Instituto de Engenharia de Sistemas e Computadores de Coimbra

Institute of Systems Engineering and Computers

INESC - Coimbra

Nicolas Gomes, Pedro Cerqueira and Luís Alçada-Almeida

Determinants of Worlwide software piracy losses

No. 2 2014

ISSN: 1645-2631

Instituto de Engenharia de Sistemas e Computadores de Coimbra

INESC - Coimbra

Rua Antero de Quental, 199; 3000-033 Coimbra; Portugal

www.inescc.pt

Determinants of worldwide software piracy losses

Nicolas Dias Gomes[a], Pedro André Cerqueira[a,b] and Luís Alçada-Almeida[a,c]

Abstract

This paper studies the determinants of software piracy losses along four major macroeconomic dimensions: Labor force, Technological, Educational and Access to Information using a large dataset available from 1994 to 2010, comprising 109 countries.

The results show that, regarding the labor dimension, employment in services has a deterrent effect while labor force with higher education and youth unemployment have positive effects on piracy losses. As for the technological dimension, more patents by residents have a positive effect while the effect of R&D is negative. In terms of the Educational dimension the results obtained show that more spending on education increase the piracy losses but, at the same time, more schooling years have the opposite effect. Finally regarding the Access to Information it seems that access to Internet diminishes the losses while the share of Internet broadband subscriptions has no effect.

Key Words: Piracy Losses, Software Piracy, Copyright, System GMM

JEL: C12, C23, C51, L86

[a] (Corresponding author), Faculty of Economics, University of Coimbra; e-mail: nicolasdiasgomes@gmail.com
[b] GEMF, University of Coimbra.
[c] INESC – Coimbra.

1. Introduction

In the past decades we have witnessed a huge development of the hardware and software industries, some examples of these developments were the "MS-DOS" and the Apple "Macintosh" with its friendly user interface. These operating systems were used in working environments, and the equipment in which they operated were very expensive. Both Mac OS and Windows operating systems evolved over several releases. For example Windows 8 evolved from its predecessor, the Windows 7; bringing new features such as a new UI (user interface). As the personal computers became more powerful, it made possible to develop software to an increasing array of applications, as for instance, music and digital edition.

With these developments it came also the need to improve the existing software to meet the requirements of consumers. These improvements came at a cost; it was necessary continuous research and development by companies in order to maintain quality and keep up with the changes in technology. The cost of these investments is passed to consumers in the form of a license; the consumer pays a license to use the software during a certain amount of time (normally annually) or buys a perpetual license for that particular software. This perpetual license is not really "perpetual" as software becomes obsolete quickly. Operating system updates are released within a regular period of time and sometimes these upgrades turn old software versions unusable. Investment on software must also be done into its protection against piracy because it has the characteristic of being easily distributed with virtually no cost associated.

The massive use of computers and the Internet made the problem of software piracy potentially more severe, as the pirated software can be uploaded in the Internet within increasingly shorter periods of time. Before the expansion of Internet usage, only hard copies were available which were easier to track. To prevent this phenomenon, companies must invest to avoid pirated software from being used in addition to the investment to meet the speciation of potential consumers. Pirated software starts its journey when the original program protection is bypassed by another program or action performed by hackers. The investment that the companies make must incorporate different layers of protection. Unfortunately the software protections are often hacked. To cover these costs the companies must increase *R&D*, which increases the initial prices, but these prices sometimes take away potential buyers. In recent years, and due to the need of the information being available anywhere and anytime, many software products offer online services that replace the need of the software being installed in the computer. These software can be free of charge or, to access its full capabilities, may be necessary an user registration fee. Because each time the user uses the software, he must be identified and logged-in, the risk of piracy is reduced.

Previous studies on the determinants of software piracy resorted to the software piracy rates as the dependent variable[1]. However official publications also report the software piracy losses and to our knowledge no empirical work conducted an analysis of this variable.

Both piracy rates and losses measure the illicit behavior in a country, in the first case it measures the percentage of software that is being illegally used at a given time, but it omits the importance of the software industry in the economy. We can have a low piracy rate and huge losses, example of this is the USA, on the other side countries with piracy rate above 90% may represent little impact on this industry due to the small domestic software markets. This variable allows to measure the benefit to the economy national income in lowering piracy.

Our contribution to the literature is as follows:

(i) We will examine what are the determinants of software piracy losses along four dimensions: the structure of the labor force, the technological development, the level of education and the availability of information;

(ii) We will use a panel methodology that provides consistent estimates when the dataset is persistent: the System-GMM proposed by Arellano and Bover (1995) and Blundell and Bond (1998).

The structure of the paper is the following: section two presents a brief survey of the empirical literature on software piracy that used cross sectional and panel data analysis. Section three describes the piracy rates and losses presented by the *Business Software Alliance* since 1994 in different regions of the world; section four explains the various dimensions and possible effects. Section five describes the econometric specification and the results and section six concludes.

[1] Due to data restrictions previous research used two main methodologies: cross-sectional and panel data using classical methods due to small time periods.

2. Literature review

This section reviews the empirical works that used cross-sectional and/or panel data to explain the software piracy phenomenon using the data provided by the *Business Software Alliance.* A panel data approach is more appropriated when all information is available for different countries in different years. Unfortunately when this is not the case it is only possible to make a cross-sectional analysis.

Marron and Steel (2000) used the average piracy rate from 1994 to 1997 (cross-sectional data for 72 countries). Some of the variables of interest were Gross Domestic Product per capita (*GDPpc*) in 1994, expenditure on R&D as a percentage of GDP and the average years of schooling in population with age over 25 years (see Barro and Lee, 2010). Education was significant at 5% with a negative coefficient and, the income was always significant and negative. They also introduced regional dummies that represent the different geographical areas of the world[2] and found that Europe and the Middle East were statistically significant with a positive impact.

Economic factors can reduce software piracy, but the distribution of income is equally important. Andrés (2006a) try to find if income inequality has some effect on the software piracy rates, to see this he used the *GINI* index. He also includes *"rule of law"* that indicates the effectiveness of the legal system. Additional to these variables the author included real GDP per capita *(GDPpc)* and the average years of secondary education aged over 25 years old (Barro and Lee, 2010). This sample covers only 35 countries for 1995, with few observations; ordinary least squares (OLS) were used. Only Inequality (measured by the *GINI* Index) and Law (measured by the rule of law) were significant at 5% level, being both negative. More equal societies and societies with a good legal system would lead to less software piracy rates.

In a recent study conducted by Goel and Nelson (2009) four additional influences were considered that could explain software piracy rates. They were: economic influences, institutional influences, technological influences and others. As a robustness analysis they analyzed the impact of legal institutions that represent the origin of the legal system (French, British, Socialist and German) and religious fractionalization, which measure how many religions are in a country. Some of the conclusions are that the size of software market makes piracy more lucrative, more corrupt nations have greater piracy. Diffusion of Internet, computer based technologies and accessibility to information networks in a

[2] These regions are: Africa, Asia, Europe, Latin America and the Middle East.

country enables both potential pirates and protectors of intellectual properties to do their jobs more effectively. Greater literacy enhances the piracy of software. Additional variables such as urbanization were used that served as control variables. Literacy rate omits the different levels of education; a person can be literate and lack the ability to use computers and software. French and Socialist legal systems have a positive effect on software piracy. Ethnic fractionalization has a negative effect on piracy.

While several authors use a large group of countries, Chen, Chen and Yeh (2010) used a panel of 11 countries from the "Far East Countries"[3] from 1995 to 2006, totalizing 132 observations. The methodology used was pooled OLS, random effect and fixed effect, being the later the most appropriate based on the Hausman test (Hausman, 1978). The main conclusion was that more income reduces software piracy, but at the same time, more unemployment led to less piracy. This sign must be viewed with caution; unemployment leads to less disposable income that would lead to less consumption, this can shift the consumption from legal software to illegal software. On the other hand, the psychological conditions normally associated with unemployment situations often lead to disinterest in the use of information technology. This variable is meaningless without an in deep analysis of the education of the labor force and the availability of technology that allow the consumption of both legal and illegal software.

More recently Boyce (2011) used a panel data approach to analyze some international determinants of software piracy, having in the sample between 278 and 340 observations for 105 countries from 2006 until 2009. Various econometric methods were used: fixed effects, pooled OLS and random effects. They have different assumptions, being the more appropriate the fixed effects. This study finds that Internet has a negative effect, but broadband has a positive effect; people with Internet may or may not use illegal software, while the speed is what determines the amount of information that can be retrieved at the same time. Higher Internet speeds increase the availability of both legal and illegal software. A greater IPR (*intellectual property rights index*) leads to a decrease in piracy rates. To take into account the non-linearity of some variables he introduces the squared HDI and GINI (HDI^2,$GINI^2$), these variables have a negative coefficient with statistically significant signs. This study tried to improve existing works with the introduction of a large

[3] These countries are Taiwan, South Korea, Japan, Singapore, China, Indonesia, Vietnam, Thailand, Malaysia, Hong Kong and Philippines.

group of countries over the years, overcoming some drawbacks that a cross-national study may have.

However, the studies that used a panel data have not considered that the dependent variable shows considerable persistence over time and we should use a dynamic framework. In this context as noticed for example by Baltagi (2008), the OLS estimator is biased and inconsistent even if the error terms are not serially correlated, the random effects estimator is also biased and even if the fixed effects estimator becomes consistent when T gets large, Judson and Owen (1999) show that for T=30 the bias can be as much as 20% of the true value of the coefficient of interest. Furthermore Soto (2009) reinforces these findings by showing that even when the N is not very large the GMM estimators proposed by Arellano and Bover (1995) and Blundell and Bond (1998) display the best features in terms of small sample bias and precision.

3. Evolution of the software piracy losses and rates over the years

This section describes the evolution of the piracy rates and losses since 1994 for different regions of the world using the data provided by the Business Software Alliance (BSA)[4].

Figure 1 presents the evolution of the piracy losses over the last 17 years measured at current prices (BSA, 2011) in millions of dollars; piracy losses for each region are the result of the sum of the losses for all countries in that region[5]. The piracy rates and losses are annually published by the *Business Software Alliance*, with the help of an external consultant[6]. In the period of 2002-2003, *Business Software Alliance* changed its consultant from IPRC *(International Planning and Research Corporation)* to IDC and, consequently, the results were substantially different. Furthermore in 2002 *Business Software Alliance* included other types of software such as operating system and games; as previously the reports only included business applications. This graph clearly shows this changed occurred in 2002 as there is a big jump from 2002 to 2003.

In the majority of regions the amount of losses grew at a constant rate since 2002 with the exception of Asia Pacific, North America and Western Europe where losses grew

---

[4] *Business Software Alliance* is a group pressure that estimates annually the software piracy losses and rates across different groups of countries.

