Joint Discussion Paper
Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 08-2016

Seo-Young Cho

A Crime 2.0 –
Cybercrime, e-Talent, and Institutions

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

MACIE PAPER SERIES

Marburg Centre for
Institutional Economics

MACIE

Nr. 2016/4

A Crime 2.0 –
Cybercrime, e-Talent, and Institutions

Seo-Young Cho
MACIE, Philipps-Universität Marburg

Marburg Centre for Institutional Economics • Coordination: Prof. Dr. Elisabeth Schulte
c/o Research Group Institutional Economics • Barfuessertor 2 • D-35037 Marburg

Phone: +49 (0) 6421-28-23196 • Fax: +49 (0) 6421-28-24858 •
www.uni-marburg.de/fb02/MACIE • macie@wiwi.uni-marburg.de

Philipps Universität Marburg

A Crime 2.0
- Cybercrime, e-Talent, and Institutions

Seo-Young Cho
(University of Marburg)

February 2016

Abstract:
Cybercrime is typically profiled as a skill-intensive crime committed by educated, young criminals. This observation raises the controversial question of whether advanced knowledge and skills are a pull factor of cybercrime. In this paper, the linkage between e-skills and cybercrime is investigated using statistics from up to 28 European countries. Through the investigation, it is shown that electronic skills induce more cybercrime under weak institutions where the rules of law do not provide protection and incentives for productive entrepreneurial activities. This compound effect between e-skills and institutions suggests that institutional factors are crucial to allocating human capital between productive and criminal activities in cyberspace.

Keywords: cybercrime; e-skills; institutions; entrepreneurship
JEL-codes: E24; K42; O15

Acknowledgements: The author thanks Lukas Stroemsdoerfer and Noelani Boise for their research assistance. Usual disclaimer applies.

* Contact: Research Group of Empirical Institutional Economics, School of Business and Economics, Philipps-University of Marburg. Barfuessertor 2, D-35037 Marburg, Germany.
Tel. 49 (0)6421-28-23996. Fax. 49 (0)6421-28-21740. Email. seo.cho@wiwi.uni-marburg.de
Web. www.uni-marburg.de/fb02/empinsti

1. Introduction

Today the internet is an essential tool for communication, information gathering, and other daily activities (e.g. banking, shopping, and networking). Accordingly, crimes in cyberspace deeply affect various aspects of daily life. The United Nations Office on Drugs and Crime (UNODC 2013) estimates that in the near future, almost 70% of all crimes will involve cyberspace to some extent via online operation and/or transactions leaving electronic traces of criminal activities.

Despite the fact that cybercrime has more presence in our society than ever before, the causes and consequences of cybercrime are rarely studied in economic research. Indeed, most studies investigating cybercrime have been conducted by information scientists (see Anderson et al. 2013; Böhme and Moore 2012; Moore et al. 2009). While these studies provide meaningful analysis of the nature of cybercrime with technological insights, it is necessary to view the problem through a socioeconomic perspective which takes into account various economic and institutional factors otherwise overlooked.

Cybercrime is a skill-intensive form of criminality committed by educated, skilled, young persons – so-called *'Yahoo boys'* (UNODC 2013; CSIS 2014). This observation leads to the question regarding whether a high level of IT skills is a pull factor of cybercrime. Indeed, it is a controversial issue revealing unexpected side effects of knowledge and skills. However, one should also note that e-skills are not always wasted in cybercrime but can also be utilized in more productive entrepreneurial activities. What then makes a difference in the utilization of e-skills? To respond to this question, the role of institutions must be analyzed.

According to Baumol (1990), talent ('entrepreneurship') can be used in either productive or destructive sectors and the allocation of talent between the two depends on institutions that determine incentive structures and the rules of the game. Following his argument, individuals with e-skills have two choices to invest their talent in either a productive IT sector or cybercrime. Which is selected depends on the pay-offs each sector offers. Under well-developed institutions where the outcomes of innovation and property rights are protected, talented individuals are incentivized to invest their skills in developing IT products and technologies for legal activities. However, if institutions are not well-established, such entrepreneurial activities are not rewarded properly as property protection is not ensured. In this case, skilled entrepreneurs would be more susceptible to taking part in underground cyber-activities that may offer immediate pay-offs.

With this argument in mind, the relationships between cybercrime, e-skills, and institutional quality are investigated in this paper and in particular, the role of institutions in allocating e-talent between productive and criminal activities is addressed. Investigating the determinants of cybercrime is an important issue because prevailing cybercrime forgoes benefits of online activities such as e-commerce, online banking, information sharing and participation via the internet (Böhme and Moore 2012). An analytical study of cybercrime would facilitate a systematic understanding of the nature of the problem. Such understanding can, in turn, contribute to minimizing losses in social welfare triggered by detrimental effects of cybercrime.

For an empirical analysis, the European Statistics (European Commission 2010) from up to 28 countries are employed in this paper. The results of the empirical investigation show that e-skills increase the prevalence of cybercrime, however, the positive effect of e-skills declines as a country improves its law and order and protection of property rights. Also, when institutional quality reaches the level of Finland (i.e. a score of approximately 2 measured by the Rule of Law Index, Kaufmann et al. 2010), the pulling effect of e-skills on cybercrime disappears. In other words, the better legal protection a country provides, the more e-skills are invested in legitimate sectors instead of illicit activities in cyberspace. Robustness of the findings is tested with respect to potential omitted variable biases and the choice of estimation model by employing an instrument (reading skills measured by PISA scores) and extending the model with different interaction terms and estimation techniques, respectively.

This paper is organized as follows. Section 2 presents a theoretical framework of modeling cybercrime, e-skills, and institutions and proposes testable working hypotheses. Data sources and empirical estimation models are described in section 3. In this section, identification issues in defining e-skills are also discussed. Section 4 shows the empirical results and studies the role of e-skills and institutions in cybercrime. In this empirical analysis, the endogeneity of the model is discussed and the results of an instrumental variable estimation are presented. Section 5 concludes.

2. Modeling for Cybercrime

2.1. Cybercrime, e-Skills, and Institutions

Cybercrime is often described as a crime of high skills, low risks and high returns (CSIS 2014). According to a study conducted by the UNODC (2013), the common profiles of cybercriminals are young men with relatively high levels of education and technical skills (‘*Yahoo Boys’*).[1] These profiles are similar to those of other 'white collar' criminals.

Such findings trigger a controversial issue that high education and advanced skills can be a pull factor of inducing cybercrime. However, IT skills are not always used for destructive activities in cyberspace; they are also invested in productive entrepreneurial fields. Furthermore, there are substantial differences in the prevalence of cybercrime across countries that cannot be fully explained by cross-country variations in education and e-skills. This observation leads to a surmise that there is another essential factor that determines the relationship between cybercrime and e-skills.

