SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2023
ORGANIZATION AND MANAGEMENT SERIES NO. 179

ASSESSING THE IMPACT OF CYBER RISK PERCEPTION ON CYBER INSURANCE PURCHASE DECISIONS

Michał THLON[1*], Grzegorz STRUPCZEWSKI[2]

[1] Uniwersytet Ekonomiczny w Krakowie, thlonm@uek.krakow.pl, ORCID: 0000-0001-9627-7773
[2] Uniwersytet Ekonomiczny w Krakowie; strupczg@uek.krakow.pl, ORCID: 0000-0002-7882-120X
* Correspondence author

Purpose: The goal of this paper is to investigate how cyber risk perception influences medium and large companies' decisions to purchase cyber insurance.
Design/methodology/approach: The study collected data from 386 managers in medium and large Polish enterprises through a questionnaire. It examined managerial perceptions of cyber risk, considering firm attributes like size, age, and type. Various statistical methods, including Pearson's chi-square test, multiple correspondence analysis, and the random forests classifier, were employed for comprehensive data analysis.
Findings: The study highlighted the pivotal role that perceptions of cyber risk play in shaping decisions concerning cyber insurance. Managers' perceptions regarding the gravity and probability of cyber threats had a significant impact on the choices made by their organizations. Furthermore, the study identified the presence of the availability heuristic as a noteworthy factor influencing decision-making in the realm of cyber insurance. Moreover, specific determinants emerged as influential in a company's decision to invest in cyber coverage. These determinants encompassed the size of the employee base, annual turnover, the severity of previous cyber losses, and the frequency of successful cyber-attacks experienced by the firm over the preceding five years.
Research limitations/implications: The study recognizes potential biases, including non-response and sampling frame bias, as well as the limitations of self-reported data. However, it offers valuable insights for policymakers in enhancing cyber-attack resilience through cyber insurance. Future research should explore factors influencing cyber insurance purchases, examine the complexity of cyber risk perception, and consider context-specific studies and multidisciplinary approaches.
Practical implications: The outcomes of this research have practical implications for both businesses and policymakers. It provides insights into enhancing cyber-attack resilience through cyber insurance, helping businesses make informed decisions regarding risk management. This research can impact industry policy by guiding the development of tailored insurance offerings.
Originality/value: The article fills a gap in the literature concerning the analysis of the relationship between cyber risk perception and the decisions of medium and large companies regarding the purchase of cyber insurance in the Polish market. It provides valuable insights for policymakers, insurance providers, and businesses looking to improve their cybersecurity practices and resilience to cyber threats.

Keywords: cyber risk, cyber risk management, random forests, cyber insurance, cyber risk perception.

Category of the paper: research paper.

1. Introduction

Information technology, interconnectedness, and moving businesses to cyberspace are cornerstones of the modern digital economy. The digital realm not only presents tremendous opportunities, but also is a substantial source of cyber risks. Attacks are growing increasingly sophisticated, and the severity of their financial consequences has been profound. Privacy protection is among the major issues globally; media report data breaches almost every day (Munich Re, 2020). Cyber incidents can threaten the financial stability of national economies (DTCC 2020; Bank of Canada, 2021; Bank of England, 2022). Thus, cyber risk is of concern at both the micro- and macro-prudential levels; however, measuring the consequences of cyber risk on the overall financial system remains at an early stage (Brando et al., 2022).

The scale of financial losses resulting from cyber risk is illustrated by the SolarWinds case, which was one of the most serious incidents of its kind in the US. The costs for the US government alone amounted to hundreds of millions of dollars, even though the attack was likely not aimed at destroying IT infrastructure but instead was likely espionage (Nolan, Fixler 2021). Another example is the NotPetya malware infection in 2017, which is considered to have been the most destructive in history. It originally targeted Ukraine but spread to dozens of countries and contributed to losses estimated at $10 billion (Alladi et al., 2020).

Insurance can be considered as a risk management tool for cyber risk (Alladi et al., 2020). The cyber insurance market has been growing rapidly for a decade and is predicted to continue its 20-30% annual growth rate in the near future (Greenwald, 2020). Risk aversion determines the demand for insurance among individuals and small firms where the owner makes most decisions. For larger businesses, risk aversion among company owners and managers is considered insufficient to explain motivations for purchasing property insurance (Main, 1983; MacMinn, 1987; Mayers, Smith, 1990). Thus, other motives are being investigated, such as preserving a company's liquidity in case unfortunate events occur (Main, 1983); reduction of bankruptcy costs and financial distress (Main, 1982; MacMinn, 1987); tax optimization (Main, 1982); compliance with regulations, in some industries (Mayers, Smith, 1990); and demonstration of good corporate risk management practices (Main, 1982; Grace, Rebello, 1993). These motives can be represented by various characteristics of a given company, including size of employment, annual turnover, industry type, and legal status of a company; taken together, these characteristics can be considered to comprise a 'company profile' (Krummaker, 2019). Ultimately, two predominant factors may influence a company's decision

to purchase insurance: risk aversion of its owners or managers (which is the result of risk perception) and the company profile. The interplay of these forces in the real-life environment and their impacts on decisions to buy insurance policies is worth empirical investigation. A research gap remains in this area. Moreover, because cyber insurance is a relatively new product on the market, the factors associated with firms' decisions to purchase it have not been sufficiently investigated. This marks another research gap worth addressing.

Poland is a leading economy in the Central and Eastern European (CEE) region. The Polish insurance market is growing rapidly. The CAGR of the non-life insurance market for the 2013-2022 period is 6.15% (EIOPA, 2023). Poland's share of the non-life-insurance gross premiums written in the EU is the greatest in comparison with other post-communist countries that accessed the EU (EIOPA, 2023).

This explanatory study translates feedback from practice into theory. Our goal is to investigate how cyber risk perception influence medium and large companies' decisions to purchase insurance. As control variables, a range of features describing company profile are considered. More specifically, this study addresses two primary research questions:

- RQ1. Are cyber risk perception and company profile significantly associated with companies' purchase of cyber insurance? Which elements of risk perception and company profile are most influential in a multidimensional approach?
- RQ2. Is managerial cyber risk perception influenced by the availability heuristic?

The following section presents the literature review for theory-based development of hypotheses. Next, we specify the methodology of the questionnaire survey and the research methods utilized. The results acquired are then presented and discussed. The last section concludes.

