The Role of Enterprise Systems Sourcing on Cybersecurity Breaches

Scott Chu
The University of British Columbia
scott.chu@sauder.ubc.ca

Hasan Cavusoglu
The University of British Columbia
cavusoglu@sauder.ubc.ca

Abstract

This study examines the relationship between enterprise systems (ES) sourcing and cyber breaches. We argue that the greater degree of compatibility resulting from sourcing ES modules from fewer vendors (i.e., higher ES standardization) enables easier management of the implemented modules and reduces cybersecurity risks. We find that ES standardization is negatively related to cyber breach occurrence. When analyzing different breach types, we find that ES standardization is negatively related to both cyber breaches caused by external perpetrators and those caused by organizational insiders. While there is no conclusive evidence to suggest whether sourcing ES modules from fewer vendors is generally advantageous or disadvantageous, our findings indicate that it is advantageous from a cybersecurity standpoint.

Keywords: Enterprise Systems, Enterprise Systems Standardization, IT Investment, Cybersecurity, Data Breach

1. Introduction

Enterprise systems (ES) implementations are major information technology (IT) initiatives, as evidenced by the estimated 913 billion U.S. dollars spent worldwide on enterprise software in 2023.[1] Despite their potential benefits, ES implementations can pose substantial cybersecurity risks. In recent years, multiple warnings have highlighted the potential for cyberattacks stemming from security vulnerabilities in ES applications. For example, in February 2022 the U.S. Cybersecurity and Infrastructure Security Agency (CISA) issued a warning related to security flaws in unpatched SAP applications.[2] A few months later these SAP vulnerabilities were added to the U.S. CISA *Known Exploited Vulnerabilities Catalog*, suggesting that hackers were attempting to exploit them. Such cases underscore the vulnerability of ES to malicious cyber activity, which should motivate firms to configure their ES in a manner that can mitigate cybersecurity risks.

One aspect of a firm's ES strategy that we consider is the number of vendors from which the firm sources its ES. Selecting appropriate vendors is a challenging yet critical aspect of ES implementations since the off-the-shelf solutions offered by different vendors can vary in terms of system architecture, user interfaces, the modeling of business processes, and specific functionalities (Lee et al., 2022). In addition, it is unclear whether sourcing modules from fewer vendors is generally advantageous or disadvantageous (Angst et al., 2017b). On one hand, sourcing ES modules from more vendors can offer greater flexibility and better alignment of the ES modules with the firm's business processes (e.g., Light et al., 2001). Indeed, firms may adopt a "best of breed" strategy that involves sourcing ES modules from different vendors and integrating them. For example, a firm may purchase a customer relationship management (CRM) module from Oracle and an accounting module from SAP. In such a scenario, it would be necessary to integrate these modules to allow for the seamless flow of sales information, for example, from the CRM module to the accounting module. However, relying on more vendors can increase compatibility issues between dissimilar ES modules, which would require more expertise and attention from IT staff (Angst et al., 2017b). By specifically focusing on the cybersecurity implications of ES sourcing choices, we can contribute to the debate surrounding whether it is advantageous to source ES modules from fewer vendors.

This study is driven by the following research question: How does the number of vendors used by a firm for its ES modules impact its cyber breach

[1] https://www-statista-com.eu1.proxy.openathens.net/statistics/203428/total-enterprise-software-revenue-forecast/ (last accessed May 26, 2024)

[2] https://www.cisa.gov/news-events/alerts/2022/02/08/critical-vulnerabilities-affecting-sap-applications-employing (last accessed May 26, 2024)

URI: https://hdl.handle.net/10125/109560
978-0-9981331-8-8

occurrence? Previous research suggests the number of ES vendors is an important aspect of firms' ES strategy (Du, 2015; Angst et al., 2017b; Cao et al., 2022a, 2022b; Gong et al., 2023). This study stands in contrast to previous firm-level studies that focus on how the number of vendors that a firm uses for its ES modules affects outcomes more related to value generation, such as firm performance (Gong et al., 2023) or mergers and acquisitions (M&A) outcomes (Du, 2015; Cao et al., 2022a, 2022b). By focusing on the cybersecurity implications of ES sourcing, we contribute unique insights into the ongoing academic discussion surrounding the advantages and disadvantages of relying on fewer ES vendors.

To perform our empirical analyses, we use a panel dataset from 2007 to 2017 constructed from the Aberdeen Computer Intelligence Technology Database (CITDB), Advisen Cyber Loss Database (CLD), and Compustat. We build on previous studies to construct a measure of the degree to which a firm relies on fewer vendors for its ES modules across business locations (i.e., sites), which is referred to as *ES standardization* (Du, 2015; Cao et al., 2022a, 2022b). Higher ES standardization corresponds to the firm relying on fewer vendors for its ES modules across sites, which can reduce ES complexity. Overall, we find that ES standardization has a mitigating effect on the occurrence of cyber breaches. In addition, to gain a more comprehensive picture of how ES standardization relates to cyber breach occurrence, we differentiate cyber breaches based on whether the perpetrator responsible is an organizational insider or external to the breached firm. Using this classification scheme, we find that ES standardization is negatively related to both cyber breaches that are caused by organizational insiders and those caused by perpetrators external to the firm.