With this in mind, one should raise the following question, *in which conditions do skilled-individuals use their IT skills for cybercrime instead of legitimate activities?,* in order to identify what causes such differences. In other words, what leads e-skills to be invested in destructive activities in cyberspace? According to Becker (1968), the decision of committing a crime depends on the function of pay-offs for underground activities, probabilities for apprehension, the level of punishment, and opportunity costs (pay-offs in legitimate sectors). Many governmental and private sector reports point out that the probabilities of being caught are low and punishment is often lenient in many countries with respect to cybercrime. This is mainly because law enforcement is not yet well-established in this emerging type of crime and awareness of its significance is low (CSIS 2014; UNODC 2013). Given that punishment is taken into account in the cost function of (potential) cybercriminals with low probabilities, those with relevant e-skills would commit cybercrime if the payoff for underground activities exceeds that of legitimate

[1] On the other hand, technical skills that are required to commit cybercrime are not necessarily very high anymore because technical toolkits and manuals can be purchased. But, committing cybercrime certainly requires at least some high levels of skills to install, implement, and manipulate such programs.

sectors.[2] This implies that the lack of opportunities for productive and profitable economic activities pull skilled-individuals to be engaged in cybercrime (Kshetri 2010).

The environments in which the underground sector is more profitable for skilled-individuals are primarily determined by institutional conditions of a country. There is a seminal work by Baumol (1990) that studies the allocation of talent ('entrepreneurship')[3] between productive and destructive activities. This work updates the discussions of Schumpeter (1934) who viewed entrepreneruship as mainly productive. Baumol argues that not all entrepreneurial activites are necessarily productive; rather entrepreneurship is allocated into two types – productive and unproductive/destructive. This allocation is determined by institutional conditions that give incentives to a certain type of activities. For instance, if accumulating wealth via productive activity A is constrained in society (due to the lack of legitimate opportunities and/or complications in rules that regulate such an activity), talented invidiauals (entrepreneurs) may shift their talent (entrepreneurship) to an unproductive activity B that bypasses law – say an underground activity. Baumol provides historical evidence from ancient Rome where the accomulation of wealth by commerce was stigmatized, and confusionistic China where the wage level of public offcials was low despite their high education and competitiveness. Both cases led high profiled officials to renk-seeking and corruption – i.e. accoumulation of wealth via unproductive activities.

This argument can be further investigated to explaining the allocation of cyber-entreprenship. Individauls with relevant e-skills choose to engage in legitimate IT business or cybercrime, depending on expected pay-offs. As previously stated, institutional conditions are a significant factor in shaping the pay-off function. Iinstitutions that offer opportunities and incentives for innovation and protect property rights of such outcomes encourage legitimate entrepreneurial activities. On the other hand, under weak institutions where entrepreneurial efforts are not properly rewarded and individuals with relevant IT skills do not find opportunities to utilize their

[2] In addition, there are psychological factors such as feeling of guilt and risk aversion factors that are not taken into account in this argument. These factors vary significantly across individuals and are widely unobservable.

[3] Baumol (1990) defined 'entrepreneurs' as persons who are capable and creative in finding ways that add to their own wealth, power, and/or prestige.

talent for innovation, talent is diverted to unproductive, underground cyber-activities which offer more immediate pay-offs.

This framework offers an explanation on cross-country differences in the prevalence of cybercrime. In some countries, the human capital of e-skills is invested in the legitimate digital economy because institutions promote productive entrepreneurial activities, while in some other countries with a similar level of e-human capital, this talent is wasted in the shadow economy and criminal activities due to weak institutions. The difference in the utilization of e-skills arises from different institutions that offer dissimilar incentives.

Anecdotal evidence also supports the compound interaction effect between e-skills and institutional quality. A number of studies (Krebs 2014; Symantec 2015; UNODC 2013) name Eastern Europe, Russia, China and some other Southeast Asian countries as main hubs of cybercrime. These countries are abundant with individuals with high levels of e-skills, but have weak institutions that fail to protect property rights and undermine innovation. Particularly, some of these countries lack opportunities in productive industries where e-skills can be invested, and furthermore, organized criminal groups are present in cybercriminal operations on a large scale, reflecting flaws in institutional protection.

2.2.Working Hypotheses

As discussed above, the prevalence of cybercrime can be explained by interacting effects of e-skills and institutional quality. In this section, several working hypotheses are proposed that can be empirically tested. To do so, the function of cybercrime is modeled below in order to formulate the relationships between cybercrime, e-skills, and institutions.

$$\text{Cybercrime} = f\ (\text{e-skills, Institution, } X)$$

, with vector X: control factors that determine cybercrime

The main focus of the investigation is the (net) effects of e-skills and institutions on cybercrime. IT skills are expected to have a positive effect on cybercrime because electronic skills and technological knowledge are a necessary condition for committing such a crime. However, as discussed in section 2.1, the investment of e-talent in unproductive activities is determined by the quality of institutions that govern the rules of the game and incentives. Thus, e-skills combined with weak institutions where underground activities can flourish create a sufficient condition for cybercrime to progress. Thus, a net effect of e-skills is conditional on the quality of institutions of a country. There are additional factors that potentially influence cybercriminal activities – such as internet infrastructures, employment conditions, and economic, demographic and cultural conditions of a country. These factors are listed in vector X.

With the arguments of e-skills and institutions in mind, the following hypotheses are derived. First, a higher level of e-skills increases the prevalence of cybercrime, given that having relevant e-skills is required to commit cybercrime.

> H_0: The prevalence of cybercrime increases in the level of e-skills.
>
> First order derivative 1: $\frac{d\ \text{Cybercime}}{d\ \text{e−skills}} = f'_{\text{e-skills}} > 0$

Second, better institutions that reward productive entrepreneurial activities constrain criminal activities in cyberspace. Institutions can constrain cybercrime through two channels: (i) protecting property rights and innovation that increases the benefits of productive activities; and (ii) implementing law enforcement against crime that raises risk costs of (potential) cybercriminals.

> H_0: The prevalence of cybercrime decreases in the level of institutional quality.
>
> First order derivative 2: $\frac{d\ \text{Cybercime}}{d\ \text{Institution}} = f'_{\text{Institution}} < 0$

As discussed in section 2.1, the allocation of e-skills between productive and unproductive entrepreneurial activities depends on institutions. This leads to a third hypothesis below that formulates a sufficient condition for cybercrime.