2. Literature review for hypotheses development

Cyber risk perception refers to people's beliefs, attitudes, judgments, and feelings toward cyber risk, and incorporates the wider social and cultural values, as well as outlook, people adopt toward cyber threats (Van Schaik et al., 2017; Wei et al., 2021). According to Slovic (2000), individual risk perception can be an important factor in making decisions about risk. The dominant view in the literature on cyber insurance is a positive assessment of insurance as a form of corporate response to cyber risks. Talesh (2018) points to the growing role of insurance companies in supporting businesses in adapting to the world of cyber threats by providing insurance coverage and unique risk management services that affect how organizations comply with privacy regulations.

Notably, cyber risk has different characteristics in comparison to other insurance lines (Böhme, Kataria, 2006). First, cyberattacks directly or indirectly affect all users of a certain type of technology. Secondly, both the business continuity and information security of a single enterprise depend heavily on the efforts of other market players with which that enterprise interacts. Anderson and Moore (2006) have concluded that these considerations impede both the development and the application of cyber insurance.

Nevertheless, in prior literature, insurance companies are seen as bridging the gap for companies that see themselves as unprepared for the risks of data breaches or IT system compromises (Herr, 2019). Cyber insurance is pointed to as creating a strong incentive to invest in cybersecurity (Bolot, Lelarge, 2009). Partial cyber insurance can motivate reluctant insurance customers to invest more efficiently in self-defense (Pal, Golubchik, 2010), and cyber insurance premiums can thus be estimated more fairly (Herath, Herath, 2011).

Cybersecurity researchers also highlight the insurance industry's ability to create the attitudes that motivate customers to implement adequate cyber risk protection tools (Talesh, 2017). In this context, it is important that insurance companies can collect data on breaches and then compile and share insights on the factors shaping a risky environment, acting as a central repository of particularly relevant IT security-related data (Levite et al., 2018). As a result, the field of cyber insurance should be viewed in a much broader context than is the case for the traditional insurance market. In contrast, the literature also points out that although the transfer of cyber risk to insurance companies can be an effective tool for managing risk and is increasingly offered by global insurance carriers like AXA, Generali, and Allianz, the market remains in an early stage of development (Marotta et al., 2017).

In our research, we follow the approach of de Smidt and Botzen (2017) in studying individual risk perception and its covariates, particularly in developing the questionnaire. Thus, risk perception is broken down into three components: risk awareness, perceived probability of successful cyber-attack, and its impact. The proxy for cyber risk awareness is the question of how possible (or not possible) is a successful cyber-attack on the respondent’s own company (variable PROB). Perceived probability is measured via a question that asks the respondent to estimate the frequency (which is the inversion of probability) of successful cyber-attacks on their own company (variable FREQ). Using frequency format instead of probability format is justified by the results of Schapira et al. (2001); they argue that using discrete frequencies in estimating risk magnitude provides greater salience and understanding, compared to probability format, in communication of probabilistic outcomes. Finally, the perceived financial impact of cyber risk is measured by a question that asks the respondent to choose from a range of estimated monetary cyber losses if a successful cyber-attack were to occur in their company (variable IMPACT). As Barberis (2013) argues, research on individual risk perception should not focus only on probability, but also on subjective estimation of potential losses in monetary terms. Both influence the protective behavior of an individual. Therefore, we hypothesize that:

H1. The purchase of cyber insurance is associated with individual cyber risk perception.

The over- or underestimation of risk can be explained by the so-called availability heuristic. According to this heuristic, people perceive hazardous events as high-risk if such events are easy to imagine, recall, or conceptualize the occurrence of (e.g., Tversky, Kahneman, 1973). In this respect, personal experience of risky situations becomes extremely important. Following de Smidt and Botzen (2017), salience of cyber threats determines the level of individual risk perception in the cybersecurity context. The proxy of this factor in the current study is a question regarding personal experience of a successful cyber-attack against the respondent's own organization in the last five years (variable COUNT). Therefore, we hypothesize that:

H2. The availability heuristic is related to cyber risk perception.

Finally, the current study includes a set of control variables characterizing a company profile, in order to investigate how specific features of a company are associated with its decision to purchase cyber insurance. Prior literature suggests that the following features of an enterprise are related to the demand for corporate insurance (Main, 1982; Mayers, Smith, 1990; Hoyt, Khang, 2000; Krummaker, 2019): firm size measured by employment and annual turnover as a proxy; firm age; type of business; the firm's legal form; origin of equity; share of intangibles in total assets; and the firm's share of equity in total liabilities (leverage). Therefore, we hypothesize that:

H3. The purchase of cyber insurance is associated with a company profile.

Hypotheses H1 and H3 address the research question RQ1, and hypothesis H2 addresses RQ2.

3. Materials and Methods

3.1. Survey design

Our survey was carried out in 2019 using the CATI method. The questionnaire comprises 17 questions covering cyber risk perception, cyber insurance, and key characteristics of the surveyed companies (the firm's size in terms of employment and annual turnover, the firm's type of business, the firm's age, the firm's legal form, the firm's equity structure, and the firm's origin). The research sample is a stratified random sample encompassing medium and large enterprises (MLEs) operating in Poland in multiple industries, excluding financial services and public administration. The stratified random sample method involves dividing the entire group into layers, and then randomly selecting independent samples from each layer; the size of each stratified sample is proportional to the size of its respective layer. The layers for the current study's sample were determined according to the industry type (which resulted in nine layers), size of business (medium or large), and headquarters location.

3.2. Research methods

First, we examine whether cyber risk perception and company profile are associated with cyber insurance purchases. The statistical method used is one-dimensional analysis that enables a description of basic relations between variables, using Pearson's chi-square test of independence.

Next, using multiple correspondence analysis (MCA), we attempt to answer the question which components of cyber risk perception and company profile influence decisions about insuring against cyber-attacks. MCA allows analysis of the pattern of relationships among several categorical dependent variables. Technically, MCA is obtained by using a standard correspondence analysis on an indicator matrix (i.e., a matrix whose entries are 0 or 1). MCA can thus define the structure of a particular data set and in this study's context, aid in identifying the significant contributing factors to firms' decisions regarding insurance purchase.

