How Do Consumers React to Cybercrime?

Rainer Böhme
Department of Information Systems
University of Münster
Münster, Germany
Email: rainer.boehme@uni-muenster.de

Tyler Moore
Computer Science and Engineering Department
Southern Methodist University
Dallas, TX, USA
Email: tylerm@smu.edu

***Abstract*—We conduct a secondary analysis of data collected to survey EU citizens' experiences and concerns with cybercrime. We devise a series of logistic regressions that measure how exposure to cybercrime can inhibit online banking, shopping and other activities. We consider three forms of exposure: directly falling victim, expressing concern about security, and reading news reports. We find that directly experiencing cybercrime decreases the likelihood of shopping and banking online by 4-5 percentage points. We find that expressing concern about cybercrime has nearly twice as much negative impact on online behavior than directly experiencing cybercrime. People who have not heard anything about cybercrime in news reports or from colleagues are more likely to bank online than those who have heard such reports. We conclude by reviewing limitations of existing survey approaches and make recommendations for improving questions in future cybercrime surveys.**

I. Introduction

Cybercrime has attracted more attention in recent years. As more daily activities are migrated online, from banking and commerce to communication and recreation, the potential for exploitation by criminals has increased. Indeed, a sophisticated underground economy has emerged, where criminals exploit the openness and scale of the Internet to defraud consumers in online banking and commerce [1].

While cybercrime is a very real threat, only a minority of consumers are directly victimized. Substantially more are made aware of the threat, and many people remain hesitant to participate online due to the perceived risk of cybercrime. But what factors affect that decision, and what countervailing factors might increase participation?

To answer these questions, we perform a secondary analysis of survey data collected on behalf of Eurobarometer in Spring 2012 [2]. This survey asked many questions about EU citizens' concerns about and reactions to cybercrime. The accompanying report presented descriptive statistics on how experiences with cybercrime varied across all 27 EU Member States.

In this paper, we instead focus on the relationship between experiences and concerns over cybercrime and the resulting actions taken by consumers. This allows us to quantify the impact of cybercrime in a statistically robust manner. In particular, we construct a set of logistic regressions that examine how exposure to cybercrime can lead people to reduce their online participation in banking, commerce and in general.

We use survey questions asking whether people intend to bank or shop less because of cybercrime as response variables. We then create explanatory variables in four groups. The first group measures *experience* with cybercrime, such as falling victim to identity theft or receiving phishing emails. The second group captures *concern* over cybercrime, such as concern about the security of online payments. The third group of predictors measure *exposure to news* about cybercrime, broken down by media source. Finally, the fourth group of predictors measure *proficiency*. This includes not only general indicators such as educational attainment, but also more specific ones reflecting online expertise. We also use a number of questions about security practices such as running antivirus and changing passwords as indicators of security proficiency.

We hypothesize that each of these groups of predictors impact people's decision to reduce online participation. Indeed, we find evidence that many of the collected predictors do impact consumer behavior. We find that general concerns about cybersecurity have the strongest negative effect. For instance, holding all other factors constant, concern over the security of online payments decreases the likelihood of banking online by 12 percentage points. By contrast, actually experiencing cybercrime has a smaller effect: falling victim to identity theft reduces the likelihood of banking online by 5 percentage points. We conclude that concern over cybercrime imposes greater opportunity costs in terms of reducing online participation than actually experiencing cybercrime.

We also investigate positive factors that correspond to increased online participation. General indicators of proficiency such as higher educational attainment can help, increasing the likelihood of buying goods online by 3 percentage points, for example. The biggest positive factor, perhaps unsurprisingly, is confidence about one's own Internet skills (increasing the likelihood of banking online by 12 percentage points and shopping by 8 percentage points).

We hope that the results presented in this paper can be used to improve our understanding of what drives the indirect costs of cybercrime. It has long been recognized anecdotally that cybercrime can inhibit online participation. However, from the models presented here, we now understand better which factors cause the biggest impact, both in reducing online participation and in bolstering it. Furthermore, the findings should also ultimately help to prioritize policy responses.

The paper is organized as follows. We begin by reviewing related work on understanding the impact of cybercrime on consumer behavior in Section II. Next, in Section III, we

Pre-publication copy, November 2012. To appear in the proceedings of the 7th APWG eCrime Researchers Summit (eCrime)

describe the key research questions and how we selected relevant questions from the survey for use in our model. We describe the survey demographics and regression model in Section IV, followed by results in Section V. We discuss limitations and opportunities for improving future survey questions in Section VI. Finally, we conclude in Section VII.

II. Related work

Much of what we know about the prevalence of cybercrime comes from the direct observations of researchers [1]. Unlike physical crimes, cybercrimes can sometimes be remotely monitored by security researchers over the Internet. This has occasionally yielded insights on victimization rates, such as the one estimate suggesting that 0.4% of the Internet population falls for phishing attacks annually [3].

However, while direct observation has been crucial in understanding the behavior of cybercriminals, it has been less useful when the goal is to understand the behavior of victims or those not directly affected. In these circumstances, surveys can be helpful. For example, the long-running CSI survey asks firms questions about their experiences with various attacks, the resulting losses, and their security investments [4]. Meanwhile, a few surveys target consumers for their experiences and concerns with online crime. One example is the Eurostat ICT survey that asks a few security-specific questions [5].

Companies from the information security industry occasionally commission surveys of consumers to gauge the impact of cybercrime. For example, Symantec surveyed consumers worldwide about their experiences with cybercrime, finding that 54% of surveyed adults have “experienced virus/malware attacks”, and that 58% of these incidents occurred within the last year [6]. One limitation of industry-sponsored surveys is that they can be presented in ways designed to hype fears of cybercrime. For instance, the Symantec report takes the survey data and extrapolates loss estimates using estimated average losses. As explained by Florêncio and Herley [7], multiplying the victim rate by average losses can lead to grossly inflated estimates when the loss distribution is highly skewed, as is typically the case for cybercrimes.

The dataset we rely on in this paper was conducted on behalf of the European Commission, so there should be no corresponding incentive problem. Furthermore, we were given access to the microdata on survey responses, which enabled the analysis presented in this paper.

It has been recognized that the indirect costs of cybercrime, including a reduction in online participation, can be very high [8]. But what exactly leads people to limit their activities in favor of improved security?

