Empirically Evaluating the Effect of Security Precautions on Cyber Incidents

Neil Gandal
Berglas School of Economics
Tel Aviv University, Israel
gandal@tauex.tau.ac.il

Tyler Moore*
School of Cyber Studies
The University of Tulsa, USA
tyler-moore@utulsa.edu

Michael Riordan
Department of Economics
Columbia University, USA
mhr21@columbia.edu

Noa Barnir
Berglas School of Economics
Tel Aviv University, Israel
noabarnir@gmail.com

Abstract

Available data on firm cybersecurity often exhibits a positive correlation between investment in security precautions and cyber attacks since investments are often made after a firm has been breached. Using survey data from Israeli firms about their cyber defenses, we overcome the endogeneity obstacle using an instrumental variable (IV) drawn from questions about a cybersecurity directive. The resulting regressions examine the causal relationship between security precautions potentially undertaken by enterprises and the likelihood of experiencing a cyber incident. Once suitably instrumented and controlling for characteristics that make some firms more attractive attack targets than others, we find robust evidence that increased adoption of security controls does in fact reduce the likelihood of being breached.

Keywords: Cybersecurity, Precautions, Cyber Incidents, Empirical Study

*Corresponding author. 800 South Tucker Drive, Tulsa, Oklahoma, 74114, United States

1 Introduction

Cybersecurity is widely recognized as essential to the functioning of modern economies. Firms and governments are spending ever greater sums on countermeasures designed to mitigate risks, yet we know surprisingly little about which of these investments, if any, reduce the risk of experiencing a cyber incident[1], and by how much. There has been virtually no econometric evidence showing that firm investment in cybersecurity defenses reduces the likelihood of being attacked.

In this paper, we leverage a dataset gathered by the Israeli Central Bureau of Statistics (CBS) at the behest of the Israeli National Cyber Directorate (INCD). Our objective is to empirically examine how security precautions undertaken by enterprises affect the likelihood of experiencing a cyber incident. We use a very detailed firm-level data set from an ICT use and cybersecurity survey undertaken in 2020-2021 by the INCD and CBS. In our assessment, this is by far the most comprehensive firm-level cybersecurity survey ever undertaken at a national level. We utilize this dataset to establish, for the first time, a causal link between firm-level cybersecurity investments and outcomes (i.e., whether firms experience incidents).

In addition to the value of the findings themselves, we hope that this may serve as motivation and guide to efforts such as the Bureau of Cyber Statistics, whose establishment was recommended by the United States Cyberspace Solarium Commission to "collect and provid[e] statistical data on cybersecurity and the cyber ecosystem to inform policymaking" (King and Gallagher, 2020). To begin to answer such questions empirically, a broader

[1]Throughout the paper, when we refer to "cyber incidents", we follow the definition of incidents set out by Howard and Longstaff (1998), namely a grouping of attacks on computer systems that can be associated by "the distinctiveness of the attackers, attacks, objectives, sites, and timing". Note that malicious intent by attackers is required to be deemed a cyber incident. For brevity, we refer to these as cyber incidents rather than cybersecurity incidents.

population sample is needed. As we wait for such an effort to materialize, it is instructive to begin to answer these questions with suitable data from other sources. Establishing an empirical basis for cybersecurity could improve the efficiency of future investments, which have been growing steadily in recent years with little to show for it. The results presented here should be complemented by new efforts to study cybersecurity effectiveness, such as additional surveys and field experiments.

A huge problem with empirical work on this topic is that the timing of security investments is unknown. Naturally, firms who experience cyber incidents are more likely to invest in security precautions afterwards. If we cannot distinguish when the investment is made, we are likely to find a positive correlation between spending on security and being attacked. The survey enables us to overcome this endogeneity obstacle because it asked whether respondents were aware of a directive to implement security controls and if they implemented them. Since that guidance was issued in June 2017, firms who became aware of it had ample time to respond by adjusting security investments before being asked about it in the survey for the period of 2019–20.

We now summarize the key findings from our analysis. When we run a regression of whether the firm had to handle a cyber incident on the number of security precautions and firm and industry characteristics without instrumenting for precautions, we find that the estimated parameter (coefficient) on the security precautions is, not surprisingly, positive and statistically significant. This is because many firms likely adopted the security measures after suffering a cyber incident and/or were attractive targets for attackers.

Once we instrument for the number of security precautions employed by the firm (using implementation of the directive as the instrument), the estimated parameter on security

precautions is negative and statistically significant. This means that employing more security precautions indeed reduces the probability of suffering a cyber incident. When we use six easy to implement (hereafter, basic) cybersecurity precautions as the security variable, the results are qualitatively unchanged.

Importantly, in our regressions, we include many control variables that take into account the attractiveness of the target to attackers. These variables include firm size (both revenues and employees), as well as dummy variables for high-tech firms, firms that use cloud services, firms in high-risk industries, and firms that use information about visitors' behavior on its website. The estimated parameters on these variables are positive and statistically significant as well.

We do not draw any conclusions about the causal effects of particular security precautions. Rather, we focus on the number of controls as a measure of security posture. We find that for large firms with significant revenues using e-commerce and cloud services (the riskiest firms), adopting all six basic security precautions reduces the probability of experiencing a cyber incident from 80% to 42%. This shows that these six basic security measures make a difference. Moreover, adopting even more controls also has a positive impact. We find that for large firms with significant revenues using cloud services using 18 out of 20 security precautions rather than 13 security precautions reduces the probability of experiencing a cyber incident from 81% to 58%.