We also utilized the "random forests" method, which is essentially a generalization of the idea of decision trees and belongs to the so-called ensemble methods. Random forests work by performing classification using a group of decision trees. The final decision on classification is made by majority voting on the classes indicated by each decision tree. Each decision tree is constructed based on a bootstrap sample, which is formed by drawing with return N objects from a learning set of N items. In addition, at each node of a given tree, the division is made only on the basis of k randomly selected features. In addition to choosing the appropriate type of method, a necessary challenge is choosing an appropriate type of model. The data collected for the current study forms a complete set, so there are no problems with gaps, but the dataset is unbalanced in that there are many more uninsured than insured companies. The lack of balance issue warrants further analysis. In other words, the distribution of the variable representing the purchase of cyber insurance is uneven, and one class dominates in terms of quantity; in our case, far more companies do not have a cyber policy than have one. This is a skewed distribution and a predictive model may thus not be well fitted. There are two main approaches to solving this problem at the data level— methods that modify the available data to balance the dataset. The most popular methods are oversampling, which generates artificial occurrences of a less frequent class, and undersampling, which is the opposite approach and reduces the dominant observations to compensate for imbalances. In this study, both methods have been used, to find the optimal model fit. In general, random oversampling duplicates values that occur less frequently in the learning dataset and can result in over-fitting some models. Random undersampling, in contrast, removes values that occur more frequently and can result in the loss of in-formation that significantly affects the model. Optimal parameter values were assumed in the simulations: $N = 350$ for oversampling and $N = 100$ for undersampling.

Another issue is the selection of the optimal parameters for a given random forest model. In a random forest algorithm, two parameters are important: the number of decision trees used in the forest (parameter *ntree*) and the number of random variables used in each tree (parameter *mtry*). A common approach is to set *mtry* to a default value, in our case the square root of the total number of all predictors, and search for the optimal *ntree* value. To find the number of decision trees that satisfy a useful classifier, random forests with different tree counts were built. We built 10 classifiers of the random forest type for each value of the *ntree* parameter along with the OOB error rate. As a result, we obtained the number of trees, and hence the optimal number of predictors, in which the error rate stabilized and reached a minimum. For the selection of the final random forest model, a two-step approach was used. First, the most optimal set of random forest parameters was selected, and then - for this set of parameters - the standard model and the models with oversampling and undersampling were developed and compared with each other, selecting the model that yielded the best results. The data was split into two datasets: a training dataset and a test dataset, at a ratio of 75% to 25%, where the test dataset was used to verify the results obtained.

The next step in analysis is to evaluate the results for each model, which consists of comparing correctly and incorrectly matched results in a confusion matrix; identifying the variables that most affect the classification result; and finally evaluating the classifier's quality using ROC curves. The best classification models are those that maximize the parameters of the ROC curve: sensitivity and specificity. Sensitivity is the proportion of the model's accurate predictions of "ones" (indicating the occurrence of an event) to all the ones observed in a sample (the actual occurrence of an event).

To assess the quality of a model, the area under the graph of the ROC curve (denoted as AUC) can be calculated and taken as a measure of the goodness and accuracy of fit of a given model. The classification quality of a model is good when the curve is above the diagonal $y = x$, that is, when the parameter AUC has a value greater than 0.5. A general rule for assessing the classification quality of models is as follows (Hosmer, Lemeshow, 2000; Kumari, Rajnish, 2015):

- AUC = 0.5: classification is not good (it is comparable to a random classifier);
- 0.5 < AUC < 0.6: poor classification;
- 0.6 < AUC < 0.7: acceptable classification;
- 0.7 < AUC < 0.8: good classification;
- 0.8 < AUC < 0.9: very good classification; and
- AUC > 0.9: excellent classification.

Calculations and graphs have been done in RStudio v. 1.4.1717 software, equipped with the latest versions of the randomForest, h2o, and ROCR libraries.

4. Results and Discussion

4.1. One-dimensional analysis

The dataset collected through the survey is expressed in the form of qualitative variables. The categories of these variables are measured on a nominal or ordinal scale. Table 1 shows the structure of the received responses, where possessing cyber insurance (*INSURANCE* variable) acts as a grouping variable.

Table 1.
Structure of the sample grouped by the cyber insurance purchase criterion (n = 386)

Variable	Categories of variable and their codes	Does a firm have cyber insurance? (INSURANCE)		Total
		No	Yes	
TYPE	Manufacturing (1)	147	24	171
	Trade (2)	68	8	76
	Services (3)	114	15	129
EMPL	Up to 250 (1)	309	2	311
	More than 250 (2)	20	45	65
YEARS	Up to 10 years (1)	155	11	166
	More than 10 years (2)	174	36	210
FORM	Corporation (1)	238	39	277
	Other (2)	91	8	99
CPTL	Domestic (1)	301	32	333
	Foreign (2)	28	15	43
EQUITY	0-25% (1)	211	43	254
	26-50% (2)	50	4	54
	Above 50% (3)	68	0	68
TURN-OVER	Up to 50 (1)	284	14	298
	51-100 (2)	32	15	47
	More than 100 (3)	13	18	31
INTANG	Up to 25% (1)	289	41	330
	More than 25% (2)	40	6	46
COUNT	No (1)	274	4	278
	Yes (2)	55	43	98
PROB	Low (1)	30	2	32
	Medium (2)	245	20	265
	High (3)	54	25	79
IMPACT	Up to PLN 100k (1)	278	25	303
	More than PLN 100k (2)	51	22	73
FREQ	Less than once a year (1)	297	20	317
	Once a year or more (2)	32	27	59

Note: The numbers in parentheses next to the variable categories indicate the variable category codes used in the statistical analysis.

Source: the authors.

The size of a company, as measured by both its number of employees (*EMPL*) and the volume of annual revenue (*TURNOVER*), shows the strongest relationship with the purchase of cyber insurance. In terms of employment size, almost all medium-sized companies (50-250 employees) in our sample are uninsured (99.4%). In contrast, large companies have a high penetration of cyber insurance (69%) and only 31% of those surveyed are uninsured.

Turning to the second measure of company size, annual turnover, observations indicate that as the volume of turnover increases, the percentage of companies with cyber insurance increases.

While in the category up to PLN 50 million only 4.7% of respondents have insurance, in the turnover group of PLN 51-100 million the share of insured companies rises to 32%, and in the highest turnover category of over PLN 100 million 58% of respondents have purchased a cyber risk policy. These results are consistent with previous studies. Large organizations are more vulnerable to uncertain large losses caused by cyber-crime. This aspect of cyber risk motivates their decisions to prepare by buying related insurance (De Smidt, Botzen, 2017).