2. Literature review

There has been limited attention devoted towards investigating the influence of non-security IT-related initiatives on cyber breach occurrence. Previous studies have indicated that cyber breach mitigation is not influenced by overall IT budget (Li et al., 2023), digitalization and network embeddedness (Li et al., 2021). While simply adopting IT may not play a significant role in mitigating cyber breaches, prior studies suggest that certain IT strategies may be effective. For instance, Pang and Tanriverdi (2022) find that increased IT modernization and migration of legacy IT systems to the cloud each play a role in reducing the number of security incidents. In another study, Liu et al. (2020) find that the mitigating influence of centralized IT decision making on the breach likelihood of educational institutions becomes amplified when there are more heterogeneous operating systems and hardware systems. Similarly, Tanriverdi et al. (2020) find that internal IT complexity and external IT complexity increases breach occurrence in multihospital healthcare systems.

Overall, these empirical studies suggest that simply implementing technology is not effective in deterring cyber breaches. Rather, the technology's effectiveness is dependent on its alignment with certain strategic choices that stem from reducing system complexity. We seek to extend this literature stream by examining how cyber breach occurrence is affected by the degree to which a firm relies on fewer vendors for ES modules across its sites (i.e., ES standardization). To the best of our knowledge, previous research has not empirically examined this aspect of firms' ES strategy in the context of cybersecurity.

3. Hypothesis development

ES are a type of information systems that encompass a variety of modules supporting end-to-end business processes. While each module supports specific business capabilities, these seemingly separate systems are often integrated to some degree to function coherently. Generally, a common integration layer is at the core of firms' ES, which collects data from and feeds data into different modules to allow for the real-time view of business processes (Davenport, 1998). Firms tend to select ES from vendors that offer an acceptable level of compatibility to minimize integration costs (Lee et al., 2022). However, it is plausible that systems offered by the same vendor are more compatible and easier to collectively maintain compared to a collection of systems from multiple vendors.

We propose various arguments consistent with the extant literature that support the notion that higher ES standardization can lead to reduced cyber breach occurrence. We define ES standardization as the degree to which a firm's sites source ES from fewer vendors, where fewer vendors correspond to higher standardization (Du, 2015; Cao et al., 2022a, 2022b). We expect the greater degree of compatibility arising from higher standardization is the primary mechanism that relates ES standardization to cyber breach occurrence. First, standardization may facilitate better cybersecurity planning, as it may be easier to establish unified security standards and access management policies (Liu et al., 2020). Second, when ES modules are sourced from a few vendors, the implementation of new software releases/updates may be easier (Ford et al., 2010; Angst et al., 2017b). Moreover, there may be fewer and less disparate security fixes to implement. Coordinating and deciding a strategy to deploy security patches is a significant challenge for enterprises (Cavusoglu et al., 2008), partly because vendors release security patches at varying frequencies. For instance, SAP has an *SAP Security Patch Day* that coincides with the second Tuesday of the month during which security fixes are published. By contrast, Oracle announces security fixes in their *Critical Patch Updates Advisories* on the third Tuesday of January, April, July, and October.[3] When firms use more vendors, this can increase the challenges associated with navigating disparate security fixes released at varying frequencies. As such, using more vendors may present additional hurdles that can delay the realization of enhanced cybersecurity across a firm's sites. Taken together, higher ES standardization may better position the firm to respond to potential weaknesses and repel attacks. Therefore, we hypothesize the following.

Hypothesis: ES standardization is negatively related to cyber breach occurrence.

4. Data and sample

Our panel dataset was constructed from three primary data sources and covers the period 2007 to 2017. Our first data source is the Standard and Poor's Compustat database, which provides accounting and financial data for public firms that file annual 10-K reports with the U.S. Securities and Exchange Commission. To form our sample, we combine the Compustat data with information on cyber breaches from the Advisen CLD and IT-related information from the Aberdeen CITDB.

The CITDB provides various technology- and business-related information at the site level. A *site* is a physical business location that belongs to a firm, where a firm can have multiple sites. The concept of sites is analogous to establishments in the U.S. census data. The technology-related information in the CITDB includes a "snapshot" of a site's IT stock detailing components such as personal computers, servers, network infrastructure, software, and IT employees. The CITDB has been recognized as an authoritative source of IT data and has been used in recent studies that account for the role of various firm-level IT-related initiatives (e.g., Nagle, 2019; Sambhara et al., 2022; Huang et al., 2022).

We obtain information on cyber breaches from the Advisen CLD, which has been used in prior empirical studies (Romanosky, 2016; Aldasoro et al., 2022). Advisen, a subsidiary of Zywave, is a for-profit organization that provides data, media, and technology solutions for the commercial property and casualty insurance market. Advisen maintains a team of analysts who regularly update the CLD and collect information regarding cyber incidents from various public sources including online news websites, online data breach clearinghouses, specialized legal information services, and websites belonging to state or federal governments and agencies. It is worth noting that the primary purpose of this database is to sell credible information that insurance companies can use. As such, there is an incentive for the provider to accurately maintain this data since major discrepancies may be detected by insurance professionals.

It is worth noting the possibility that the dataset may not include recently discovered breaches due, in part, to the likely time lag between breach discovery, public reporting of the event, and then inclusion of the event in the dataset by the provider. Indeed, Tonn et al. (2019) find an average gap of 3.4 years between the event date and the entry date of the event in the dataset. To avoid possible biases arising from recent breaches not being immediately included in the dataset, they remove incidents with dates that are within the four preceding years of when the dataset was acquired. In our case, we acquired the dataset in 2021 and we follow Tonn et al. (2019) by removing breaches in the four preceding years, which is why our sample of breaches ends in 2017.