The rest of the paper is organized as follows. We review prior work in Section 2. Section 3 describes the survey design, discusses relevant questions and presents summary statistics. Section 4 introduces the instrumental variable and examines its theoretical and empirical properties. Section 5 presents the econometric analysis. Finally, we conclude in Section 6.

2 Background and Prior Work

Few would dispute that cyber risk is a very serious problem for the global economy and for society. But there is a "disconnect" between acknowledgment of the problem and action to address the problem. What is the relationship between preventive measures and cyber incidents, like being targeted by ransomware or experiencing a data breach? Surprisingly little is known about the relationship among these variables, and even less is known at the micro level, that is, at the level of the firm. This should not come as a surprise, given that cybersecurity has long suffered from market failures that undermine the effectiveness of defensive investments (Anderson, 2001; Anderson and Moore, 2006). Information asymmetries are rife. Firms cannot easily discern whether the technological controls they purchase are effective. Additionally, the relationship between investment in security controls and incidents depends on other factors as well, including firm characteristics and attacker effort.

Hence, as Moore et al. (2016) found, firms often simply adopt frameworks of one kind (e.g., the NIST Cybersecurity Framework, COBIT, SANS Critical Controls). While the frameworks direct investments, they do not explicitly evaluate how taking precautions affects the security level of their organization, and ultimately whether those precautions make a breach less likely to occur. Weishäupl et al. (2018) also interview executives about their cybersecurity investment approach, finding that it is often driven by frameworks imposed by compliance obligations and that learning about which investments were effective occurs in an ad hoc manner. A study of UK firms found that those with existing cyber capabilities were more likely to make strategic investments in defensive controls (Fernandez De Arroyabe et al., 2023). Toftegaard (2022) finds mixed results about the effects of the ISO/IEC 27001

certification on the technical security of Norwegian grid operators. Using an international sample of firms, Malliouris and Simpson (2019) find that becoming compliant with ISO/IEC 27001 is associated with negative abnormal stock returns. Meanwhile, a study of Finnish information security managers found evidence of herding behavior in deciding how to invest in cybersecurity controls, deploying a "let's follow others" strategy (Shao et al., 2020).

Fortunately, the research community and industry have begun to fill in some of these gaps. For a comprehensive theoretical treatment, we refer the reader to the systematization undertaken by Woods and Böhme (2021). We describe the most relevant empirical work here. Liu et al. (2015) gather publicly-observable data on organizations' network misconfigurations and data on observed malicious activities originating from a network (e.g., spam and phishing) to construct a classifier to predict whether a data breach is subsequently reported. These crude external measures of security levels were found to be predictive of subsequent adverse outcomes. Sarabi et al. (2016) employed a similar approach, but gathered additional data on business sectors and breach types in order to identify the relative risk of incidents for different industries. Nagle et al. (2017) identified a correlation between a firm's network exposure (as measured by counting the number of open ports) and the incidence of botnet-related activity on that network.[2]

Most similar to our work, a few studies have also examined the relationship between enterprise security and the likelihood of experiencing an incident. For example, Li et al. (2023) compares reports of IT investment at 311 publicly-traded US firms to publicly disclosed data breaches at those firms. Using regulatory filings to the US Securities and Exchange

[2]Furthermore, a number of risk-rating services have appeared (e.g., offerings from SecurityScorecard, QuadMetrics, and BitSight) that commercialize the results found by researchers.

Commission, the authors estimate security awareness by counting the frequency of security-related terms in the filings. Our work differs from Li et al. (2023) because we have data on particular security precautions employed, our sample covers the whole Israeli economy, and we explicitly address the timing issue.

A number of studies have examined cybersecurity investment in the US healthcare sector. Angst et al. (2017) find a positive association between security investment at US hospitals and experiencing a breach. Kwon and Johnson (2014) differentiate between proactive investments that occur before a breach and later investments. Using a proportional hazards model, they observe that hospitals investing before experiencing a breach are less likely to experience a subsequent breach. Liu et al. (2020) study centralized governance of IT functions at universities. Similar to our work, they employ instrumental variables to control for timing issues associated with when cybersecurity breaches take place. Our work differs from these papers because they focus on one specific industry, while, as noted, our work covers broader sectors of the economy.

In our research, we are interested in the causal relationship between security precautions undertaken by firms and outcomes (i.e., incidents). We thus estimate a causal model, where ex-ante firm investment in security precautions and firm characteristics (including attractiveness to attackers) affect the likelihood of experiencing a cyber incident. The survey questions cover these variables. The survey reports characteristics of Israeli firms, the security precautions they claim to adopt, and the experience of cyber incidents. While the survey does include questions about the harms resulting from incidents, only about ten percent of the firms that had to handle a cyber incident indicated that they suffered harm. The small parentage could be attributed to the fact that it is typically hard for firms to estimate the

costs associated with a cyber incident. Hence, we cannot examine the reported harms in this paper.

