RIETI Discussion Paper Series 13-E-086

An Empirical Analysis of the Effectiveness of Information Security Measures

IIDAKA Yuki
Information-technology Promotion Agency

HANAMURA Kenichi
Information-technology Promotion Agency

KOMATSU Ayako
RIETI

SAITO Yukiko Umeno
RIETI

TSUKADA Naotoshi
RIETI

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 13-E-086
October 2013

An Empirical Analysis of the Effectiveness of Information Security Measures[1]

IIDAKA Yuki, HANAMURA Kenichi
Information-technology Promotion Agency, Japan

KOMATSU Ayako
Information-technology Promotion Agency, Japan and Research Institute of Economy, Trade and Industry

SAITO Yukiko Umeno, TSUKADA Naotoshi
Research Institute of Economy, Trade and Industry

Abstract

We examine whether the adoption of information security measures can reduce the probability of computer virus infection by using firm-level survey data and probit regression analysis. We find that implementing two security measures—Web content filtering (WCF) and restriction of bringing in/out storage media or PCs (R_in/out)—can result in a statistically significant reduction of the probability. Calculating the average partial effect, we also indicate that the adoption of each of these measures decreases the estimated probability of infection by about 10% on average. In addition to these analyses, we show that the effectiveness of some security measures differs by firm size or by sector.

Keywords: Information security measures; Computer virus infection; Probit
JEL classification: Y8, C25

[1] We are grateful to seminar participants at RIETI. The usual disclaimers apply.
This is a product of joint research between Information-technology Promotion Agency, Japan(IPA) and Research Institute of Economy, Trade and Industry (RIETI)

1. Introduction

The importance of Information technology (IT) is beyond doubt. It is a basis of socio-economic activities in a wide range of fields such as commerce, the public sector, education, and medical fields, to name a few. It has come to pervade even areas such as fishery or agriculture which seem to have previously had nothing to do with IT. Also, IT is used by a large number of people and organizations. According to a 2012 White Paper Information and Communication in Japan, nearly 80% of the population or 98.8% of firms with more than 100 employees that use the Internet rely on IT.

Because of IT's importance, it is all the more disturbing that IT has been under threat from various dangers. According to the Internet Security Threat Report 2013 by Symantec, in 2012 on a global scale, there were 116 targeted attacks on average per day, one in 291 emails contained a virus, and there were about 247,350 Web-based attacks on computers per day (which were blocked). Results of the "Survey of the state of IT use" which the Ministry of Economy, Trade, and Industry published in July 2012, stated that 28.2% of firms in Japan experienced information security incidents in 2011, most of which affected their IT.

Many organizations are aware of the danger and take some precautionary measures to protect their IT from information security incidents. The Information-technology Promotion Agency (IPA hereafter) released in December 2011 the results of a "2011 Survey of Information Security Incidents and Damages". The results state that all but six respondents took some information security measures.

However, it is quite likely that they used such measures with little knowledge about how effective they are because not much is known about

their effectiveness in general. To our knowledge, no third party evaluation of security products and solutions has been carried out with feedback from users. As far as antivirus software is concerned, there are a few independent organizations which conduct tests to assess the performance of the software. However, such organizations do not provide information about the actual effectiveness of the software. As AV-comparatives, one of the organizations, admits that a 100% protection rate of software *in test* does not guarantee that its users will be perfectly protected *in their offices*.

Only a few of those who do some research on information security and economics, which is now called security economics[2], provide clues about their effectiveness. Takemura et al. (2009) investigated effectiveness of information security measures in Internet service providers in Japan. They classified the measures into two types, technical and non-technical. Then, they analyzed whether implementing more technical measures or each of the non-technical measure could decrease the likelihood of security incidents such as illegal access, viruses and worms, and system trouble. Liu et al. (2006) considered two situations: one where firms put three security measures (a preventive measure, an information security policy, and an information security education) into place, and the other where less than three measures were put into place. Then, they examined whether firms were less likely to get infected with computer viruses in the former than in the latter. In Kwon and Johnson (2012), each of the security measures belonged to one of the following three categories: security resources, security capabilities, and security audit capabilities. They analyzed whether installing more measures in each category could reduce risks of data breach

[2] For a brief introduction of security economics, See Anderson (2002).

in healthcare organizations.

This paper is in line with these previous studies in that it, too, aims to empirically analyze the effectiveness of information security measures. However, it differs from the previous studies in two aspects. First, the previous studies bundled some or all security measures and, thus, blurred the impact of the implementation of each measure on the risks of security incidents. In this paper, we treat each security measure separately and attempt to identify the measure which could reduce the likelihood of the incidents. Second, we investigate whether there are security measures which could be effective for reducing the risks in one group, but not in the other. The two groups which we consider are (1) large firms and small/medium firms, (2) manufacturing firms and non-manufacturing firms, and (3) firms which did not adopt two particular measures, and the firms which adopted, at least, either of the measures. As far as we know, no such comparative analysis has been conducted to date.

This paper is structured as follows. Section 2 describes our data. Section 3 shows our empirical models. Section 4 presents the results. Section 5 concludes.

2. Data

We use firm-level data of “2011 Survey on Information Security Incidents and Damages". The survey was conducted in 2012 by Mitsubishi Research Institute, Inc., on behalf of IPA. Its aim was to obtain basic data on (1) which information security measures were in place for respondents between April 2011 and March 2012, (2) whether, during the period, they experienced information security incidents such as computer virus infection,

cyber-attacks, unauthorized access/abuse by an insider, and (3) the details of damages victim organizations suffered and their recovery jobs, e.g., duration of server stoppage, manpower and time needed to restore data or to resume servers and so on. The survey also collected information about characteristics of respondents such as the number of employees, their business sector, their type (listed/unlisted) and so on.

The total number of respondents was 1767.[3] Table 1 gives a breakdown of respondents, excluding 40 of them which did not give an answer about their sector. "Small&Medium" in the table indicates a respondent with less than 300 employees, whereas "Large" represents a respondent with greater than or equal to 300.

Among the various security incidents mentioned earlier, this paper deals with computer virus infection, the most common security incident in our survey. This requires us, first, to filter out 114 firms which did not answer the question on virus infection. That leaves us with (1) 523 firms which did not encounter a virus between April 2011 and March 2012, (2) 831 firms which encountered a virus, but did not get infected during the period, (3) 299 firms which got infected during the period. We use the data of only the last two groups because they alone can tell us whether their security measures protected them from a virus or failed in that attempt.

[3] Though most of respondents are a firm, some non-profit organizations such as schools or foundations are also among them. In what follows, we call all respondents a firm for simplicity.