One possible explanation is behavioral. Some economists have argued that emotional reactions to risky behavior can lead people to take decisions that are not purely rational [9]. Another possibility is that media attention to cybercrime issues might influence people’s decisions to limit their online participation. To that end, an investigation of Dutch media reports on card skimming found that debit card use fell significantly on days when articles on card skimming appeared in national newspapers [10]. The present work offers additional explanations. It helps to quantify the factors (experience, concern and media exposure) that contribute to consumers changing behavior that drives up indirect costs.

III. Analytical Approach

We first articulate the research questions we hope to answer, and then we explain how we selected questions from the Eurobarometer survey to construct our statistical model.

A. Research questions

We are interested in how exposure to cybercrime, be it from direct experiences, latent concern or news reports, could lead people to reduce their online participation. In particular, we hypothesize the following:

- **H1**: Falling victim to cybercrime reduces online participation, in particular online banking and shopping.
- **H2**: Expressing concern over cybercrime reduces online participation, in particular online banking and shopping.
- **H3**: Exposure to cybercrime in the news media reduces online participation, in particular online banking and shopping.
- **H4**: Falling victim to one form of cybercrime reduces participation in unrelated forms of online activity.

Because we expect that proficiency moderates the hypothesized relationships, we test these hypotheses after controlling for the influence of third variables measuring the individuals’ proficiency on various levels of specificity. Otherwise, variability of proficiency in the surveyed population might preclude a statistical identification of the relationships of interest, or exaggerate effects in the case of spurious correlation. For similar reasons, we control for possible level shifts in the variables of interest emerging from different legal and cultural environments as well as differences in technology adoption across the 27 EU Member States.

B. Operationalization

The Eurobarometer series of population surveys in Europe are carried out by private sector market research firms on request of the Directorate-General for Communication, “Research and Speechwriting” of the European Commission, the EU’s executive arm. Its purpose is to inform policy makers about the public opinion on current policy areas. Here we use data of the Eurobarometer Special 390 on Cyber-Security, collected in wave 77.2 of the regular Eurobarometer [2].

The survey asked a wide range of questions. We selected questions as indicators to help answer the hypotheses just raised. We grouped the indicators into the following categories:

- Effects of cybercrime
- Experiences with cybercrime
- Concerns about cybercrime
- Exposure to news about cybercrime
- Proficiency indicators

We now review the questions that we subsequently use to build our statistical model.

TABLE I
DESCRIPTIVE STATISTICS

Indicator	%
Effects of cybercrime (dependent variables)	
Less likely to buy goods online	17.5
Less likely to bank online	14.4
Less likely to participate online (summary of:)	63.0
– Less likely to give personal information on websites	36.3
– Only visit websites you know and trust	33.5
– Do not open emails from people you don't know	42.8
Experience with cybercrime	
Personal experience (at least "occassionally") with ...	
– Identity theft	8.0
– Phishing/advance-fee fraud spam	37.4
– E-commerce fraud	12.2
Concerns about cybercrime	
Personally (at least "fairly") concerned about ...	
– Identity theft	63.3
– Phishing/advance-fee fraud spam	50.2
– E-commerce fraud	51.7
Generally concerned about ...	
– Security of online payments	37.1
– Misuse of personal data	39.7
Exposure to news about cybercrime	
On television	66.5
On radio	22.9
In the newspapers	33.3
On the Internet	33.9
From friends, family or colleagues	25.5
Not heard anything about cybercrime (spontaneous)	14.8
Proficiency indicators (control variables)	
Internet access more than once a day	54.2
Bank online	47.8
Buy goods or services online	52.0
Feel confident about Internet skills	67.7
Feel informed about the risks of cybercrime	51.1
Changed at least one password in the past 12 months	48.4
Use different passwords for different sites	24.8
Antivirus installed	50.7
Higher education	46.5
Perceived social status above median	51.3

$N = 18133$ EU residents, Internet users, age 15+

a) Effects of cybercrime: The Eurobarometer survey asked respondents if they had changed any behaviors as a causal consequence from concern over cybercrime. Here is the question asked, along with a sample of available answers:

> Has concern about security issues made you change the way you use the Internet in any of the following ways?
>
> - Less likely to shop online
> - Less likely to bank online
> - Less likely to give personal information on websites
> - Only visit websites you know and trust
> - Do not open emails from people you don't know

While these changes do reduce the likelihood of being harmed by cybercrime, the actions also introduce opportunity costs, since people forgo the benefits of online participation.

Table I presents summary statistics for the responses to these questions. As can be seen in the table, 18% of users claim that they are less likely to purchase goods online, while 14% report to use online banking less often. Interestingly, caution over web browsing and email viewing is at least twice as frequently expressed: 36% intend to share less personal information and 42% say they do not open email from strangers.

Benefits to society of online participation can take many forms. First, e-commerce enables more choice, better price transparency and (for many goods) more efficient distribution. Second, online banking reduces transaction costs as financial services are intangible by nature and therefore best processed in an electronic system. In both bases, foregone benefits due to the risk of cybercrime can be estimated by the number of people claiming they are less likely to shop and bank online, respectively.

In addition to these specific branches of commercial and financial activities, we aim to include a broader indicator of online participation which encompasses non-monetary benefits such as information sharing, collaboration, and political and social activities. To construct an operable summary indicator from the answers to questions available in the survey, we combine the remaining three questions. Our summary indicator for online participation takes value 'yes' if the respondent answers 'yes' to at least one of the three questions. We are aware that this summary indicator is not perfect (see a discussion of its limitations in Section VI), but each of its components captures a relevant aspect of online participation in general. Reasonable disclosure of personal information, such as one's preferences and skills, is a prerequisite for personalization and matching. Sharing information is required for a range of valuable applications, ranging from recommender systems to initiating and maintaining collaborations with others. Visiting unknown websites is important for forming independent opinions and a prerequisite for comparing different options. Similarly, the ability to get in touch with new people via email substantially decreases communication and coordination costs in a society.

We use these three indicators as dependent variables in independent analyses, because they articulate changes in behavior that arise from experience with cybercrime. We believe it is reasonable to infer that how respondents answer other questions regarding their experiences, concerns and awareness will influence whether or not they take the actions described by these questions.