3 Firm Survey Data

In 2020-2021, the Israeli National Cyber Directorate (INCD), in conjunction with the Israeli Central Bureau of Statistics (CBS) constructed and implemented the most comprehensive survey regarding cybersecurity ever undertaken at the level of the firm. The survey, entitled "Survey of Information and Communication Technologies Usage and Cyber Security in Businesses 2020", includes detailed questions about the types of security controls adopted by organizations and whether the firm had to handle a cyber security attack, as well as questions about Internet use, e-commerce, and other firms characteristics.

3.1 Survey Design and Data Access

The survey population is representative and is drawn from the 29 825 Israeli private sector businesses with more than 10 employees, excluding the following industries: agriculture, finance, education and health. These industries are excluded because they are not overseen by the INCD because they are regulated separately. 2 500 of these firms received the survey, of which 2 020 (81%) responded. This very high response rate is because the survey was an official state survey and firms were required to complete it. Firms were selected at random in layers by industry, corporate structure (multinational and domestic), and firm size in order to mirror the Israeli economy.

The survey was conducted from July 2020 to March 2021. The carefully constructed

survey is very helpful, not only because of the number of firms that have responded, but also for the important detailed information about the security questions within.

We accessed anonymized survey response data in a secure room at the Central Bureau of Statistics. Data remains with the CBS; only the outputs of the statistical scripts that were cleared for public disclosure are available to authors.

3.2 Key Questions and Summary Statistics

The survey asks many questions related to ICT. For our purposes, we focus on a few key questions. First, we are interested in the questions involving awareness to the INCD directive. Firms were asked the following:

1. Is your enterprise aware of the cyber directives and instructions (e.g. the Cyber Defense Methodology for an Organization) published by the Israeli National Cyber Directorate (INCD)?
2. For those enterprises that were aware of the cyber directives and instructions, a follow up question was asked about whether ther the firm implemented the directives (with possible responses of full, partial, or no implementation).

Fully 993 firms indicated that they were aware of the directives and instructions from the INCD. These firms comprise our dataset for analysis, since we can only use our instrumental variable for these firms. Of the 993 firms in the sample that were aware of the cyber directives and instructions, 384 (or 39%) replied that they fully implemented the cyber directives and instructions. These numbers suggest that the firms took the survey seriously and to a large extent answered honestly. One might have expected virtually all firms to indicate that they

were aware of the survey and had implemented the directives. Again, this is likely due to the fact that the survey was an official state survey.

The second key question is whether the firm had to handle cyber incidents in the past 12 months. The specific question in the survey was "Did your organization have to handle any cyber security attacks in the past 12 months?" As explained in the survey, "handling incidents" involves the activity of security teams and involves both attack attempts and successful breaches. The survey makes it clear that incidents must involve malicious actors and not be simply accidents.

Answering "yes" (regardless of whether there was damage) means the variable *incident* takes on the value one. Overall, 46% of the enterprises reported that they had to handle a cyber attack or incident in the last 12 months.

The third key question asks firms about the security measures that they have adopted. This is covered in a single 20-part question, whose English translation appears in Figure 1. Firms simply answered yes or no for each control type. On average, firms employed 15 of the 20 security measures covered by the survey.

By looking at the controls more closely, we identified six basic ones that are relatively easy to implement and often recommended. They are:

- Strong password policy (e.g. periodical password reset, minimum password length policy, requiring combination of letters, numbers and characters)
- Keeping systems up-to-date per manufacturer's recommendation, or at a more frequent rate
- Employing two-factor authentication
- Means of detecting and responding to malware at endpoints and servers (e.g. antivirus

system)

- Use of operating systems and applications which are under full manufacturer's support (not in "end of life" state)
- Data, files and email encryption

37 percent of the firms in the data set employed all six of the basic security precautions.

10.5 Which of the following security measures has been implemented by the enterprise? Please check √ the appropriate answer in each row.

		Yes	No
10.5.1	Strong password policy (e.g. periodical password reset, minimum password length policy, requiring combination of letters, numbers and characters)		
10.5.2	Keeping systems up-to-date per manufacturer's recommendation, or at a more frequent rate		
10.5.3	Biometric identifiers (e.g. fingerprint, facial or voice recognition)		
10.5.4	Data, files and email encryption		
10.5.5	Having multiple updated backup copies, with one offsite and usually offline		
10.5.6	Control over access to enterprises' network, by user filtering		
10.5.7	The enterprise has a virtual private network (VPN) to transfer data safely through public domains		
10.5.8	Conducting a periodic survey on cyber security risks		
10.5.9	Security checks (e.g. system invasion attempts, warning system checks, security procedure assessments, etc.)		
10.5.10	Means of detecting and responding to malware at endpoints and servers (e.g. antivirus system)		
10.5.11	File typing of attachments (e.g. DOC, DOCX, PDF) to be accepted from outside the enterprise ("Whitelist")		
10.5.12	Authentication of Sender ID (DMARC) for handling spam		
10.5.13	Security protection services for the enterprises' email systems, including: detecting and responding to spam and malware (e.g. SEG, Mail Relay)		
10.5.14	URL filtering by means of firewall, Secure Web Gateway (SWG) or cloud-based security software		
10.5.15	User malware download prevention (e.g. Firewall, SWG – Secure Web Gateway) or cloud-based security software		
10.5.16	Use of Multi Factor Authentication		
10.5.17	Periodic recovery tests to ensure proper recovery if needed		
10.5.18	Use of operating systems and applications which are under full manufacturer's support (not in "end of life" state)		
10.5.19	Specifications regarding disaster recovery protocols (e.g. RTO, RPO)		
10.5.20	Cyber insurance		

Figure 1: Survey question dealing with security precautions adopted by the firm.