Table 1: Breakdown of respondents

			Small&Medium	Large	Total
All			848	879	1727
	Manufacturing		245	261	506
		Beverages, tobacco and feed	28	26	54
		Textile mill products	7	3	10
		Pulp, paper and paper products	10	5	15
		Chemical and allied products	16	22	38
		Petroleum, coal and plastic products	14	10	24
		Ceramic, stone and clay products	10	9	19
		Iron and steel	13	13	26
		Fabricated metal products and non-ferrous metals and products	25	25	50
		Electrical machinery, equipment and supplies	36	38	74
		Information and communication electronics equipment	4	5	9
		Transportation equipment	13	25	38
		Miscellaneous machinery	19	29	48
		Miscellaneous manufacturing	50	51	101
	Non-manufacturing		603	618	1221
		Agriculture, fishery, forestry, and mining	14	9	23
		Construction	93	64	157
		Electricity, gas, heat supply and water	7	5	12
		Video picture, sound information, and broadcasting	5	5	10
		Character information production and distribution	1	3	4
		Information services	72	66	138
		Transport and postal activities	62	69	131
		Wholesale	107	47	154
		Retail	40	79	119
		Finance and insurance	12	59	71
		Medical and other health services	25	23	48
		Education and learning support	26	27	53
		Miscellaneous services	139	162	301

The following 9 information security measures will be considered below:

(1) Antivirus software for network servers (M_1)

(2) Antivirus software for client PCs (M_2)

(3) Virus checks by Internet service providers (M_3)

(4) Web content filtering (M_4)

(5) Quarantine system (M_5)

(6) Restriction of bringing in/out storage media or PCs (M_6)

(7) Security patches (M_7)

(8) Information security education (M_8)

(9) Security audit (M_9).

M_{ik} (where k = 1,…,9) takes a value of 1 if firm i adopted measure k and 0 otherwise.[4] It is worth mentioning that, below, we call Web content filtering WCF and Restriction of bringing in/out storage media or PCs R_in/out.

In addition to information security measures, we consider in the analysis the following observed characteristics of firms which may affect the likelihood of virus infection.

(1) No. of Emp: natural logarithm of the number of employees of a firm. The more employees a firm has, the more employees with low awareness about information security it probably has and, thus, the more likely it is to get infected with a virus.

(2) Degree of IT: natural logarithm of the number of PCs per employee of a firm. This is meant to capture the extent to which IT is used in firms [5,6] The

[4] To be more specific, if firm i is an *uninfected* firm, M_{ik} = 1 only if it implemented measure k by the end of March 2012. If the firm is an *infected* firm M_{ik} = 1 only if it did so before its infection.

[5] This is based on Kurokawa and Minetake (2007). As an index for the degree of IT use in a firm, they calculated the total number of hardware (such as mainframe, workstation, PCs and mobile devices) per employee in a firm.

[6] One might think that there is no need to use both No. of Emp and Degree of IT as they are likely to be highly correlated. However, correlation coefficient between them is -0.0044 and, thus, not high at all.

use of more computers makes them more vulnerable to cyber threats.
(3) List: whether or not a firm is listed. Judging from anecdotal evidence, high profile information security incidents of the past few years such as Gumblar or targeted attacks seem to have been concentrated in listed firms. Thus, we cannot exclude the possibility that infections of low-level computer viruses, too, have occurred to listed firms more often than unlisted firms.
(4) Overseas: whether or not a firm does business overseas. It is not unusual for e-mails with computer viruses to be sent to a firm in Japan from its subsidiaries or affiliated companies in a foreign country where virus infection is more common than in Japan. Hence, firms with overseas operation encounter a virus more frequently than those without it and, thus, are more likely to get infected with a virus.
(5) Finance: whether or not a firm operates in the finance/insurance sector. Firms in these sectors are considered highly vigilant against cyber threats as they are concerned with reputational damages which information security incidents could cause. If this is the case, firms there might be less likely to get infected with a virus than those in other sectors.
(6) Education: whether or not a firm operates in the education sector. Firms there (especially universities) are said to be prone to security incidents largely due to people with low awareness about information security.

The characteristics (3)-(6) will be represented by a dummy variable. For instance, List takes a value of 1 if a firm is listed and 0 otherwise. Overseas, Finance, and Education will also be represented in a similar manner. The characteristics, (1)-(6), will be used as control variables in the analysis below.

2.1 Descriptive statistics[7]

Table 2 shows the percentage of firms which installed each measure. In the table, the second and third columns present the results by firm size and the fourth and fifth columns by sector.

In the first column, antivirus software for client PCs has the highest percentage while quarantine system has the smallest percentage. Thus, the former is the most widely used measure and the latter the least widely used.

If we compare small/medium firms with large firms, the percentage is higher for large firms than for small firms in all measures but virus checks by an Internet service provider. A comparison between manufacturing firms and non-manufacturing firms tells that the percentage is higher for manufacturing firms than for non-manufacturing firms in all cases but in antivirus software for client PCs, R_in/out and security audits.

Table 3 presents the rate of virus infection. It shows that 26.01% of firms got infected with a virus. The row "Firm size" shows that the rate is about 8% higher for large firms than for small/medium firms. The row "Sector" suggests that there is no large difference in the rate between manufacturing firms and non-manufacturing firms.

Table 4 shows the infection rate of users and non-users of each measure. In each column, take a look at the measures where the rate for non-users is larger than the one for users. We find that either WCF or R_in/out has the largest difference between the two rates in all columns but "Small&Medium" where antivirus software for client PCs has the largest difference.

Table 5 shows correlation coefficients among security measure binary variables. All coefficients are less than 0.5 in absolute value. This means that

[7] In all the tables below, the firms with missing values in information security measures or business sectors are excluded.

a relationship between any two variables is not strong on the whole. In other words, adoption of one measure is not strongly related to adoption/non-adoption of another measure.

Finally, Table 6 summarizes the statistics for control variables.

Table 2: Percentage of firms which use each security measure.