We next consider a series of questions that might help explain why people have decided to limit their online activities.

b) Experiences with cybercrime: One natural reason why users might choose to limit their online participation is if they have personally been victimized by cybercrime. The Eurobarometer survey asked respondents the following question:

> Cybercrimes can include many different types of criminal activity. How often have you experienced or been a victim of the following situations?
>
> - Identity theft (somebody stealing your personal data and impersonating you, e.g. shopping under your name)
> - Received emails fraudulently asking for money or personal details (including banking or payment information)

- Online fraud where goods purchased were not delivered, counterfeit or not as advertised
- Not being able to access online services (e.g. banking services) because of cyber attacks

Respondents were asked to answer "often", "occasionally", "never", or "don't know". We can see that phishing emails are encountered most frequently, by far: 37% of users report to encounter them "often" or "occasionally". Respondents report identity theft 8% of the time, compared to 12.2% for e-commerce fraud. We mention the last situation for completeness but do not consider its responses in the analysis because we believe it is almost impossible for consumers to tell and recall for what reasons a service was unavailable.

c) Concerns about cybercrime: Apart from experience, concern over cybercrime could drive people to take precautions online. The survey asked about concerns in two questions. First, after asking if respondents had ever experienced various detailed classes of cybercrimes, they asked whether respondents were personally concerned about each cybercrime category:

> And how concerned are you personally about experiencing or being a victim of the following cybercrimes?

Respondents were asked to answer "very concerned", "fairly concerned", "not very concerned", or "not at all concerned". Overall, concern was much more common than experience: 63% expressed concern over being victimized by identity theft, compared to 8% having experienced it. The gap between concern and experience was smaller for receiving phishing and advance-fee fraud emails: half of respondents are concerned while 37% receive such messages at least occasionally.

Second, the survey asked the following general question about concern over cybercrime before it talked about experience or individual forms of cybercrime:

> What concerns do you have, if any, about using the Internet for things like online banking or buying things online?

The answers to this open question were categorized by the interviewer who checked corresponding categories, of which we include the following two in our analysis:

- Security of online payments
- You are concerned about someone taking/misusing your personal data.

Spontaneous expression of concern is somewhat lower than in response to the explicit question after recalling actual experience, but still at close to 40% for both categories.

d) Exposure to news about cybercrime: What might explain the huge gap between concern and experience of cybercrime? Exposure to news reports on cybercrime could certainly drive up concern among non-victims. Fortunately, the survey also asked questions about media exposure:

> Cybercrimes can be defined as any crimes which are committed via the Internet. In the last 12 months, have you seen or heard anything about cybercrime from any of the following?

- Television
- Radio
- Newspapers
- Internet
- Friends, family or colleagues
- Not heard anything about cybercrime

Two thirds of respondents have heard stories about cybercrime on television, compared to one quarter from Internet resources. Only 14% claim to have not heard anything about cybercrime in the media or from others.

e) Proficiency indicators: Finally, apart from exposure through the media, people might become aware of cybercrime risks and adjust their behavior due to their own personal educational attainment. The survey collected traditional demographic indicators related to proficiency, along with questions about ICT and questions that can be interpreted as indicators of security proficiency.

General proficiency indicators include higher educational attainment and perceived social status. Indicators of online proficiency include daily Internet access, using online banking, buying goods online and confidence in Internet skills. Indicators of security proficiency include using antivirus, changing passwords, and feeling informed about cybercrime risks.

We include these proficiency indicators for two reasons. First, some indicators may influence the response variable. Second, we include them to control for natural variation in survey demographics so that we might describe the effect the other indicators have on a representative consumer.

IV. Data analysis

A. Survey demographics

The data contains responses to face-to-face interviews with 26,593 residents of 27 EU Member States conducted in the appropriate national language. The sampling method involved stratification by country, followed by random route and closest birthday rules to identify target households and survey participants, respectively. The dataset comes with weights, calculated as inverses of the sampling probability, to adjust the stratified sample against a universe description of socio-demographic indicators including gender, age, region and size of locality derived from the Eurostat population data, which is based on the official national censuses. Therefore, the weighted raw data can be regarded as a representative sample for all EU residents above the age of 15.

Because we are primarily interested in Internet users, we exclude all records from respondents who state that they do not use the Internet either at work, home, or in public places like libraries or Internet cafes. This leaves us with 18,133 responses representative for the EU Internet users above the age of 15. Table II summarizes the differences in key socio-demographic variables between the overall population and the population of Internet users. It shows the typical pattern of slightly better educated, younger, and more urban Internet users compared to the total population.

TABLE II
Socio-demographics of Internet users vs total population

	Total population (%)	Internet users (%)
Gender		
Male	48.3	50.9
Female	51.7	49.1
Age		
15–24	14.4	19.3
25–39	24.6	31.3
40–54	25.8	28.6
55+	35.5	20.8
Full-time education		
Up to the age of 15	20.3	8.7
16–19	42.6	43.6
20+	26.4	33.8
Still studying	9.2	12.7
Size of locality		
Rural area or village	32.8	30.0
Small or middle sized town	41.5	42.0
Large town	25.6	27.8

B. Statistical model

We construct logistic regression models to understand how concerns, experiences and awareness of cybercrime affect consumer behavior. We use the indicators of cybercrime effects as dependent variables, and use the experience, concern, media exposure and proficiency indicators as explanatory variables. The estimated model parameters quantify the extent to which experience, concern and exposure to cybercrime reduces online participation in banking, commerce and beyond.

Specifically, we estimate parameter vector $\hat{\boldsymbol{b}} = (\hat{b}_0, \hat{b}_1, \ldots)$ using the following equation per response record,

$$\overbrace{\log\left(\frac{p}{1-p}\right)}^{\text{Behavior}} = b_0 + \overbrace{b_1 D_1 + \ldots +}^{\text{Experience}} \overbrace{b_i D_i + \ldots +}^{\text{Concerns}} \underbrace{b_j D_j + \ldots +}_{\text{Exposure}} \underbrace{b_k D_k + \ldots +}_{\text{Proficiency}} \varepsilon, \tag{1}$$

where

- the dependent variable p is the probability of a respondent saying that he is less likely to do some activity online (shopping, banking, participation in general),
- b_0 is a constant intercept,
- $(D_1, \ldots, D_{i-1})$ are binary values of the explanatory variables derived from questions on the respondent's experience with cybercrime,
- likewise, $(D_i, \ldots, D_{j-1})$,$(D_j, \ldots, D_{k-1})$, and $(D_k, \ldots)$ are binary values derived from the respondent's answers to questions on concerns, exposure, and proficiency,
- and ε is an error term capturing the difference between the true relationship and the specified linear model for the logit-transformed dependent variable.

