range of industry experiences are associated with a lower breach likelihood. Additionally, findings suggest that firms take proactive measures when operating in cyber-risky environments. Specifically, firms seem to decrease the frequency of their scientific publications when the risk of data breaches increases.

Despite its limitations, this dissertation offers valuable insights for academics and practitioners and lays the foundation for further research. Future studies could provide additional insights by identifying whether the results of this dissertation continue to hold in other institutional environments (for example in Europe or in China). Other topics worth exploring include the impact of adopting technologies such as artificial intelligence (AI), cloud computing, and the Internet of Things (IoT) on firms' cybersecurity risks. Additionally, future studies could use qualitative approaches to understand firms' decision-making strategies regarding data security, considering the interplay between financial, legal, technical, and reputational factors. I believe that global interconnectedness, increased severity and frequency, reputational risks, and heightened regulatory scrutiny underscore the need for further research in this area.

Conclusión

En conclusión, esta tesis ha abordado un tema de gran relevancia en el contexto empresarial actual, especialmente a la luz de la creciente importancia de las filtraciones de datos y los ciberataques. Los tres capítulos evidencian que el efecto de las filtraciones de datos no se limita a las empresas afectadas, sino que se propaga a otras empresas. Esta tesis identifica los efectos secundarios de las filtraciones de seguridad y sugiere que los consejos de administración que tienen consejeros vinculados y las oficinas de las firmas de auditoría actúan como canales de redes a través de los cuales se transfiere el riesgo de las filtraciones de datos. Sin embargo, ciertos atributos ayudan a disminuir la probabilidad de que ocurra una filtración vía contagio. Por ejemplo, las empresas con presencia de consejeras, consejeros miembros del comité de auditoría, o consejeros ejecutivos, todos con experiencia en filtraciones, y las empresas que contratan auditores que pertenecen a oficinas de firmas de auditoría con experiencia en múltiples industrias, están asociadas con una menor probabilidad de filtración. Además, los resultados indican que las empresas toman medidas proactivas al operar en entornos de riesgo cibernético. Específicamente, las empresas parecen reducir la frecuencia de sus publicaciones científicas a medida que aumenta el riesgo de filtraciones de datos.

A pesar de sus limitaciones, esta tesis ofrece importantes contribuciones tanto para académicos como para profesionales, y sienta las bases para investigaciones futuras. Estudios futuros podrían contribuir al conocimiento investigando si los resultados de esta tesis se reproducen en otros entornos institucionales (por ejemplo, en Europa o en China). Otros temas que merecen ser analizados incluyen el impacto de la adopción de tecnologías como la inteligencia artificial (IA), la computación en la nube y el Internet de las Cosas (IoT) sobre los riesgos de ciberseguridad de las empresas. Además, los estudios futuros podrían utilizar enfoques cualitativos para comprender mejor las estrategias de toma de decisiones de las empresas en cuanto a la seguridad de los datos, teniendo en cuenta la interacción entre factores

financieros, legales, técnicos y reputacionales. La interconexión global, el aumento de la gravedad y frecuencia de los incidentes vinculados a la ciberseguridad, los riesgos reputacionales y el mayor escrutinio regulatorio ponen de manifiesto la necesidad de continuar investigando en esta área.