[55] All estimations from countries take into account the exchange rate of the dollar against national currencies

[6] IDC *(International Data Corporation)*

rapidly. Some of this growth can be explained by the introduction of new countries in the sample, technological growth or the variations in the exchange rates.

In addition to this graph we calculated the losses per GDP in these regions (figure 2). Piracy losses per GDP was obtained as follows: piracy losses and GDP for each region is the simple average of all countries belonging to a region; then we divided Losses over GDP. As in figure 1 there exists a break in the series in 2002, but now these losses in terms of GDP are higher in Latin America, Central and Eastern Europe and Asia Pacific.

Figure 1: Evolution of Software Piracy Losses

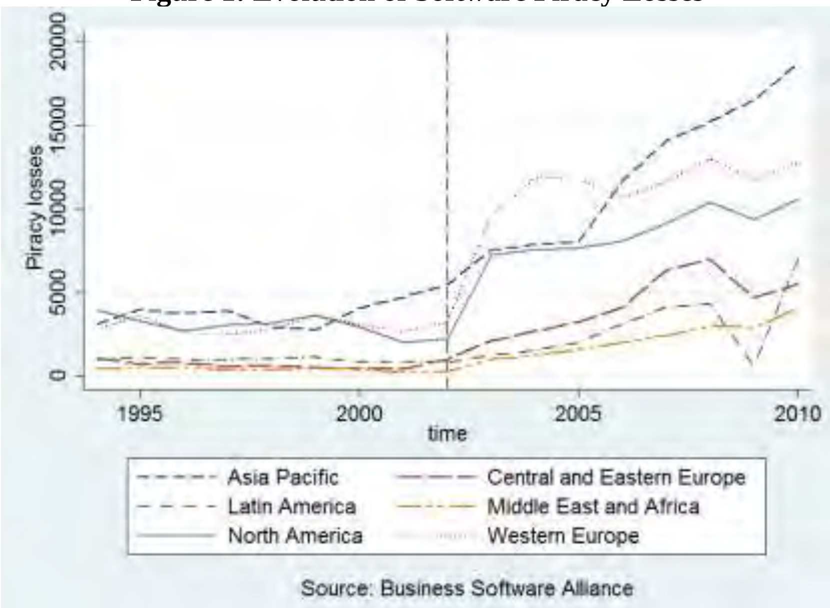


Figure 2: Losses / GDP on all regions

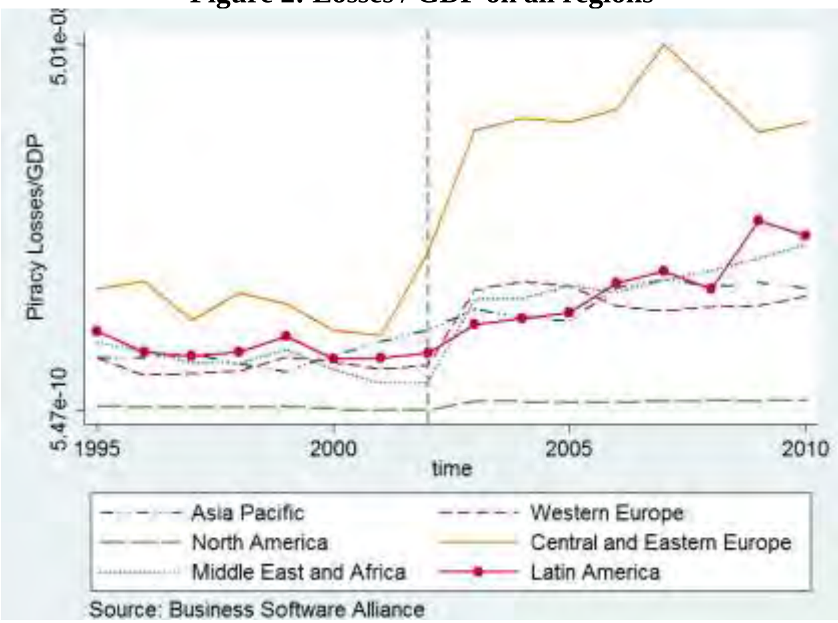


Figure 3: Evolution of Software Piracy Rates

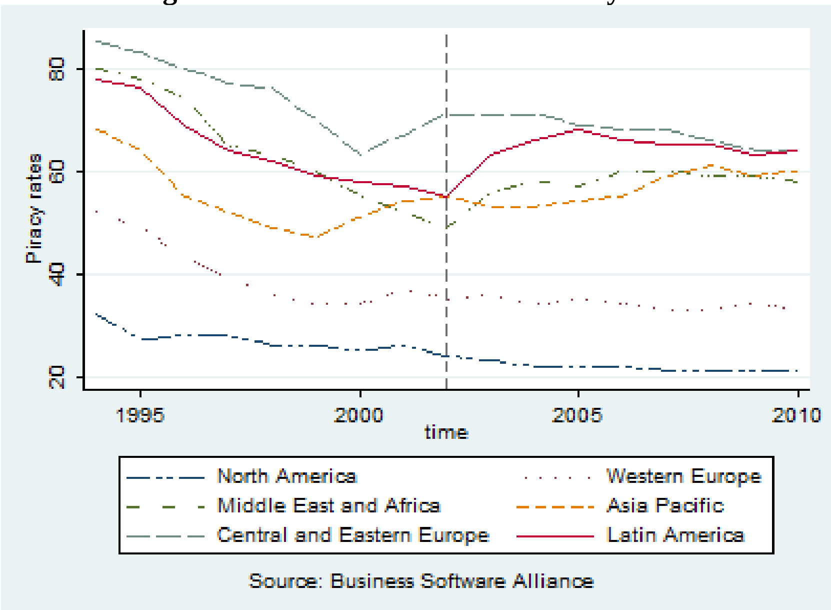


Figure 3 presents the evolution of the piracy rates. The relationship between the piracy rates and losses is not linear; losses increased over time while piracy rates fell. This decrease was not so drastic in 2002 compared to the losses increases. North America and Western Europe have the lowest piracy rates. These rates have been decreasing more rapidly until 2002 and, after this period, the rates had a smother pattern. In the case of Middle East, Africa and Latin America, software piracy decreased rapidly until 2002, then increased slowly. Png (2010) found that the change in methodology lead to a decrease from 2 p.p. per year to 1.1 p.p. points per year. We also report the software piracy rates on three regions of the world: BRIC[7], United States and European Union from 2003 until 2010 (Figure 4). The piracy rates in European Union and United States have been more or less constant over the last 8 years. In the BRICs the piracy rates dropped, nevertheless representing more than 70% of the used software.

[7]Brazil, Russia, India and China

Figure 4: Software Piracy rates on selected regions

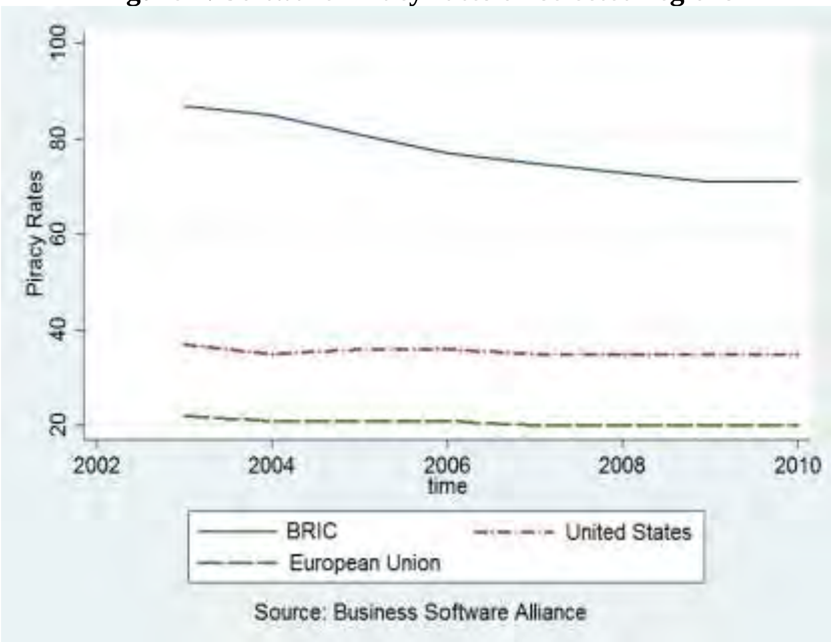
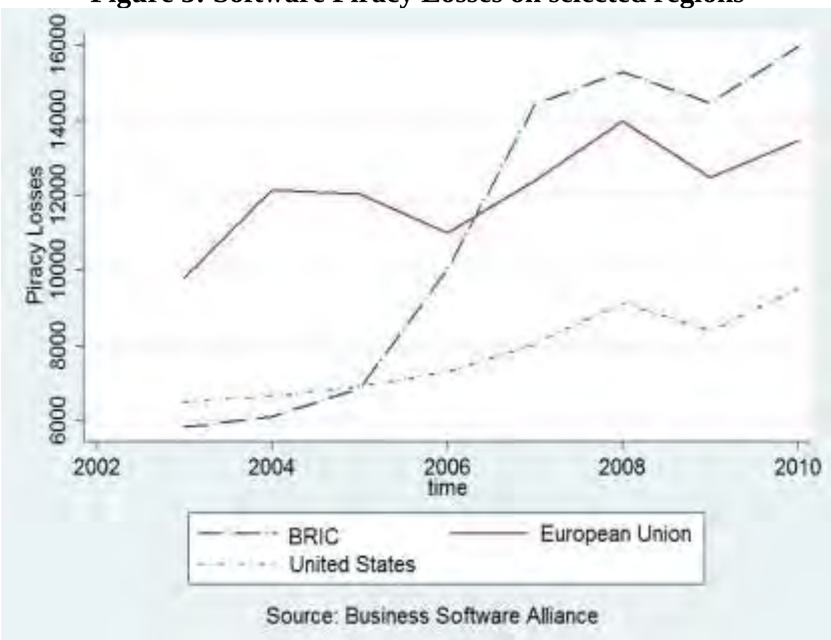


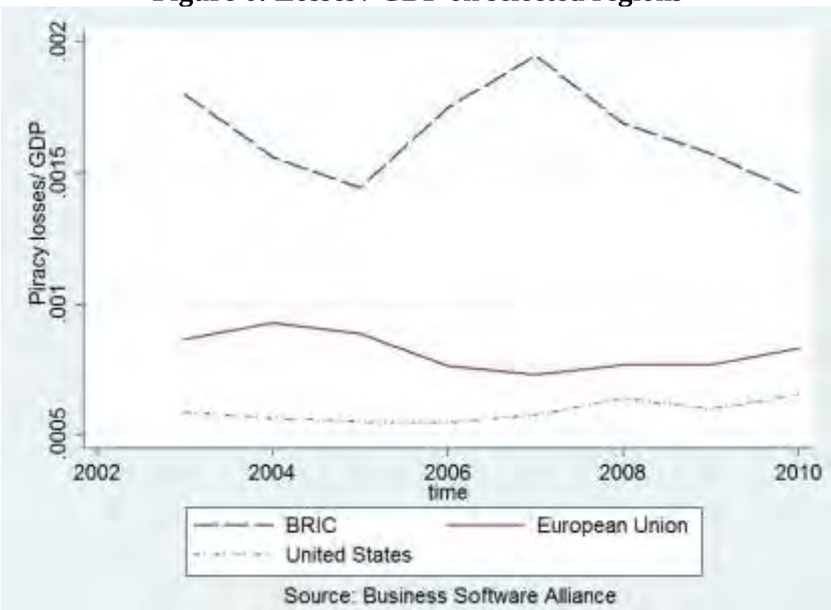
Figure 5: Software Piracy Losses on selected regions



The existence of low piracy rates don't necessarily means lower losses. Figure 5 and 6 shows the piracy losses on the three regions; piracy losses have been growing at a more or less constant rate on United States and European Union. The amount of losses almost tripled in eight years in the BRICs. When losses are analyzed as a proportion of GDP, the importance of the losses in the BRICs is even higher when compared with the EU and USA. This can represent the ineffective implementation of copyright laws.