The final grouping of key questions relate to firm characteristics including size, ICT use, and industry.

Incidents vary significantly by firm characteristics. Many large firms with significant

revenues and many employees, firms in high-risk industries, and with a significant Internet presence have had to handle a cyber incident: More than 80 percent of these firms had to handle a cyber attack in the last 12 months. These firms are attractive targets to attackers and we control for them using variables that measure the attractiveness of targets to attackers. Given the prominence of the high-tech sector in the Israeli economy (and the US economy as well), this raises concerns. Small firms with relatively low income without a significant Internet presence suffered far fewer cyber incidents. 20% of these firms had to handle a cyber attack in the last 12 months. Descriptive statistics on all of the variables used in the analysis appear in Table 1.

	Full implementation (1)	Partial or no implementation (2)
Incident	0.42	0.49
	(0.49)	(0.5)
Number of controls	15.8	14.7
	(4)	(4.1)
Average revenue of low revenue firms	8.29	9.42
	(5.12)	(5.41)
Average revenue of medium revenue firms	59.6	66.7
	(31.175)	(35.131)
Average revenue of high revenue firms	1280.4	897
	(3113.98)	(1255.32)
Average employees in small-size firms	24.96	26.28
	(11.68)	(10.75)
Average employees in medium-size firms	119.86	122.86
	(53.41)	(58.49)
Average employees in large-size firms	1394.42	1041.34
	(2399.52)	(1278.9)
High-tech	0.24	0.25
	(0.43)	(0.43)
International firm	0.41	0.47
	(0.49)	(0.5)
Manufacturing industry	0.35	0.42
	(0.48)	(0.49)
Construction or Food activities industry	0.15	0.14
	(0.36)	(0.35)
Trade industry	0.14	0.15
	(0.35)	(0.36)
Information & communication industry	0.16	0.13
	(0.37)	(0.34)
Real estate, Administrative activities industry	0.11	0.08
	(0.32)	(0.27)
Professional, scientific and technical industry	0.09	0.08
	(0.28)	(0.27)
Cloud	0.73	0.76
	(0.45)	(0.43)
Website	0.42	0.46
	(0.49)	(0.5)
E-commerce	0.2	0.2
	(0.4)	(0.4)
Observations	384	609

Notes: The average revenue is given in million shekels.
std. dev. in parentheses.

Table 1: Summary statistics by implementation of directive

Given the rich data and the availability of an exogenous instrumental variable, the INCD/CBS survey enables us to measure the impact of firm characteristics and security precautions on the likelihood of compromise/incidents. The observation is at the level of the firm.

4 Instrumental Variable Selection and Evaluation

We first motivate the need for an instrumental variable and identify a suitable candidate in Section 4.1. We next discuss both why the instrument works empirically in Section 4.2 and why it is a theoretically valid instrument in Section 4.3.

4.1 Endogeneity and our Instrumental Variable

As discussed in the introduction, a huge problem with empirical work on this topic is that the timing of security investments is unknown. Indeed, the survey data we examine exhibits a positive correlation between (i) whether the firm employed the six basic security precautions and incidents as well as between (ii) the number of security precautions the firm employed and incidents. This is because many such investments are made *ex-post*, i.e., after a firm has suffered a cyber incident. Thus, there is an endogeneity issue.

We address the endogeneity issue with an instrumental variable regression. Fortunately, we are able to construct an appropriate instrument for security precautions from the survey data. The instrumental variable is "implementation", which is a dummy variable that takes on the value one for those enterprises that report both awareness and full implementation of cyber directives and instructions with reference to a government guidance document. The

instrumental variable takes on the value zero for those enterprises that reported awareness of the government guidance document, but did not implement them at all or reported only partial implementation.

The directives document was issued in June 2017. We interpret a yes response to implementation to mean that the firm is intentionally vigilant to risks of cyber attacks. The guidance document is more than 140 pages long, and we do not interpret a yes response to the implementation question necessarily to mean that the firm literally followed all of the detailed directives and instructions in the document. We hypothesize that a firm that implemented the directives was vigilant and more likely to have adopted more security measures prior to the 12-month period for which incidents were reported in the survey (2019-2020) than firms that did not implement the directives.

Of the 993 firms that answered yes to being aware of the directives (and hence are included in our data set), 46 percent had to handle a cyber incident in the 12 months preceding the 2020-2021 survey. The CBS received responses to the 2020 survey from firms beginning in July 2020 and ending in April 2021. Hence, the 12-month period for reporting cyber incidents was no earlier than 2019.

4.2 Why the Instrument Works Empirically

Empirically, the implementation of the directives variable works as an instrument for security precautions for the following reasons.

The implementation of the directives variable is positively correlated with the number of security precautions and whether the firm employed all six basic security measures. Addi-

tionally, the implementation variable is negatively correlated with incidents.