	All (N=742)	Firm size (N = 742)		Sector (N = 723)	
		Small/Medium (N=267)	Large (N=475)	Non Manufacturing (N=481)	Manufacturing (N=242)
Antivirus software for network server	90.97%	85.39%	94.11%	89.40%	93.80%
Antivirus software for client PC	97.71%	96.25%	98.53%	97.71%	97.52%
Virus check by internet service provider	50.40%	53.56%	48.63%	49.27%	51.24%
WCF	59.30%	40.82%	69.68%	57.80%	61.16%
Quarantine system	14.82%	10.11%	17.47%	14.55%	15.70%
R_in/out	75.88%	63.67%	82.74%	76.72%	73.55%
Security patches	75.34%	61.42%	83.16%	73.60%	78.93%
Information security education	61.19%	43.82%	70.95%	60.71%	61.57%
Security audit	42.32%	32.96%	47.58%	44.49%	36.78%

Table 3: Infection rate

All (N = 742)		26.01%
Firm size (N = 742)	Small/Medium (N = 267)	20.97%
	Large (N = 475)	28.84%
Sector (N = 723)	Non-manufacturing (N = 481)	26.20%
	Manufacturing (N = 242)	26.45%

Table 4: Infection rates for users/non-users of information security measures

		All (N = 742)	Firm size (N = 742)		Sector (N = 723)	
			Small/Medium (N = 267)	Large (N = 475)	Non-manufacturing (N = 481)	Manufacturing (N = 242)
Antivirus software for network server	non-user	23.88% (16)	20.51% (8)	28.57% (8)	25.49% (13)	20% (3)
	user	26.22% (177)	21.05% (48)	28.86% (129)	26.28% (113)	26.87% (61)
Antivirus software for client PC	non-user	29.41% (5)	40% (4)	14.29% (1)	36.36% (4)	16.67% (1)
	user	25.93% (188)	20.23% (52)	29.06% (136)	25.96% (122)	26.69% (63)
Virus check by internet service provider	non-user	27.17% (100)	25% (31)	28.28% (69)	26.64% (65)	25.49% (35)
	user	24.87% (93)	17.48% (25)	29.44% (68)	25.74% (61)	26.28% (29)
WCF	non-user	30.13% (91)	24.68% (39)	36.11% (52)	32.02% (65)	26.60% (25)
	user	23.18% (102)	15.60% (17)	25.68% (85)	21.94% (61)	26.35% (39)
Quarantine system	non-user	26.11% (165)	21.25% (51)	29.08% (114)	27.01% (111)	25.49% (52)
	user	25.45% (28)	18.52% (5)	27.71% (23)	21.43% (15)	31.58% (12)
R_in/out	non-user	31.84% (57)	24.74% (24)	40.24% (33)	35.71% (40)	26.56% (17)
	user	24.16% (136)	18.82% (32)	26.46% (104)	23.31% (86)	26.40% (47)
Security patches	non-user	21.31% (39)	22.33% (23)	20% (16)	22.05% (28)	21.57% (11)
	user	27.55% (154)	20.12% (33)	30.63% (121)	27.68% (98)	27.75% (53)
Information security education	non-user	26.74% (77)	22.67% (34)	31.16% (43)	31.22% (59)	19.35% (18)
	user	25.55% (116)	18.80% (22)	27.89% (94)	22.95% (67)	30.87% (46)
Security audit	non-user	25.93% (111)	24.02% (43)	27.31% (68)	26.97% (72)	25.49% (39)
	user	26.11% (82)	14.77% (13)	30.53% (69)	25.23% (54)	28.09% (25)

*The number of firms is in parentheses

Table 5: Correlation coefficients of information security measure binary variables

	M1	M2	M3	M4	M5	M6	M7	M8	M9
Antivirus software for network server (M1)									
Antivirus software for client PC (M2)	0.1727								
Virus check by internet service provider (M3)	0.086	0.0454							
WCF (M4)	0.1746	0.0189	-0.0152						
Quarantine system (M5)	0.1193	-0.0118	-0.0149	0.1922					
R_in/out (M6)	0.1	0.0821	-0.0201	0.2404	0.1292				
Security patches (M7)	0.131	0.1444	-0.0522	0.2146	0.1584	0.3117			
Information security education (M8)	0.1699	0.0818	-0.0295	0.2949	0.168	0.412	0.458		
Security audit (M9)	0.1427	0.0578	0.0096	0.2078	0.226	0.2989	0.2446	0.3919	

Table 6: Summary statistics for control variables

	Mean	Stand. Dev
No of Emp	6.225	1.4055
Degree of IT	-0.2928	0.896
List	0.2616	0.4398
Overseas	0.3169	0.4656
Finance	0.036	0.1861
Education	0.03	0.17

3 The model

To examine which information security measures could reduce risks of computer virus infection, a standard probit model is defined as follows:

$$P_i^* = \alpha_1 + \sum_{k=1}^{9} \beta_{1k} M_{ik} + \gamma_1 Z_i + \varepsilon_{1i} \qquad P_i = 1\ (0)\ if\ P_i^* > (\leq) 0 \qquad (1)$$

P_i^* is a latent variable corresponding to the probability of virus infection. P_i is a binary variable and denotes the observed outcome about virus infection. $P_i = 1$ if firm i got infected with a virus between April 2011 and March 2012, and $P_i = 0$ otherwise. M_{ik} is a security measure dummy variable, as defined earlier.[8] Z_i is a vector of control variables. α_1, β_{1k}, and γ_1 are parameters to be estimated. ε_{1i} is an error term, which is assumed to be normally distributed

[8] The security dummy variables indicate which measure was implemented, but not which actually blocked a virus or which was breached. Thus, to be exact, (1) examines whether *implementation of* a measure, not a measure itself, could decrease the infection probability.

with mean 0 and variance 1.

We also analyze whether adoption of a measure could result in a statistically significant fall in the risk of infection in one group but not in the other. The two groups which we examine are (1) large firms and small/medium firms, (2) manufacturing firms and non-manufacturing firms, and (3) the firms which adopted neither WCF nor R_in/out, and the ones which adopted, at least, either of them. For example, take a look at Columns "Small/Medium" and "Large" in Table 4. As to antivirus software for client PCs, the difference of an infection rate between non-users and users is about 20% for Small/Medium firms whereas it is about -15% for large firms. Also, in R_in/out, the difference is about 14% for large firms while it is just about 6% for Small/Medium firms. These differences in infection rate imply the possibility that implementation of antivirus software for client PCs may result in a statistically significant decrease in the risks of infection for Small/Medium firms but not for large firms, while the opposite may hold for R_in/out.[9] As another example, consider one group which put WCF and/or R_in/out in place and the other which put neither of them in place. As will be explained below, the former group has already blocked (one of) the main routes of infection whereas the latter has not. Therefore, adoption of a measure other than WCF and R_in/out may not have a significant impact on infection probability for the former group, but it may for the latter group.

To examine whether some measures can cause a significant reduction in infection probability for small/medium firms but not for large firms and *vice versa*, we divided the data into one for small/medium firms and the other for large firms. Then, using (1) as an estimation equation, we conducted a probit

[9] As far as R_in/out is concerned, an argument similar to the above one holds for manufacturing firms and non-manufacturing firms, too.

analysis for small/medium firms and large firms. A similar analysis was carried out for manufacturing firms and non-manufacturing firms, too, by splitting the date into one for the former and the other for the latter. To investigate whether some measures can lead to a significant decrease in infection probability only if neither WCF nor R_in/out is in place, we divided the data into one for firms which had adopted neither WCF nor R_in/out and the other for firms which had adopted, at least, either of them. Then, slightly modifying (1) as below

$$P_i^* = \alpha_2 + \sum_{k \neq 4,6} \beta_{2k} M_{ik} + \gamma_2 Z_i + \varepsilon_{2i} \qquad (2)$$

we carried out a probit analysis for these two groups. Here, α_2, β_{2k}, and γ_2 are parameters to be estimated. The same assumptions as the ones about ε_{1i} apply to ε_{2i}.