Anyway the losses in terms of GDP have in any of the three regions been constant for the analyzed period.

Figure 6: Losses / GDP on selected regions



4. Variables and possible effects

This section describes the several dimensions discussed in previous papers to explain software piracy rates, namely the labor force, technological, educational and the access to information dimensions and the variables used to measure them.

4.1. Labor force dimension

Computer skills are acquired at school or at the workplace; these can range from browsing the Internet, sending e-mails or working on business applications such as word processors or spreadsheets. Different jobs require different types of software; some include imaging suits, others productivity or econometric tools, etc. We will consider three variables that reflect the structure of the labor force of the population: employment in the primary sector (Agriculture), employment in the secondary sector (Industry) and employment in the tertiary sector (Service).

In agriculture when the production is intensive, software helps to improve efficiency, controlling various elements of a greenhouse such as the temperature or humidity. In the industry sector, the use of specialized software is "normal", it comes with the machine and in many situations is developed and used by the firms internally, thus is not for sale. The software cannot be used outside of the environment that was intended to work. Big firms develop the software, or commission its development (outsource) to a specialized company (due to smaller costs). The services sector uses specialized software of accounting, taxation and productivity. Depending on the different needs, the software can cost thousands of dollars, but it comes with technical support, extremely valuable in order to maintain productivity and prevent failures. The costs associated with the acquisition of these types of software's can be deduced during a certain amount of time, reducing taxable profits. Employment in these sectors may have impact on software piracy losses. Firms want to maximize profit; in some cases due to budgetary restrictions, employers can introduce some illicit software that will benefit both employers and employees. The introduction of illicit software has risks associated that are the result from external audits that can result in fines. In spite of this, some firms may be willing to take them. Certain types of jobs, namely in the service sector can be done from home, for example in market research. In many cases workers wanting to do their jobs at home due to reduced costs may seek illicit software to implement their research. In this case there is no problem of internal audits finding illicit software. This may lead an increase of software piracy losses.

Additional to these variables we introduce the education of the labor force. We will consider labor force with primary, secondary and tertiary education. The labor force of a country plays an important role on the growth of the economy. If it is constituted by labor force with low education, this will lead to low productivity and, consequently, to small economic growth. Also, if the labor force is constituted by highly qualified people, this will lead to increased productivity which improves the standards of living. These highly qualified employees will use computers and software. More education of the labor force characterizes a double edge sword context: on one side there are more users of computers and software but, at the same time, some of the consumers will use illicit software[8].

Another measure that reflects both an income dimension and social dimension is the unemployment. We will use the total and youth unemployment that reflects people within

[8] We should note that the education of the labor force represents both a labor force dimension and educational dimension.

15-24 years without a work and the total unemployment. Both variables are expected to have a positive effect on piracy losses. An unemployed person has less disposable income and spends more time at home. Sometimes it is necessary to use certain software to start working (in the case of self-employment), but the lack of money can shift consumers from legal to illegal copies to fill their needs. Chen, Chen and Yeh (2010) found that unemployment has a negative effect on software piracy rates. Their sample was small and reflected a small group of homogeneous countries where the psychological aspects could be determinant. In this paper we use a large sample data to confirm their results.

4.2. Technological dimension

Technological dimensions can affect levels of software piracy. To measure this we introduce two variables that capture this dimension. They are the patent and the trademark applications done by residents and non-residents. Additional to this we introduce a measure of quality that is represented by the expenditure on research and development as a percentage of GDP. This variable makes also part of the educational dimension. Marron and Steel (2000) found that more research leads to less piracy.

Patent applications are worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office for exclusive rights for an invention (this can be a product or a process). The protection of this product can reach up to 20 years. Patents can be filed by domestic or by foreign applicants. Trademark applications filed are application to register a trademark with a national or regional Intellectual Property (IP) office. A trademark is a distinctive sign that identifies certain goods or services as those produced or provided by a specific person or enterprise. A trademark provides protection to the owner of the mark by ensuring the exclusive right to use it to identify goods or services, or to authorize another to use it in exchange for a license or royalty. The period of protection varies. Direct resident trademark applications are those filed by domestic applicants directly at a given national IP office while those that are filed by applicants from abroad are called “direct nonresident trademark applications”. The registration of a patent or a trademark has costs for the firm, but these are necessary in order to protect their products. All these variables are expected to have a negative effect on piracy losses although the existence of a patent is not sufficient to prevent piracy; the enforcement trough strict regulation is also necessary. More technological advanced countries have more legal protection and at the same, firm’s owner of the technology are also more close to the market and can detect more easily illegal software.

4.3. Educational dimension

In all countries there are a predetermined number of years of schooling that a child must complete, and these vary from one country to another. During this period children have specific subjects that use the computers and Internet. This early introduction to new technologies will improve productivity of future workers. In some subjects professors introduce the concept of illegal software and the risks associated with their use. The introduction of other concepts such as copyright can prevent the future use of illegal software. MacDonald and Fougere (2003) studied this effect analyzing MIS (Management Information Systems) textbooks. More years of schooling indicate that the children are aware and understand what happens if they use illegal software. This variable may or may not have a negative effect. A measure that indicates years of schooling of primary education based on the ISCED 1997 (ISCED 1) and secondary education based on the ISCED 1997 (ISCED 2 and 3) will be used. This indicator reveals the total education that a country offers. Other measures that represent the education attained measured by the total years of schooling (see Barro and Lee, 2010) could be used, but there is only data for five period years, which would substantially limit our analysis.

A measure that reflects the expenditure that is made on education will be introduced. More public expenditure on education as a percentage of GDP can reduce illegal software that students use; this will also result in more quality of education. This financial help can go both to public or private institutions. This variable is expected to affect negatively software piracy losses.

4.4. Access to information

Technology has evolved over the years. Today it is difficult or even impossible to live without it. Hardware and software industry have profited with these developments, but with the dissemination of the Internet it was also possible to download huge amounts of information, some of which not legal such as pirated software. Most modern mobile phones uses an operating system, some of which may even replace the computer (the case of smartphones). We will introduce four variables that measure the availability of information (Internet, mobile, telephone and fixed broadband Internet subscriber users).

Telephone lines are physical and fixed lines that connect a subscriber's terminal equipment to the public switched telephone network and that have a port on a telephone exchange. Integrated services digital network channels and fixed wireless subscribers are included. Fixed broadband Internet subscribers are the number of broadband subscribers with a digital subscriber line, cable modem, or other high-speed technology. Internet users are people with access to the worldwide network. Mobile cellular telephone subscriptions are subscriptions to a public mobile telephone service using cellular technology, which provide access to the public switched telephone network. Post-paid and prepaid subscriptions are included. All these variables affect the availability of software; more usage of these devises would lead to an increase in software piracy losses. Boyce (2011) found that broadband penetration rate and Internet access has a negative impact (in the fixed effect model).

5. Empirical evidence

5.1. Data, econometric specification and summary statistics

Previous econometric studies relied on cross-section or panel data analysis. In the panel data models, periods of study were relatively short; this work is the first to introduce a large time span in the analysis. Figure 1 shows that software piracy losses are highly persistent over time and that its value follows closely the GDP (see figure 2); its value is always increasing over time. Soto (2009) examined the properties of System GMM when the sample is small and the series is persistent, which is applicable to our dataset. This estimator was found to have lower bias and higher efficiency than the OLS or the fixed-effects estimator, furthermore the gain in efficiency from the two-step estimator is almost inexistent; both the one and two step distributions are virtually the same. Based on these results we will report the one-step System GMM.

Our dataset is constituted by macroeconomic variables retrieved from the World Development indicators available for the countries present in the publications provided by the *Business Software Alliance* and comprising 109 countries from 1994 until 2010. Due to the persistence of the piracy losses, we will use a dynamic panel data analysis, namely the Arellano-Bover (1995) and Blundell-Bond (1998) estimator. This estimator was developed because the lagged-level instruments of the original Arellano and Bond (1991) estimator become weak when the autoregressive process becomes too persistent or the ratio of the variance of the panel-level effects to the variance of the idiosyncratic error becomes too large. The System GMM uses both level and first-difference of the lagged dependent variable as instruments.

The dependent variable is the piracy losses due to pirated software, and it’s measured in millions of dollars. The independent variables measure various dimensions of a country:

labor force dimension, technological dimensions, educational dimension, and access to information. In our analysis we will use the nominal GDP (*GDP*)[9] as a control variable.

The estimator used poses some problems, namely in the case of too many instruments, when the instrument count is high they may fail to expunge their endogenous components and biasing coefficient estimates toward those from non-instrumenting estimators as discussed by Roodman (2009a). With the limitation of lags we overcome this problem, e.g. the number of instruments higher than the number of countries. In the end of each regression we report the number of instruments used and also through our analysis, the number of instruments will be smaller than the number of countries following Roodman (2009b).

The econometric specification is given in equation 1 as follows:

$$\ln(\text{Losses})_{it} = \beta_{it} + \theta_1 \ln(\text{Losses})_{i,t-1} + \alpha_1 \ln(\text{GDP})_{it} + X_{it}\alpha_X + Y_{it}\alpha_Y + Z_{it}\alpha_Z + W_{it}\alpha_W + \\ + \alpha_6 \text{Change}_t + \upsilon_i + \varepsilon_{it} \quad (1)$$

i =1,…,109 represents the countries and *T =1994,…,2010* the time periods

The variable *Losses* is the piracy losses measured in millions of dollars and *GDP* is the Gross Domestic Product at current prices. Additional to this we could also consider $\frac{\ln(losses)}{\ln(GDP)}$ as a dependent variable that represents the relative importance of the piracy losses in relation to GDP[10].