With $\boldsymbol{D}$ and p known for all N response records, $\hat{\boldsymbol{b}}$ can be estimated with the maximum likelihood method. We use the implementation provided by the survey extension for *R*, a statistical package, to account for the weighting required to correct known coverage errors of the sampling strategy [11].

Our main results are estimates for three logistic regressions, one for each dependent variable: (1) less likely to shop online, (2) less likely to bank online, and (3) our summary indicator on reduced online participation. Almost all explanatory variables are included in all three regression with the exception of two proficiency indicators. We believe that information on whether an individual actually shops or banks online is a sufficiently specific indicator of online business experience. But *actual* online shopping is obviously endogenous in a regression with the likelihood to *reduce* online shopping as dependent variable, in particular since the survey could not be very specific about the timing of changes in behavior. Therefore, it is safer to exclude this variable. Instead we decided to include actual online *banking* as a proxy which seems still specific enough, but does not cause ambiguity in the interpretation. We mirror this procedure for the second regression, and we did not find a reasonably specific substitute for the third regression explaining broader online participation.

Missing values in one variable (e.g., due to declined response) led to an exclusion of the entire record. We chose list-wise exclusion over imputation because missing values of multiple variables are largely clustered in specific records. We could not identify a systematic pattern behind the distribution of the altogether 859 ($< 5\%$) incomplete records in the dataset.

A final methodological remark concerns the inclusion of country fixed effects. Our sample spans 27 different countries with their own culture, language, jurisdiction, and path of technology adoption. These differences in the institutional environment certainly affect behavior of consumers, criminals, businesses, and law enforcement. They cause level shifts in many variables between countries. For example, the percentage of the first dependent variable (less likely to buy goods online) varies among the large and statistically relevant EU Member States in the range between 13% for Germany and 28% for Spain. (For more details, see the country breakdowns in [2].) Since our interest is not on aggregate effects of differences between countries, but on the representative consumer's individual reaction to cybercrime, we attenuate the variation between countries by including additive dummies for 29 regions.[1] The associated coefficients represent a linear approximation of the specifics in a country. Another way to think of this is to allow the intercept b_0 in Eq. (1) to take one value per region.

V. Results

Table III (in the appendix) reports the estimated coefficients $\hat{\boldsymbol{b}}$, one regression per column. The coefficients describe the effect of a change in the predictor on the (log odds ratio) of the dependent variable. A *positive value* denotes that answering 'yes' to the predictor question *decreases* the likelihood of buying, shopping or participating online and vice versa. Along

[1] Eurobarometer samples East Germany and Northern Ireland as independent regions to reduce sampling errors and, in the case of Germany, to facilitate the construction of long-running historical time series.

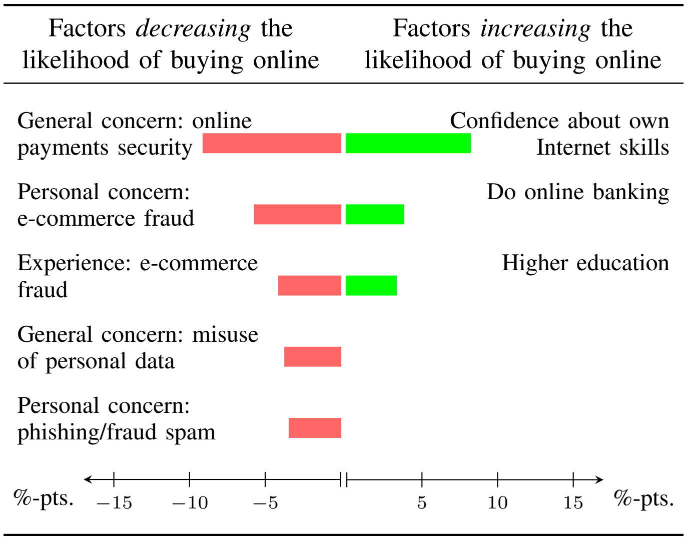

Fig. 1. Effect sizes of the most relevant predictors of the decision (not) to buy online despite (because of) fear of cybercrime, given as partial effect for a representative consumer after controlling for third variables (in %-pts.).

with the coefficients we report the standard errors of the estimation (in brackets, adjusted for the stratified sampling) and we use the convention to annotate coefficients which are statistically significantly different from zero with one, two, or three stars if the probability of error (i.e., reporting a spurious correlation) is less or equal to 5%, 1%, and 0.1%, respectively.

We also report Nagelkerkes' pseudo-R^2 as an overall measure of model fit. Broadly speaking, one can say that our model specifications explain between 70% and 86% of the variation in the dependent variable, although this analogy from least squares regressions of Gaussian variables does not fully hold in our case of a logistic regression with many binary predictors. Also note that the pseudo-R^2 values summarize the explanatory power of the reported predictors as well as the (suppressed) country fixed effects. Since a large share of variation is between countries and therefore easily explainable by the fixed effects, these extremely good fitness measures should be interpreted with caution. As we are not striving for predictive power, but rather for statistically robust identification of relationships between variables, any value for pseudo-R^2 above 0.1 (10%) is sufficient to justify interpretation.

Unlike for linear regressions where the dependent variable is numerical, the log odds ratio produced in logistic regression tables do not have a directly intuitive interpretation (apart from their sign). Instead, in the following subsections we compute partial effects for a representative consumer when controlling for all other explanatory variables. This lets us state the percentage point change in the dependent variable that corresponds to the predictor being set.

A. *Impediments to online shopping*

Figure 1 presents the factors that correspond to changes in the likelihood of shopping online. We can immediately see that expressing concerns have a more negative impact than actually experiencing cybercrime. People who fret over the security of online payments exhibit a 10 percentage point increase in the likelihood of reducing online shopping. Given that 18% of consumers overall claim to shop online less due to cybercrime concerns (see Table I), this corresponds to a more than 50% increase in the cautious behavior. General concerns about personal data misuse also have a strong negative impact on the decision to shop online.

By contrast, victims of e-commerce fraud are 5 percentage points less likely to bank online, substantially less than those who are just concerned. This is a striking result. One might suspect that being defrauded would deter future participation. While there is a negative effect, and victims indeed express concern slightly more often than unaffected indviduals (43% vs 37%), one reason why the partial effect of experiencing fraud is less substantial than merely expressing concern could be that those who experience e-commerce fraud realize that the experience is not overly taxing. Banks and credit card companies often reimburse customers in disputes, whereas those who have not experienced fraud might envision a more terrible outcome.