4 Estimation Results

Table 7 presents the estimation results for the case where the full data set was used. In Column (A), which shows the results without control variables, three variables turn out to be statistically significant, WCF, R_in/out, and security patches. The sign of parameters on the first two are negative, which means that implementation of these two could make virus infection less likely. By contrast, the sign of security patches is positive, which suggests that applying security patches could raise the risk of virus infection.[10]

[10] One may suspect multicollinearity between security patches and other security dummy variables. However, multicollinearity is unlikely to be present. As shown in Table 5, the largest coefficient in absolute value between security patches and another measure is 0.458 (a coefficient between security patches and information security education).

The results on WCF and R_in/out are likely to do with the fact that they block the possible main routes of virus infection. In our survey, about 62% of firms think that the virus came from Web pages. Also, about 46% think that the source of virus is external storage devices such as USB memory. The result indicates that placing a measure in such routes could lower the likelihood of virus infection.

As mentioned earlier, in addition to information security measures, there are other factors which could affect infection probability. If such factors are omitted from an estimation equation, if they are correlated with security measure dummy variables, and if parameters on such factors are statistically significant, there would be an omitted variable bias problem. Responding to such a concern, we conducted a probit regression with such factors, i.e., control variables.

The results are presented in Columns (B) to (E). In all cases, both WCF and R_in/out are statistically significant and negative. This and the result in Column (A) suggest the robustness of the results on WCF and R_in/out.

Note that there is a difference in a coefficient on security patches between Column (A) and Columns (B) to (E). It is significant and *positive* in the former, but insignificant in the latter. A comparison between Column (A) and Column (B) implies that two control variables, No. of Emp and Degree of IT, make this difference. Our preliminary research results (not presented here) suggested that there was a positive correlation between the number of employees in a firm and use of security patches and, also, that firms with more employees were more likely to get infected with a virus. Here, the number of employees presumably acted as a confounding factor and produced a spurious relationship in which adoption of patches could raise the

risk of infection. An analogous argument holds for Degree of IT, too. It seems that inclusion of these control variables in (1) solved this spurious relationship problem and led to the results in (B) to (E).[11]

Using the results in Table 7, we can estimate the impact on the infection probability of implementing WCF or R_in/out. As is often the case with a probit or logit model, we can do so by calculating the average partial effect (APE). Since we use a dummy variable, APE for WCF is defined as

$$n^{-1}\sum_{i=1}^{n}\{\mathrm{G}(\widehat{\alpha_1}+\widehat{\beta_{14}}+\sum_{k\neq 4}\widehat{\beta_{1k}}M_{ik}+\widehat{\gamma_1}Z_i)-\mathrm{G}(\widehat{\alpha_1}+\sum_{k\neq 4}\widehat{\beta_{1k}}M_{ik}+\widehat{\gamma_1}Z_i)\} \quad (3)$$

where $\widehat{\alpha_1}$, $\widehat{\beta_{14}}$, $\widehat{\beta_{1k}}$, and $\widehat{\gamma}_1$ are estimated parameters and G(・) denotes the standard normal cumulative distribution function. APE for R_in/out can be obtained by replacing 4 with 6 in (3). Estimated APE are shown in Table 8.

When calculating APEs in Column (j) of Table 8, we used parameters in Column (j) of Table 7 where j = (A,…,E). Table 8 illustrates that putting either of these measures in place could lower estimated infection probability by around 10% on average.

4.1 A comparison between large firms and small/medium firms.

Table 9 shows the estimation results for small/medium firms and large firms. Let us start with small/medium firms. The parameter on antivirus software for client PC is statistically significant and negative but only in Column (B),

[11] Another possible cause which makes the sign of security patches in Column (A) positive is a positive bias resulting from endogenous regressors. We have treated all security dummy variables as exogenous so far. However, it is not unreasonable to think that whether a firm adopts measures depends on, say, whether it or other firms experienced security incidents in the past. If this is the case, exogeneity of the dummy variables does not hold and we are likely to have a positive bias on the parameters in this case.

not in Columns (A) and (C). It is hardly a robust result and, thus, we cannot say anything further about the software. As to all other security measures, their parameters turn out to be statistically insignificant. Next, let us turn to large firms. The parameters on WFC, R_in/out, and security patches are statistically significant in all columns. As to the sign of the parameters, it is negative for WCF and R_in/out but positive for security patches. This means that installing the first two measures could reduce the risk of infection.

The above results suggest that the effectiveness of WFC and R_in/out differs between small/medium firms and large firms. They could lower the likelihood of infection for large firms but there is no statistical evidence that the same holds true for small/medium firms.

4.2 A comparison between manufacturing firms and non-manufacturing firms.

Table 10 presents the estimation results for manufacturing firms and non-manufacturing firms. We begin with non-manufacturing firms. The parameters on WCF, R_in/out, and information security education are statistically significant and negative. Thus, these three measures could reduce the likelihood of infection. When it comes to manufacturing firms, the only security measure whose parameter is statistically significant and negative is WCF, and it is so only if control variables are included in the estimation equation.

The above results suggest that effectiveness of R_in/out and information security education unambiguously differs between manufacturing firms and non-manufacturing firms. The two measures could reduce the risk of infection for non-manufacturing firms whereas they would not cause a

statistically significant fall in the risk for manufacturing firms.

4.3 A comparison between firms which adopted neither WCF nor R_in/out and firms which adopted, at least, either of them.

In Table 11 are the estimation results for the case where neither WCF nor R_in/out was in place and the other where, at least, either of them was in place. As to the former case, antivirus software for client PCs is statistically significant and negative in all columns. Hence, installing the software could reduce the likelihood of virus infection. By contrast, all other security measures do not have a statistically significant and negative parameter in any columns. Thus, they cannot be expected to lower the infection risks. As to the latter case, there turn out to be no security measures whose parameters are statistically significant and negative in either of the three columns.

The above results indicate that adoption of antivirus software for client PCs could reduce the risk of infection if neither WCF and/or R_in/out is in place. By contrast, such adoption does not have a statistically significant and negative impact on the risk otherwise.