X_{it} is a vector of labor force, and it reflects the labor force dimension and it's constituted by the labor force, type of employment and unemployment.

$$X_{it} = [\text{Labp}_{it}\ \text{Labs}_{it}\ \text{Labt}_{it}\ \text{Empagri}_{it}\ \text{Empind}_{it}\ \text{Empserv}_{it}\ \text{Unempyouth}_{it}\ \text{Unemp}_{it}] \quad (2)$$

were *Labp* is the labor force with primary education, *Labs* is the labor force with secondary education, *Labt* is the labor force with tertiary education, *Empagri* is the employment in agriculture, *Empind* is the employment in industry, *Empserv* is the

[9] GDP is measured in current US dollars, this variable will be considered as endogenous. This variable is used to control the market dimension.

[10] To assess the validity of both assumptions we performed regressions. In Both cases the variables maintain the same coefficient, being the only difference the magnitude of these. We opted by the absolute value of Losses as it provided best estimates, maintaining significance.

employment in the services sector. *Unempyouth* is the unemployment of people from 15 to 24 years old and *Unemp* is the total unemployment.

Y_{it} is a vector that represents the technological dimension and it's constituted by patents, trademarks and the research and development.

$$\mathrm{Y_{it}} = [\mathrm{R\&D_{it}} \ln(\mathrm{Patres})_{\mathrm{it}} \ln(\mathrm{Patnon})_{\mathrm{it}} \ln(\mathrm{Tradres})_{\mathrm{it}} \ln(\mathrm{Tradnon})_{\mathrm{it}}] \quad (3)$$

R&D represents the research and development expenditure as a percentage of GDP, *Patres* are the patent applications done by residents, *Patnon* are the patent applications done by nonresidents. *Tradres* is the trademark applications done directly by residents and the *Tradnon* is the trademark applications done directly by nonresident. Both patents and trademarks are in logarithms.

Z_{it} is a vector of the education dimensions. It combines years of schooling and public expenditure in the different levels of education (primary, secondary and tertiary) as a percentage of *GDP*.

$$\mathrm{Z_{it}} = [\mathrm{Yschoolpri_{it}}\ \mathrm{Yschoolsec_{it}}\ \mathrm{Pubexp_{it}}] \quad (4)$$

$Yschoolpri$ is the duration in years of primary education, *Yschoolsec* is the duration in years of secondary education. *Pubexp* represents public expenditure on education as a percentage of GDP.

The vector W_{it} represents the various variables that represent access to information.

$$W_{it} = [\ln(Fbis)_{it} \ln(Mobile)_{it}\ \ln(Phone)_{it}\ \ln(Net)_{it}] \quad (5)$$

Fbis is fixed broadband Internet subscribers, *Mobile* is mobile cellular subscriptions, $Phone$ is the phone lines and *Net* is the access to the Internet. These variables are measured per 100 people. We introduce logarithms in this dimension.

Additional to these variables, a dummy variable *(Change)* will be introduced that reflects the change in methodology provided by the *Business Software Alliance.* Before 2003 it will have a value of 0 and of 1 afterwards. We will also introduce a set of time dummies.

Table 1 presents the descriptive statistics of the various dimensions. In 2010 data for software piracy losses was available for 109 countries. Every year the *Business Software Alliance* improves its estimates performing more surveys; initially surveys were made on

15 countries[11] (BSA 2003); in 2010 surveys were made on 32[12] countries, a total of 15000 computer users were inquired. Since 1994 a total of 1522 observations are available.

Table 1: Summary statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
Losses	1522	289.63	841.48	0.254	9515
GDP	1838	366323.5	1225590	984.28	14447100
Labor force dimension					
Empagri	1316	16.60	16.43	0	72.2
Empind	1317	24.41	6.51	5.8	51.8
Empserv	1317	58.36	14.61	13.3	87.4
Labp	811	30.06	17.55	0	89
Labs	805	42.67	17.11	2	80.2
Labt	811	23.70	10.68	0	66.1
Unempyouth	1104	17.84	9.93	1.6	70.9
Unemp	1368	8.93	5.77	0.5	37.3
Technological dimension					
Patres	1298	11177.49	47273.17	2	384201
Patnon	1323	6337.01	20975.35	1	248249
Tradres	1323	21402.75	60719.95	1	973460
Tradnon	1322	5846.57	7345.99	33	67838
R&D	939	1.06	0.96	0.01	4.80
Educational dimension					
Pubexp	882	4.59	1.57	2.20e-06	9.66
Yschoolpri	1850	5.51	0.97	3	8
Yschoolsec	1514	6.48	0.97	4	9
Acess to information dimension					
Net	1763	20.26	24.26	0	95.63
Fbis	1207	5.82	8.98	0	38.10
Phone	1831	25.02	19.21	0.23	74.69
Mobile	1833	44.49	44.77	0	205.28

5.2. Empirical application

This section presents the empirical results for the various dimensions. Some of our variables in education don't vary over time and consequently the fixed effect was not the best choice as it omits invariant regressors. Based on the Hausman test; Chen, Chen and Yeh (2010), Andrés (2006b), Boyce (2011) found that the fixed effect was more appropriated in panel data analysis. Traditional methods such as fixed effects or random effects produce inconsistent estimates when the lagged dependent variable enters as a

[11] Bolivia; Brazil; Chile; China; Colombia; Costa Rica; Dominican Republic; Guatemala; Kuwait; Malaysia; Mexico; Romania; Spain; Taiwan and the United States.

[12] Emerging markets include: Argentina; Brazil; China; Chile; Colombia; Czech Republic; India; Indonesia; Korea, Republic of; Malaysia; Mexico; Nigeria; Poland; Russian Federation; Saudi Arabia; South Africa; Thailand; Turkey; Ukraine; Vietnam. Mature markets include: Australia; Canada; France; Germany; Italy; Japan; Netherlands; Spain; Sweden; Switzerland; United Kingdom; United States.

regressor. Due to the nature of our dataset (extremely unbalanced when considering certain dimensions), we will split our analysis into 14 regressions in the dynamic model. In all of them there will be control variables for each dimension. Our results will be conducted using the one-step System GMM[13].

For the System GMM to be applicable, it is necessary that there is no evidence of second order autocorrelation AR(2). Additional to this, instrument must be valid. To test this we report the Hansen test for validity of instruments (Hansen, 1982). This test assumes under the null hypothesis that instruments are valid and it's a robust version of the Sargan test; one problem that may occur is that it can be weakened by instruments proliferation. We also report the number of instruments following Roodman (2009b).

Tables 2, 3 and 4 presents the regressions within each dimension, all of them have control variables for the remaining dimensions. The control variable in the educational dimension is the result of the sum of primary and secondary schooling years, "*School*". In the Technological dimension, variables were constructed that provide the best estimates; the sum of patents from residents and nonresidents, "*ln(Patents)*"*;* the sum of trademark from residents and nonresidents, "*ln(Trademark)*". When necessary we also summed the total patents and trademarks, "*ln(Legal)*"*;* this variable give us a general idea of the overall demand for this kind of protection.

Columns 1 through 5 summarize the results from the labor force. From these results we conclude that the higher the share of people working in the services sector the lower will be the piracy rate. This can be seen in columns 1, 2 and 4 where the base sector (omitted variable)[14] was the share of people working in the agriculture sector or in the industrial sector. In either case the share of people working in the services sector has a negative significant impact in the piracy losses. If we use the services sector as the base sector (column 3 or 5) the industrial and agricultural sectors have positive and significant impact (although the significance of the agricultural sector is not robust across specifications) pointing to the same conclusion, this is, the higher the share in these sectors

[13] The two-step System GMM is presented in Annex because the gains from the one-step to the two-step System GMM are marginal (Judson e Owen, 1999; Soto, 2009). In the two-step System GMM we take into account the corrected covariance matrix proposed by Windmeijer (2005).

[14] We should note that the sum of the three sectors adds up to 1, so we cannot have the three variables simultaneously in the regression due to multicolinarity. In these case we consider one sector as the base one (and omit that from the regression) and the coefficients of the others sectors is the differential impact between that sector and the base one.

(and lower in the service sector) the higher will be the losses due to software piracy. This result was unexpected, compared with what we expected.

As for the labor qualifications, as before, we have to omit one of the variables and use it as the base case. In this case the higher the share of workers with tertiary education the higher will be the losses due to piracy. Furthermore, the results indicate that is the division between the share of the workers with the tertiary education and the others that matters. This can be seen in columns 1 and 2 when we consider the share with primary education, the estimated coefficient of the share with secondary education is close to zero and non-significant. The same resuly is obtained in column 4 when we use the secondary education share as the base and analyze the estimate of the share with primary of education. This result is in line with what we expected.

Another important variable present in the labor force is the youth unemployment. This variable combined with the level of education of the labor force and access to technology will determine the use of software at home. This variable has always a positive impact, but it was significant only in regressions 1 through 5 in which labor force and type of employment was present. There are other types of unemployment. Some examples are the long-run unemployment and total unemployment. The first variable was not suited as the reduced number of observations made difficult to estimate with this methodology. As an additional robustness check we included total unemployment; significance was not present, nevertheless it maintains the positive coefficient.

Columns 6 through 9 show the various variables in the Technological dimension. Patents and trademarks offer protection for those who innovate; this protection can be done by residents or by non-residents that will protect their product. In columns 8 and 9 *ln(Patents)* has a significant positive impact on the losses, as *ln(Trademarks)* is non-significant. When a disaggregated analysis is made on the origin of patents and trademarks applicants (columns 6 and 7) the trademarks continue to not have a significant impact in losses, as in terms of the number of patents is the number of patents done by residents that have a significant impact on losses. A final variable that was found to have a strong effect in deterring software piracy losses was *R&D* which has always a negative coefficient.

The positive coefficient of patents and the negative coefficient of *R&D* at first may appear odd, but it can be explained as follows: a company makes a breakthrough after many years of research; this will allow increased productivity, efficiency and protection for different components of the company products which can also be extended to other products from other companies. In order to protect this discovery, the company will file a patent of the discovery that will allow some level of protection from other companies and from potential pirates. The existence of the patent by itself it's not synonym of protection; national Intellectual Property offices must also be able to enforce and protect them. The positive coefficient can be explained by the existence of patents that are the result of research, but the lack of power exerted by national IP offices will not prevent piracy in spite of the existence of patents.