Further, for firms that employed all six basic security measures and implemented the directives, there was a much lower likelihood of a cyber incident (47%) than there was for firms that employed all six basic security measures but did not implement the directives (65%). This addresses the timing issue. Further, for firms that did not employ all six basic security measures, there is virtually little difference in the probability of a cyber incident between (i) firms that implemented the directives (38%) and (ii) firms that did not implement the directives (41%).

Furthermore, on average, firms in the data set employed 15 security precautions. For firms that employed more than 15 precautions and implemented the directives, there was a lower likelihood of a cyber incident (47%) than firms that employed more than 15 precautions, but did not implement the directives (59%). Again, this addresses the timing issue. Further, for firms that employed less than 15 security measures, there is a smaller difference in the probability of a cyber incident between (i) firms that implemented the directives (28%) and (ii) firms that did not implement the directives (33%).

Finally, the first stage OLS regressions of the endogenous security variables on the instrument (and all other variables used in the IV regressions) shows that for both specifications, the estimated coefficient on the instrument is indeed positive and statistically significant in explaining our endogenous security variables.

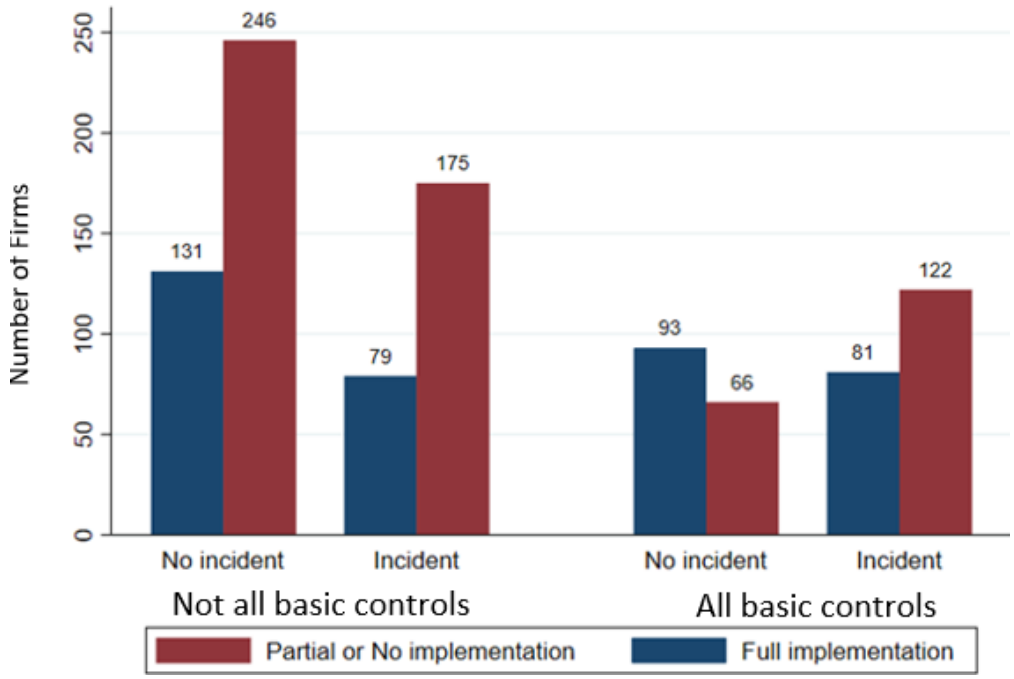

Figure 2: Number of firms with and without security incidents by use of all basic controls and implementation of directives. Note that the heading "Not all basic controls" means that the firm did not employ all six of the basic controls, while the heading "All basic controls" means that the firm employed all six basic controls

4.3 Theoretical Validity of the Instrument

The instrument is the "implementation of directives" question on the survey. We believe that it is valid theoretically for the following reasons:

First, and most importantly, it is very likely, other things being equal, that firms that implemented the directives employed more security precautions prior to the 12 month period for reporting incidents than firms that did not implement the directives. This is especially true for six basic cybersecurity precautions that can be immediately employed.

Additionally, an instrument also has to be randomly assigned or as good as randomly assigned in the sense of being unrelated with any omitted variables we would like to control for. Our instrument is as good as randomly assigned since we interpret the instrumental

variable as an indicator of vigilance, which is an unobserved characteristic of the firm and firms with similar observable characteristics may or may not be vigilant. This is consistent with random assignment. A more vigilant firm is one that is more concerned about cybersecurity threats and is more likely to be an early adopter of security precautions to discourage or thwart attacks. Empirically, the summary statistics, which we reported separately for firms that implement the directives and firms that did not implement the directives, show that there is little difference between the observable variables for the two groups.

Further, an instrument has to affect the outcome of interest only through one channel, namely through the endogenous variable, which is the adoption of security measures. it is quite reasonable that, in this setting, vigilance affects the outcome of interest only through one channel, namely through the endogenous variable (security precautions).

Finally, an ideal instrument should be uncorrelated with attacks. Since vigilance and the implementation responses were unobservable to potential attackers for most of the security precautions, these variables presumably are uncorrelated with attacks conditional on the adoption of the basic security precautions and observed characteristics of the firm. This is especially true for the six basic security precautions.