The result on antivirus software for client PCs has an important implication: there could exist substitutability between the software and WCF and/or R_in/out. If neither WCF nor R_in/out is in place, the software plays a role of blocking a virus. However, if either WCF or R_in/out is adopted, the software is replaced by WCF and/or R_in/out in a sense.[12]

[12] There is an alternative method which enables us to examine whether adoption of a measure affects infection probability differently according to adoption/non-adoption of WCF and/or R_in/out and whether there could exist substitutability between WCF and/or R_in/out and the measure. See Appendix .

5. Conclusion

The purpose of this paper was to examine which information security measures could be effective for reducing risks of computer virus infection. To carry out this investigation, we used a probit regression. Among nine security measures we considered, it turned out that WCF and R_in/out were effective for lowering the risks

Using the regression results, we calculated the impact of implementing each of these two measures on infection probability by computing the average partial effect. We found that the estimated probability would fall by about 10% on average with implementation.

In addition to the above analyses, we investigated whether there would be security measures which could be effective in one group but not in the other. The two groups we compared were (1) large firms and small/medium firms, (2) manufacturing firms and non-manufacturing firms, and (3) the firms which had adopted neither WCF nor R_in/out and the ones which had adopted, at least, either of them. Comparing results for large firms and small/medium firms, we found that installation of WCF and R_in/out could cause a statistically significant reduction in the risk of infection only for large firms. A comparison of results for manufacturing firms with the results for non-manufacturing firms showed that adoption of R_in/out and information security education could lead to a statistically significant fall in the risk only for non-manufacturing firms. Finally, we compared results for firms which adopted neither WCF nor R_in/out with the results for firms which adopted, at least, either of them. We found that adoption of antivirus software for client PCs could bring about a statistically significantly fall in infection probability only if neither WCF nor R_in/out was in place. This

result implies that there could exist substitutability between the software and WCF and/or R_in/out.

References

Anderson, R. (2002) *Why Information Security is Hard – An Economic Perspective*,

http://www.cl.cam.ac.uk/~rja14/Papers/econ.pdf

D'Agostine L M. and Santangelo G D. (2012). *Does the global fragmentation of R&D activities pay back? The home region perspective* paper presented at DRUID (Danish Research Unit for Industrial Dynamics) Academy 2012

Cassiman, B. and Veugelers, R (2006) *In search of complementarity in the innovation strategy: internal R&D and external knowledge acquisition* Management Science vol. 52, no.1 pp. 23-32

Kodde, D.A. and Palm, F.C. (1986) *Wald criteria for jointly testing equality and inequality restrictions*, Econometrica, vol. 54, no.5, pp. 1243-1248.

Kurokawa, F. and Minetaki, K., (2007) *How can IT Raise Productivity Linked with Workplace Re-organization and Human Capital in Japan?* Mimeo.

Kwon, J. and Johnson, E.M. (2012) Security resources, capabilities and cultural values: links to security performance and compliance. Paper presented at WEIS (Workshop on the Economics of Information Security) 2012

Liu, W., Tanaka, H., and Matsuura, K. (2006) *Am empirical analysis of security investment in countermeasures based on an enterprise survey in Japan*. Paper presented at WEIS (Workshop on the Economics of

Information Security) 2006

Mohnen, P. and Röller L.-H (2005) *Complementarities in innovation policy* European Economic review vol.49 1431-1450

Takemura, T., Osajima, M., and Kawano, M. (2009) *Economic analysis on information security incidents and the countermeasures: the case of Japanese internet service providers* K. Jayanthakumaran (Ed.), Advanced Technologies, INTEH, Chapter 5, pp.73-89, December.

Table 7 : Estimation results

Dependent : Infection dummy	(A)	(B)	(C)	(D)	(E)
Antivirus software for network server	0.1397	0.0818	0.0620	0.0785	0.0629
	(0.1864)	(0.1965)	(0.1973)	(0.1976)	(0.1987)
Antivirus software for client PC	-0.2085	-0.4592	-0.3049	-0.4500	-0.3032
	(0.3391)	(0.3532)	(0.3758)	(0.3530)	(0.3755)
Virus check by internet service provider	-0.0677	0.0119	0.0103	0.0165	0.0102
	(0.1010)	(0.1077)	(0.1090)	(0.1089)	(0.1098)
WCF	-0.2280**	-0.3831***	-0.3950***	-0.3552***	-0.3682***
	(0.1095)	(0.1208)	(0.1223)	(0.1221)	(0.1234)
Quarantine system	-0.0140	-0.0341	-0.0139	-0.1121	-0.0933
	(0.1478)	(0.1556)	(0.1570)	(0.1595)	(0.1613)
R_in/out	-0.2868**	-0.3704***	-0.3704***	-0.3270**	-0.3210**
	(0.1307)	(0.1384)	(0.1408)	(0.1404)	(0.1431)
Security patches	0.3596**	0.2289	0.2317	0.2115	0.2215
	(0.1407)	(0.1543)	(0.1557)	(0.1547)	(0.1559)
Information security education	-0.0455	-0.1633	-0.1448	-0.1653	-0.1513
	(0.1291)	(0.1395)	(0.1427)	(0.1401)	(0.1433)
Security audit	0.0719	0.0229	0.0096	0.0448	0.0289
	(0.1137)	(0.1204)	(0.1219)	(0.1218)	(0.1231)
No of Emp		0.2343***	0.2161***	0.2351***	0.2155***
		(0.0443)	(0.0505)	(0.0451)	(0.0519)
Degree of IT		0.1662**	0.1378**	0.1529**	0.1226
		(0.0684)	(0.0717)	(0.0709)	(0.0750)
List			0.0072		0.0144
			(0.1429)		(0.1456)
Overseas			0.1112		0.1212
			(0.1320)		(0.1339)
Finance				-0.4603	-0.4067
				(0.3216)	(0.3256)
Education				0.3880	0.4904
				(0.2760)	(0.3180)
_cons	-0.4641	-1.2924***	-1.3503***	-1.3319***	-1.3398***
	(0.3433)	(0.4082)	(0.4403)	(0.4096)	(0.4419)
obs	742	693	680	681	669
Pseudo R2	0.019	0.0585	0.0582	0.063	0.0622
Log likelihood	-417.212	-374.69628	-367.1876	-368.5265	-361.5158

Standard errors in parentheses $^{*}P<0.1$, $^{**}P<0.05$, $^{***}P<0.01$

Table 8 : Average Partial Effect

	(A)	(B)	(C)	(D)	(E)
WCF	-0.0725	-0.1169	-0.1203	-0.1084	-0.1122
R_in/out	-0.0912	-0.1130	-0.1128	-0.0998	-0.0979

Table 9: Estimation results[13]