Regarding education we included financial and non-financial measures. The first perspective was never considered in previous studies and in the second case; variables used on previous research were literacy rate and the average years of schooling of people age 25 and over (Barro and Lee, 2010). Both of these variables have the problem of data availability. To overcome this problem we will introduce a proxy variable that indicates the years of schooling of both primary and secondary education offered by the educational system of a country. This variable is not perfect as it omits the education attained but it offers us a benchmark. Additional to this, we will introduce a variable that indicates spending on education. Columns 10 through 14 show the results.

Columns 10 through 12 present the years of schooling. As in the labor force dimension, we include the education of the labor force; results are robust. An increase of years of primary and secondary education appears to have a negative impact on piracy losses, but only in the first case this variable is significant at 1% across all regressions with a coefficient of around -0.250.

The financial aspects of education are presented in columns 13 and 14. Public spending on education can go both to public or private institutions and depending on the different levels of education different resources are allocated. Public expenditure on education has a positive effect and is significant at 1% (column 14). This public spending can also go directly to students through direct help in the form of scholarships. There are many ways a student can use this help, some examples are: acquisition of computer, software, access to the Internet, etc. This will increase the availability to digital content such as music, software and movies. We were expecting a negative impact; nevertheless this may indicate that more access to digital content can also increase the availability of illegal software. Only with increase awareness is that this problem can be mitigated.

Several alternative hypotheses were considered with different variables within the dimension that reflect the access to information. The access to Internet – *ln(Net)* – has a negative and significant effect on losses (columns 13 and 14), while the access to a broadband connection has no impact on losses (columns 5 through 9). This is clearly unexpected, but seems that access to Internet reduces the losses because of increased awareness of the problem by the consumers, because countries with a higher share of people connected to the Internet are able to track those who use illicit software, or simply because you can download and buy the software directly from the original company and not from local intermediaries, many of them may be selling pirated software. These results are in accordance with the findings of Boyce (2011).

To assess the validity of our findings we also performed regressions in which the dependent variable was a fraction of GDP, *ln(Losses/GDP)* (see Annex B, tables B4 to B9), in these regressions the dimension related to the labor force was statistically significant being robust across all regressions. In the technological dimension, patents applicants from residents maintained significance, although R&D loss some of this significance. Variable that represents access to information, namely Internet users, and mobile subscriptions maintained statistical significance.

Table 2: Dynamic model using one-step System GMM

VARIABLES	1	2	3	4	5
L.ln(losses)	**0.593*** **	**0.514*** **	**0.632*** **	**0.610*** **	**0.459*** **
	(6.899)	**(4.914)**	**(7.196)**	**(7.300)**	**(4.117)**
ln(GDP)	**0.333*** **	**0.311*** **	**0.306*** **	**0.309*** **	**0.408*** **
	(3.404)	**(2.995)**	**(3.191)**	**(3.395)**	**(2.695)**
ln(Mobile)	-0.017	-0.049			
	(-0.468)	(-1.251)			
*ln(Mobile)*Change*		**0.382** **			
		(2.345)			
ln(Phone)			0.130	0.023	
			(0.691)	(0.148)	
*ln(Phone)*Change*			0.022	-0.108	
			(0.129)	(-0.628)	
ln(Fbis)					0.027
					(0.773)
*ln(Fbis)*Change*					0.069
					(1.431)
Empagric			0.011	-0.019	**0.019*** **
			(1.423)	(-1.474)	**(3.204)**
Empind	0.013	0.015	**0.024** **		**0.028** **
	(1.367)	(1.331)	**(2.450)**		**(2.502)**
Empserv	**-0.012** **	**-0.011* **		**-0.031*** **	
	(-2.133)	**(-1.735)**		**(-2.742)**	
Labp			**-0.013* **	0.004	**-0.015* **
			(-1.797)	(0.809)	**(-1.704)**
Labs	-0.003	-0.003	**-0.016** **		**-0.013* **
	(-0.603)	(-0.603)	**(-2.179)**		**(-1.801)**
Labt	**0.016** **	0.008		**0.022** **	
	(2.023)	(0.779)		**(2.378)**	
Unempyouth	**0.012** **	**0.014** **	**0.013** **	**0.012** **	**0.010** **
	(2.036)	**(2.458)**	**(2.312)**	**(2.166)**	**(2.035)**
ln(Patents)	-0.018	0.000	-0.021	-0.014	-0.002
	(-0.712)	(0.013)	(-0.505)	(-0.379)	(-0.068)
ln(Trademarks)	0.045	**0.087** **	0.027	0.039	0.052
	(1.036)	**(2.111)**	(0.429)	(0.754)	(0.919)
School	0.021	0.057	-0.113	0.012	-0.036
	(0.240)	(0.573)	(-1.077)	(0.146)	(-0.273)
Change	**0.392*** **	**-1.130* **	0.300	0.735	**0.352*** **
	(4.452)	**(-1.840)**	(0.453)	(1.124)	**(3.537)**
Constant	**-7.625*** **	**-8.008*** **	**-5.506*** **	**-5.760*** **	**-8.843*** **
	(-2.982)	**(-2.997)**	**(-2.578)**	**(-2.601)**	**(-2.872)**
Observations	507	507	507	513	415
Countries	62	62	62	62	59
AR1	-4.767	-4.352	-4.477	-4.666	-3.593
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
AR2	-0.173	0.424	-0.419	-0.432	0.293
p-value	[0.862]	[0.672]	[0.675]	[0.666]	[0.770]
Instruments	53	53	53	53	51
Hansen	43.41	36.60	41.78	38.65	43.24
p-value	[0.185]	[0.394]	[0.200]	[0.308]	[0.109]

Notes: Dependent variable is *ln(losses)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(losses)* and *L.ln(GDP)* were considered as two endogenous instruments. Only one lag was used as instrument.

Table 3: Dynamic model using one-step System GMM (cont.

VARIABLES	6	7	8	9
L.ln(losses)	**0.509*** **	**0.610*** **	**0.544*** **	**0.538*** **
	(5.276)	**(6.410)**	**(5.193)**	**(4.872)**
ln(GDP)	**0.349*** **	**0.294** **	**0.389*** **	**0.392*** **
	(2.984)	**(2.143)**	**(2.861)**	**(2.763)**
ln(Mobile)	-0.021	0.075	**0.170** **	**0.146* **
	(-0.245)	(0.679)	**(2.252)**	**(1.651)**
*ln(Mobile)*Change*	0.248	0.009		0.083
	(1.622)	(0.054)		(0.555)
ln(Fbis)	-0.003	-0.016	0.004	0.003
	(-0.091)	(-0.617)	(0.158)	(0.118)
Unempyouth			0.009	0.011
			(1.381)	(1.511)
Unemp	0.003	-0.019		
	(0.173)	(-1.245)		
ln(Patents)			**0.055* **	**0.060** **
			(1.845)	**(2.134)**
ln(Trademarks)			-0.037	-0.039
			(-0.676)	(-0.689)
School	0.086	-0.062	0.053	0.066
	(1.003)	(-0.614)	(0.519)	(0.616)
R&D	**-0.326*** **	**-0.134* **	**-0.256** **	**-0.274** **
	(-3.048)	**(-1.901)**	**(-2.128)**	**(-2.352)**
ln(Patres)	**0.138*** **			
	(2.742)			
ln(Patnon)	0.020			
	(0.602)			
ln(Tradres)		0.123		
		(1.505)		
ln(Tradnon)		-0.082		
		(-1.429)		
Change	-0.566	0.311	**0.342*** **	0.004
	(-0.933)	(0.498)	**(3.165)**	(0.007)
Constant	**-8.578*** **	**-5.597** **	**-9.327*** **	**-9.513*** **
	(-3.381)	**(-2.561)**	**(-3.355)**	**(-3.184)**
Observations	479	473	412	412
Countries	75	75	66	66
AR1	-2.999	-3.277	-3.715	-3.588
p-value	[0.003]	[0.001]	[0.000]	[0.000]
AR2	0.623	0.195	-0.230	-0.123
p-value	[0.533]	[0.846]	[0.818]	[0.902]
Instruments	51	51	51	51
Hansen	36.34	44.43	35.00	34.12
p-value	[0.406]	[0.132]	[0.516]	[0.511]

Notes: Dependent variable is *ln(losses)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(losses)* and *L.ln(GDP)* were considered as two endogenous instruments. Only one lag was used as instrument.

Table 4: Dynamic model using one-step System GMM (cont.)

Variables	10	11	12	13	14
L.ln(losses)	**0.553*** **	**0.550*** **	**0.543*** **	**0.624*** **	**0.633*** **
	(6.512)	**(7.465)**	**(7.021)**	**(6.673)**	**(6.165)**
ln(GDP)	**0.440*** **	**0.374*** **	**0.401*** **	**0.322*** **	**0.268*** **
	(4.892)	**(4.753)**	**(5.284)**	**(3.643)**	**(2.755)**
ln(Mobile)	-0.029	-0.041	-0.028		0.034
	(-0.823)	(-1.181)	(-0.769)		(0.459)
*ln(Mobile)*Change*					**0.360*** **
					(2.646)
ln(Net)				**-0.102***	**-0.167***
				(-1.876)	**(-1.741)**
R&D				**-0.184****	**-0.181****
				(-2.518)	**(-2.010)**
Yschoolpri	**-0.294*** **	**-0.234*** **	**-0.255*** **		
	(-3.106)	**(-3.372)**	**(-3.455)**		
Yschoolsec	-0.070	-0.110	-0.027		
	(-0.571)	(-1.248)	(-0.253)		
Pubexp				**0.142****	**0.177*** **
				(2.327)	**(2.751)**
Labp	**-0.016****	0.002			
	(-2.495)	(0.457)			
Labs	**-0.022****		**-0.009***		
	(-2.558)		**(-1.812)**		
Labt		**0.012****	**0.014*** **		
		(2.048)	**(2.725)**		
Unemp	0.000	-0.009	-0.005	-0.008	0.003
	(0.005)	(-0.991)	(-0.495)	(-0.518)	(0.152)
ln(Legal)	**-0.034***	-0.011	-0.019	0.017	**0.049****
	(-1.692)	(-0.730)	(-1.021)	(0.878)	**(2.094)**
Change	**0.400*** **	**0.438*** **	**0.420*** **	**0.473*** **	**-1.029***
	(4.374)	**(5.473)**	**(4.710)**	**(4.445)**	**(-1.895)**
Constant	**-5.347*** **	**-5.711*** **	**-6.344*** **	**-7.115*** **	**-6.524*** **
	(-2.951)	**(-3.396)**	**(-3.662)**	**(-3.952)**	**(-3.212)**
Observations	520	526	520	442	442
Countries	64	64	64	66	66
AR1	-4.271	-4.389	-4.314	-3.682	-3.416
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
AR2	-0.438	-0.465	-0.294	-0.549	0.122
p-value	[0.661]	[0.642]	[0.768]	[0.583]	[0.903]
Instruments	55	55	55	47	47
Hansen	43.81	47.60	44.44	44.70	35.36
p-value	[0.313]	[0.191]	[0.290]	[0.104]	[0.313]

Notes: Dependent variable is *ln(losses)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(losses)* and *L.ln(GDP)* were considered as two endogenous instruments. Only one lag was used as instrument.