Another set of variables having a statistically significant association with buying cyber insurance are the probability of a successful cyber-attack on a company (*PROB*), the expected frequency of cyber-attacks on a company in the future (*FREQ*), and the potential losses resulting from a cyber-attack (*IMPACT*). These variables are based on respondents' subjective assessments of the scale of cyber threats and they are proxies for cyber risk perception. Their role is to measure the perception of cyber risk by an individual making key financial decisions in a company, including the decision to purchase insurance. The strongest association with the purchase of cyber insurance comes from the frequency of cyber-attacks (*FREQ*). In cases where respondents estimate that their company is likely to experience a cyber-attack no more than once a year (i.e., low frequency of incidents), the share of insured companies is only 6.3%. In contrast, when respondents expect cyber-attacks several times a year (i.e., high frequency of incidents), the percentage of insured companies is significantly higher, 45.8%. Subjective assessment of potential losses due to a cyber-attack on a company (*IMPACT*) is another factor differentiating whether a population of companies has cyber insurance.

Thus, in the group of respondents who believe that their company's maximum possible cyber loss will not exceed PLN 100,000, only 8% of companies are insured against cyber risk. If the anticipated losses are higher than PLN 100,000, the share of insured firms rises to 30%. When asked how respondents perceived the likelihood of their company becoming a victim of a cyber-attack in the future (*PROB*), 8.5% answered "low," 70.5% answered "medium", and 21.0% answered "high". However, the distribution of responses varies depending on whether a company has cyber insurance. Interestingly, in the group of insured companies, the percentages of individual responses are 0.5%, 5.3%, and 6.6%, respectively, while among uninsured companies the responses arranged as follows 8.0%, 65.2%, and 14.4%. These results confirm the fundamental principle that the propensity to purchase insurance is strongly influenced by risk aversion, as well as by subjective risk assessment (risk perception).

For insured companies, 14.4% of respondents perceived cyber risk as high, while among managers of uninsured companies half as many, 6.6% of respondents, indicated they perceived the highest assessment of this probability. Pearson's chi-square test of independence has been used to verify whether relationships between the purchase of cyber insurance and explanatory variables are statistically significant. Table 2 provides a summary of these results.

Table 2.
Pearson's Chi2 test of independence between the INSURANCE variable and selected explanatory variables

Description	**Chi2 test statistic**	**p-value**
H1: The purchase of cyber insurance is associated with the cyber risk perception of a managerial decision-maker in a company		
FREQ ↔ INSURANCE	73.73	Less than 0.0001
IMPACT ↔ INSURANCE	51.28	Less than 0.0001
PROB ↔ INSURANCE	48.86	Less than 0.0001
H2: The availability heuristic is related to cyber risk perception of a managerial decision-maker in a company		
COUNT ↔ FREQ	142.01	Less than 0.0001
COUNT ↔ IMPACT	18.78	Less than 0.0001
COUNT ↔ PROB	89.05	Less than 0.0001
H3: There are specific characteristics of a company that distinguish insured companies from uninsured companies against cyber risks		
EMPL ↔ INSURANCE	230.569	Less than 0.0001
YEARS ↔ INSURANCE	18.786	0.0009
TURNOVER ↔ INSURANCE	92.353	Less than 0.0001
EQUITY ↔ INSURANCE	15.550	0.0014
CPTL ↔ INSURANCE	26.371	Less than 0.0001
TYPE ↔ INSURANCE	0.7288	0.6946
FORM ↔ INSURANCE	9.0085	0.1155
INTANG ↔ INSURANCE	0.0524	0.9742

Source: the authors.

Regarding H1, findings indicate a strong association between cyber insurance purchase and cyber risk perception. How firm managers reported perceiving the probability and potential impact of a cyber-attack has affected their decision to insure against this risk. The same pattern applies to cyber risk awareness (i.e., the higher the anticipated possibility of a successful cyber-attack, the bigger the share of insured companies in the population of respondents). These findings align with De Smidt & Botzen (2017), who demonstrated that cyber risk perception is driven by risk awareness, perceived probability, and perceived damage.

Regarding H2, based on the assumption that the availability heuristic explains the risk perception of professional decision-makers, De Smidt & Botzen (2017) found that the experience of a successful cyber-attack positively impacts cyber risk awareness (PROB) and perceived cyber risk probability (FREQ), but is not associated with assessing the cyber risk's potential impact (IMPACT). Our results are partially consistent, indicating that the availability heuristic is related to cyber risk perception in all three aspects, thus confirming H2.

Findings indicate that firm size measured by employment and annual turnover, equity ownership, firm age, and leverage (the share of equity in total liabilities) are significantly related to existing cyber risk coverage in MLEs. Other factors, such as the type of business, a firm's legal form, and a firm's share of intangibles in total assets did not differentiate the population of insured and uninsured companies. In other words, these were not good indicators of a firm deciding to purchase cyber insurance. Hence, hypothesis H3 is confirmed.

4.2. Multiple Correspondence Analysis (MCA)

In this step, we investigate the factors that primarily influence cyber policy purchase decisions if all variables are analyzed jointly. In other words, we intend to create the profile of a company that insures against cyber risk. We have run MCA with three clusters in two dimensions. Table 3 presents the results of this analysis.

We applied a scaling method that combines MCA with dimension reduction, with k-means for clustering. Parameter *alpha_k* = 0.85 was set. This is a non-negative scalar to adjust for the relative importance of MCA (*alpha_k* = 1) and *k*-means (*alpha_k* = 0) in the solution. The chosen two dimensions explain 91.49% of the total inertia. If the analysis were expanded to three dimensions, only 0.08% more of the total inertia would be explained, but the ability to graphically present the results on a plane would be lost. Therefore, we decided to limit the analysis to two dimensions.