Dependent: Infection dummy	(A)		(B)		(C)	
Firm Size	small/medium	large	small/medium	large	small/medium	large
Antivirus software for network server	0.2381 (0.2694)	0.0327 (0.2754)	0.1059 (0.2800)	0.0941 (0.2909)	0.0391 (0.2874)	0.0941 (0.2934)
Antivirus software for client PC	-0.5231 (0.4456)	0.4463 (0.6109)	-1.0768** (0.4949)	0.2541 (0.6129)	-0.8493 (0.5333)	0.2445 (0.6131)
Virus check by internet service provider	-0.2833 (0.1821)	0.0294 (0.1246)	-0.1641 (0.1948)	0.0801 (0.1324)	-0.1388 (0.1979)	0.0569 (0.1362)
WCF	-0.3029 (0.1921)	-0.2787** (0.1408)	-0.3060 (0.2065)	-0.4181*** (0.1519)	-0.2664 (0.2138)	-0.4017** (0.1568)
Quarantine system	0.0620 (0.3178)	-0.0567 (0.1700)	0.1103 (0.3453)	-0.0746 (0.1772)	0.0933 (0.3515)	-0.1410 (0.1848)
R_in/out	-0.0655 (0.2084)	-0.4673*** (0.1745)	-0.2562 (0.2279)	-0.4759** (0.1838)	-0.2528 (0.2332)	-0.4440** (0.1930)
Security patches	0.0566 (0.2166)	0.5401*** (0.1974)	-0.2165 (0.2428)	0.4946** (0.2141)	-0.2574 (0.2466)	0.5019** (0.2166)
Information security education	0.0262 (0.2326)	-0.1160 (0.1593)	-0.1040 (0.2526)	-0.1603 (0.1722)	-0.1236 (0.2584)	-0.1253 (0.1782)
Security audit	-0.3219 (0.2262)	0.1952 (0.1357)	-0.3043 (0.2457)	0.1218 (0.1417)	-0.2622 (0.2542)	0.1094 (0.1452)
No of Emp			0.3883** (0.1592)	0.2260*** (0.0648)	0.4078** (0.1638)	0.1817** (0.0738)
Degree of IT			0.5286*** (0.1723)	0.0893 (0.0785)	0.5412*** (0.1918)	0.0354 (0.0875)
List					0.1022 (0.4083)	0.0141 (0.1592)
Overseas					-0.0613 (0.2841)	0.1880 (0.1573)
Finance					(omitted)	-0.3311 (0.3373)
Education					0.0716 (0.5253)	0.4312 (0.4346)
_cons	-0.1686 (0.4341)	-0.9331 (0.6308)	-1.0485 (0.8483)	-2.1755*** (0.7339)	-1.3003 (0.8893)	-1.9818*** (0.7561)
obs	267	475	250	443	241	426
Pseudo R2	0.0375	0.0364	0.0924	0.0629	0.0854	0.0696
Log likelihood	-131.986	-274.9591	-116.0168	-250.4285	-113.7519	-239.877

Standard errors in parentheses $^{*}P<0.1$, $^{**}P<0.05$, $^{***}P<0.01$

[13] In Column (C) for small/medium firms, a result for Financial/insurance sector was omitted as the variable was dropped in a probit regression due to a complete separation.

Table 10: Estimation results[14,15]

Dependent: Infection dummy	(A)		(B)		(C)	
Sector	Non-manu	Manu	Non-manu	Manu	Non-manu	Manu
Antivirus software for network server	0.1647 (0.2180)	0.1487 (0.4000)	0.0721 (0.2260)	0.3037 (0.4635)	0.0237 (0.2285)	0.3100 (0.4685)
Antivirus software for client PC	-0.3885 (0.4178)	0.2894 (0.6352)	-0.5971 (0.4402)	-0.1716 (0.6366)	-0.3751 (0.4754)	-0.2424 (0.6462)
Virus check by internet service provider	-0.0220 (0.1266)	-0.1623 (0.1785)	0.0272 (0.1340)	0.0334 (0.1934)	0.0277 (0.1368)	0.0203 (0.1947)
WCF	-0.2600* (0.1351)	-0.1947 (0.2012)	-0.3557** (0.1470)	-0.4957** (0.2273)	-0.3348** (0.1505)	-0.5310** (0.2294)
Quarantine system	-0.1545 (0.1902)	0.1547 (0.2558)	-0.1645 (0.1981)	0.1310 (0.2727)	-0.2233 (0.2064)	0.1392 (0.2754)
R_in/out	-0.3286** (0.1633)	-0.2403 (0.2340)	-0.4549*** (0.1746)	-0.2988 (0.2437)	-0.3895** (0.1827)	-0.3308 (0.2472)
Security patches	0.4887*** (0.1698)	-0.0205 (0.2698)	0.3464* (0.1859)	-0.1170 (0.2991)	0.3243* (0.1885)	-0.0718 (0.3000)
Information security education	-0.2960* (0.1562)	0.5039** (0.2505)	-0.3770** (0.1695)	0.3778 (0.2672)	-0.3494** (0.1746)	0.3785 (0.2706)
Security audit	0.1019 (0.1421)	-0.0315 (0.2026)	0.1135 (0.1498)	-0.1824 (0.2152)	0.1437 (0.1549)	-0.1819 (0.2165)
No of Emp			0.2336*** (0.0563)	0.2119*** (0.0763)	0.2431*** (0.0658)	0.1505* (0.0879)
Degree of IT			0.1495* (0.0786)	0.3251** (0.1580)	0.1089 (0.0875)	0.2616 (0.1645)
List					-0.1375 (0.1892)	0.1668 (0.2454)
Overseas					0.1349 (0.1781)	0.2252 (0.2289)
Finance					-0.4528 (0.3345)	(omitted)
Education					0.4652 (0.3281)	(omitted)
_cons	-0.2202 (0.4142)	-0.9967 (0.6722)	-1.0470** (0.5072)	-1.6827** (0.7626)	-1.3619** (0.5571)	-1.4021* (0.7931)
obs	481	242	452	229	441	228
Pseudo R2	0.0384	0.0267	0.0739	0.0724	0.0803	0.0817
Log likelihood	-266.000	-136.0668	-243.1864	-121.2075	-234.6476	-119.7163

Standard errors in parentheses $^{*}P<0.1$, $^{**}P<0.05$, $^{***}P<0.01$

[14] In the second row of Table 10, Non-Manu and Manu indicate non-manufacturing firms and manufacturing firms respectively.

[15] In Column (C) for non-manufacturing firms, results for Financial/insurance sector and Education sector were omitted as the variables were dropped in a probit regression due to collinearity.