Furthermore, expressing a more personal concern—fears over falling victim to e-commerce fraud—decreases the likelihood of reducing online shopping somewhat less than the more general concern about payment security.

On the positive side, confidence about one's own Internet skills substantially increases the likelihood of shopping online. We must be careful how to interpret this finding, however. Nearly 70% of people already exhibit confidence in Internet skills, and so it is unlikely that even substantial additional investment in skills training has sufficient leverage to cure the problem. Unsurprisingly, banking online makes it more likely to also shop online. Finally, higher education correlates with an increase in online shopping despite security issues.

Notably, none of the media exposure indicators are significant. That is, hearing about cybercrime in the news or in conversation does not appear to deter online shopping. This suggests that consumers may not be making the connection between reports of cybercrime and e-commerce risks.

Another seemingly missed connection is between identity theft and e-commerce. Falling victim to identity theft does not correlate with a reduction in shopping online.

B. *Impediments to online banking*

Figure 2 presents the factors that correspond to changes in the likelihood of banking online. In many respects, the effects are similar to those found for online shopping. Once again, concerns are quantitatively more inhibiting than actual experience. In particular, general concerns are more influential than specific personal concerns.

Among negative experiences, identity theft scores highest, intuitively the most detrimental of all reported experiences with regard to online banking. However, when we examine the effect of experiencing other cybercrime, we can now see a difference from the results for online shopping. Recall that for online shopping, only experiencing e-commerce fraud, not identity theft, deterred future shopping. Instead, for online

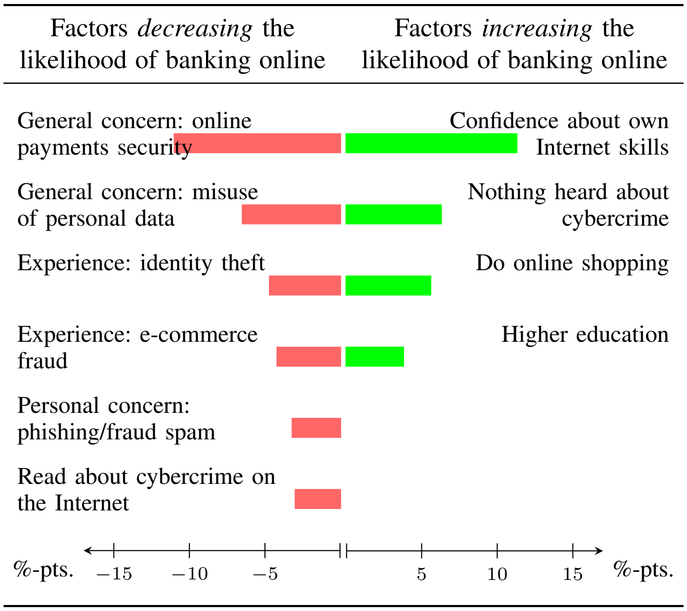

Fig. 2. Effect sizes of the most relevant predictors of the decision (not) to bank online despite (because of) fear of cybercrime, given as partial effect for a representative consumer after controlling for third variables (in %-pts.).

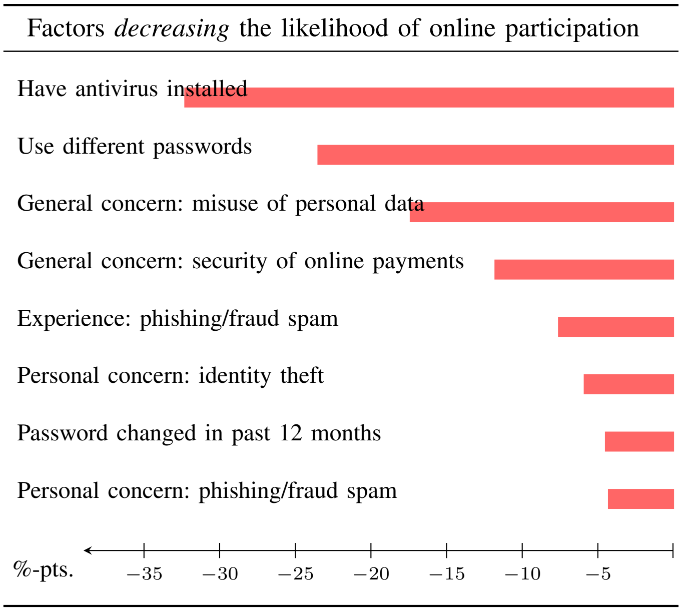

Fig. 3. Effect sizes of the most relevant predictors of the decision not to participate online because of fear of cybercrime, given as partial effect for a representative consumer after controlling for third variables (in %-pts.).

banking we observe a negative "spillover" effect from negative experiences in e-commerce fraud. In principle, being victimized by one form of cybercrime should not inhibit online participation in other unrelated areas. An experience with a dodgy online seller, for instance, should not affect someone's propensity to use online banking. However, we can see from the data that this does in fact happen: experiencing e-commerce fraud decreases the likelihood of banking online by 4 percentage points. This is only slightly less than the effect of experiencing identity theft.

Similar to the case of online shopping, we see positive influence from confidence about one's own Internet skills (+11 %-pts. vs +8 %-pts. for shopping), and for higher education. Online shopping creates a positive spillover, as anticipated by the inclusion of this variable as proficiency indicator.

Unlike for online shopping, the effects of media exposure on online banking are statistically significant: ignorance of cybercrime increases the likelihood of banking online, and conversely, having heard about cybercrime decreases the likelihood of online banking—in particular when consulting Internet sources. At this point we can only speculate if this difference of media effects is because of the media's tendency to primarily cover the most scandalous threats, e.g., banking rather than shopping fraud, or whether direct positive experience with uncontested online purchases overshadows media reports about individual cases of fraud. What we can state with more certainty is that the positive effect of ignorance about cybercrime only affects a small minority of people (15%). Therefore, it does not seem to be a winning strategy for banks to hush up cybercrime in order to move more retail business online. Nevertheless, it is interesting to note that banks' interest on the prominence of cybercrime in the media is exactly opposite to that of the security industry, who tries to play up the significance of the problem [8].