While major cyber breaches and attacks receive the majority of public attention, there are a noteworthy number of events in the dataset that we do not view as being related to firms' IT systems, networks, or attempts to access thereof. For example, events that involve business practices related to unauthorized data collection or contacting individuals without their consent do not align with our conceptualization of cyber breaches. To address this concern, we filter out events from our sample that stem from the following: unauthorized telephone communication or fax transmissions; lack of firm compliance with privacy laws; unauthorized access or physical theft of a firm-issued laptop/desktop where the firm's information or IT systems are unaffected; point of sale, automatic teller machine, or smartphone/tablet misuse that does not involve the potential compromise of firms' IT systems; and the mishandling of printed documents, portable hard drives, magnetic data storage tapes, thumb drives, or CD-ROM discs. Excluding such events from our sample

[3] See here for more information on the *SAP Security Patch Day* https://support.sap.com/en/my-support/knowledge-base/security-notes-news.html (last accessed May 28, 2024) and here for more information on *Oracle's Critical Patch Updates Advisories* https://www.oracle.com/ca-en/security-alerts/#CriticalPatchUpdates (last accessed May 28, 2024).

is advantageous since the remaining events are plausibly related to characteristics of firms' IT.

After reducing the number of firm-year observations to those that have non-null values for the variables used in our empirical analyses, we are left with a sample of 8,446 firm-year observations from years 2007 to 2017 that are comprised of 1,783 firms. There are 100,700 unique CITDB sites that belong to these 1,783 firms in our sample. In addition, 242 firms experience at least one cyber breach, where 81 firms are breached in multiple years during this period.

5. Variables and measurement

5.1. Dependent variables

The main dependent variable in our study is the total number of cyber breaches that a firm experienced in a given year. Similar to previous IS research that differentiates breach types (e.g., Haislip et al., 2021), we also differentiate externally-initiated and internally-initiated cyber breaches. An externally-initiated breach is one where the perpetrator(s) responsible is not formally affiliated with the firm at the time of the event (e.g., nation state actor, former employee). On the contrary, an internally-initiated breach is one where the breached firm is directly responsible or where the perpetrator(s) can be considered an organizational insider such as an employee or trusted third party. Differentiating externally- and internally-initiated breaches is meaningful as there may be inherent differences in their occurrence and detection. This differentiation allows us to obtain a more comprehensive picture of the impact of ES standardization on cyber breaches. We identify externally-initiated breaches if Advisen reports some sort of external perpetrator, whereas internally-initiated breaches are those that Advisen reports information related to an internal cause. Out of the 1,019 cyber breach events from 2007 to 2017 in our final sample, 764 are identified as externally-initiated and 240 are identified as internally-initiated.

5.2. Enterprise systems standardization

For our ES measures, we focus on four types of modules that are recognized as being widely used by firms: accounting, human resources management (HRM), CRM, and supply chain management (SCM) (Tian & Xu, 2015; Sambhara et al., 2022). We use an entropy measure for ES standardization that closely follows Du (2015) and Cao et al. (2022a, 2022b). For each firm, the standardization calculation for module *i* (e.g., CRM) in year *t* is made as follows:

$$-\Sigma_{j=1}^{M_{i,t}}\left(\frac{Site\ Employees_{i,j,t}}{Firm\ Employees_t} \times ln\frac{Firm\ Employees_t}{Site\ Employees_{i,j,t}}\right) \quad (1)$$

where j refers to the vendor and $M_{i,t}$ the total number of different vendors the firm uses for module *i*. $Site\ Employees_{i,j,t}$ refers to the number of employees belonging to sites that have implemented module *i* offered by the *j*th vendor. Moreover, $Firm\ Employees_t$ refers to the total number of employees across a firm's sites. The firm-level ES standardization measure is calculated by taking the average of the four ES module standardization scores. The interpretation of the measure is that a higher ES standardization value indicates that the firm used fewer vendors for each ES module type across its sites, on average. Theoretically, the maximum possible value of this measure is zero, which corresponds to the firm only using one vendor for each of the four ES module types (Cao et al., 2022b).

5.3. IT-related control variables

We construct several measures to capture other IT characteristics of a firm. First, we utilize a measure for *ES extensiveness*, which captures the average implementation of the four ES modules (Du, 2015; Cao et al., 2022a). The module implementation score is the sum of the proportions of a site's employees to the total firm employees for each site that has installed the particular module. We consider a module as implemented if it is listed by the CITDB as a software product that is implemented at the site, and zero otherwise. The final firm-level measure of ES extensiveness is obtained by taking the average across the four module implementation scores for each firm. This measure can theoretically range between zero and one.

We also control for a firm's security technology implementation by following previous studies that count the number of implemented security technologies (Kwon & Johnson, 2014; Angst et al., 2017a; Li et al., 2021). We use the CITDB to gather the necessary information where we focus on the following eight security technologies: security software, network firewall, intrusion detection system, disaster recovery software, virtual private network, identity management software, access management software, and anti-virus software. We construct our firm-level measure in two broad steps. First, for each technology, we aggregate the number of sites that implemented it where we weight each site using the ratio of site employees to total firm employees (Li et al., 2023). Then, we sum the firm-level implementation values belonging to each of the technologies, producing a final measure for firms' security investment that theoretically ranges from zero to eight.

Finally, we include a measure for total firm-level IT spending. This equals the total IT budget scaled by sales (Kobelsky et al., 2008). The IT budget includes the budget for multiple categories of IT including computer hardware, software, storage, and telecommunications infrastructure. To get the firm-level IT budget, we aggregate the site-level IT budget values from the CITDB.