> H_0: The marginal effect of e-skills on cybercrime decreases as institutional quality improves.
>
> Second order derivative: $\frac{d^2\text{Cybercime}}{d\ e{-}skills^{\ 2}} \mid \text{Institution} = f''_{\text{e-skills}} \mid \text{Institution} < 0$

This theoretically-presumed concave function of the marginal effect of e-skills explicitly assumes that e-talent is shifted towards more productive cyber-activities as institutional quality improves.

3. Research Design

3.1.Data

According to the European Commission (2012), cybercrime is defined as any crime committed via the internet. A study of the UNODC (2013) further specifies this broad definition of cybercrime by categorizing criminal acts using computers into three types. They are, namely, (i) acts against the confidentiality, integrity, and availability of computer data or systems – such as data breaches; (ii) computer-related acts for personal or financial gain or harm – such as spam emails and identity fraud in cyberspace; and (iii) computer content-related acts – such as child pornography and hatred speeches. Based on these classifications of cybercrime, data on the following six major dimensions of cybercrime that was collected from 28 European countries (European Commission 2010) is employed for the empirical analysis.

(i) A virus or other computer infection (e.g. worm or Trojan horse) resulting in loss of information or time.
(ii) Unsolicited emails sent to individual ('spam') which resulted in financial or personal harm.
(iii) Abuse of personal information sent on the internet and/or other privacy violations (e.g. abuse of pictures, videos, or personal data uploaded on community websites).
(iv) Financial loss as a result of receiving fraudulent messages ('phishing') or being redirected to fake websites asking for personal information ('pharming').
(v) Financial loss due to fraudulent payment (credit or debit) card use.
(vi) Children accessing inappropriate websites or connecting with potentially dangerous persons from a computer within the household.

The cybercrime statistics used in this paper are the percent of individuals in a country[4] who experienced one of the above six types of cybercrime in the past 12 months.

As for the measurement of e-skills, the European Statistics (European Commission 2010) that captures the percent of individuals who have a high level of proficiency in internet applications are used. Internet skills of individuals can reflect both perspectives: consumer and business competencies. On the one hand, the level of e-proficiency among consumers of internet services indicates how well users can handle problems in cyberspace (Böhme and Moore 2012). Therefore, the high proficiency of consumers can reduce the probability of falling victim to cybercrime. On the other hand, internet skills also implicate the e-talent of IT professionals that can be used for criminal businesses in cyberspace. High e-skills, in this case, are expected to increase cybercrime.

To estimate the effects of e-talent invested in cybercrime, a measurement that captures e-skills from a business perspective is needed. Accordingly, an indicator of internet skills that are required to operate services in cyberspace is selected: the percent of individuals who have experience in creating a website. This indicator is a more relevant proxy for business proficiency than other indicators assessing basic skills such as abilities to use search engines and attach files in email that are more related to consumers' competency. This chosen indicator of web operation skills measures the highest internet competency among the available indicators in the European Statistics. Additionally, any potentially remaining effects of e-skills driven from the consumers' side are further addressed by controlling for users' awareness on cybercrime and internet accessibility. Further details of these variables are explained in section 3.2.

To measure the quality of institutions that determine incentives for legal and illegal activities, the Rule of Law Index – taken from the World Governance Indicator (Kaufmann et al. 2010) – is used. The Index evaluates the institutions of a country, specifically with respect to property rights protection, contract enforcement, and legal protection against crime.[5] These evaluation criteria are directly relevant to the allocation of incentives between legal and illegal activities. The scores

[4] The statistics represent the percent of affected individuals among those living in urban areas – i.e. densely-populated area (at least 500 inhabitants /Km²).

[5] According to Kaufmann et al. (2010), the Rule of Law Index "*reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence"*.

of the Rule of Index range from approximately –2.5 (weakest) to +2.5 (strongest governance performance). In the sample of the 28 European countries, the range is between –1 and +1.98. The descriptive statistics of all variables used in this paper are presented in appendix 1.

Lastly, there is a remaining issue in data collection to be settled: whether employing data that measures the domestic levels of e-skills and institutions is an adequate choice for an analysis of cybercrime with a transnational nature. Cybercrime is widely operated on a global scale, but a large part of its operations – such as spreading spam emails and sending fraudulent messages – takes place locally. This is because such tasks require local languages and specific knowledge about local practice and societal establishments. Thus, cybercriminals who carry out operational tasks are locally hired, even if the recruiter is part of a globally organized criminal group which plays a role as an umbrella organization for local criminals. This pattern is similar to other transnational crimes such as illegal drug trade (UNODC 2013). In recruiting local criminals with e-skills and operating cybercriminal acts, countries with weaker institutions are targeted since skilled individuals are more susceptible to taking offers from criminal enterprises, and law enforcement against crime is weak in these countries. With this aspect of cybercriminal operation in mind, local e-skills and domestic institutions are relevant to the explanation of country-level variations in the prevalence of cybercrime.

3.2. Empirical Model

The hypothesized relationships between cybercrime, e-skills and institutions in section 2.2 are formally modeled below (see equation 1) and tested empirically through regression estimations using the European Statistics.

$$\text{cybercrime}_i = \beta_1\text{e-skill}_i + \beta_2\text{institution}_i + \beta_3\text{e-skill}_i{*}\text{institution}_i + X_i{}^{\prime}\text{Ɯ} + u_i \qquad (1)$$

The dependent variable *cybercrime* captures the percent of individuals who experienced one of the six types of cybercrime described in section 3.1 in the past 12 months. One of the main independent variables analyzed is *e-skill*, which represents the percent of individuals with the experience of website creation. This proxy is selected in order to reflect e-skills that are high enough for cybercriminal activities, instead of simpler skills required for ordinary internet users (see discussions in section 3.1). *Institution* is another independent variable of focal interest that measures the level of the rules of law of a country – in particular, property rights protection and contract enforcement that provide important institutional environments to facilitate entrepreneurial and innovative activities. In addition, the interaction term of e-skills and institutions is included in the model in order to identify whether more e-skills are employed in cybercriminal activities in poorer institutional environments than in better institutions. This interaction variable is used to estimate the potentially different effects of e-skills across different levels of institutional quality.