6. Conclusion

This paper examined the impact of several dimensions that might explain the phenomenon of software piracy losses. Due to the nature of our dataset and the availability of information we opted to use a dynamic panel data analysis that could track the growth of piracy losses over time. We found that, several dimensions explain this growth; labor force, educational, technological, and access to information.

The labor force was one of the dimensions considered, we controlled for both the type of employment and the education of employees. Higher levels of education resulted in more losses, but a higher share of employment in the service sector has a negative impact on losses. This is a result of more access to information by employees with higher education and the capability to track illicit content through internal audits in the service sector.

In the technological dimensions, patents and trademarks were analyzed as one of the explanations of software piracy. Patents were significant and positive; they grant a protection for those who innovate but other factors must be considered such as the effectiveness of these and the punishment for those who infringe the law. One of the blueprints presented by the *Business Software Alliance* is to "modernize IP Laws to account for new innovations", and these innovations are patented both at home and abroad, having different results. Again ours findings suggest that more protection in the form of trademarks or patents can in some cases reduce losses. Another variable introduced was *R&D* which was found to have a negative effect on software piracy losses.

In 2012 the ninth edition of the BSA (2012) global piracy study presented some "blueprints" to reduce or at least mitigate this problem. One of these solutions is the increase of public education and awareness. The results from the education dimensions support these statements; more years of schooling have a deterrent effect on piracy. When our analysis turns to financial aspects of education; more spending means more piracy losses. Further research introducing additional variables must be followed.

More Access to information has mixed results on piracy losses. This can be explained by the nature of the different devices used to access digital content, for example the Internet.

7. References

[1] Andrés, A.R. (2006a), “Software Piracy and Income Inequality”, *Applied Economics Letters*, Taylor and Francis Journals, 13(2), 101-105.

[2] Andrés, A.R. (2006b), “The Relationship Between Copyright Software Protection and Piracy: Evidence from Europe”, *European Journal of Law and Economics*, Springer, 21(1), 29-51.

[3] Arellano, M.; Bond, S. (1991), “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations”, *Review of Economic Studies,* 58(2), 277-297.

[4] Arellano, M., and Bover, O. (1995), “Another Look at the Instrumental Variable Estimation of Error-components Models”, *Journal of Econometrics*, 68, 29-51.

[5] Baltagi, B., 2008. Econometric Analysis of Panel Data, 4th Ed. Chichester, UK. John Wiley & Sons, Ltd.

[6] Barro, Robert and Lee, Jong-Wha, "A New Data Set of Educational Attainment in the World, 1950-2010." Journal of Development Economics, forthcoming,

[7] Blundell, R., and S. Bond (1998), “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”, *Journal of Econometrics*, 87, 115-43.

[8] Boyce, J.A. (2011), “International determinants of software piracy”, Thesis (M.A., Economics), California State University, Sacramento.

[9] BSA (2003), Eighth annual BSA global software piracy study: Trend in software piracy 1994-2002, Business Software Alliance.

[10] BSA (2011), Eight annual BSA and IDC global software piracy study. Business Software Alliance.

[11] BSA (2012), Ninth annual BSA and IDC global software piracy study, “Shadow Market”, Business Software Alliance.

[12] Chen, Chen and Yeh (December 2010), “Determinants of Software Piracy: Evidence from Far East Countries”, *Journal of International and Global Economic Studies*, 3(2), 53-62.

[13] Goel, R.K., and Nelson, M.A. (2009), “Determinants of Software Piracy: Economics, Institutions and Technology”, *The Journal of Technology Transfer*, 34(6), 637-658.

[14] Hansen, L. (1982), “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica* 50(3), 1029-54.

[15] Hausman, J. A. (1978) *"Specification tests in econometrics."* Econometrica, 46, 1251–1271.

[16] Judson, R., Owen, L., (1999). "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists". *Economics Letters*, 65, 9-15.

[17] MacDonald Laurie E. & Fougere, Kenneth T. (2003). "Software piracy: A study of the extent of coverage in introductory MIS textbooks", *Journal of Information Systems Education*, 13(4).

[18] Marron, D.B., and Steel, D.G. (2000), "Which Countries Protect Intellectual Property? The case of Software Piracy", *Economic Inquiry*, 38(2), 159-174.

[19] Png, I.P.L. (September 2010), "On the Reliability of Software Piracy Statistics", *Electronic Commerce Research and Applications*, 9(5), 365-373.

[20] Roodman, D. (2009a), "A Note on the Theme of too Many Instruments", *Oxford Bulletin of Economics and Statistics,* 71(1), 135-58.

[21] Roodman, D. (2009b), "How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata", *Stata Journal*, 9(1), 86-136.

[22] Soto, M. (2009). "System GMM estimation with a small sample," Working Papers 395, Barcelona Graduate School of Economics.

[23] Windmeijer, F. (2005), "A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators", *Journal of Econometrics,* 126, 25-51.

[24] Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Annex A: Methodology used by the Business Software Alliance

To determine the software piracy rates and losses the *Business Software Alliance* has at his disposal huge amounts of information, being able to conduct extensive surveys in the population. In the estimates of 2010 presented in the eighth annual *Business Software Alliance* study, the *Business Software Alliance* relied on Ipsos Public Affairs that conducted more than 15000 surveys on business and consumer PC users. A brief description of the methodology is presented in the report and reproduced here.

The methodology was the following to obtain piracy rates:

$$Piracy\ rate = \frac{unlicensed\ software\ units}{total\ software\ units\ installed} \quad (A.1)$$

To obtain the total software units installed it was used the following:

$$Total\ Software\ Units\ Installed = \#PCs\ Getting\ Software \times Software\ Units\ perPC \quad (A.2)$$

The legitimate software units and the unlicensed software units are given by the following expressions:

$$Legitimate\ Software\ Units = \frac{Software\ Market\ Values}{Average\ Software\ Unit\ Price} \quad (A.3)$$

$$Unlicensed\ Software\ Units = Total\ Software\ units\ Installed - Legitimate\ Software\ Units \quad (A.4)$$

Finally the commercial value of unlicensed software is given by:

$$Commercial\ Value = \#Unlicend\ Software\ Units\ \times Average\ Software\ Unit\ Price \quad (A.5)$$

Business Software Alliance uses confidential information to achieve these results; it does provide us in their annually reports the basic methodology, omitting many variables that are used.

Annex B: Additional regressions

Table B 1: Two-step System GMM

VARIABLES	15	16	17	18	19
L.ln(losses)	**0.627*** **	**0.487*** **	**0.576*** **	**0.587*** **	**0.465*** **
	(6.761)	**(4.195)**	**(4.435)**	**(5.727)**	**(3.210)**
ln(GDP)	**0.326*** **	**0.303*** **	**0.361** **	**0.358*** **	**0.375** **
	(3.044)	**(3.035)**	**(2.433)**	**(3.111)**	**(2.161)**
ln(Mobile)	-0.014	**-0.061** **			
	(-0.334)	**(-1.990)**			
*ln(Mobile)*Change*		**0.443** **			
		(2.472)			
ln(Phone)			0.054	-0.105	
			(0.283)	(-0.811)	
*ln(Phone)*Change*			-0.058	-0.131	
			(-0.310)	(-0.766)	
ln(Fbis)					0.041
					(0.917)
*ln(Fbis)*Change*					0.075
					(1.229)
Empagric			0.009	-0.024	**0.020***
			(0.930)	(-1.555)	**(1.808)**
Empind	0.013	0.012	0.022		**0.022***
	(0.764)	(0.991)	(1.303)		**(1.843)**
Empserv	**-0.015***	**-0.013***		**-0.033** **	
	(-1.831)	**(-1.767)**		**(-2.335)**	
Labp			-0.012	0.005	-0.007
			(-1.219)	(0.757)	(-0.665)
Labs	-0.003	-0.003	**-0.016***		-0.008
	(-0.450)	(-0.514)	**(-1.701)**		(-1.079)
Labt	**0.019***	0.008		**0.023** **	
	(1.661)	(0.625)		**(2.350)**	
Unempyouth	**0.011** **	**0.013*** **	**0.010***	**0.011** **	**0.008***
	(2.012)	**(3.713)**	**(1.788)**	**(2.329)**	**(1.721)**
ln(Patents)	-0.028	-0.000	-0.009	-0.008	0.002
	(-0.901)	(-0.006)	(-0.212)	(-0.204)	(0.048)
ln(Trademarks)	0.054	**0.119*** **	0.009	0.025	0.057
	(1.053)	**(3.048)**	(0.126)	(0.395)	(0.790)
School	0.027	0.083	-0.154	0.059	-0.085
	(0.188)	(0.977)	(-0.883)	(0.500)	(-0.737)
Change	**0.342*** **	**-1.340** **	0.618	0.794	**0.341*** **
	(3.312)	**(-2.112)**	(0.887)	(1.239)	**(3.186)**
Constant	**-7.532** **	**-8.383*** **	-5.706	**-6.710*** **	**-7.852*** **
	(-2.572)	**(-3.872)**	(-1.636)	**(-2.633)**	**(-2.644)**
Observations	507	507	507	513	415
Countries	62	62	62	62	59
AR1	-3.520	-3.154	-3.217	-3.327	-2.452
p-value	[0.000]	[0.002]	[0.001]	[0.001]	[0.014]
AR2	-0.141	0.458	-0.645	-0.661	0.145
p-value	[0.888]	[0.647]	[0.519]	[0.509]	[0.885]
Instruments	53	53	53	53	51
Hansen	43.41	36.60	41.78	38.65	43.24
p-value	[0.185]	[0.394]	[0.200]	[0.308]	[0.109]

Notes: Dependent variable is *ln(losses)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(losses)* and *L.ln(GDP)* were considered as two endogenous instruments. Only one lag was used as instrument