C. Impediments to online participation

Finally, Figure 3 presents the factors that correspond to changes in the likelihood of online participation as measured by our summary indicator. The first observation is that unlike for the previous regressions, all significant factors are negative, that is, they decrease the likelihood of online participation.[2]

A second observation is that the impact of many factors is higher in terms of percentage point changes than for online banking and shopping. This can be attributed to the fact that 63% of people claim to be less likely to participate online, compared to 18% for banking and 14% for shopping. Hence, the 32 %-pts. decrease in the likelihood of participation corresponds to a roughly 50% reduction compared to the average consumer. By comparison, the 12 %-pts. reduction in the likelihood of banking online due to concern about online payment security decreases that likelihood by two thirds.

The two biggest negative factors (installing antivirus software and using different passwords across websites) reflect general security awareness and therefore a more skeptical attitude towards online safety and security.

Once again, we see that general concerns about security trump more personal concerns, suggesting that less specific risks have more of a chilling effect on participation than do particular risks such as identity theft and phishing emails. Finally, it is worth noting that achieving higher education is now associated with a reduction in online participation (by as little as 3.3 %-pts., therefore not included in Fig. 3), unlike for

[2] There is one small positive effect not included in the figure. Frequent Internet use increases participation by 3.8 %-pts., but this is to be expected given that frequent Internet users are quite likely to continue participation.

online banking and shopping. The likely explanation for this result is that higher education achievement is also correlated with more secure behavior.

D. Robustness

We conducted some robustness and sanity checks before we settled for the model specification presented in this paper. Stepwise inclusion of groups of predictors did not reveal noteworthy suppression effects. Certainly, some predictors are mildly correlated; partly by definition, as for the case of the media channels and the "nothing heard" indicator. Since all predictors are binary, we do not have to worry about outliers driving the results. The only robustness check worth documenting is an alternative specification without inclusion of country fixed effects. Table IV (in the appendix) reports the estimated coefficients. We are pretty confident in the robustness of our results because none of our relevant predictors change sign or substantially in magnitude.

Quite expectedly, the goodness of fit measure is smaller when the between-country variation is averaged out in the alternative specification without country fixed effects. Comparing all six regressions, we see that the additional explanatory power of the country fixed effects is much higher for online shopping and online banking than for general online participation. This is plausible if one recalls that the institutional environment, which the country fixed effects try to control for, affects specific online businesses more (e.g., by direct regulation) than it affects online participation in general.

VI. Discussion

We now discuss the results of the analysis, its limitations, and opportunities for improving future cybercrime survey questions.

A. Revisiting the hypotheses

We can return to the research questions posed in Section III-A in order to assess the extent to which they are supported by the data analysis.

- **H1**: *Supported with evidence*
- **H2**: *Supported with strong evidence*
- **H3**: *Supported only for online banking*
- **H4**: *Some support for e-commerce fraud*

Regarding H1, we found for each indicator that experiencing cybercrime decreased the likelihood of online participation. Furthermore, for H2, we found lots of evidence that expressing concern about cybercrime corresponds to reducing online participation. We found mixed support for hypothesis H3: exposure to cybercrime in the news media reduces online banking activity, but does not have a discernible impact on online shopping or other forms of participation. Finally, for H4, we find some evidence for negative spillover from experiencing one form of e-crime onto other activities. Falling victim to e-commerce fraud hinders online banking participation, but suffering identity theft does not seem to reduce online shopping activity.

B. Overarching observations

One important and unexpected finding from the analysis is that *concern* about cybercrime inhibits online participation more than direct *experience* with cybercrime does. One explanation for this result is that people might find the actual experience of cybercrime to be less painful than their conception of what might be possible. Regardless of what is driving the result, its implications are clear: assuaging society's concerns over cybercrime could make a greater impact than allocating further resources on assisting victims. Note that we are *not* suggesting to reduce current levels of support for cybercrime victims; rather, current forms of support (e.g., credit-card fraud reimbursement policies) leave victims with a modestly negative effect on their future participation.

For all forms of participation, we found that general concerns about cybercrime had a bigger impact than more personal concerns. These general concerns were asked in the form of open questions. People who voluntarily express concerns without being prompted seem to have internalized the concern to the extent that it actually affects behavior. Part of this effect can also be attributed to known biases in people's perception of risk: abstract and unknown risks appear more threatening than risks which people can understand, have experience with, and feel in control of [12].

A number of predictors were not significant across all forms of participation. For example, feeling informed about cybercrime risks was not significant when country fixed effects were included in the regression. Without country fixed effects, this indicator became significant, which suggests that the level of security awareness varies substantially between countries. This is an area that could benefit from further investigation.

Finally, it is worth noting that media exposure to cybercrime proved less relevant to people's behavior than we had expected. In particular, differentiation between channels (TV vs newspaper, etc.) seems irrelevant. Of course, the concerns people have about cybersecurity must have come from somewhere, and it remains possible that these concerns were triggered by media exposure even if it was not accounted for in the survey.

C. Limitations

First, recall the general limitations of cross-sectional population survey data. The formulation of hypotheses suggests causality although the data strictly allow only statements about correlation. In addition, even a representative sample and face-to-face interviews, the most accurate (and expensive) mode of data collection, does not make our results immune to biases from response effects. Most importantly, our hypotheses stipulate change in behavior as a causal consequence of cybercrime. What we actually measure is *stated behavior* in response to the question wording "security issues made you change the way you use the Internet". The responses are susceptible to well-known cognitive biases, such as selective retrieval from memory, or response editing to meet expectations of social desirability [13]. Since a series of eight items was asked in a single block of questions, it also remains uncertain if all respondents have correctly remembered the causal part of the question for all their responses. (Fortunately, our main indicators shopping and banking were asked first.)

Another limitation, common to all secondary analyses, is that the indicators we used as response variables were limited by the questions asked in the survey. We are mostly pleased with the questions on online banking and shopping (though see suggestions for improvement in the next section below). However, we wanted to include a broader indicator that encompassed participation beyond these two important areas. Because no questions were directly asked about changes in general online participation in response to cybercrime, we had to construct a composite indicator of online participation based on three questions that are instances of limited participation. Consequently, the indicator does not completely cover all forms of online participation. Furthermore, treating people who answered positively to at least one of these questions as less likely to participate online is a simplification that loses information, since someone answering affirmatively to all three surely participates less than someone who answers affirmatively only once. Yet another difficulty with this indicator is that the questions were asked in a block together with questions we used as security proficiency indicators. This increases the risk that response pattern effects are responsible in part for the quantitatively largest negative effects of security proficiency on online participation in general (cf. Fig. 3).