Table 3.
Multiple Correspondence Analysis results

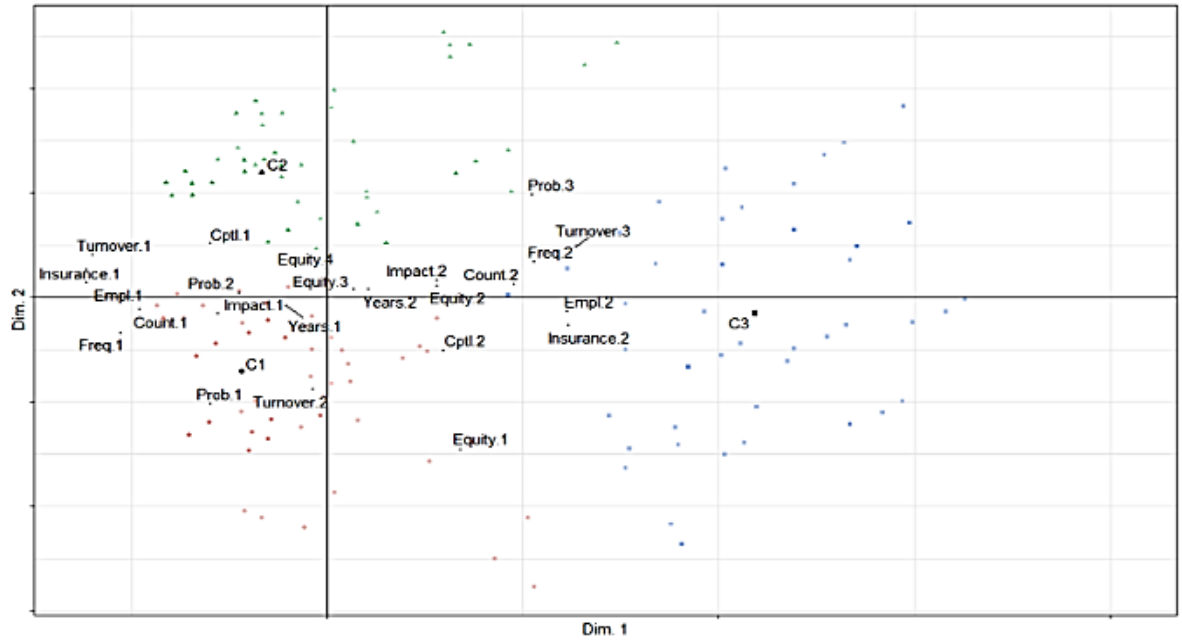
Dimension	**1**	**2**	**3**	
Singular value	0.327	0.025	0.020	
Eigenvalue	0.1071	0.0006	0.0004	Total inertia = 0.1446
Explained inertia (%)	91.11	0.38	0.08	
Cumulative explained inertia (%)	91.11	91.49	91.57	

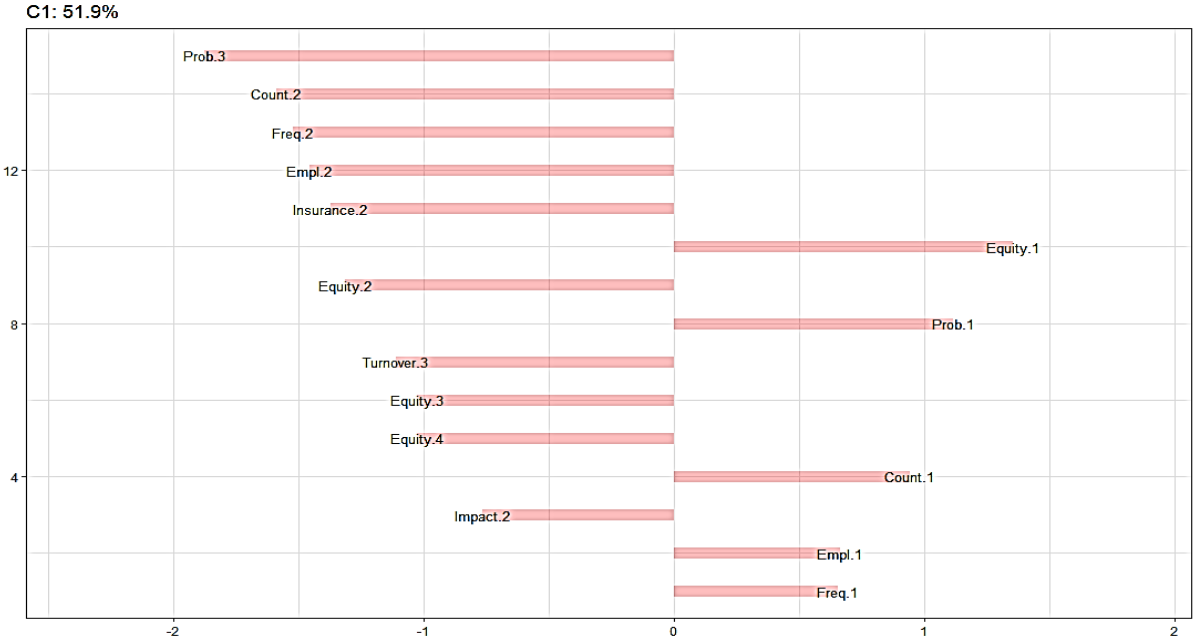
Source: the authors.

Fig. 1 presents the results of MCA in detail. The first graph shows a set of points representing all variables and their categories in a two-dimensional factor space. The points clearly arrange themselves into three groupings (clusters), denoted by points C1, C2 and C3. The other three bar graphs show the relative importance of the variable categories in each cluster. The longer the bar on the graph, the greater the impact of a given category of variable on forming a cluster, and the stronger relation with other variables in the cluster. The impact of a variable can be either positive (bar directed to the right) or negative (bar directed to the left). Cluster 1 (C1) includes 51.9% of objects, cluster 2 (C2) includes 33.2% of objects, and cluster 3 (C3) includes 14.9% of objects. C3 provides the most relevant insights into factors associated with the purchase of cyber insurance (*INSURANCE 2* variable category is the second most influential). They are as follows:

- business profile factors: employment above 250, annual turnover above PLN 100 million, and foreign equity capital;
- cyber risk perception: experience of at least one successful cyber-attack against the business in the last 5 years, high level of perceived frequency of cyber-attacks on the company "Once a year or more," high level of cyber risk awareness, and high anticipated impact of a cyber-attack on the company "More than PLN 100,000."

Thus, we have demonstrated that managerial cyber risk perception is associated with cyber insurance purchase decisions. Moreover, firm size and origin of equity are significant markers of firms deciding to insure against cyber risk.