Table B 2: Two-step System GMM

VARIABLES	20	21	22	23
L.ln(losses)	**0.497*** **	**0.603*** **	**0.555*** **	**0.547*** **
	(4.364)	**(5.889)**	**(4.110)**	**(3.930)**
ln(GDP)	**0.367*** **	**0.273* **	**0.410** **	**0.420** **
	(2.636)	**(1.761)**	**(2.351)**	**(2.411)**
ln(Mobile)	-0.045	0.047	0.091	0.047
	(-0.495)	(0.345)	(0.985)	(0.442)
*ln(Mobile)*Change*	**0.259* **	0.032		0.095
	(1.752)	(0.170)		(0.572)
ln(Fbis)	0.003	-0.027	0.008	0.011
	(0.069)	(-0.798)	(0.187)	(0.248)
Unempyouth			0.008	0.010
			(1.086)	(1.145)
Unemp	0.006	-0.020		
	(0.290)	(-1.162)		
ln(Patents)			0.053	**0.061* **
			(1.529)	**(1.738)**
ln(Trademarks)			-0.040	-0.045
			(-0.548)	(-0.598)
School	0.155	-0.049	0.107	0.123
	(1.381)	(-0.415)	(0.528)	(0.553)
R&D	**-0.402*** **	-0.091	**-0.305* **	**-0.336* **
	(-2.804)	(-0.881)	**(-1.700)**	**(-1.830)**
ln(Patres)	**0.163*** **			
	(2.931)			
ln(Patnon)	0.018			
	(0.544)			
ln(Tradres)		0.150		
		(1.346)		
ln(Tradnon)		-0.108		
		(-1.558)		
Change	-0.586	0.251	**0.357** **	-0.028
	(-0.920)	(0.344)	**(2.542)**	(-0.041)
Constant	**-9.845*** **	**-5.173* **	**-10.115*** **	**-10.392*** **
	(-2.941)	**(-1.764)**	**(-3.529)**	**(-3.635)**
Observations	479	473	412	412
Countries	75	75	66	66
AR1	-2.681	-2.704	-2.414	-2.336
p-value	[0.007]	[0.007]	[0.016]	[0.020]
AR2	0.583	0.0593	-0.254	-0.130
p-value	[0.560]	[0.953]	[0.799]	[0.896]
Instruments	51	51	51	51
Hansen	36.34	44.43	35.00	34.12
p-value	[0.406]	[0.132]	[0.516]	[0.511]

Notes: Dependent variable is *ln(losses)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(losses)* and *L.ln(GDP)* were considered as two endogenous instruments. Only one lag was used as instrument

Table B 3: Two-step System GMM (cont.)

Variables	24	25	26	27	28
L.ln(losses)	**0.548***** **(4.531)**	**0.590***** **(5.795)**	**0.566***** **(5.306)**	**0.620***** **(5.787)**	**0.554***** **(3.938)**
ln(GDP)	**0.437***** **(3.625)**	**0.320***** **(3.593)**	**0.373***** **(3.636)**	**0.328***** **(2.694)**	**0.312**** **(2.039)**
ln(Mobile)	-0.024 (-0.560)	-0.036 (-0.905)	-0.025 (-0.577)		-0.045 (-0.613)
*ln(Mobile)*Change*					**0.436***** **(2.637)**
ln(Net)				**-0.103**** **(-2.402)**	-0.069 (-0.755)
R&D				**-0.174*** **(-1.950)**	**-0.230**** **(-2.344)**
Yschoolpri	**-0.302***** **(-2.850)**	**-0.227***** **(-2.767)**	**-0.247***** **(-3.230)**		
Yschoolsec	-0.034 (-0.227)	-0.086 (-0.757)	0.008 (0.058)		
Pubexp				**0.144*** **(1.777)**	**0.192**** **(2.532)**
Labp	**-0.015**** **(-1.975)**	0.004 (0.602)			
Labs	**-0.024**** **(-2.265)**		**-0.013*** **(-1.656)**		
Labt		0.014 (1.399)	**0.013**** **(2.001)**		
Unemp	-0.000 (-0.038)	-0.010 (-0.908)	-0.004 (-0.401)	-0.015 (-0.832)	0.006 (0.339)
ln(Legal)	-0.027 (-1.187)	-0.001 (-0.064)	-0.013 (-0.588)	0.007 (0.295)	**0.060*** **(1.894)**
Change	**0.407***** **(3.539)**	**0.423***** **(4.282)**	**0.415***** **(3.709)**	**0.456***** **(4.947)**	**-1.315*** **(-1.930)**
Constant	**-5.536***** **(-2.630)**	**-4.971**** **(-2.561)**	**-5.965**** **(-2.486)**	**-7.041***** **(-2.919)**	**-7.526**** **(-2.527)**
Observations	520	526	520	442	442
Countries	64	64	64	66	66
AR1	-3.381	-3.589	-3.440	-3.147	-2.387
p-value	[0.001]	[0.000]	[0.001]	[0.002]	[0.017]
AR2	-0.453	-0.304	-0.220	-0.633	-0.352
p-value	[0.651]	[0.761]	[0.826]	[0.527]	[0.725]
Instruments	55	55	55	47	47
Hansen	43.81	47.60	44.44	44.70	35.36
p-value	[0.313]	[0.191]	[0.290]	[0.104]	[0.313]

Notes: Dependent variable is *ln(losses)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(losses)* and *L.ln(GDP)* were considered as two endogenous instruments. Only one lag was used as instrument.

Table B 4: One-Step System GMM for Losses per GDP

Variables	29	30	31	32	33
L.ln(Losses/GDP)	**0.662***** **(6.750)**	**0.639***** **(6.305)**	**0.672***** **(7.703)**	**0.667***** **(7.839)**	**0.561***** **(3.766)**
ln(Mobile)	-0.017 (-0.401)	-0.042 (-0.921)			
*ln(Mobile)*Change*		**0.207**** **(2.007)**			
ln(Phone)			-0.095 (-0.256)	-0.180 (-0.524)	
*ln(Phone)*Change*			0.117 (0.545)	0.044 (0.198)	
ln(Fbis)					-0.004 (-0.073)
*ln(Fbis)*Change*					0.059 (0.912)
Empagric			0.008 (0.699)	-0.003 (-0.141)	0.013 (1.301)
Empind	-0.004 (-0.209)	-0.012 (-0.540)	0.021 (1.379)		0.028 (1.520)
Empserv	**-0.014**** **(-2.138)**	**-0.015**** **(-2.409)**		-0.014 (-0.896)	
Labp			-0.011 (-1.214)	0.006 (0.873)	**-0.020***** **(-2.672)**
Labs	-0.007 (-1.061)	-0.004 (-0.589)	**-0.018**** **(-2.210)**		**-0.023**** **(-2.645)**
Labt	0.010 (1.297)	0.008 (0.888)		**0.020**** **(2.343)**	
Unempyouth	**0.027***** **(2.940)**	**0.025**** **(2.571)**	**0.027***** **(3.592)**	**0.023***** **(3.110)**	**0.025***** **(3.154)**
ln(Patents)	0.033 (0.929)	0.024 (0.663)	0.027 (0.570)	0.017 (0.337)	0.009 (0.236)
ln(Trademarks)	-0.047 (-0.895)	-0.022 (-0.460)	-0.044 (-0.574)	-0.021 (-0.251)	-0.025 (-0.278)
School	-0.065 (-0.357)	-0.098 (-0.501)	-0.029 (-0.159)	0.055 (0.343)	0.213 (0.911)
Change	**0.311***** **(2.977)**	-0.572 (-1.430)	-0.132 (-0.164)	0.059 (0.072)	0.232 (1.584)
Constant	-0.876 (-0.312)	-0.605 (-0.209)	-1.469 (-0.647)	-2.628 (-0.959)	-5.452 (-1.492)
Observations	507	507	507	513	415
Countries	62	62	62	62	59
AR1	-4.241	-4.215	-4.190	-4.051	-3.124
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]
AR2	-0.507	-0.191	-0.446	-0.448	0.375
p-value	[0.612]	[0.849]	[0.656]	[0.654]	[0.708]
Instruments	51	51	51	51	50
Hansen	36.28	36.29	35.83	36.84	37.40
p-value	[0.409]	[0.362]	[0.383]	[0.339]	[0.274]

Notes: Dependent variable is *ln(losses/GDP)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(Losses/GDP)* was considered as an endogenous instruments. Lags 1 through 3 were used (columns 29 to 33).

Table B 5: One-Step System GMM for Losses per GDP

Variables	34	35	36	37
L.ln(Losses/GDP)	**0.643*** **	**0.687*** **	**0.611*** **	**0.609*** **
	(7.485)	**(9.005)**	**(5.504)**	**(5.352)**
ln(Mobile)	0.046	0.114	**0.289***	0.273
	(0.491)	(0.944)	**(1.646)**	(1.458)
*ln(Mobile)*Change*	0.123	0.049		0.041
	(0.867)	(0.401)		(0.240)
ln(Phone)			**0.611*** **	**0.609*** **
			(5.504)	**(5.352)**
*ln(Phone)*Change*			**0.289***	0.273
			(1.646)	(1.458)
ln(Fbis)	-0.028	**-0.046** **		-0.033
	(-0.927)	**(-2.054)**		(-0.980)
Unempyouth			**0.015** **	**0.016** **
			(2.051)	**(2.117)**
Unemp	0.011	-0.002		
	(0.562)	(-0.121)		
ln(Patents)			0.047	0.049
			(1.612)	(1.525)
ln(Trademarks)			-0.050	-0.050
			(-1.123)	(-1.119)
School	-0.007	-0.074	0.017	0.010
	(-0.084)	(-0.938)	(0.131)	(0.075)
R&D	**-0.188** **	-0.053	-0.179	-0.187
	(-2.239)	(-0.798)	(-1.638)	(-1.642)
ln(Patres)	**0.061** **			
	(2.060)			
ln(Patnon)	-0.003			
	(-0.129)			
ln(Tradres)		0.057		
		(1.333)		
ln(Tradnon)		-0.040		
		(-0.817)		
Change	-0.219	0.060	**0.236***	0.067
	(-0.381)	(0.122)	**(1.881)**	(0.100)
Constant	**-3.147*** **	**-2.134***	**-4.194** **	**-4.091** **
	(-3.027)	**(-1.785)**	**(-2.178)**	**(-2.007)**
Observations	479	473	412	412
Countries	75	75	66	66
AR1	-3.875	-4.168	-3.373	-3.271
p-value	[0.000]	[0.000]	[0.001]	[0.001]
AR2	0.586	0.259	-0.0911	-0.0522
p-value	[0.558]	[0.796]	[0.927]	[0.958]
Instruments	50	61	50	50
Hansen	45.21	54.98	40.05	39.13
p-value	[0.116]	[0.171]	[0.295]	[0.290]

Notes: Dependent variable is *ln(losses/GDP)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(Losses/GDP)* was considered as an endogenous instruments. Lags 1 through 3 were used (column 34, 36 and 37) and lags 1 through 4 were used in column 35.

Table B 6: One-Step System GMM for Losses per GDP (cont.)