Vector X comprises other economic, technological, developmental and demographic factors that are potentially important in explaining the prevalence of cybercrime. Since being connected via the internet is a necessary condition for participating in cyberspace, internet infrastructures (percent of individuals who have access to the internet at home) are taken into account. Besides, unemployment rates among university graduates aged 25 to 34 are controlled for. This demographic group (young and educated) represents the typical profiles of cybercriminals whose lack of opportunities to be legitimately employed increases their probabilities of committing cybercrime. Furthermore, a country's endowments of technical innovation – measured by the number of patents certified in a country – are taken into account. The patent variable reflects the culture and practice of innovation in the scientific community that can have constraining effects on cybercriminal activities. In addition, people's awareness and concerns about cybercrime is also controlled for in order to find whether perceived risks predict real occurrences of crime. The awareness and concern variable also captures cultural and psychological factors regarding how people in a country are anxious about crime. Also, the levels of economic wealth (*income*) and human capital (*tertiary education*) are included in the model because these factors affect the payoff functions of criminal acts. Additionally, a demographic factor – population sizes – is included as larger countries tend to have higher rates of criminal activities. Other uncorrelated unobserved factors are denoted in equation 1 as u_i – an idiosyncratic error term.

To estimate the model, both linear and non-linear regression methods are employed. First, as the dependent variable is a round percent (because the European Statistics provide round numbers only), these numbers are treated as integers (e.g. 27, 30, 45) and thus, a negative binomial estimation method is applied accordingly. Second, the model is transformed into a log-linear model by taking the logarithm of the dependent variable and estimated by an ordinary least squares method. This log-linear estimation provides two advantages: securing the normal distribution of the dependent variable and a more straightforward interpretation of the magnitudes of coefficients. By applying both linear and non-linear estimations, one can test whether results are robust to the choice of estimation methods or driven by (changes in) linearity assumptions. In order to account for potential heteroscedasticity, robust standard errors clustered at the country level are applied. Due to data availability, cross-section data without time-variations (using the information from 2010) is exploited for the estimations. Potential problems caused by omitted variables (unobserved country heterogeneity) are further addressed in section 4.3 by employing an instrumental variable method.

3.3. Identification of e-Skills and Extended Model

As discussed in section 3.1, an adequate choice of the measurement of e-skills that captures professional proficiency is crucial to single out the effect of entrepreneurial e-talent on cybercrime. To distinguish this effect from other effects created by the consumers' side, a variable measuring a high level of e-proficiency – website creation and management – is selected.

However, it is still possible under this specification for the chosen measurement of e-skills to partially account for consumers' competencies (as the skill level of ordinary users has improved since the internet has become a daily working and communicational tool for many people). This would result in an underestimation of the (presumably) positive effect of professional e-talent on cybercrime. To account for these potential overlaps in e-skills between experts and ordinary users, an alternative model that further distinguishes the e-skills of potential cybercriminals is proposed below.

$$\begin{aligned} \text{cybercrime}_i = {} & \alpha_1\text{e-skill}_i + \alpha_2\text{institution}_i + \alpha_3\text{unemployment}_i + \alpha_4\text{e-skill}_i{*}\text{institution}_i \\ & + \alpha_5\text{e-skill}_i{*}\text{unemployment}_i + \alpha_6\text{unemployment}_i{*}\text{institution}_i \\ & + \alpha_7\text{e-skill}_i{*}\text{institution}_i{*}\text{unemployment}_i + X_i{}'\Psi + u'_i \qquad (2) \end{aligned}$$

In this specification, three additional interaction terms are newly included: (i) e-skills and unemployment rates among university graduates aged 25 to 34; (ii) unemployment rates and institutions; and (iii) e-skills, unemployment rates, and institutions. The interaction term between e-skills and unemployment specifically accounts for the effect of e-proficiency of those who are young and well-educated but lack opportunities in legal sectors – i.e. individuals whose profiles are similar to those of typical cybercriminals. Furthermore, the interaction term of e-skills, unemployment and institutions is taken into account to identify whether better institutions can constrain educated but unemployed young people from using their e-skills for cybercrime. Lastly, the interaction between unemployment and institutions is incorporated into the equation to find any additional effect of institutions constraining the young, educated unemployed from committing cybercrime.

These interaction effects may cause the equation to be excessively rigid because this extended model tries to identify whether institutions can constrain the unemployed – who are already deprived of opportunities in the legal sectors – from using their skills for cybercrime. In fact, the baseline specification (equation 1) that is modeled to find in which institutional conditions e-skills are invested in legal or illegal sectors addresses the conceptual arguments of this paper more adequately (see discussions in section 2). However, in this alternative specification, the term, e-talent of potential cybercriminals, is more strictly defined and tested in order to find any additional evidence supporting the role of institutions in constraining cybercrime.

4. Results

4.1.Baseline Results

The descriptive statistics presented in figures 1 and 2 show tentative evidence regarding the relationships between e-skills and cybercrime (figure 1) and institutions and cybercrime (figure 2). The binary correlation between e-skills and cybercrime suggests a clear positive association, while the correlation between institutions and cybercrime does not indicate a specific direction.

In order to investigate these relationships systematically, a regression analysis is employed. Table 1 shows the results of the baseline estimations. Columns 1–3 present the results estimated by a negative binomial regression and columns 4–6 by a log-linear estimation. As many variables are presumably correlated to each other, the regression is first run with five major factors (e-skills, rules, unemployment rates, internet connection, and patents). More variables are added later in order to identify the effect of each variable. Columns 1 and 4 show the results of the regressions with the five variables; columns 2 and 5 include the interaction of e-skills and institutions; and columns 3 and 6 additionally control for economic, developmental, risk, and demographic factors.

The coefficient of the e-skill variable is statistically insignificant without its interaction with institutions (*rule*), but becomes positive with the 1–5% level of significance when the interaction term is included. This implies that the effect of e-skills becomes meaningful when its relationship to institutional quality is accounted for.

The compound effect of e-skills and institutions can be more clearly understood when one looks at the interaction term. The interaction of the two variables captures a partial effect of e-skills in different levels of institutional quality for a representative country. In all regression models presented in table 1, the coefficient of the interaction term is negative with the 1–5% level of significance. Its magnitude becomes larger after including further control variables (columns 3 and 6). The positive effect of e-skills together with the negative effect of the interaction indicates that the function of cybercrime is increasing in e-skills but the size of the effect of e-skills declines as institutional quality improves. For example, increasing e-skills in the Netherlands (with a score of the Rule of Law Index being 1.81) or the United Kingdom (1.76) increases the prevalence of cybercrime by a lesser amount than that of Italy (0.38) or Greece (0.61). Furthermore, the positive effect of e-skills loses its statistical significance when the institutional quality of a country reaches a very high level – a score around 2 (that is calculated based on the results presented in column 6). This means that in countries like Finland (1.98) and Sweden (1.96), increasing e-skills hardly exacerbates the prevalence of cybercrime. The decreasing marginal effect of e-skills in the level of institutional quality is graphically illustrated in figure 3 that presents a clear downward direction. On the other hand, the coefficient of the *rule* variable itself is mostly insignificant. This result suggests that institutional quality alone without accounting for a country's e-talent is not an important determinant of cybercrime.