Table 11-1: Estimation results[16,17]

Dependent : Infection dummy	(A)		(B)		(C)	
adoption/non-adoption	non-adoption	adoption	non-adoption	adoption	non-adoption	adoption
Antivirus software for network server	0.6575*	0.0074	0.6270	-0.0588	0.6982	-0.0880
	(0.3910)	(0.2217)	(0.4578)	(0.2332)	(0.4863)	(0.2344)
Antivirus software for client PC	-1.8294**	0.4839	-2.4665***	0.3412	-2.3581***	0.3461
	(0.7345)	(0.5275)	(0.7942)	(0.5484)	(0.8647)	(0.5492)
Virus check by internet service provider	0.0012	-0.0855	0.0595	-0.0210	0.0541	-0.0347
	(0.2594)	(0.1109)	(0.2952)	(0.1177)	(0.3070)	(0.1202)
Quarantine system	(omitted)	-0.1035	(omitted)	-0.1136	(omitted)	-0.1659
		(0.1526)		(0.1596)		(0.1654)
Security patches	-0.0210	0.5325***	-0.4564	0.4496**	-0.5178	0.4564**
	(0.2624)	(0.1719)	(0.3164)	(0.1879)	(0.3273)	(0.1901)
Information security education	0.2598	-0.1326	-0.2400	-0.2113	0.0351	-0.1945
	(0.3256)	(0.1398)	(0.3840)	(0.1515)	(0.4280)	(0.1549)
Security audit	0.6333*	0.0151	0.6774	-0.0397	0.6476	-0.0218
	(0.3687)	(0.1198)	(0.4391)	(0.1261)	(0.4738)	(0.1295)
No of Emp			0.3947***	0.1996***	0.4374***	0.1769***
			(0.1335)	(0.0459)	(0.1658)	(0.0545)
Degree of IT			0.6809***	0.0663	0.6415***	0.0218
			(0.2032)	(0.0735)	(0.2321)	(0.0807)
List					-0.0753	-0.0329
					(0.6642)	(0.1514)
Overseas					-0.3320	0.2003
					(0.4579)	(0.1423)
Finance					(omitted)	-0.4595
						(0.3281)
Education					0.4282	0.3263
					(0.7970)	(0.3829)
_cons	0.7568	-1.4902***	-0.0648	-2.4567***	-0.4108	-2.3542***
	(0.6624)	(0.5385)	(0.8965)	(0.6084)	(0.9908)	(0.6240)
obs	112	631	104	590	99	571
Pseudo R2	0.0757	0.0193	0.2098	0.0488	0.2227	0.0536
Log likelihood	-68.483	-340.521	-54.757603	-308.59704	-51.596552	-297.75155

Standard errors in parentheses $^{*}P<0.1$, $^{**}P<0.05$, $^{***}P<0.01$

16 Some results were omitted in the table below due to a complete separation (Quarantine system) or collinearity (Finance).

17 In the second row of Table 11, non-adoption indicates the firms which adopted neither WCF nor R_in/out whereas adoption refers to the firms which adopted, at least, either WCF or R_in/out.

Appendix

In Section 4-3, using the estimation equation (2), we investigated whether it could happen that adoption of a measure could result in a statistically significant reduction in infection probability if neither WCF nor R_in/out was in place and not otherwise. We showed that it could occur to antivirus software for client PCs and suggested from the result that there could exist substitutability between the software and WCF and/or R_in/out.

This appendix presents an alternative approach which examines the same issue as above and whether there could exist substitutability between WCF and/or R_in/out and other measures. For this purpose, we rely on an empirical method adopted by Mohnen and Röller (2005). They analyzed whether one obstacle to a firms' decision to innovate would be complementary to or substitute for another obstacle. In their paper, two obstacles were substitute if the condition called submodularity held.

How a test for submodularity is conducted goes as follows. Suppose that there are two variables x_1 and x_2 and a function W. Then, these two variables are considered substitutes in the function W if and only if W satisfies the following submodularity condition

$$W(x_1,x_2+1) - W(x_1+1,x_2+1) \geq W(x_1,x_2) - W(x_1+1,x_2) \quad \forall x_1, x_2 \qquad (A.1).^{18}$$

Applying this method to our case is straightforward. Now let F(m) denote the function corresponding to infection probability where $m = (M_1,\ldots, M_9)$. For example, given $M_{k'}$ where $k' \neq 4,5$, M_4 and M_5 are substitutes if the following inequality conditions hold and at least one of them does so with

[18] By contrast, two variables are considered complementary if and only if the opposite inequality of (A.1) holds.

strict inequality.

$$F(m_{123}, 0, 1, m_{67}, m_{89}) - F(m_{123}, 1, 1, m_{67}, m_{89}) \leq F(m_{123}, 0, 0, m_{67}, m_{89}) - F(m_{123}, 1, 0, m_{67}, m_{89}) \quad \text{(A.2)}$$

where $m_{123} = (M_1, M_2, M_3) = (0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (1,0,1), (0,1,1), (1,1,1)$, $m_{67} = (M_6, M_7) = (0,0), (1,0), (0,1), (1,1)$, $m_{89} = (M_8, M_9) = (0,0), (1,0), (0,1), (1,1)$. Note that the direction of inequalities in (A.2) is opposite to that in (A.1). This is simply due to a difference in a *dependent* variable. In this paper, 0 in a dependent variable represents success (no infection) and 1 in the variable represents failure (infection), while it is the other way around in Mohnen and Röller (2005). If we swap 0 and 1 in our data, the directions of inequalities in (A.1) and (A.2) correspond.

At this point, there arises one problem. (A.2) suggests that we have 128 (=8*4*4) inequality conditions to examine. Unfortunately, checking all of these conditions is cumbersome.

To simplify our analysis, when analyzing substitutability between M_k and $M_{k'}$ ($k \neq k'$), we exclude from the infection probability function all security measure dummy variables but these two.[19] This reduces the number of an inequality condition to 1 in examining such substitutability.

Following Cassiman and Veugelers (2006), we define a probit model as follows:

$$P_i^* = \theta_{00}^k(1 - M_{ik})(1 - M_{i46}) + \theta_{10}^k M_{ik}(1 - M_{i46}) + \theta_{01}^k(1 - M_{ik})M_{i46} + \theta_{11}^k M_{ik} M_{i46} + \gamma_5 Z_i + \varepsilon_{5i} \quad \text{(A.3)}$$

[19] There is a caveat. Due to this assumption, the submodularity condition used here differs from that in the literature. Thus, the results presented below should be taken with some caution.

where k = 1,2,3,5,7,9. θ_{00}^{k}, θ_{10}^{k}, θ_{01}^{k}, θ_{11}^{k} and γ_5 are parameters to be estimated. As in Cassiman and Veugelers (2006), there is no constant term solely to make the interpretation of coefficients easy. The same assumptions as the ones for ε_{1i} apply to ε_{5i}.