5 Regression Analysis

We now describe the regression analysis that incorporates the instrumental variable. We describe the estimation equation and included variables in Section 5.1, followed by the estimation and results in Section 5.2. In Section 5.3, we discuss additional diagnostics concerning the instrumental variable including a first stage regression.

5.1 Estimation Equation

The equation we estimate is

$$Y_i = X_i\alpha + \beta S_i + \epsilon_i. \tag{1}$$

The dependent (or response) variable, denoted Y_i, is a dummy variable that takes on the value one if the firm had to handle a cyber incident and zero otherwise. The primary independent variable of interest is security precautions, denoted S_i. For robustness, we use two different measurements (and consequently, two regressions):

- First Case: How many of the twenty security precautions did the firm employ?
- Second Case: Did the firm employ the six basic security precautions? (yes/no):
- In both cases, as noted, the security variable is denoted S_i
- Our instrumental variable for S_i is the dummy variable implementation, denoted I_i, where this variable takes on the value one if the firm implemented the 2017 directives and zero otherwise.

We include additional independent variables that take into account the attractiveness of the target. These exogenous variables include firm size (revenues and employees), as well as dummy variables for high-tech firms, firms that use cloud services, firms in high-risk industries, and firms that use information about visitors' behavior on its website. These independent variables, denoted X_i, are defined as follows:

- Firm size and firm revenue variables:
  - **High Revenue** is a dummy variable that takes on the value one if the firm's income is above 400 million shekels and zero otherwise.

  - **Large Firm** is a dummy variable that takes on the value one if the firm has 250 or more employees and zero otherwise.

- **High-tech** is a dummy variable that takes on the value one if the firm is in the high-tech sector and zero otherwise.
- **High-risk industries** is a dummy variable that takes on the value one if the firm is in one of the following high-risk (of attack) industries: mining and quarrying, manufacturing, electricity, water supply, and information and communication, [3]
- **Cloud** is a dummy variable that takes on the value one if the enterprise purchases any cloud computing services used over the internet (including membership, pay by use or any other payment agreement) and zero otherwise.
- **Website** is a dummy variable that takes on the value one if the enterprise (including the company group) use information about visitors' behavior on its website (e.g. clicks, items viewed) for improving user experience and zero otherwise.
- **E-commerce** is a dummy variable that takes on the value one if during 2019 the enterprise received orders for goods or services placed via a website or app and zero otherwise.

Note that for each of these variables, we are following the categorizations utilized by the survey.

The main parameter of interest (β) in equation (1) is the one that measures the impact of the security precautions (S_i). The parameters associated with the target variables (X_i) are denoted α. They measure the attractiveness of the target. The term ϵ_i is the error term.

[3] Our results are virtually unchanged if we exclude this variable from the regressions.

5.2 Estimation and Results

We now estimate the relationship between incidents (the dependent variable) and firm characteristics, Internet use, and the precautionary security posture of the firm using equation (1). We use two alternative specifications measuring a firm's security posture: the number of security controls (from the list specified in the survey) adopted by the firm, or whether all of six basic controls were adopted by the firm.

Following the empirical conventional wisdom (Angrist (2001) and Angrist and Pischke (2009)), we estimate linear regressions with robust standard errors. In the initial OLS regressions, we do not instrument for our security precautions variable. We then estimate instrumental variable (IV) regressions where we instrument for the security precautions.[4]

[4]In the case of our first specification, where the security measure is the number of controls, using an IV probit yields virtually identical results. In the case of our second specification, where the security measure is a dummy variable itself, an IV probit is not appropriate – and linear models are always employed. See Angrist (2001) for more details.

(S_i) =	**Number of controls**		**Basic controls**	
	OLS	IV	OLS	IV
Dep. variable (Y_i)=Incident	(1)	(2)	(3)	(4)
Security Measures (S_i)	0.023	-0.047	0.0951	-0.379
	(0.003)	(0.028)	(0.034)	(0.221)
High revenue	0.113	0.232	0.143	0.188
	(0.04)	(0.064)	(0.04)	(0.048)
Large firm	0.037	0.094	0.052	0.069
	(0.036)	(0.046)	(0.036)	(0.04)
High-tech	0.0337	0.105	0.0365	0.137
	(0.04)	(0.053)	(0.041)	(0.065)
High risk industries	0.099	0.140	0.110	0.120
	(0.033)	(0.041)	(0.033)	(0.037)
Cloud	0.075	0.170	0.096	0.144
	(0.036)	(0.055)	(0.036)	(0.044)
Website	0.060	0.104	0.069	0.098
	(0.033)	(0.042)	(0.034)	(0.04)
E-commerce	0.049	0.045	0.05	0.037
	(0.04)	(0.047)	(0.041)	(0.045)
Constant	-0.08	0.786	0.185	0.263
R2	0.110		0.089	
Observations	993	993	993	993

This table includes results from estimating equation (1).
The dependent variable (Yi) is whether an incident occurred. The independent variables include the "security precautions" variable (Si) and the other independent variables (denoted X_i).
Columns (1) & (3) are the initial ordinary least squares (OLS) regressions for each specification, where Si is the number of controls in column (1), and whether all of the six basic controls were adopted is the security variable in column (3).
In columns (2) & (4) we estimate instrumental variable (IV) regressions for each specification.
Standard errors clustered by firm in parentheses.

Table 2: Regression results.