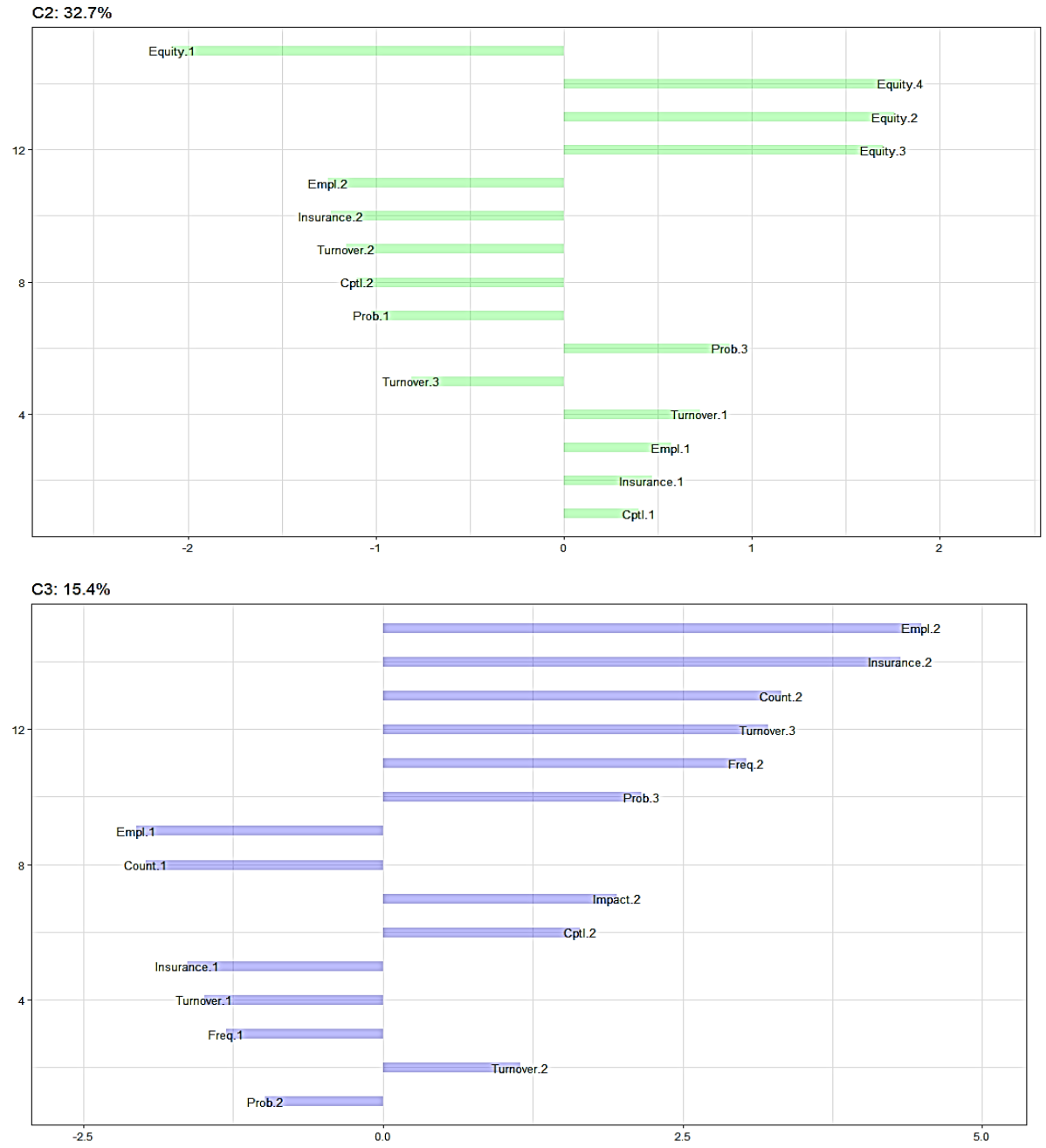


Figure 1. Clusters resulting from the MCA

Source: the authors

4.3. Random forests

Our purpose was to explore which elements of a company's business profile and cyber risk perception are related to having cyber insurance (so *INSURANCE* is a grouping variable). The analysis began with fitting an optimal set of random forest parameters, known as tuning. One hundred different parameter sets were tested. Table 4 shows the six best-fit models with their parameters.

Table 4.
Results of tuning the parameters of the random forest model – TOP 6 fitted models

Parameter *mtry*	Parameter *ntree*	Share of training dataset	Mean Squared Prediction Error (MSE)
2	200	0.75	0.0495
2	300	0.63	0.0498
2	500	0.80	0.0499
6	200	0.80	0.0499
6	300	0.75	0.0499
2	200	0.80	0.0504

Source: the authors.

Next, the parameter values were determined of the model that proved best in terms of minimizing MSE (*mtry = 2*, *ntree = 200*). Then, the accuracy of the model fit was com-pared using confusion matrices for the test dataset in three different variants: the baseline model (dataset unchanged, denoted STD), the model with oversampling (denoted OS), and the model with undersampling (denoted US). This step is shown in Table 5.

Table 5.
Comparison of confusion matrices of three variants of the data set of the selected model

Model	STD	OS	US
TP*	73	73	71
FP*	7	7	9
FN*	1	0	0
TN*	13	14	14
Accuracy	91.49%	92.55%	90.43%

Source: the authors.

The model with oversampling (OS) was chosen because it yields a maximum accuracy of 92.55%. Table 6 presents goodness-of-fit measures of this model. The model accurately identified companies with cyber insurance in most cases but was wrong in 7 out of 94 test data records.

Table 6.
Measures of goodness of fit of the estimated model

Parameter	Value
Accuracy = $\frac{TP+TN}{P+N}$	92.55%
Accuracy - 95% CI	(0.853; 0.970)
Sensitivity = $\frac{TP}{P}$	1
Specificity = $\frac{TN}{N}$	0.913
Negative Predicted Value = $\frac{TN}{TN+FN}$	1
Prevalence = $\frac{P}{P+N}$	0.223
Balanced accuracy = $\frac{Sensitivity+Specifity}{2}$	0.956

Notes: T - Positives (number of occurrences of an event), N - Negatives (number of non-occurrences of an event), TP - True Positives (number of correct indications of positive events), TN - True Negatives (number of correct indications of negative events), FP - False Positives (number of false indications of positive events), FN - False Negatives (number of false indications of negative events).

Source: the authors.

The high quality of the model fit to our dataset confirms the ROC curve (Figure 2). The value of the area under the curve (AUC) index is 99.6%, so the random forest with oversampling for the *INSURANCE* variable can be considered a very good classifier.