Following Mohnen and Röller (2005) and bearing in mind the difference in the direction of inequality mentioned earlier, we examine whether the following inequality is satisfied.

$$\theta_{00}^{k} - \theta_{10}^{k} \geq \theta_{01}^{k} - \theta_{11}^{k} \qquad \text{(A.4)}$$

If (A.4), i.e., the submodurality condition holds, there is substitutability between M_k and WCF and/or R_in/out (WCF/R_in/out hereafter). To test this, first, we run a probit regression of (A.3) for measure k. (See Table A1 for the result). Next, we conduct the substitutability test by using the estimation results and (A.4) as the null hypothesis. The test procedure adopted by D'Agostine & Santangelo (2012) is applied here.[20]

The test results are shown in Table A2. The measures which could have substitutability with WCF/R_in/out turn out to be antivirus software for client PCs and security patches.[21]

Though these two measures satisfy (A.4), this does not mean that adoption of the software or patches could lead to a statistically significant reduction in infection probability if neither WCF nor R_in/out is in place and not

[20] Their test procedure goes as follows. First, we test the null hypothesis (A.4) of equality. If it is rejected, then we test the null of submodularity versus supermodularity (the opposite inequality of (A.4)). If the null cannot be rejected at the second stage, we conclude that there is substitutability between measure k and WCF/R_in/out.

[21] As for the other measures, they did not pass even the first equality test mentioned in the previous footnote. The exception is quarantine system. It passed the equality test, but we obtained an unexpected result: it would be c*omplementary* to WCF/R_in/out. As our interest is substitutability, we will not discuss this result further.

otherwise. For instance, suppose that $\theta_{00}^k = \theta_{10}^k$ and $\theta_{01}^k = \theta_{11}^k$. These two equalities imply that, regardless of adoption of WCF/R_in/out, adoption of measure k does not affect infection probability. However, given these equalities, (A.4) holds. Installation of the software or patches could result in a statistically significant fall in the risk of virus infection in the same way discussed in Section 4 only if the following two hold *simultaneously*

$$\theta_{00}^k - \theta_{10}^k \geq 0 \text{ and } \theta_{01}^k - \theta_{11}^k = 0 \text{ where k = 2,7} \qquad \text{(A.5)}$$

The first constraint implies that implementing measure k could reduce the risk of infection if neither WCF nor R_in/out is in place. The second one means that adoption of measure k has no impact on infection probability if WCF/R_in/out is in place. We can test (A.5) as the null hypothesis by following the test procedure suggested in Kodde & Palm (1986).[22,23]

Table A3 presents the test results. (A.5) is satisfied for antivirus software for client PCs, but not necessarily for security patches. This indicates that adoption of the software could cause a statistically significant fall in infection probability if neither WCF nor R_in/out is in place while it could not otherwise. As to security patches, it is unclear.

A relationship between (A.4) and (A.5) is worth pointing out. If (A.4) holds, so does (A.5). That is, suppose that the adoption of antivirus software for client PCs could result in a statistically significant reduction in the risk of infection if neither WCF nor R_in/out is in place but not otherwise. Then, it

[22] Before doing the test, we checked whether both constraints in (A.5) would hold with equality simultaneously. Such a possibility was rejected with the significance level of 5% when k = 2 and 10% when k = 7.

[23] In the test, first, we calculate the Wald test defined in (2.16) in their paper. Then, using their table for a critical value, we reject (do not reject) the null hypothesis if the test is larger (smaller) than the value.

verifies that there is pairwise substitutability between the software and WCF/R_in/out.

In Section 4-3, we showed the same result about the impact of the adoption of the software. However, as to substitutability between the software and WCF/R_in/out, we did not verify it but inferred it from the results.

Table A1: Estimation result[24]

Dependent : Infection dummy	$M_k=M_1$	$M_k=M_2$	$M_k=M_3$	$M_k=M_5$	$M_k=M_7$	$M_k=M_8$	$M_k=M_9$
$(1-M_k)(1-M_{46})$	-1.4575***	0.1274	-1.3312***	-1.4233***	-1.1670***	-1.4357***	-1.5036***
	(0.3450)	(0.6201)	(0.2641)	(0.2135)	(0.2310)	(0.2155)	(0.2161)
$M_k(1-M_{46})$	-1.3789***	-1.5354***	-1.3294***	-0.288	-1.5265***	-1.4791***	-1.0856***
	(0.2334)	(0.2150)	(0.2398)	(0.6556)	(0.2564)	(0.3141)	(0.3652)
$(1-M_k)M_{46}$	-2.1678***	-2.2633***	-2.03215***	-2.0453***	-2.2131***	-2.080***	-2.1049***
	(0.2902)	(0.4658)	(0.2514)	(0.2300)	(0.2472)	(0.2306)	(0.2308)
M_kM_{46}	-2.0876***	-2.1327***	-1.9894***	-2.1032***	-1.9524***	-2.2280***	-2.2203***
	(0.2443)	(0.2294)	(0.2409)	(0.2706)	(0.2420)	(0.2448)	(0.2487)
Scale	0.2178***	0.2262***	0.2093***	0.2150***	0.2043***	0.2319***	0.2299***
	(0.0361)	(0.0344)	(0.0362)	(0.0349)	(0.0356)	(0.0354)	(0.0352)
Degree of IT	0.1371**	0.1789***	0.1775***	0.1614***	0.1560***	0.1961**	0.1724***
	(0.0547)	(0.0514)	(0.0551)	(0.0521)	(0.0546)	(0.0544)	(0.0516)
Obs	846	932	825	904	908	932	921
Log likelihood	-460.903	-499.243	-451.121	-491.293	-492.34023	-499.954	-497.210

Standard errors in parentheses $^{*}P<0.1$, $^{**}P<0.05$, $^{***}P<0.01$

[24] Here, No. of Emp and Degree of IT are used as control variables because they are the only variables which had statistical significance in Tables 7,9,10, and 11.

Table A2: Substitutability test results

	Substitutability
$M_k=M_1$	No
$M_k=M_2$	Yes (0.992)
$M_k=M_3$	No
$M_k=M_5$	No
$M_k=M_7$	Yes (0.991)
$M_k=M_8$	No
$M_k=M_9$	No

P value for (A.4) in parentheses

Table A3: Wald tests for equality and inequality restrictions[25,26]

Antivirus software for client PC	Security Patches
0.094	2.959/5.884

At the 10% significance level, the critical value is 3.808

[25] As there are one equality constraint and one inequality constraint, the number of degree of freedom is 2. Moreover, what Kodde and Palm (1986) call lower bound critical value and upper bound critical value coincide.

[26] To calculate the Wald test, first, we ran quadratic optimization with inequality constraints. In optimization algorithm, we tried three initial values. In the case of antivirus software for client PCs, the optimization results and the consequent Wald test were almost the same regardless of the initial values. However, for security patches, the results for quadratic optimization and, thus, the Wald test differed according to the initial values.