The results from the regressions are shown in Table 2. The regressions show the estimated parameters and (in parentheses) the standard errors. In the initial ordinary least squares (OLS) regression for each specification, we do not instrument for the cybersecurity precaution variable. Table 2 shows that in both specifications, the estimated parameter on cybersecurity precautions is positive and statistically significant. That does not mean that taking precautions leads to incidents, but rather that firms more prone to attack are more likely to adopt precautions, or those who suffer an incident might install controls following the incident. These possibilities illustrates the "endogeneity" problem, and are why we need to run an instrumental variable regression.

In the second regression for each specification, we instrument for our security posture variable using the implementation variable. Table 2 shows that in such a case, for both specifications, in the instrumental variables (IV) regression, the estimated parameter on cybersecurity precautions is negative and statistically significant with a p-value of 0.09. Further, for both IV specifications, a one-tailed test rejects the the null hypothesis that the parameter on security precautions is non-negative with p-values below 0.05.

Importantly, in our regressions, we included many variables that take into account the attractiveness of the target. These variables include firm size, revenues, employees, as well as dummy variables for high-tech firms, firms that use cloud services, firms in high-risk industries, and firms that use information about visitors' behavior on its website. The estimated parameters on these variables are positive and statistically significant. In particular, Table 2 shows that large firms with many employees and significant revenues, firms in high-risk industries, high-tech firms, firms that use cloud services and firms that use information about visitors' behavior on its website are more likely to suffer an incident than other firms. These

effects are statistically significant. The findings makes sense, given that firms with such characteristics are attractive targets to attackers.

Quantitative Significance of our Results We use the parameter estimates on security measures from the Instrumental Variable (IV) regressions (Equation 1) to get a quantitative measure of the difference in the estimated probability of suffering an incident using both specifications.

We find that for large firms with significant revenues using e-commerce and cloud services, using all six basic security precautions reduces the probability of experiencing a cyber incident from 80% to 42%. This suggests that these six basic security precautions are quite effective.

For firms that employed all six basic security precautions, the average number of security measures employed is 18, while for firms that did not employ all six basic security precautions, the average number of security measures employed is 13. We find that for large firms with significant revenues using e-commerce and cloud services (the riskiest firms), using 18 out of 20 security precautions rather than 13 security precautions reduces the probability of experiencing a cyber incident from 81% to 58%. Running an IV probit regression, using 18 out of 20 security precautions rather than 13 security precautions reduces the probability of experiencing a cyber incident from 76% to 58%.

Dep. Variable (S_i) =	**Number of controls** (1)	**Basic controls** (2)
Implementation (I_i)	1.242	0.154
	(0.244)	(0.0303)
High revenue	1.674	0.0916
	(0.264)	(0.039)
Large firm	0.861	0.0415
	(0.285)	(0.036)
High-tech	0.985	0.207
	(0.28)	(0.04)
High risk industries	0.651	0.0285
	(0.268)	(0.032)
Cloud	1.409	0.105
	(0.322)	(0.034)
Website	0.673	0.0672
	(0.252)	(0.033)
E-commerce	-0.0670	-0.0284
	(0.306)	(0.041)
F-statistic (instrument)	25.89	25.70
Constant	11.91	0.0943
R2	0.105	0.164
Observations	993	993

Notes: This table includes results from estimating equation (2).
The dependent variable (S_i) is number of controls in column (1), and whether all of six basic controls were adopted in column (2).
The independent variables include the instrument (I_i) and all other independent variables (denoted X_i).
Standard errors clustered by firm in parentheses.

Table 3: First stage regressions results

5.3 Additional Diagnostics and Discussion

To provide additional empirical justification for this instrument, we conduct first stage OLS regressions of the endogenous security variables on the instrument (I_i) and all the independent variables (X_i) used in the IV regressions. The first-stage equation is

$$S_i = X_i\gamma + \delta I_i + u_i. \tag{2}$$

The key parameter of interest is δ. The error term of the equation is denoted u_i. All of the right-hand side variables are exogenous. Hence we use OLS regressions for the first stage. See Table 3 for the results.

The results of this estimating equation (2) show that for both specifications, after controlling for the target variables, the estimated coefficient on the instrument is positive and statistically significant. In fact, the estimated coefficients associated with the instrument are five standard deviations beyond zero for both specifications. That means our instrument is highly statistically significant in explaining our security variables. In other words, it shows that more vigilant firms are much more likely to adopt security precautions. This is an additional confirmation that the instrument works well.

Like most instruments, we doubt that our instrumental variable is a perfect one, because some firms in the sample could have become aware of the guidance document and implemented its directives and instructions in response to suffering a cyber incident. Intuitively, this imperfection is likely to bias our estimate of the effect of security precautions on incidents toward zero. Using the Nevo and Rosen (2012) analysis of imperfect instruments, this is indeed the case for our data. That is, using the Nevo and Rosen (2012) analysis,

our instrumental variable estimate will be biased toward zero if the Covariance of $\tilde{x}$ and our instrument is greater than zero, where $\tilde{x}$ are the residuals from an OLS regression of our endogenous variable on all other right hand variables (all of which are exogenous). This is indeed the case for both specifications. Hence, the true parameter likely is more negative than our estimate in both regressions. Consequently, we expect our quantitative estimate of the effect of basic security precautions on preventing incidents to be conservative. We thus conclude that, all else equal, a stronger security posture reduces the frequency of cyber incidents.