5.4. Non-IT-related control variables

We control for various firm- and industry-specific non-IT factors that are commonly included in existing firm-level IS studies. First, we control for a firm's diversification to account for the possibility that firm operations across different business segments may require diverse IT solutions to support the various data and workflow requirements (McElheran, 2014). We use the Compustat Historical Segments database to construct the entropy measure of total diversification as described in Palepu (1985). We also control for firm size through the natural logarithm of total assets and for R&D investment, measured as the ratio of R&D expenses to sales. In addition, we include three additional control variables to account for the possibility that a firm's financial situation can affect the timing and magnitude of IT-related expenditures (Kobelsky et al., 2008). First, we include a measure for profitability, defined as operating income before depreciation scaled by total assets. Second, we control for *financial slack*, defined as cash and short-term investments scaled by total assets (Kim & Bettis, 2014). Third, we also control for the leverage ratio, measured as the sum of a firm's long-term debt and debt in current liabilities scaled by total assets. Furthermore, we follow the approach in prior studies to control for *dynamism* and *munificence*. *Dynamism* is measured as the volatility of industry sales, whereas *munificence* is measured as the growth in industry sales (Keats & Hitt, 1988; Cao et al., 2022a).

6. Empirical strategy

We follow prior IS research that uses a Poisson model when the dependent variable is a count variable (Mallipeddi et al., 2021; Cao et al., 2022a). Furthermore, cyber breach occurrence takes a limited range of non-negative integer values, with the outcome equaling zero for a nontrivial fraction of the population, which makes the Poisson regression model a suitable option (Wooldridge, 2010). We model the number of cyber breaches experienced by firm *i* in year *t* using a Poisson regression as follows:

$$NumBreach_{i,t} = G(\beta \times ESStandardization_{i,t} + \gamma \times Controls + \alpha_j + \lambda_t + u_{i,t}) \qquad (2)$$

where $G(.)$ is an exponential function of the covariates. Equation (2) consists of our focal independent variable, $ESStandardization_{i,t}$, and a vector of the control variables, $Controls$. The impact of ES standardization on cyber breach occurrence is captured through the coefficient β, where we predict its estimate will be negative. The coefficient estimates of all control variables are denoted by the vector γ. We include 2-digit SIC industry- and year- fixed effects that are represented by α_j and λ_t, respectively. The error term is represented by $u_{i,t}$. To further examine the antecedents of cyber breach occurrence, $NumBreach_{i,t}$ is re-analyzed with a focus on externally-initiated and internally-initiated breaches.

Even though we utilize industry-fixed effects to control for any industry-level time-invariant omitted variables and year-fixed effects to control for time-related factors, the main endogeneity concern is that there could be pertinent factors not absorbed by the fixed effects that influence ES standardization and cyber breaches. To address the possible contamination of time-variant omitted variable bias, we employ an instrumental variable (IV) approach. Our IV captures the degree of diversity related to the use of different vendors by peer sites that belong to the same 4-digit SIC industry as the focal firm. We use an entropy measure to capture this diversity. In our calculation, we use all site-year observations in the CITDB that are assigned to a 4-digit SIC industry. For each module (e.g., CRM), the peer vendor standardization of a focal firm's peer sites in 4-digit SIC industry *i* during year *t* is defined as follows:

$$-\Sigma_{j=1}^{N_i}\left(\frac{Sites_{i,j,t}}{Total\ Sites_{i,t}} \times ln\frac{Total\ Sites_{i,t}}{Sites_{i,j,t}}\right) \qquad (3)$$

where $Sites_{i,j,t}$ is the number of sites in the same industry *i* as the focal firm, excluding the focal firm's sites, that have adopted vendor *j* for that module. N_i is the total number of vendors that have a presence in industry *i* for that module. $Total\ Sites_{i,t}$ is the total number of peer sites in industry *i*. Using the equation, we calculate the peer vendor standardization scores for each of the four modules and use the average as the measure of a firm's peer vendor standardization. This approach produces a unique industry-level measure that applies to the firm-year, since the focal firm's sites are removed from the calculation above. The interpretation is that a higher value corresponds to peer sites' ES sourcing being more concentrated among a smaller number of vendors, on average.

A suitable instrumental variable needs to satisfy two conditions: (1) it needs to directly influence ES standardization, and (2) it should be unrelated to the error in the main model (i.e., $u_{i,t}$) and thus only affect cyber breach occurrence through its influence on ES standardization. Regarding the first condition, we believe that peer vendor standardization directly affects the focal firm's ES standardization. We posit that firms may be more inclined to standardize the ES modules across sites with the prevalent vendors in the firm's industry. A vendor's industry expertise can be a key selection criterion for evaluating potential ES solutions from that vendor. This is due, in part, to the widespread belief that vendors try to design ES in an attempt to reflect industry best practices (Davenport, 1998). Since it is the vendor that is defining what "best" means (Davenport, 1998) a firm may trust solutions that are adopted by more competitors in the industry. It is possible that firms believe popular vendors in a given industry have received more opportunity to elicit feedback that could be used to improve these vendors' products, which may make these products more attractive to potential client firms. In addition, firms may find it more valuable to favor software with a larger market share since there may be benefits associated with compatibility both within a firm and with the firm's business partners (Chen et al., 2011).