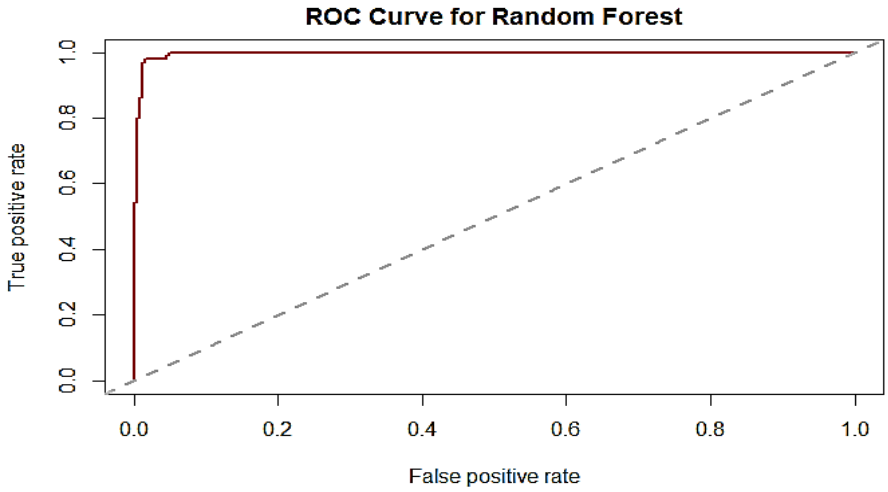


Figure 2. ROC curve.

Source: the authors.

The next step after the model estimation was to examine the impact of each variable on the classification result. The Mean Decrease Gini (MDG) was calculated for each variable. The higher the MDG, the greater the influence of the variable on the classification result. Figure 3 shows the highest influence of the *EMPL* variable. Moreover, the high impacts of *INC_MAXLOSS*, *COUNT* and *TURNOVER* are visible. Meanwhile, the variables *IMPACT* and *FREQ* have relatively lower impacts in the classification result.

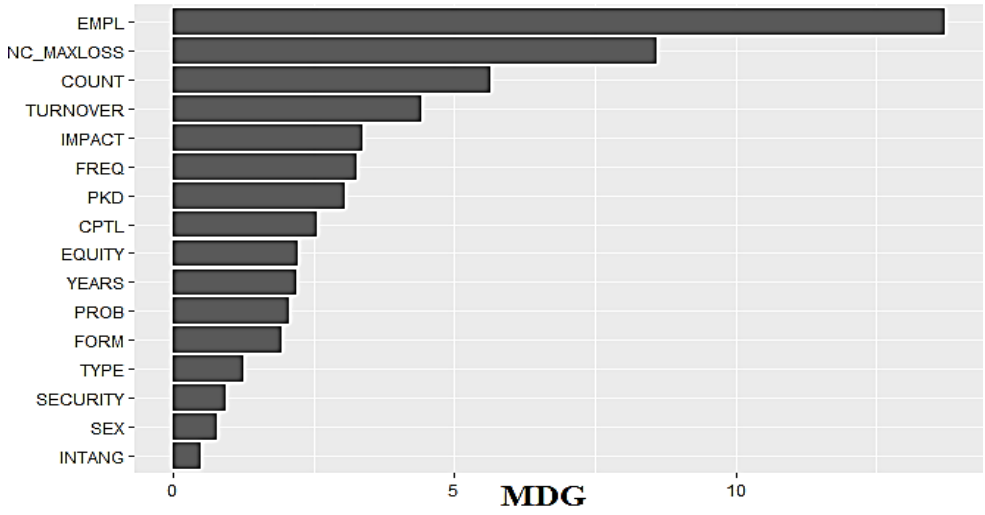


Figure. 3. MDG for explanatory variables in the estimated random forest model.

Source: the authors.

In other words, the factors that distinguish insured companies from uninsured ones in terms of cyber risk protection are primarily company size expressed as number of employees, and annual turnover. These are followed by the severity of experienced loss due to a cyber

incident, and the number of successful cyber-attacks experienced by a given company during the last 5 years. These are also the factors that make up a company's profile (i.e., the objective characteristics of an entity).

Lower on the list of factors influencing the decision to insure against cyber risk come cyber risk perception, that is, the expected negative consequences of a cyber incident and the subjectively perceived frequency of possible cyber-attacks in the future.

5. Conclusions

We found an association between managerial cyber risk perception and the decision to purchase corporate cyber insurance purchase. Negative experience with cyber threats, which shapes managers' risk perception, drives them to purchase insurance. Notably, the availability heuristic influences cyber risk perception. However, confidence in a manager's own company's cybersecurity capabilities does not affect their perception of cyber risk. Moreover, firm size, industry type, firm age, and equity ownership are significant markers of firms that have decided to insure against cyber risk. Therefore, we prove that the decision to purchase cyber insurance is also associated with some company profile elements.

Our MCA and random forests analyses show that decisions to purchase cyber risk coverage are mostly determined by factors related to company size, such as employment size and annual turnover. The number of successful cyber-attacks against a company, along with the maximum value of cyber losses a company has incurred in the last 5 years, are also important decision-making factors related to purchasing cyber insurance. This negative feedback from the past influences the cyber risk perceptions of company owners and managers, and thereby stimulates demand for cyber coverage.

Thus, we conclude that the propensity for company management to buy cyber insurance is driven by the interplay between cyber risk perception and company profile that defines firm-specific cyber risk exposure and insurance needs.

The reader should be aware of the potential biases and limitations of this study:

- Non-response bias: Despite efforts to reach a representative sample of individuals, it is always possible that individuals who did not participate in the survey are substantially different, in terms of their underlying beliefs, from those who completed the questionnaire.
- Sampling frame bias: The accuracy is based on contact information and the degree to which the list is representative of medium and large Polish enterprises that participated in the survey. We also acknowledge that the results may be biased by external events such as media coverage.
- Self-reported results: We are aware of the possibility that subjects may not provide accurate responses.

However, our study provides relevant feedback for policy-makers responsible for cyber security, particularly regarding incentives to improve cyber-attack resilience through cyber insurance. Moreover, the investigation of factors determining cyber insurance purchase can help insurance carriers target their offers on the market.

Cyber risk perception and insurance purchase decision-making are complex research-areas where both determinative factors and other cognitive processes can be influenced by each other. This can indicate that the dimensions differ across populations, industries and countries, creating grounds for further context-specific studies.

Further research may benefit from more multidisciplinary approach, and contextual studies within demand for cyber insurance can contribute to develop targeted tools for cyber risk management to enhance resilience of businesses and other organizations.

Acknowledgements

The research was funded by the University of Economics in Cracow research funds. The authors declare no potential conflicts of interest with regard to the research, authorship, and publication of this article.