The model specification presented here does not include interaction effects. So we cannot comment on what factors moderate the strength of relationship between each predictor and the dependent variable. Moreover, some predictors (and groups of predictors) are clearly correlated. Our model specification does not account for this collinearity explicitly; implicitly, collinearity causes noise and hampers the identification of statistically significant effects. In this sense, our analysis method is conservative. A reasonable next step would be to formulate hypotheses about multi-stage effects between (groups of) predictors. For example, media exposure and experience shape concerns, while only concerns ultimately affect behavior; all this is moderated by proficiency. Path models, like structural equation modeling (SEM), are the tools of choice to investigate such systems of hypotheses.

D. *Improving survey questions*

Overall, the survey authors are to be commended for the careful design of questions. Nonetheless, we observe several opportunities for improving future surveys. First, it would be very helpful to ask follow-up questions to victims of cybercrime in order to assess the impact of the crime. Asking victims to state how much money and time was lost due to the crimes would be a good start.

Second, the questions we used in the response variables could be improved with additional differentiation. For instance, in addition to asking people if they are "less likely" to participate online, one could add further differentiation to help quantify what "less likely" actually means. For example, one could ask if respondents intend to "avoid it completely", "use it half as often", etc.

Experience with cybercrime is already measured on an ordinal scale, but it remains unclear how participants interpret its categories. For example, being a victim of identity theft "often" (maybe once or twice a year) is certainly different from receiving spam messages "often" (several times a day). Here it seems feasible to use a quantitative scale without overloading the cognitive capacity of respondents. Another idea to get an anchor point for the interpretation of the scales is to include some questions about offline crimes, ideally those recorded accurately in official police statistics. Similarly, if the scales measuring concerns were anchored with a question on concerns about widespread offline crimes (e.g., burglary, car theft, or vandalism), policy makers would be better informed on the right balance of law enforcement online and offline.

Finally, some questions could be improved to clarify their interpretation. For example, the security proficiency indicators would become more informative by asking if a secure behavior (e.g., password change) was made on the respondent's own initiative or triggered externally, for instance by force of a password policy for work-related email. More importantly, some of the cybercrime-specific questions could be revised to clarify the crime being considered. For instance, in addition to asking if respondents have "received emails fraudulently asking for money", one could ask if respondents had actually responded and been defrauded as a result. One question we did not include in our analysis asked if respondents had ever recalled "not being able to access online services (e.g. banking services) because of cyber attacks". This is a suboptimal question for two reasons. First, denial-of-service attacks on online services are rare, especially online banking services. Second, consumers are not in a good position to ascribe the reason for a service being inaccessible. They can only observe that the site is down, and ascribing the outage to cybercrime would almost always be speculation.

VII. Conclusion

The Internet brings huge benefits to society, from improved global communications to more efficient means of commerce. Unfortunately, cybercrime poses a substantial threat to all these positive outcomes. In addition to the direct harm it can cause, cybercrime threatens to impose even greater opportunity costs by deterring online participation.

Using a secondary analysis of a very large and representative survey of European citizens, we empirically examine how several factors affect online participation in activities such as banking and shopping. The factors we examine include experiencing cybercrime, concerns over cybercrime, and exposure to news reports about cybercrime.

We find that concern and experience with cybercrime both reduce online participation, but that the effect is stronger for those expressing concern. This matters since falling victim is relatively rare, compared to those who express concerns but are not victimized. We also find evidence that experiencing one form of cybercrime sometimes spills over and inhibits participation in unrelated areas. In particular, victims of e-commerce fraud tend to reduce their online banking, even though the activities are unrelated. This suggests that the prevalence of cybercrime could pose a problem for all parties who rely on the Internet, not just the relatively limited areas where cybercrime occurs.

To the best of our knowledge, we have presented in this paper the first multivariate analysis of a representative sample on the effects of cybercrime, in particular on what factors drive the indirect effects of cybercrime. None of this would have been possible without sharing of microdata from the organization originally carrying out the survey. We hope that in future organizations will see the benefits of cooperation in sharing data on cybersecurity, so that we might improve our understanding and countermeasures.

Acknowledgment

The authors would like to thank the Eurobarometer team at the European Commission for making the data available.

References

[1] T. Moore, R. Clayton, and R. Anderson, "The economics of online crime," *Journal of Economic Perspectives*, vol. 23, no. 3, pp. 3–20, Summer 2009.

[2] European Commission, "Special Eurobarometer 390 Cyber security," 2012, http://ec.europa.eu/public_opinion/archives/ebs/ebs_390_en.pdf.

[3] D. Florêncio and C. Herley, "Evaluating a Trial Deployment of Password Re-Use for Phishing Prevention," in *eCrime Researchers Summit*, ser. ACM International Conference Proceeding Series, L. F. Cranor, Ed., vol. 269. ACM, 2007, pp. 26–36.

[4] R. Richardson, "2010/2011 CSI computer crime and security survey," 2011, http://reports.informationweek.com/abstract/21/7377/Security/research-2010-2011-csi-survey.html.

[5] Eurostat, "Information society statistics," http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Information_society_statistics, 2011.

[6] Symantec Corporation, "Norton 2011 cybercrime report," 2011, http://www.symantec.com/content/en/us/home_homeoffice/html/cybercrimereport/.

[7] D. A. F. Florêncio and C. Herley, "Sex, lies and cyber-crime survey," in *Workshop on the Economics of Information Security (WEIS)*, 2011.

[8] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. van Eeten, M. Levi, T. Moore, and S. Savage, "Measuring the cost of cybercrime," in *Workshop on the Economics of Information Security (WEIS)*, 2012.

[9] G. Loewenstein, E. Weber, C. Hsee, and N. Welch, "Risk as feelings," *Psychological Bulletin*, vol. 127, no. 2, pp. 267–286, 2001.

[10] A. Kosse, "Do newspaper articles on card fraud affect debit card usage?" Netherlands Central Bank, Research Department, DNB Working Papers 339, Mar. 2012. [Online]. Available: http://ideas.repec.org/p/dnb/dnbwpp/339.html

[11] T. Lumley, "Analysis of complex survey samples," *Journal of Statistical Software*, vol. 9, no. 1, pp. 1–19, 2004.