Beyond simply being attracted to solutions offered by popular vendors, we believe that firms will also be inclined to standardize their ES modules with popular vendors in their industry. According to institutional theory, firms seek to establish and maintain legitimacy (i.e., acceptance of a firm by its environment) (Kostova & Zaheer, 1999). The firm's characteristics and actions are key factors that shape legitimacy (Kostova & Zaheer, 1999), leading firms to adopt managerial practices or mimic the strategies of others within their peer group in an attempt to gain legitimacy from these peers (Angst et al., 2017b). As such, if fewer vendors are prevalent within an industry, then a firm may be more inclined to use fewer vendors as well. This tendency may be amplified by the common belief among IT practitioners that ES implementation is a complex undertaking due to considerations such as change management and technical integration. However, easier implementation may be realized by sourcing modules from fewer vendors due to a reduction in integration costs (Angst et al., 2017b; Cao et al., 2022b). To summarize, if a firm is going to source their modules from fewer vendors, it may be more advantageous to utilize vendors that are more popular in the focal firm's industry since these vendors are more attuned to industry best practices. Based on these arguments, we believe that a firm will be more inclined to standardize their ES if peer sites tend to favor a smaller number of ES vendors. Therefore, we expect a positive relationship between the peer vendor standardization and the firm's ES standardization.

Regarding the second condition for a suitable IV (i.e., instrument exogeneity), we believe peer vendor standardization is not correlated with time-variant omitted variables that exist in the error of the main model. Moreover, it is plausible that peers' ES sourcing does not directly influence the cyber risk of the focal firm. The notion of constructing an IV based on the attributes of industry peers aligns with prior firm-level IS research that suggests these IVs can be considered exogenous to firms and can influence firms' IT investment (e.g., Cao et al., 2022a). We adopt this perspective into our IV approach by attempting to exploit exogenous industry-related characteristics to correct for an endogenous firm-level IT-related variable.

To use these IVs to correct for the endogeneity of ES standardization, we follow prior IS studies (Mallipeddi et al., 2021; Cao et al., 2022a) by using the two-step control function approach (Wooldridge, 2015). In the first step, we run an ordinary least squares (OLS) regression of ES standardization on peer vendor standardization and the control variables. Next, we obtain the residual from this first-stage OLS regression. In the second step, we run the Poisson regression of cyber breach occurrence on ES standardization, our control variables, and the first-stage OLS residual to obtain our parameter estimates. In the control function approach, the first-stage residual can be thought of as proxying for factors in the error ($u_{i,t}$ in Equation (2)) that are correlated with the endogenous ES standardization variable (Wooldridge, 2015). The second stage estimation takes the following form:

$$NumBreach_{i,t} = G(\beta \times ESStandardization_{i,t} + \gamma \times Controls + \alpha_j + \lambda_t + \rho \times \widehat{v_{i,t}} + \varepsilon_{i,t}) \quad (4)$$

where $\widehat{v_{i,t}}$ is the first-stage residual obtained from the regression of ES standardization on peer ES standardization and the control variables. The new resulting error term, $\varepsilon_{i,t}$, is uncorrelated with ES standardization by virtue of including $\widehat{v_{i,t}}$ in the regression. As a result, this estimating equation addresses concerns related to the endogeneity of ES standardization.

7. Results

Table 1 presents the descriptive statistics for the sample of firm-year observations used in our empirical analyses. Notably, the number of cyber breaches in firm-year observations range from 0 to 45. The majority of these events refer to separate instances of fraudulent customer account log-ins that potentially compromised customers' sensitive information.

Table 1. Descriptive statistics

Variable	Mean	Std Dev	Min	Max
Num Breaches	0.12	1.12	0.00	45.00
Num External Breaches	0.09	1.01	0.00	44.00
Num Internal Breaches	0.03	0.23	0.00	6.00
ES Standardization	-0.22	0.18	-1.86	0.00
ES Extensiveness	0.30	0.24	0.00	1.00
IT Investment	0.04	0.79	-1.47	40.91
Security Implementation	1.05	1.16	0.00	6.35
Total Diversification	0.24	0.36	-0.14	1.69
Log of Assets	21.19	2.27	9.39	27.71
R&D Investment	5.77	314.40	-38.38	25,684.40
Profitability	0.08	0.36	-20.33	1.70
Financial Slack	0.16	0.18	0.00	0.97
Leverage	0.28	0.74	0.00	56.35
Dynamism	1.03	0.03	1.00	1.33
Munificence	1.03	0.09	0.45	1.81
Peer Vendor Standardization	-0.21	0.17	-1.29	0.00

7.1. Main analyses

One benefit of the control function approach is that the coefficient estimate for ρ, as shown in Equation 4, produces a heteroskedasticity-robust, regression-based Hausman test of whether ES standardization is endogenous (Wooldridge, 2015). The coefficient of the first-stage residual in the second-stage regressions is statistically significant, as shown in Columns 5-7 of Table 2. This suggests that ES standardization is endogenous which justifies the use of an IV approach. When we examine the first-stage results in Column 1 of Table 2, we see a statistically significant positive association between peer vendor standardization and ES standardization ($p < 0.05$). Additionally, the F-statistic associated with the coefficient estimate on peer vendor standardization is 6.22. This validates our argument that when peer sites concentrate their ES sourcing among fewer vendors, a focal firm is more inclined to select their ES modules from fewer vendors (i.e., higher standardization).