References

1. Alladi, T., Chamola, V., Zeadally, S. (2020). Industrial control systems: Cyberattack trends and countermeasures. *Computer Communications, No 155*, pp. 1-8.
2. Anderson, R., Moore, T. (2006). The economics of information security. *Science, vol. 314(5799),* pp. 610-613.
3. Bank of Canada (2021). *Financial System Survey – Spring 2021,* https://www.bankofcanada.ca/2021/05/financial-system-survey-highlights-spring-2021, 12.12.2022.
4. Bank of England (2022). *Systemic Risk Survey Results – 2021 H2,* https://www.bankofengland.co.uk/systemic-risk-survey/2021/2021-h2, 12.12.2022.
5. Barberis, N.C. (2013). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspectives vol. 27(1),* pp. 173-196.
6. Böhme, R., Kataria, G. (2006). *Models and Measures for Correlation in Cyber-insurance*. Workshop on the Economics of Information Security, 26–28 June 2006. UK: University of Cambridge.

7. Bolot, J.C., Lelarge, M. (2009). Cyber Insurance as an Incentive for Internet Security. In: M.E. Johnson (ed.), *Managing Information Risk and the Economics of Security.* , New York, USA: Springer.
8. Botzen, W., Kunreuther, H., Michel-Kerjan, E. (2015). Divergence between individual perceptions and objective indicators of tail risks: Evidence from floodplain residents in New York City. *Judgment and Decision Making, vol. 10(4),* pp. 365-385.
9. Brando, D., Kotidis, A., Kovner, A., Lee, M., Schreft, S.L. (2022). Implications of cyber risk for financial stability. *Fed Notes, 12 May 2022*.
10. De Smidt, G., Botzen, W.J. (2017). Perceptions of corporate cyber risks and insurance decision-making. *Working Paper, vol. 18*. Philadelphia: University of Pennsylvania, pp. 1-33.
11. DTCC (2020). *Systemic risk barometer*, https://www.dtcc.com/-/media/Files/Downloads/Thought-Leadership/26362-Systemic-Risk-2020.pdf, 12.12.2022.
12. EIOPA (2023). *Insurance statistics*, https://www.eiopa.europa.eu/tools-and-data/statistics-and-risk-dashboards/insurance-statistics_en#Premiums,claimsandexpenses, 22.02.2023.
13. Grace, M.F., Rebello, M.J. (1993). Financing and the demand for corporate insurance. *Geneva Papers on Risk and Insurance Theory, vol. 18(2),* pp. 147-172.
14. Greenwald, J. (2020). Cyber insurance premiums climbing rapidly: S&P. *Business Insurance, 4 September 2020*.
15. Herath, H.S., Herath, T.C. (2011). Copula-based actuarial model for pricing cyber-insurance policies. *Insurance Markets and Companies, vol. 2(1),* pp. 7-20.
16. Herr, T. (2019). Cyber Insurance and Private Governance: The Enforcement Power of Markets. *Regulation & Governance – Early View*, DOI:10.1111/rego.12266.
17. Hosmer, D.W., Lemeshow, S. (2000). *Applied Logistic Regression*. New York, USA: Wiley.
18. Hoyt, R.E., Khang, H. (2000). On the demand for corporate property insurance. *Journal of Risk and Insurance, vol. 67(1),* pp. 91-107.
19. Krummaker, S. (2019). Firm's demand for insurance: An explorative approach. *Risk Management and Insurance Review, vol. 22(2),* pp. 279-301.
20. Kumari, D., Rajnish, K. (2015). Investigating the effect of object-oriented metrics on fault proneness using empirical analysis. *International Journal of Software Engineering and its Applications, vol. 9,* pp. 171-188.
21. Levite, A., Kannry, S., Hoffman, W. (2018). *Addressing the Private Sector Cybersecurity Predicament: The Indispensable Role of Insurance*. Carnegie Endowment for International Peace, https://carnegieendowment.org/files/Cyber_Insurance_Formatted_FINAL_WEB.PDF, 23.11.2022.
22. MacMinn, R.D. (1987). Insurance and corporate risk management. *Journal of Risk and Insurance, vol. 54(4),* pp. 658-677.

23. Main, B.G.M. (1982). The firm's insurance decision. Some questions raised by the Capital Asset Pricing Model. *Managerial and Decision Economics, vol. 3(1),* pp. 7-15.
24. Main, B.G.M. (1983). Why large corporations purchase property/liability insurance. *California Management Review, vol. 25(2).*
25. Marotta, A., Martinelli, F., Nanni, S., Orlando, A., Yautsiukhin, A. (2017). Cyber-insurance survey. *Computer Science Review, vol. 24,* pp. 35-61.
26. Mayers, D., Smith, C.W. (1990). On the corporate demand for insurance: Evidence from the reinsurance market. *The Journal of Business, vol. 63(1),* pp. 19-40.
27. Munich Re (2020). *Cyber insurance: Risks and trends 2020,* https://www.munichre.com/topics-online/en/digitalisation/cyber/cyber-insurance-risks-and-trends-2020.html, 30.11.2022.
28. Nolan, C., Fixler, A. (2021). *The economic costs of cyber risk, Foundation for Defense of Democracies,* https://www.fdd.org/analysis/2021/06/28/the-economic-costs-of-cyber-risk/, 12.12.2022.
29. Pal, R., Golubchik, L. (2010). *Analyzing self-defense investments in internet security under cyber-insurance coverage.* Proceedings of the IEEE 30th International Conference on Distributed Computing Systems, Genova, Italy, 21-25 June 2010, pp. 339-347.
30. Schapira, M.M., Nattinger, A.B., McHorney, C.A. (2001). Frequency or Probability? A Qualitative Study of Risk Communication Formats Used in Health Care. *Medical Decision Making, vol. 21(6),* pp. 459-467.
31. Slovic, P. (2000). *The perception of risk*. London, UK: Earthscan Ltd.
32. Talesh, S.A. (2017). Insurance Companies as Corporate Regulations: The Good, The Bad, and The Ugly. *DePaul Law Review, vol. 66(2),* pp. 463-502.
33. Talesh, S.A. (2018). Data Breach, Privacy, and Cyber Insurance: How Insurance Companies Act as "Compliance Managers" for Businesses. *Law & Social Inquiry, vol. 43(2),* pp. 417-440.
34. Tversky, A., Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, vol. 5(2),* pp. 207-232.
35. Van Schaik, P., Jeske, D., Onibokun, J., Coventry, L., Jansen, J., Kusev, P. (2017). Risk perceptions of cyber-security and precautionary behavior. *Computers in Human Behavior, vol. 75,* pp. 547-559.
36. Wei, X., Finbarr, M., Xian, X., Wenpeng, X. (2021). Dynamic communication and perception of cyber risk: Evidence from big data in media. *Computers in Human Behavior, vol. 122(106851),* DOI: 10.1016/j.chb.2021.106851.