[12] P. Slovic, "Perception of risk," *Science*, vol. 236, no. 4799, pp. 280–285, 1987.

[13] R. Tourangeau, L. Rips, and K. Rasinski, *The Psychology of Survey Response*. Cambridge: University Press, 2000.

Appendix

TABLE III

Main regressions including country fixed effects

	Dependent variable: Less likely to . . .		
Predictor	Buy online	Bank online	Participate online
Experience: identity theft	0.166 (0.1275)	0.293* (0.1321)	−0.162 (0.1247)
Experience: phishing/fraud spam	−0.128 (0.0765)	0.135 (0.0856)	0.321*** (0.0707)
Experience: e-commerce fraud	0.240* (0.1106)	0.267* (0.1192)	0.055 (0.1040)
Personal concern: identity theft	0.042 (0.0902)	0.082 (0.1010)	0.243** (0.0786)
Personal concern: phishing/fraud spam	0.215* (0.0839)	0.215* (0.0921)	0.178* (0.0778)
Personal concern: e-commerce fraud	0.362*** (0.0846)	−0.023 (0.0917)	0.037 (0.0758)
General concern: security of online payments	0.558*** (0.0676)	0.752*** (0.0757)	0.502*** (0.0666)
General concern: misuse of personal data	0.231*** (0.0691)	0.437*** (0.0788)	0.753*** (0.0665)
Heard about cybercrime on television	0.110 (0.0914)	−0.123 (0.1019)	0.076 (0.0810)
Heard about cybercrime on radio	0.108 (0.0866)	0.121 (0.0972)	−0.068 (0.0805)
Read about cybercrime in the newspapers	−0.115 (0.0810)	−0.053 (0.0886)	0.098 (0.0762)
Read about cybercrime on the Internet	0.115 (0.0787)	0.196* (0.0859)	0.143 (0.0731)
Personal communication on cybercrime	0.097 (0.0783)	−0.056 (0.0868)	0.086 (0.0752)
Nothing heard about cybercrime	0.039 (0.1284)	−0.480** (0.1521)	−0.143 (0.1089)
Internet access more than once a day	−0.058 (0.0748)	−0.093 (0.0809)	−0.159* (0.0652)
Do online banking	−0.240** (0.0873)		
Do online shopping		−0.382*** (0.0923)	
Confidence about own Internet skills	−0.493*** (0.0811)	−0.738*** (0.0876)	0.068 (0.0698)
Informed about cybercrime risks	−0.104 (0.0730)	0.105 (0.0813)	−0.123 (0.0660)
Password changed in past 12 months	0.024 (0.0742)	0.010 (0.0828)	0.188** (0.0645)
Use different passwords	0.042 (0.0845)	0.149 (0.0978)	1.113*** (0.0922)
Have antivirus installed	−0.047 (0.0705)	−0.014 (0.0812)	1.444*** (0.0640)
Higher education	−0.206** (0.0716)	−0.263** (0.0809)	0.137* (0.0632)
Social status above median	−0.086 (0.0682)	0.041 (0.0779)	−0.024 (0.0627)
Country fixed effects	yes	yes	yes
N (unweighted)	17274	17274	17274
Nagelkerkes' pseudo-R^2	0.79	0.86	0.70

TABLE IV

ROBUSTNESS CHECK: ALTERNATIVE SPECIFICATION EXCLUDING COUNTRY FIXED EFFECTS

	Dependent variable: Less likely to …		
Predictor	Buy online	Bank online	Participate online
Experience: identity theft	0.217 (0.1227)	0.341** (0.1293)	−0.183 (0.1218)
Experience: phishing/fraud spam	−0.096 (0.0743)	0.166* (0.0832)	0.353*** (0.0685)
Experience: e-commerce fraud	0.156 (0.1076)	0.249* (0.1166)	0.004 (0.1009)
Personal concern: identity theft	0.066 (0.0881)	0.159 (0.0986)	0.226** (0.0769)
Personal concern: phishing/fraud spam	0.180* (0.0813)	0.174 (0.0901)	0.173* (0.0769)
Personal concern: e-commerce fraud	0.339*** (0.0823)	−0.002 (0.0887)	−0.009 (0.0743)
General concern: security of online payments	0.559*** (0.0656)	0.779*** (0.0723)	0.497*** (0.0643)
General concern: misuse of personal data	0.219** (0.0671)	0.416*** (0.0746)	0.758*** (0.0649)
Heard about cybercrime on television	0.159 (0.0891)	−0.143 (0.0987)	0.090 (0.0781)
Heard about cybercrime on radio	0.140 (0.0842)	0.091 (0.0944)	−0.017 (0.0788)
Read about cybercrime in the newspapers	−0.142 (0.0778)	−0.142 (0.0859)	0.120 (0.0732)
Read about cybercrime on the Internet	0.104 (0.0765)	0.183* (0.0838)	0.136 (0.0713)
Personal communication on cybercrime	0.083 (0.0766)	−0.076 (0.0847)	0.119 (0.0740)
Nothing heard about cybercrime	0.027 (0.1270)	−0.417** (0.1504)	−0.178 (0.1066)
Internet access more than once a day	−0.002 (0.0728)	−0.063 (0.0793)	−0.130* (0.0638)
Do online banking	−0.230** (0.0806)		
Do online shopping		−0.444*** (0.0867)	
Confidence about own Internet skills	−0.580*** (0.0795)	−0.774*** (0.0858)	0.059 (0.0667)
Informed about cybercrime risks	−0.107 (0.0709)	0.174* (0.0794)	−0.170** (0.0644)
Password changed in past 12 months	0.014 (0.0721)	0.034 (0.0806)	0.184** (0.0630)
Use different passwords	0.019 (0.0834)	0.131 (0.0950)	1.090*** (0.0905)
Have antivirus installed	−0.053 (0.0676)	−0.144 (0.0788)	1.515*** (0.0620)
Higher education	−0.154* (0.0685)	−0.275*** (0.0784)	0.168** (0.0611)
Social status above median	−0.091 (0.0646)	0.072 (0.0736)	−0.035 (0.0594)
(Intercept)	−1.687*** (0.1172)	−1.748*** (0.1393)	−1.222*** (0.1062)
N (unweighted)	17274	17274	17274
Nagelkerkes' pseudo-R^2	0.16	0.21	0.58