The baseline results shown in Columns 2-4 from Table 2 indicate that ES standardization only has a negative relation with internally-initiated cyber breaches ($p < 0.01$; Table 2, Column 4). Columns 5-7 in Table 2 show the control function approach second-stage regressions, which correct for the endogeneity of ES Standardization. These results show that ES standardization has a negative impact on overall cyber breach occurrence ($p < 0.01$; Table 2, Column 5), externally-initiated cyber breach occurrence ($p < 0.01$; Table 2, Column 6), and internally-initiated cyber breach occurrence ($p < 0.05$; Table 2, Column 7). It is worth noting that by using industry- and year- fixed effects, the Poisson fixed-effects regressions drop observations that belong to industries that do not observe any variation in the dependent variable across the sample time period. Overall, these results are consistent with our hypothesis, which predicts a negative relationship between ES standardization and cyber breach occurrence.

It is worth noting that security technology implementation generally exhibits a negative and statistically significant association with cyber breach occurrence. Even though the influence of security technology implementation is not our primary interest, this finding aligns with a priori expectations of how greater security investments can improve cybersecurity. However, this contrasts with prior empirical findings that do not find a negative relation between the level of security investments and data breaches (Sen & Borle, 2015; Angst et al., 2017a; Li et al., 2021; Li et al., 2023). In addition, ES extensiveness generally exhibits a positive and statistically significant relation with the number of cyber breaches a firm experiences. This indicates that higher digitization can introduce increased cybersecurity risks. As such, the positive relation between ES extensiveness and cyber breach occurrence highlights the importance of implementing ES strategies that can mitigate the firm's cybersecurity risk.

7.2. Split-sample analysis

Firms in IT-intensive industries are more prone to cyber-attacks compared to firms in non-IT intensive industries (Haislip et al., 2021). This motivates us to perform a split-sample analysis where we separately examine the relationships between ES standardization and total cyber breach occurrence for IT-intensive and non-IT intensive industries. Following Haislip et al. (2021), a firm is in an IT-intensive industry if it belongs to any of the following 3-digit SIC industries: 357, 367, or 737. Table 3 shows the results of the control function approach second-stage regressions for the IT-intensive industries sample (Column 1) and the non-IT intensive industries sample (Column 2). The coefficient that indicates the relationship between ES standardization and total cyber breaches is negative and statistically significant ($p < 0.05$) for both subsamples. However, the coefficient on ES standardization for the IT-intensive subsample is of greater magnitude than the coefficient on ES standardization for the non-IT intensive subsample. We find a Chow F-statistic of 281.50 against an F critical value of 2.13 at a significance level of 0.01. This indicates that the coefficients on ES standardization from the two subsamples are statistically different. This suggests that using fewer ES vendors to mitigate cyber breaches is a more effective strategy for firms operating in IT-intensive industries. Indeed, firms in IT-intensive industries typically have significant technical knowledge and IT resources (Haislip et al., 2021). Regardless of whether a firm operates in an IT-intensive or non-IT intensive industry, using fewer ES vendors can mitigate cyber breaches.

Table 2. Main analyses

	First Stage	**Baseline**			**Control Function Approach**		
	ES Standardization (1)	**Num Breaches (2)**	**Num External Breaches (3)**	**Number of Internal Breaches (4)**	**Num Breaches (5)**	**Num External Breaches (6)**	**Number of Internal Breaches (7)**
Peer Vendor Standardization	0.07**						
	(0.03)						
ES Standardization		0.22	0.62	-0.98***	-25.30***	-28.10***	-20.52**
		(0.39)	(0.49)	(0.32)	(6.14)	(7.78)	(8.25)
ES Extensiveness	0.17***	1.22***	1.47***	0.41	5.36***	6.15***	3.52**
	(0.01)	(0.45)	(0.51)	(0.39)	(1.13)	(1.44)	(1.42)
Security Implementation	-0.00	-0.24**	-0.28**	-0.16	-0.29***	-0.33***	-0.20
	(0.00)	(0.10)	(0.13)	(0.13)	(0.08)	(0.11)	(0.14)
IT Investment	-0.00***	3.61***	3.74**	3.19***	3.41**	3.48*	3.07**
	(0.00)	(1.26)	(1.54)	(0.89)	(1.66)	(1.98)	(1.47)
Total Diversification	-0.01	-0.28	-0.30	-0.33	-0.54***	-0.59**	-0.54**
	(0.01)	(0.25)	(0.33)	(0.23)	(0.20)	(0.26)	(0.25)
Log of Assets	-0.02***	0.76***	0.77***	0.79***	0.33***	0.27*	0.46***
	(0.00)	(0.04)	(0.04)	(0.07)	(0.11)	(0.14)	(0.15)
R&D Investment	0.00***	-0.27***	-0.28**	-0.25***	-0.26*	-0.27*	-0.24
	(0.00)	(0.10)	(0.12)	(0.07)	(0.15)	(0.15)	(1.06)
Profitability	-0.01	5.36***	5.82***	3.74***	5.20***	5.63***	3.56***
	(0.01)	(1.79)	(1.88)	(0.83)	(1.59)	(1.76)	(1.03)
Financial Slack	0.14***	2.01**	2.26**	0.68	5.40***	6.10***	3.26**
	(0.02)	(0.88)	(1.05)	(0.76)	(1.04)	(1.31)	(1.37)
Leverage	-0.00	-0.86	-0.99	-0.84	-0.95**	-1.08*	-0.94*
	(0.00)	(0.61)	(0.81)	(0.56)	(0.46)	(0.61)	(0.52)
Dynamism	-0.14	-8.56*	-9.64*	-9.69*	-10.38***	-11.23**	-11.68**
	(0.10)	(4.74)	(5.79)	(5.17)	(3.71)	(4.77)	(5.25)
Munificence	-0.02	1.60	1.39	2.53**	0.00	0.69	2.12*
	(0.03)	(1.24)	(1.49)	(1.05)	(0.93)	(1.13)	(1.24)
Stage 1 Residual					25.58***	28.78***	19.58**
					(6.24)	(7.93)	(8.31)
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit Statistics							
Observations	8,446	8,171	8,077	7,733	8,171	8,077	7,733
R^2	0.19						
Pseudo R^2		0.51	0.52	0.37	0.51	0.53	0.37
Squared Correlation		0.47	0.42	0.33	0.49	0.44	0.34