Turning to the findings of the other control factors, higher internet connectivity is positively associated with cybercrime. This is because e-infrastructures increase not only cyber activities per se but also the pool of victims of cybercrime. Quantitatively, a 1%-point increase in the share of individuals with internet connection induces additional cybercrime by a 3%-point. Also, concerns about cybercrime – public awareness about cyber-privacy infringement – are positively associated with the prevalence of cybercrime. This effect is not yet conclusive because it is significant in the non-linear model (column 3) but not in the log-linear model (column 6). Still, the results tentatively suggest that perceived risks do indeed reflect actual occurrences, insofar as cybercrime concerns. In addition, income level and high education (the group of individuals with a completion of tertiary education) increase cybercrime – an outcome somewhat different from their effects on traditional crimes (such as burglary or robbery). This finding implies that cybercrime is a crime of wealth and development.

On the other hand, the effect of unemployment rates of university graduates aged 25 to 34 is not significant, contrary to the theoretical expectations. This is possibly because this is an aggregate measurement capturing not only the unemployment rates of science and engineering graduates but also other social science and humanity majors. The unemployment rates of science and engineering graduates are a more relevant measurement for potential cybercriminals but these data are currently unavailable in the European Statistics.

Last, the effect of (log) patents, used as a proxy for the culture of innovation and technical infrastructures, is ambiguous, as the coefficient is mostly insignificant. The ambiguity likely arises from two opposing aspects that the level of patents captures; one measuring technological endowments of a country that may pull more cybercrime, and the other reflecting the culture of positive entrepreneurship that would constrain cybercrime. By combining the two effects – each of which leads in the opposite direction of the other –, the effect of patents may have been muted.

4.2.Results of the Extended Model

As discussed in section 3.3, the baseline model is extended in order to further distinguish the effect of e-skills of potential cybercriminals from those of ordinary users. To do so, three additional interaction terms are newly included in the extended model: (i) interaction between e-skills and unemployment rates of young, educated individuals; (ii) interaction between unemployment rates of this group and institutional quality; and (iii) interaction between e-skills, unemployment rates, and institutional quality. Columns 1 and 3 present the results of the regressions including (i), and columns 2 and 4 include all three interaction terms.

The results of this extended regression analysis presented in table 2 further confirm the findings of the baseline estimations. The coefficient of e-skills is still positive and significant when including all three additional interaction terms (see columns 2 and 4). The interaction of e-skills and institutions is negative and significant to a large extent, except in column 4. Seemingly, the constraining effect of institutions remains robust when interaction with unemployment is additionally accounted for.

In regards to the newly included interaction effects, there is some evidence that e-skills of young, educated unemployed individuals (*unemployment* × *e-skill*) further contribute to the positive effect of e-skills (see columns 1 and 3). However, when controlling for the two additional interactions (*unemployment* × *rule* and *e-skill* × *unemployment* × *rule*), the effect of *unemployment* × *e-skill* loses its statistical significance. This is probably because this variable has high multi-collinearity with the other interaction terms, causing it to absorb its own variations.

Interestingly, the coefficient of the *unemployment* variable has a negative sign with the 1% level of statistical significance (columns 1 and 3), which contradicts the theoretical expectations. However, the effect of unemployment has to be interpreted together with the interaction term instead of looking into the effect of the single variable alone. By taking the results of the log-linear estimation (column 3), one can calculate the quantitative effect of unemployment in relation to the level of e-skills, since the size of its effect depends on the level of e-skills. Specifically, when the share of individuals with relevant e-skills in a country is lower than 11.3%, the unemployment of young, educated individuals constrains cybercrime. However, when the share of those with the e-skills reaches 11.3% or higher in the respective country, a higher level of unemployment exacerbates cybercrime. This result suggests that there is a threshold level of e-skills (11.3% in this case) that turns young, educated, unemployed people into illicit activities in cyberspace, possibly because boosting cybercrime requires a certain level of established e-endowments that can be pulled out.

When the two additional interaction terms of *unemployment* × *rule* and *e-skill* × *unemployment* × *rule* are further included, all three newly added interaction effects in the extended model lose their statistical significance (see columns 2 and 4). This result may have arisen due to high multi-collinearity of the interaction variables. Also, the inclusion of the interaction between unemployment and institutions (i.e. *unemployment* × *rule* and *e-skill* × *unemployment* × *rule*) may have imposed an overly rigid assumption that better institutions can constrain the unemployed from committing illegal cyber activities, despite the fact that they are already excluded from legal employment opportunities and thus the choice of legitimate jobs is currently unavailable to them.

Overall, the results of the extended analysis vindicate that e-skills increases cybercrime but having better institutions in a country constrains e-skills from being invested in cybercriminal activities. Furthermore, there is tentative evidence that e-skills of the young and educated unemployed (the pool of potential cybercriminals) contribute to the exacerbation of cybercrime. This finding supports the main argument that the lack of opportunities pushes e-skills to be wasted in criminal operations. On the other hand, better institutions do not necessarily prevent those who are already out of legal markets from committing cybercrime. Through this finding, one can surmise that the institutional role of crime prevention cannot be separated from its economic role of promoting opportunities in the labor markets.

Furthermore, the effects of the other control variables in the extended model widely support the findings of the baseline estimations, with a more strongly pronounced positive effect of perceived risks (the *concern* variable).

4.3.Endogeneity Concern: An Instrumental Variable Approach

In the baseline and extended models above, the determinants of cybercrime are controlled for as much as possible. However, unless the controls are totally exhaustive, unobserved heterogeneity may still be a remaining issue. In other words, unobserved heterogeneity that affects the prevalence of cybercrime in a country could also influence the level of e-skills of that country. For example, in the models above, technical infrastructures and culture of innovation are taken into account via the inclusion of internet connection and patent variables, respectively. However, these variables may not be perfect proxies for infrastructure and innovation factors and therefore be subject to measurement errors that lead to omitted variable biases, as these factors likely affect both cybercrime and e-skill levels.