6 Concluding Remarks

We have presented econometric evidence that greater firm investment in cybersecurity does in fact yield results. Drawing from a comprehensive survey of Israeli firms, we find that organizations adopting a suite of six basic controls are less likely to subsequently experience incidents. We successfully established this relationship by carefully selecting an instrumental variable to compensate for unknowns in the timing of security investments that often plague empirical work in cybersecurity. These findings have significant public policy implications, especially as governments demand greater returns to cybersecurity investments that can be objectively evaluated. Of course, more work is needed to corroborate these findings and strengthen the empirical evidence linking investment cybersecurity defenses to experiencing cyber incidents. As good as these survey questions are, it would be desirable to collect additional evidence such as direct observations of security control employment and experiencing attacks. Such "triangulation" would strengthen the connection between cyber investment

and secure outcomes. Additionally, more work is needed to connect the experience of cyber incidents with the harm they cause. Establishing such connections is hard for many reasons, primarily because most firms cannot readily quantify the harms/costs. Finally, Israel is a country with an advanced cyber economy that can serve as a useful case study to inform future efforts in larger economies like the United States.

Acknowledgements

We are grateful to the editor, Eugene Spafford, and to two referees for comments and suggestions, which greatly strengthened the paper. We also thank Itai Benartzi, Iddo Bar Noy, Yael Lederman, Daniel Roash, Analia Schlosser, and the Israeli National Cyber Directorate (INCD) as well as the Central Bureau of Statistics, and participants at the 21st Workshop on the Economics of Information Security (WEIS) for their helpful suggestions and cooperation in conducting this research. We gratefully acknowledge support from the US National Science Foundation (NSF) Award No. 2147505 and the US Israel Binational Science Foundation (BSF) Award No. 2016622 and Award No. 2021711.

References

Anderson, R. (2001). Why information security is hard - an economic perspective. In *Proceedings of the 17th Annual Computer Security Applications Conference (ACSAC'01)*, New Orleans, LA.

Anderson, R. and Moore, T. (2006). The economics of information security. *Science*, 314(5799):610–613.

Angrist, J. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business and Economic Statistics*, 19(1):2–16.

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Angst, C. M., Block, E. S., D'Arcy, J., and Kelley, K. (2017). 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.

Fernandez De Arroyabe, I., Arranz, C. F. A., Arroyabe, M. F., and Fernandez de Arroyabe, J. C. (2023). Cybersecurity capabilities and cyber-attacks as drivers of investment in cybersecurity systems: A UK survey for 2018 and 2019. *Computers & Security*, 124:102954.

Howard, J. and Longstaff, T. (1998). A common language for computer security incidents. Technical report, Sandia National Laboratories.

King, A. and Gallagher, M. (2020). United States Cyberspace Solarium Commission Final Report. `https://cybersolarium.org/wp-content/uploads/2022/05/CSC-Final-Report.pdf`.

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

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

Liu, C.-W., Huang, P., and Lucas, H. (2020). Centralized it decision making and cybersecurity breaches: Evidence from u.s. higher education institutions. *Journal of Management Information Systems*, 37:758–787.

Liu, Y., Sarabi, A., Zhang, J., Naghizadeh, P., Karir, M., Bailey, M., and Liu, M. (2015). Cloudy with a chance of breach: Forecasting cyber security incidents. In *24th USENIX Security Symposium (USENIX Security 15)*, pages 1009–1024, Washington, D.C. USENIX Association.

Malliouris, D. and Simpson, A. (2019). The stock market impact of information security investments: The case of security standards. *WEIS 2019*. `https://ora.ox.ac.uk/objects/uuid:5de5f4cb-5fcb-46bb-9cd3-d13817d27e05`.

Moore, T., Dynes, S., and Chang, F. (2016). Identifying how firms manage cybersecurity investment. In *15th Workshop on the Economics of Information Security (WEIS)*.

Nagle, F., Ransbotham, S., and Westerman, G. (2017). The effects of security management on security events. In *Workshop on the Economics of Information Security*.

Nevo, A. and Rosen, A. (2012). Identification with imperfect instruments. *: The Review of Economics and Statistics*, 94(3):659–671.

Sarabi, A., Naghizadeh, P., Liu, Y., and Liu, M. (2016). Risky business: Fine-grained data breach prediction using business profiles. *Journal of Cybersecurity*, 2(1):15–28.

Shao, X., Siponen, M., and Liu, F. (2020). Shall we follow? Impact of reputation concern on information security managers' investment decisions. *Computers & Security*, 97:101961.

Toftegaard, O. (2022). An effect analysis of iso/iec 27001 certification on technical security of norwegian grid operators. pages 2620–2629.

Weishäupl, E., Yasasin, E., and Schryen, G. (2018). Information security investments: An exploratory multiple case study on decision-making, evaluation and learning. *Computers & Security*, 77:807–823.

Woods, D. W. and Böhme, R. (2021). Sok: Quantifying cyber risk. In *2021 2021 IEEE Symposium on Security and Privacy (S&P)*, pages 909–926, Los Alamitos, CA, USA. IEEE Computer Society.