Note: Standard errors, clustered by firm, are in parentheses. Standard errors for the estimates in Columns 5-7 are clustered by firm and based on 1000 bootstrap replications (Wooldridge, 2015). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 3. Split-sample analysis

	IT-Intensive Sample	Non-IT Intensive Sample
	Num Breaches (1)	**Num Breaches (2)**
ES Standardization	-115.66**	-14.92**
	(47.04)	(6.57)
Stage 1 Residual	117.04**	15.21**
	(47.26)	(6.66)
Full Set of Control Variables Included	Yes	Yes
Industry Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
Fit Statistics		
Observations	1,429	6,742
Pseudo R^2	0.73	0.41
Squared Correlation	0.70	0.47

Note: Standard errors, clustered by firm, are in parentheses. Standard errors for the estimates are clustered by firm and based on 1000 bootstrap replications (Wooldridge, 2015). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

8. Discussion

Our study investigates the role of ES standardization on firms' cybersecurity. We argue that using fewer vendors for ES modules across a firm's sites (i.e., higher ES standardization) enables easier management of the modules and their associated cybersecurity risks. Overall, our findings suggest that ES standardization has a mitigating effect on cyber breaches, reducing the incidence of cyber breaches caused by both organizational outsiders and insiders.

By virtue of our instrumental variable approach, one periphery finding of this study is that peer vendor standardization is positively related to a focal firm's ES standardization. This suggests when peer sites in the same industry concentrate the sourcing of their ES among fewer vendors, the focal firm is more inclined to favor fewer vendors for its ES (i.e., higher standardization). We argue that this is because firms are attracted to standardizing their ES with the popular vendors in their industry since these vendors may be more attuned to industry best practices.

It is important to acknowledge that this study has a few limitations that merit consideration. First, our study focuses on large U.S. public firms. As such, our results may not be generalizable to small and medium sized firms or firms based in other countries. Second, it has been acknowledged that there is great risk like selecting inappropriate vendors during the adoption and implementation stages of the ES life cycle (Tian & Xu, 2015). However, we are unable to control for the quality of the ES implementation as information pertaining to implementation success during the early stages of the ES initiative before system go-live is unavailable in the CITDB. Another limitation is that our analyses rely on measures constructed from publicly reported cybersecurity breaches just like prior empirical studies (e.g., D'Arcy et al., 2020; Haislip et al., 2021). Publicly reported incidents may not comprehensively represent the cybersecurity challenges faced by firms as not all cybersecurity incidents are publicly reported. However, our breach data source is the most comprehensive one on cyber incidents available (Romanosky, 2016), compared to the Privacy Rights Clearinghouse database, which is used by the majority of empirical cybersecurity studies (Schlackl et al., 2022), or other data sources like the Identity Theft Resource Center (e.g., Foerderer & Schuetz, 2022).

With ES sourcing being a key decision faced by decision-makers, our findings have practically relevant implications. While there may be advantages to using more ES vendors and allowing individual sites to choose ES modules that best fit their business processes, such an approach might overlook cybersecurity considerations, leading to potential unintended dire consequences for cybersecurity. Our findings suggest that vendors' cybersecurity practices and the interdependency of security risks between modules should be important considerations when developing ES sourcing strategies. Given that characteristics of firms' ES can inadvertently affect the occurrence of cyber breaches, we call for more attention from researchers and practitioners to focus on the interplay between ES and cybersecurity.

9. References

Aldasoro, I., Gambacorta, L., Giudici, P., & Leach, T. (2022). The drivers of cyber risk. *Journal of Financial Stability*, *60*, 100989.

Angst, C. M., Block, E. S., D'Arcy, J., & Kelley, K. (2017a). When do IT security investments matter? Accounting for the influence of institutional factors in the context of healthcare data breaches. *MIS Quarterly*, *41*(3), 893-916.

Angst, C. M., Wowak, K. D., Handley, S. M., & Kelley, K. (2017b). Antecedents of information systems sourcing strategies in U.S. hospitals: A longitudinal study. *MIS Quarterly*, *41*(4), 1129–1152.

Cao, C., Ray, G., Subramani, M., & Gupta, A. (2022a). Enterprise systems and the likelihood of horizontal, vertical, and conglomerate mergers and acquisitions. *MIS Quarterly*, *46*(2), 1227–1242.

Cao, C., Ray, G., Subramani, M., & Gupta, A. (2022b). Enterprise systems and M&A outcomes for acquirers and targets. *MIS Quarterly*, *46*(3), 1295–1322.

Cavusoglu, H., Cavusoglu, H., & Zhang, J. (2008). Security patch management: Share the burden or share the damage? *Management Science*, *54*(4), 657–670.