Variables	38	39	40	41	42
L.ln(Losses/GDP)	**0.550*** **	**0.584*** **	**0.523*** **	**0.692*** **	**0.700*** **
	(4.301)	**(5.507)**	**(4.438)**	**(7.502)**	**(6.888)**
ln(Mobile)	-0.014	-0.047	-0.027		0.038
	(-0.366)	(-1.580)	(-0.706)		(0.374)
*ln(Mobile)*Change*					**0.281** **
					(2.266)
ln(Net)				**-0.118** **	-0.171
				(-2.474)	(-1.602)
R&D				-0.034	-0.096
				(-0.522)	(-1.222)
Yschoolpri	-0.204	-0.072	-0.196		
	(-1.252)	(-0.525)	(-1.200)		
Yschoolsec	0.042	0.079	0.078		
	(0.278)	(0.655)	(0.513)		
Pubexp				0.163	0.155
				(1.586)	(1.494)
Labp	-0.013	0.003			
	(-1.497)	(0.623)			
Labs	**-0.023** **		**-0.012* **		
	(-2.250)		**(-1.943)**		
Labt		**0.014** **	**0.014** **		
		(2.275)	**(2.044)**		
Unemp	**0.027** **	**0.020* **	**0.025** **	0.015	0.023
	(2.110)	**(1.894)**	**(1.970)**	(0.903)	(1.242)
ln(Legal)	0.000	-0.000	0.002	0.026	0.032
	(0.008)	(-0.023)	(0.079)	(1.313)	(1.408)
Change	**0.371*** **	**0.342*** **	**0.390*** **	**0.365*** **	-0.831
	(3.177)	**(3.493)**	**(3.503)**	**(3.291)**	(-1.626)
Constant	-1.493	**-3.810*** **	**-3.133* **	**-3.409*** **	**-3.396*** **
	(-0.961)	**(-2.674)**	**(-1.810)**	**(-3.573)**	**(-3.486)**
Observations	520	526	520	442	442
Countries	64	64	64	66	66
AR1	-4.350	-4.764	-4.405	-4.236	-3.800
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
AR2	-0.403	-0.314	-0.335	-1.042	-0.611
p-value	[0.687]	[0.753]	[0.738]	[0.297]	[0.541]
Instruments	53	53	53	57	46
Hansen	46.51	46.31	46.19	56.00	41.62
p-value	[0.191]	[0.196]	[0.200]	[0.126]	[0.119]

Notes: Dependent variable is *ln(losses/GDP)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(Losses/GDP)* was considered as an endogenous instruments. Lags 1 through 3 were used (columns 38, 39, 40 and 42) and lags 1 through 4 (column 41).

Table B 7: Two-Step System GMM for Losses per GDP

Variables	43	44	45	46	47
L.ln(Losses/GDP)	**0.660*** **	**0.667*** **	**0.678*** **	**0.668*** **	**0.616*** **
	(7.156)	**(7.289)**	**(8.029)**	**(6.816)**	**(4.286)**
ln(Mobile)	-0.007	-0.015			
	(-0.176)	(-0.386)			
*ln(Mobile)*Change*		**0.160***			
		(1.658)			
ln(Phone)			-0.165	-0.144	
			(-0.538)	(-0.355)	
*ln(Phone)*Change*			0.113	-0.054	
			(0.731)	(-0.235)	
ln(Fbis)					0.038
					(0.869)
*ln(Fbis)*Change*					-0.009
					(-0.163)
Empagric			0.008	0.002	0.013
			(0.850)	(0.124)	(1.361)
Empind	-0.010	-0.012	0.016		0.017
	(-0.441)	(-0.628)	(0.824)		(0.666)
Empserv	**-0.015****	**-0.017****		-0.008	
	(-2.190)	**(-2.278)**		(-0.528)	
Labp			-0.011	0.004	**-0.021*** **
			(-1.332)	(0.814)	**(-2.760)**
Labs	-0.008	-0.006	**-0.017****		**-0.022****
	(-1.381)	(-0.889)	**(-2.234)**		**(-2.523)**
Labt	0.011	0.010		**0.018***	
	(1.242)	(1.375)		**(1.882)**	
Unempyouth	**0.024****	**0.023*** **	**0.023*** **	**0.020*** **	**0.022****
	(2.541)	**(2.592)**	**(3.058)**	**(2.784)**	**(2.578)**
ln(Patents)	0.031	0.027	0.036	0.020	0.017
	(0.674)	(0.572)	(0.825)	(0.531)	(0.504)
ln(Trademarks)	-0.047	-0.032	-0.051	-0.027	-0.032
	(-0.863)	(-0.578)	(-0.955)	(-0.361)	(-0.440)
School	-0.123	-0.144	-0.045	0.015	0.159
	(-0.618)	(-1.000)	(-0.278)	(0.099)	(0.753)
Change	**0.280*** **	-0.415	-0.134	0.395	0.235
	(2.904)	(-1.066)	(-0.238)	(0.467)	(1.442)
Constant	0.112	0.393	-0.806	-2.486	-3.992
	(0.035)	(0.152)	(-0.377)	(-0.764)	(-1.291)
Observations	507	507	507	513	415
Countries	62	62	62	62	59
AR1	-3.679	-3.725	-3.590	-3.338	-2.683
p-value	[0.000]	[0.000]	[0.000]	[0.001]	[0.007]
AR2	-0.636	-0.287	-0.613	-0.701	-0.145
p-value	[0.525]	[0.774]	[0.540]	[0.483]	[0.884]
Instruments	51	51	51	51	50
Hansen	36.28	36.29	35.83	36.84	37.40
p-value	[0.409]	[0.362]	[0.383]	[0.339]	[0.274]

Notes: Dependent variable is *ln(losses/GDP)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(Losses/GDP)* was considered as an endogenous instruments. . Lags 1 through 3 were used (columns 43 to 47).

Table B 8: Two-Step System GMM for Losses per GDP

Variables	48	49	50	51
L.ln(Losses/GDP)	**0.634*** **	**0.680*** **	**0.614*** **	**0.606*** **
	(6.359)	**(7.534)**	**(4.656)**	**(4.464)**
ln(Mobile)	0.048	0.155	0.223	0.194
	(0.358)	(1.329)	(1.094)	(0.942)
*ln(Mobile)*Change*	0.071	0.016		0.074
	(0.428)	(0.109)		(0.460)
ln(Fbis)	-0.022	**-0.052***	-0.019	-0.019
	(-0.612)	**(-1.757)**	(-0.485)	(-0.481)
Unempyouth			**0.015***	0.015
			(1.690)	(1.620)
Unemp	0.007	0.001		
	(0.290)	(0.051)		
ln(Patents)			0.046	0.050
			(1.532)	(1.498)
ln(Trademarks)			-0.049	-0.053
			(-1.159)	(-1.181)
School	0.011	-0.066	-0.033	-0.049
	(0.123)	(-0.799)	(-0.211)	(-0.299)
R&D	**-0.178***	-0.034	-0.204	-0.223
	(-1.719)	(-0.393)	(-1.574)	(-1.578)
ln(Patres)	**0.056***			
	(1.813)			
ln(Patnon)	-0.003			
	(-0.150)			
ln(Tradres)		0.054		
		(0.972)		
ln(Tradnon)		-0.038		
		(-0.589)		
Change	-0.028	0.200	0.208	-0.090
	(-0.045)	(0.332)	(1.542)	(-0.140)
Constant	**-3.367** **	-2.487	-3.217	-2.951
	(-2.403)	(-1.597)	(-1.437)	(-1.364)
Observations	479	473	412	412
Countries	75	75	66	66
AR1	-3.271	-3.516	-2.569	-2.490
p-value	[0.001]	[0.000]	[0.010]	[0.013]
AR2	0.372	0.314	-0.280	-0.236
p-value	[0.710]	[0.753]	[0.779]	[0.814]
Instruments	50	61	50	50
Hansen	45.21	54.98	40.05	39.13
p-value	[0.116]	[0.171]	[0.295]	[0.290]

Notes: Dependent variable is *ln(losses/GDP)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(Losses/GDP)* was considered as an endogenous instruments. Lags 1 through 3 were used (column 48, 50 and 51) and lags 1 through 4 were used in column 49.

Table B 9: Two-Step System GMM for Losses per GDP (Cont.)

Variables	52	53	54	55	56
L.ln(Losses/GDP)	**0.605*****	**0.604*****	**0.542*****	**0.699*****	**0.732*****
	(4.115)	**(4.845)**	**(4.509)**	**(5.216)**	**(6.130)**
ln(Mobile)	-0.012	-0.032	-0.026		0.078
	(-0.293)	(-0.913)	(-0.664)		(0.555)
*ln(Mobile)*Change*					0.322
					(1.595)
ln(Net)				**-0.098****	-0.228
				(-2.321)	(-1.634)
R&D				-0.050	-0.115
				(-0.589)	(-1.165)
Yschoolpri	-0.147	-0.048	-0.164		
	(-0.843)	(-0.360)	(-1.021)		
Yschoolsec	0.077	0.111	0.115		
	(0.447)	(0.817)	(0.631)		
Pubexp				0.142	0.203
				(1.143)	(1.525)
Labp	-0.014	0.003			
	(-1.537)	(0.529)			
Labs	**-0.024****		**-0.012***		
	(-2.177)		**(-1.836)**		
Labt		**0.015***	0.016		
		(1.724)	(1.460)		
Unemp	**0.027****	**0.021****	**0.024****	0.015	0.027
	(2.455)	**(2.013)**	**(2.059)**	(0.644)	(1.320)
ln(Legal)	0.009	0.009	0.011	0.022	0.047
	(0.315)	(0.548)	(0.435)	(0.722)	(1.312)
Change	**0.343****	**0.315*****	**0.375*****	**0.331****	-0.988
	(2.406)	**(2.904)**	**(3.076)**	**(2.608)**	(-1.254)
Constant	-1.718	**-4.238****	**-3.595***	**-3.205*****	**-3.624*****
	(-0.922)	**(-2.559)**	**(-1.852)**	**(-2.797)**	**(-3.320)**
Observations	520	526	520	442	442
Countries	64	64	64	66	66
AR1	-3.270	-3.348	-3.335	-3.346	-3.070
p-value	[0.000]	[0.001]	[0.001]	[0.001]	[0.002]
AR2	-0.261	-0.340	-0.352	-1.015	-0.654
p-value	[0.794]	[0.734]	[0.725]	[0.310]	[0.513]
Instruments	53	53	53	57	46
Hansen	46.51	46.31	46.19	56.00	41.62
p-value	[0.191]	[0.196]	[0.200]	[0.126]	[0.119]

Notes: Dependent variable is *ln(losses/GDP)*. Robust t-statistics in parentheses; *, ** and *** represent statistical significance at 10%, 5% and 1% respectively. *L.ln(Losses/GDP)* was considered as an endogenous instruments. Lags 1 through 3 were used (columns 52, 53, 54 and 56) and lags 1 through 4 (column 55).