To circumvent such unobserved heterogeneity and account for the endogeneity of e-skills, the use of an instrumental variable is a more viable strategy. Therefore, an instrument that reflects the level of e-skills but does not necessarily explain cybercrime directly is employed. The instrument chosen for e-skills is *reading skills*, measured by the PISA test (Programme for International Student Assessment, OECD 2006). The PISA evaluates the knowledge and skills of 15 year-old students in reading, mathematics and science. While each section measures students' competencies with its own literacy assessment, all three sections require general aptitudes and skills of learning and understanding, meaning that the evaluation outcomes are positively correlated to each other. As e-skills are typically based on mathematical and scientific knowledge, PISA mathematics and science scores can be a direct indicator of e-skills. However, they are also very likely correlated with the prevalence of cybercrime, given that mathematics and science education are prerequisites for computer science. On the other hand, reading skills are not directly required to commit cybercrime, while skill levels in reading are likely highly correlated with e-skill levels (i.e. countries with high reading skills tend to have high competencies in mathematics and science, as well). With this in consideration, reading skills are the most likely variable among the three PISA measurements of competencies to satisfy the exclusion criteria as an instrument. Therefore, the PISA reading scores – specifically, the percent of students who achieved the 4th and 5th quintiles of the scale (0–600) – are exploited for the instrumental variable estimations. While the PISA test evaluates knowledge and skills of students at the age of 15, this age group is nearing the completion of obligatory schooling in most countries and therefore can represent the competency levels of the general population.

Table 3 presents the results of tests which vindicate the selection of the instrument. The first-stage regression provides evidence that the PISA reading scores explain the level of e-skills with statistical significance at the 5%-level. Quantitatively, by increasing reading scores by a 1%-point, e-skills are improved by a 0.55%-point. Also, the F-statistics are almost 16, supporting the joint significance of the controls. Furthermore, the test for exclusion criteria (presented below the results of the first stage regression in table 3) shows that the instrument is correctly excluded in the second stage. The coefficient of the instrument is statistically insignificant when including both e-skills and the instrument, but gains significance after excluding the e-skills. This result suggests that reading skills explain cybercrime only via their linkage with the endogenous e-skill variable.

Turning to the second stage regressions, the results confirm the findings of the baseline estimations presented in table 1. There is some evidence that e-skills increase the prevalence of cybercrime (see column 2 in table 3, the log-linear estimation). On the other hand, the effect of institutions is statistically insignificant. Most importantly, the interaction between e-skills and institutions has a negative effect on cybercrime that is significant in both non-linear and log-linear models. The quantitative magnitude of the compound effect is similar to that of the baseline results. Namely, e-skills increase cybercrime until institutional quality, measured by the Rule of Law Index, reaches a score of approximately 2. Once past this threshold, e-skills do not increase cybercrime anymore. The results of the other controls are also highly similar to the baseline estimations. Cybercrime has a positive relation with e-infrastructures (*internet connection*), perceived risks (*concerns*), and wealth (*income*). However, after controlling for the endogeneity of e-skills, the coefficient of tertiary education loses its statistical significance.

5. Conclusion

As cybercrime emerges as a prevailing crime that affects our daily life, it is crucial to identify under which socioeconomic conditions skilled-individuals employ their e-talent in crime instead of productive entrepreneurial activities. This paper provides empirical evidence supporting the importance of institutions in constraining e-skills from being wasted in underground cyber-activities. The findings of this study reassert the valuable contribution of the scholarship of new institutional economics – good institutions and institutional protection are a prerequisite for development and innovation.

References

Anderson, Ross, Chris Barton, Rainer Böhme, Richard Clayton, Michel J. G. van Eeten, Michael Levi, Tyler Moore, Stefan Savage. 2013. Measuring the Cost of Cybercrime, in *The Economics of Information Security and Privacy* (ed. Rainer Böhme), pp. 265-300. Springer: Berlin Heidelberg

Baumol, William. 1990. Entrepreneurship: Productive, Unproductive and Destructive. *Journal of Political Economy* 98: 893-921.

Becker, Gary. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76: 169–217.

Böhme, Rainer and Tyler Moore. 2012. How do Consumers React to Cybercrime? *The Proceedings of the 7th APWG eCrime Researchers Summit (eCrime).*

Center for Strategic and International Studies (CSIS). 2014. Net Losses: Estimating the Global Cost of Cybercrime. Intel Security: Santa Clara (CA).

European Commission. 2010. European Statistics (EuroStat). http://ec.europa.eu/eurostat/data/database

European Commission. 2012. Cybersecurity Report. Special Eurobarometer 390. Brussels.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2010. The Worldwide Governance Indicators: A Summary of Methodology, Data and Analytical Issues. World Bank Policy Research Working Paper No. 5430. Washington D.C.

Kshetri, Nir. 2010. Simple Economics of Cybercrime and the Vicious Circle (chapter 2), in *The Global Cybercrime Industry* (N. Sshetri). Springer-Verlag: Berlin and Heidelberg.

Krebs, Brian. 2014. *Spam Nation. The Inside Story of Organized Cybercrime – from Global Epidemic to Your Front Door.* Sourcebooks: Naperville.

Moore, Tyler, Richard Clayton, and Ross Anderson. 2009. The Economics of Online Crime. *Journal of Economic Perspectives* 23(3): 3-20.

Organisation for Economic Co-operation and Development (OECD). 2007. PISA 2006 Science Competencies for Tomorrow's World. OECD: Paris.

Schumpeter, Joseph. 1934. *The Theory of Economic Development*. Harvard University Press: Cambridge (MA).

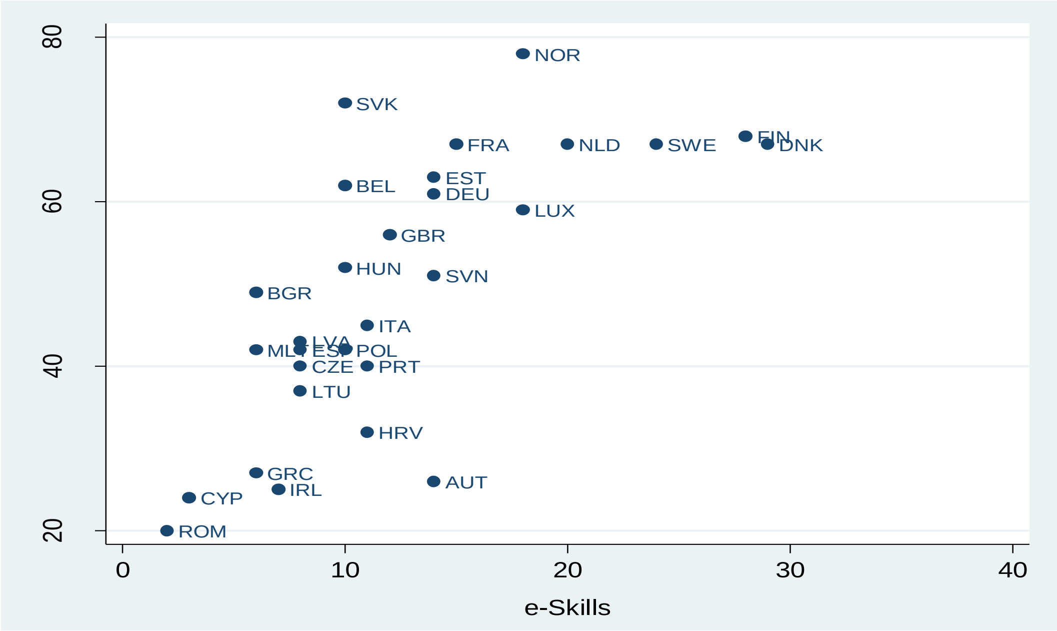
Symantec. 2015. Internet Security Threat Report Vol. 20. Symantec Corporation: Mountain View (CA).