Chen, P. Y., Kataria, G., & Krishnan, R. (2011). Correlated failures, diversification, and information security risk management. *MIS Quarterly*, *35*(2), 397-422.

Davenport, T. H. (1998). Putting the enterprise into the enterprise system. *Harvard Business Review*, *76*(4), 121-131.

D'Arcy, J., Adjerid, I., Angst, C. M., & Glavas, A. (2020). Too good to be true: Firm social performance and the risk of data breach. *Information Systems Research*, *31*(4), 1200–1223.

Du, K. (2015). Research note - parenting new acquisitions: Acquirers' digital resource redeployment and targets' performance improvement in the U.S. hospital industry. *Information Systems Research*, *26*(4), 829–844.

Foerderer, J., & Schuetz, S. W. (2022). Data breach announcements and stock market reactions: a matter of timing?. *Management Science*, *68*(10), 7298-7322.

Ford, E. W., Menachemi, N., Huerta, T. R., & Yu, F. (2010). Hospital IT adoption strategies associated with implementation success: Implications for achieving meaningful use. *Journal of Healthcare Management*, *55*(3), 175–189.

Gong, J., Liang, Y., & Ramasubbu, N. (2023). Software-vendor diversification: A source of organizational rigidity in adversity?. *Journal of Management Information Systems*, *40*(2), 338-365.

Haislip, J., Lim, J. H., & Pinsker, R. (2021). The impact of executives' IT expertise on reported data security breaches. *Information Systems Research*, *32*(2), 318–334.

Huang, P., Ceccagnoli, M., Forman, C., & Wu, D. J. (2022). IT knowledge spillovers, absorptive capacity, and productivity: Evidence from enterprise software. *Information Systems Research*, *33*(3), 908–934.

Keats, B. W., & Hitt, M. A. (1988). A causal model of linkages among environmental dimensions, macro organizational characteristics, and performance. *Academy of Management Journal*, *31*(3), 570–598.

Kim, C., & Bettis, R. A. (2014). Cash is surprisingly valuable as a strategic asset. *Strategic Management Journal*, *35*(13), 2053-2063.

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.

Kostova, T., & Zaheer, S. (1999). Organizational legitimacy under conditions of complexity: The case of the multinational enterprise. *Academy of Management Review*, *24*(1), 64-81.

Kwon, J., & Johnson, M. E. (2014). Proactive versus reactive security investments in the healthcare sector. *MIS Quarterly*, *38*(2), 451–471.

Lee, K., Han, K., Animesh, A., & Pinsonneault, A. (2022). Does IT matter to acquisitions? The impacts of IT distance on post-acquisition performance. *MIS Quarterly*, *46*(4), 2261– 2288.

Li, H., Yoo, S., & Kettinger, W. J. (2021). The roles of IT strategies and security investments in reducing organizational security breaches. *Journal of Management Information Systems*, *38*(1), 222–245.

Li, W. W., Leung, A. C. M., & Yue, W. T. (2023). Where is IT in information security? the interrelationship among IT investment, security awareness, and data breaches. *MIS Quarterly*, *47*(1), 317–342.

Light, B., Holland, C. P., & Wills, K. (2001). ERP and best of breed: a comparative analysis. *Business Process Management Journal*, *7*(3), 216-224.

Liu, C. W., Huang, P., & Lucas Jr, H. C. (2020). Centralized IT decision making and cybersecurity breaches: Evidence from US higher education institutions. *Journal of Management Information Systems*, *37*(3), 758-787.

Mallipeddi, R. R., Janakiraman, R., Kumar, S., & Gupta, S. (2021). The effects of social media content created by human brands on engagement: Evidence from Indian general election 2014. *Information Systems Research*, *32*(1), 212-237.

McElheran, K. (2014). Delegation in multi-establishment firms: Evidence from I.T. purchasing. *Journal of Economics & Management Strategy*, *23*(2), 225–258.

Nagle, F. (2019). Open source software and firm productivity. *Management Science*, *65*(3), 1191–1215.

Palepu, K. (1985). Diversification strategy, profit performance and the entropy measure. *Strategic Management Journal*, *6*(3), 239–255.

Pang, M. S., & Tanriverdi, H. (2022). Strategic roles of IT modernization and cloud migration in reducing cybersecurity risks of organizations: The case of US federal government. *The Journal of Strategic Information Systems*, *31*(1), 101707.

Romanosky, S. (2016). Examining the costs and causes of cyber incidents. *Journal of Cybersecurity*, *2*(2), 121-135.

Sambhara, C., Rai, A., & Xu, S. X. (2022). Configuring the enterprise systems portfolio: The role of information risk. *Information Systems Research*, *33*(2), 446–463.

Schlackl, F., Link, N., & Hoehle, H. (2022). Antecedents and consequences of data breaches: A systematic review. *Information & Management*, *59*(4), 103638.

Sen, R., & Borle, S. (2015). Estimating the contextual risk of data breach: An empirical approach. *Journal of Management Information Systems*, *32*(2), 314–341.

Tanriverdi, H., Kwon, J., & Im, G. *Data Breaches in Multihospital Systems: Antecedents and Mitigation Mechanisms*. In, *ICIS 2020 Proceedings*, 2020, 10.

Tonn, G., Kesan, J. P., Zhang, L., & Czajkowski, J. (2019). Cyber risk and insurance for transportation infrastructure. *Transport Policy*, *79*, 103–114.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*(2), 420-445.