United Nations Office on Drugs and Crime (UNODC). 2013. Comprehensive Study on Cybercrime. Vienna.

World Bank. 2010. World Development Indicator. http://databank.worldbank.org/data/home.aspx

Figure 1. Binary Correlation between Cybercrime and e-Skills

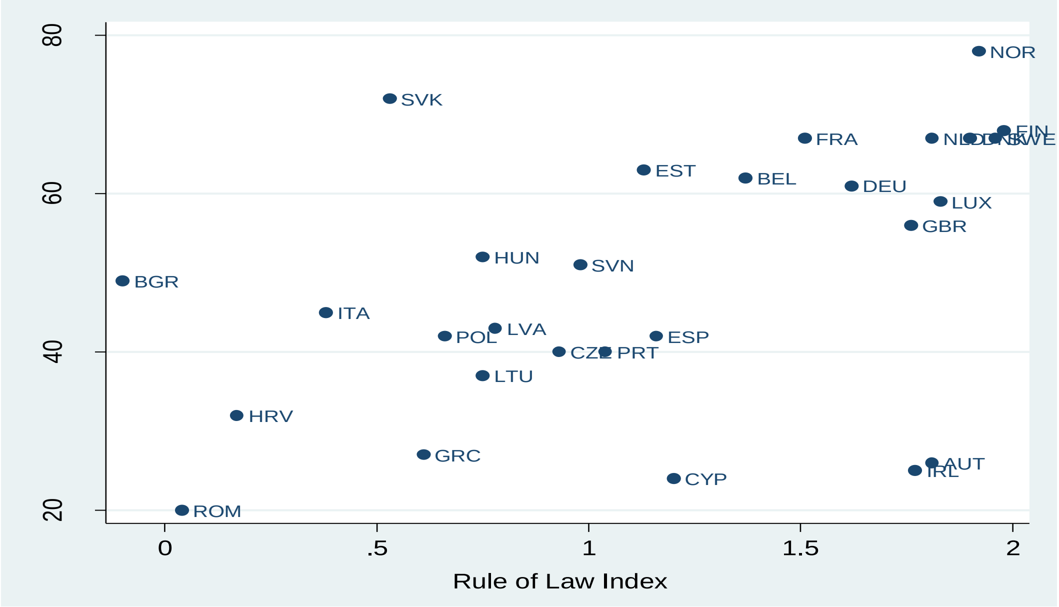
(28 European Countries, 2010)



Source: European Commission (2010)

Figure 2. Binary Correlation between Cybercrime and Institutional Quality

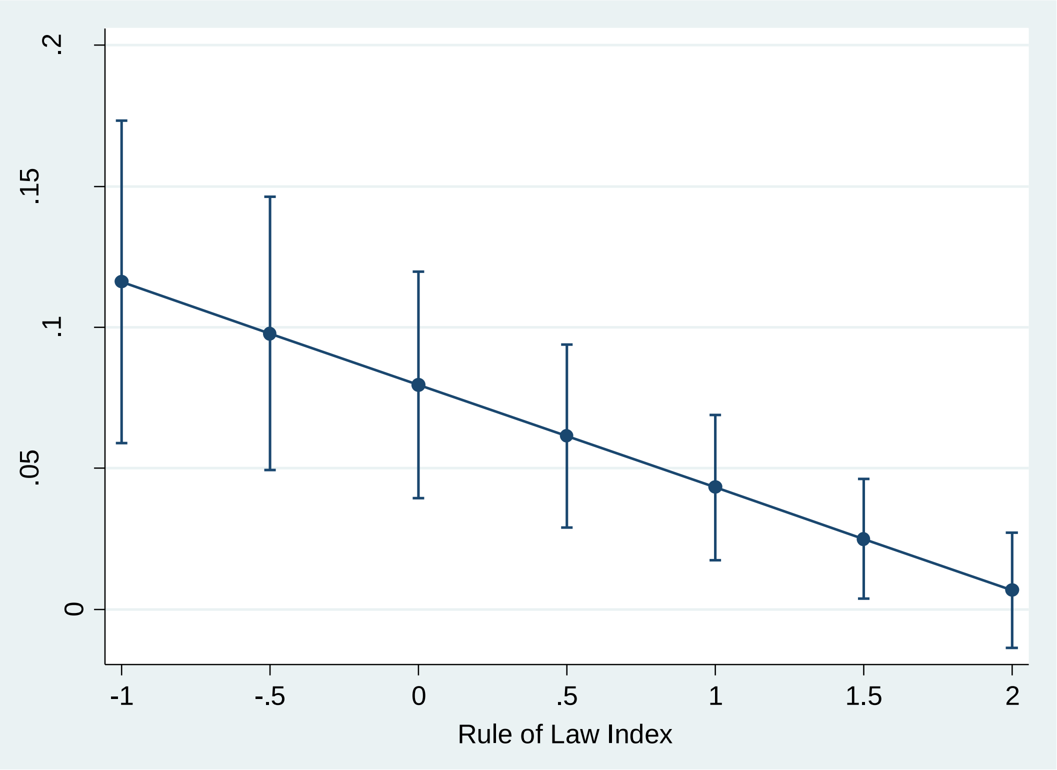
(28 European Countries, 2010)



Source: European Commission (2010) and Kaufmann et al. (2010)

Figure 3. Average Marginal Effects of e-Skills in Levels of Institutional Quality

(with 95% Confidence Interval)



Note: this figure corresponds to the log-linear regression results presented in column 6 in table 1.

Table 1. Baseline Model: Cybercrime, e-Skills, and Institutions (28 European countries, 2010)

	DV: Aggregate Cybercrime					
	Negative Binomial Regression			Log-linear Regression		
	(1)	(2)	(3)	(4)	(5)	(6)
e-Skill	0.01	0.07	0.078	0.02	0.08	0.08
	(0.01)	(0.025)***	(0.015)***	(0.013)	(0.03)**	(0.02)***
Rule	-0.29	-0.01	-0.07	-0.28	0.01	-0.10
	(0.14)**	(0.18)	(0.15)	(0.17)	(0.21)	(0.20)
e-Skill*Rule		-0.033	-0.04		-0.03	-0.04
		(0.013)**	(0.008)***		(0.015)**	(0.009)***
Unemployment	-0.01	0.01	-0.01	-0.01	-0.01	-0.01
(Univ, 25-34)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Internet Connection	0.02	0.02	0.024	0.02	0.02	0.03
	(0.007)***	(0.007)***	(0.006)***	(0.01)**	(0.01)**	(0.01)***
(log) Patents	0.02	0.02	-0.09	0.02	0.01	-0.12
	(0.02)	(0.012)	(0.06)	(0.02)	(0.02)	(0.09)
Concern-privacy			0.09			0.10
			(0.04)**			(0.06)
Tertiary Education			0.004			0.004
			(0.002)*			(0.003)
(log) Income			0.11			0.11
			(0.06)*			(0.08)
(log) Population			0.10			0.13
			(0.09)			(0.13)
Countries	28	28	26	28	28	26
Pseudo R^2	0.13	0.16	0.20	0.26	0.23	0.15

Note: Robust standard errors are clustered at the country level and presented in the parenthesis. * *p*<.10, ** *p*<.05, *** *p*<.01.

Table 2. Extended Model: Cybercrime, e-Skills, and Institutions (26 European countries, 2010)

DV: Aggregate Cybercrime				
	Negative Binomial Regression		Log-linear Regression	
	(1)	(2)	(3)	(4)
e-Skill	0.02	0.08	0.01	0.10
	(0.02)	(0.04)**	(0.03)	(0.058)*
Rule	-0.16	-0.59	-0.21	-0.56
	(0.15)	(0.34)*	(0.22)	(0.51)
e-Skill*Rule	-0.03	-0.04	-0.03	-0.055
	(0.01)***	(0.02)*	(0.01)***	(0.03)
Unemployment	-0.08	-0.08	-0.09	-0.06
(Univ, 25-34)	(0.02)***	(0.05)	(0.03)***	(0.08)
Unemployment*Rule		0.07		0.06
		(0.05)		(0.08)
e-Skill*Unemployment	0.007	-0.001	0.008	-0.004
	(0.002)***	(0.005)	(0.003)**	(0.008)
e-Skill*Unemployment		0.001		-0.003
*Rule		(0.004)		(0.006)
Internet Connection	0.03	0.03	0.03	0.03
	(0.006)***	(0.004)***	(0.01)***	(0.006)***
(log) Patents	-0.10	-0.04	-0.12	-0.06
	(0.05)*	(0.03)	(0.09)	(0.05)
Concern-privacy	0.12	0.09	0.13	0.094
	(0.04)**	(0.03)***	(0.05)**	(0.046)*
Tertiary Education	0.007	0.005	0.007	0.004
	(0.002)***	(0.002)***	(0.004)*	(0.0027)
(log) Income	0.19	0.25	0.22	0.265
	(0.06)***	(0.06)***	(0.10)**	(0.105)**
(log) Population	0.08	0.02	0.11	0.05
	(0.07)	(0.04)	(0.12)	(0.07)
Countries	26	26	26	26
Pseudo R^2	0.22	0.26	0.21	0.16

Note: Robust standard errors are clustered at the country level and presented in the parenthesis. * *p*<.10, ** *p*<.05, *** *p*<.01.

Table 3. Instrumental Variable Approach: Cybercrime, e-Skills, and Institutions

(25 European countries, 2010)

	DV: Aggregate Cybercrime	
	Negative Binomial Regression	Log-linear Regressions
e-Skill	1.83	0.075
	(1.54)	(0.035)**
Rule	-12.82	-0.09
	(8.70)	(0.20)
e-Skill*Rule	-0.96	-0.035
	(0.52)*	(0.01)***
Unemployment	-0.69	-0.01
(Univ, 25-34)	(0.63)	(0.01)
Internet Connection	1.37	0.026
	(0.50)***	(0.01)**
(log) Patents	-3.70	-0.12
	(6.12)	(0.14)
Concern-privacy	3.70	0.09
	(1.92)*	(0.04)***
Tertiary Education	0.23	0.004
	(0.14)	(0.003)
(log) Income	7.97	0.12
	(4.66)*	(0.11)
(log) Population	4.09	0.13
	(7.38)	(0.17)
Countries	25	25
R^2	0.79	0.77
	First Stage Regression	
Instrument	0.55	
	(0.22)**	
Control Variables	Yes	
Countries	25	
F (10, 14)	15.96***	
Adjusted R^2	0.86	

	Exclusion Criteria: Aggregate Cybercrime (DV)			
	Negative Binomial Regression		Log-linear Regressions	
	(1)	(2)	(3)	(4)
e-Skill	0.08		0.08	
	(0.02)***		(0.03)***	
Reading (PISA)	-0.001	0.02	-0.001	0.02
	(0.005)	(0.008)**	(0.007)	(0.009)**
Rule	-0.05	-0.26	-0.06	-0.26
	(0.18)	(0.22)	(0.24)	(0.30)
e-Skill*Rule	-0.04		-0.04	
	(0.01)***		(0.01)***	
Reading*Rule		-0.01		-0.012
		(0.005)*		(0.007)*
Control Variables	Yes	Yes	Yes	Yes
Countries	25	25	25	25
(Pseudo) R^2	0.18	0.16	0.23	0.24

Note: Robust standard errors are clustered at the country level and presented in the parenthesis. * *p*<.10, ** *p*<.05, *** *p*<.01. The instrumented variable is *e-skill* and the external instrument is *Reading (PISA)*.

Appendix 1. Descriptive Statistics and Data Sources

	Observations	Mean	Std. Dev.	Min.	Max.	Source
Aggregate Cybercrime (%)	28	49.36	16.66	20	78	European Commission (2010)
Concern-Privacy (%)	27	1.52	1.05	0	3	European Commission (2010)
e-Skills (%)	28	12.46	6.68	2	29	European Commission (2010)
Internet Connection (%)	28	69	14.04	47	93	European Commission (2010)
Unemployment Rate (Univ., 25-34)	28	5.89	3.50	1.80	14.80	European Commission (2010)
Rule (rule of law index)	28	1.15	0.63	-1	1.98	Kaufmann et al. (2010)
Tertiary Education (%)	27	67.44	16.19	18.21	108.09	World Bank (2010)
Patents (log, number)	28	6.77	1.91	1.39	10.76	World Bank (2010)
Income (log, USD)	28	10.21	0.71	8.79	11.54	World Bank (2010)
Population (log, number)	28	15.97	1.30	13.14	18.22	World Bank (2010)
PISA Reading (% of levels 4 & 5)	27	26.04	9.07	3.5	48.5	OECD (2007)