# THE IMPACT OF TECHNICAL COPY PROTECTION AND INTERNET SERVICES USAGE ON SOFTWARE PIRACY

## - AN INTERNATIONAL SURVEY ON SEQUENCER SOFTWARE PIRACY -

Djekic, Petar, Dept. of Media Management, University of Cologne, Pohligstrasse 1, 50969 Cologne, Germany, petar.djekic@uni-koeln.de, www.mm.uni-koeln.de

Loebbecke, Claudia, Dept. of Media Management, University of Cologne, Pohligstrasse 1, 50969 Cologne, Germany, claudia.loebbecke@uni-koeln.de, www.mm.uni-koeln.de

## Abstract

*Software piracy has recently gained enormous attention, not only in the context of P2P-networks. As one countermeasure against software piracy, publishers have been implementing Digital Rights Management systems such as technical copy protection measures into their software products. This paper examines the impact of different technical copy protection measures and Internet services usage on software piracy using data from an internationally organized online survey. The results show that technical protection measures fail to achieve their protection goals, as none of the studied protection measures completely avoids piracy. A higher level of copy protection does not always make a legal software installation more likely. In contrast, a low level of protection does not necessarily lead to intense illegal copying. P2P- and Chat-networks compromise the security of technical copy protections as they provide access to cracked software copies, fostering software piracy. Based on our results, we discuss the impact of our findings on the publishers' anti-piracy strategy from an economic point of view and present possible security improvements for hardware- and software-based copy protections.*

*Keywords: Software Piracy, Technical Copy Protection, Internet Services Usage, Anti-Piracy Strategy.*

# 1 INTRODUCTION

According to the Business Software Alliance (2003a), software publishers were losing USD 11 billion in 2002 due to illegal reproduction and usage of their products. This challenges designers in their efforts to develop software for legal diffusion and commercial exploitation in numerous ways. Some literature (Conner & Rumelt 1991, Givon & Mahajan & Muller 1995, Gopal & Sanders 1997, Haruvy & Mahajan & Prasad 2004) points to technical copy protection as a preventive measure against software piracy. Other authors (Devanbu & Stubblebine 2000, Kingpin 2000, Anckaert & De Sutter & De Bosschere 2004) note that technical copy protection can be cracked; the check for the copy protection token is removed from the application software, while the software's functionality remains unchanged. The cracked software can be easily copied and shared on different Internet services (e.g. P2P-networks, FTP-servers). The insecurity of technical copy protection has been acknowledged, yet the resulting effects on software piracy have barely been investigated.

In this context, this paper examines the impact of technical copy protection and Internet services usage on software piracy. Our empirical study is based on an online survey among the users of so-called sequencer software in 45 countries. As a particular kind of music software, sequencer software is used for creating and arranging music songs and controlling external music equipment like synthesizers. Sequencer software was chosen as application software for this study because its users are commonly considered as technology-friendly, using the Internet more often than the average population. Also, many different types of copy protection are applied to protect sequencer software, allowing us to analyze a broad range of technical copy protection measures.

# 2 TECHNICAL COPY PROTECTION OVERVIEW

## 2.1 Levels of technical copy protection

There are two levels of technical copy protection (Gopal et al. 1997): software-based and hardware-based protection. Software-based protections include software tokens, watermarking, code partitioning (Devanbu et al. 1991) or encryption. Hardware-based protection consists of hardware tokens. A token describes an item which is checked for by the application software (here the sequencer software) upon installation or launch. If the application software cannot find the token, further execution of the program is impossible. The check for the token is referred to as 'authorization' or 'activation'. Table 1 shows the copy protection measures studied in this paper. As water marking, encryption and code partitioning do not directly affect the ability to copy software, they are not taken into account in this study.

The measure 'Serial Number' uses a long random number as a token which has to be entered by the user during installation or launch of the application software. In case of 'CD-Check', the token is a file placed on the Installation CD-ROM. During installation or launch of the application software, the software checks the presence of the CD-ROM that contains the token. When the application software uses the measure 'Multiple CD-Checks', the protection is the same, as CD-Check except that the token is placed across two or more CD-ROMs which all have to be present for successful authorization. The measure 'Challenge-Response' also uses a long number as a token. Different from the measure Serial Number, with the measure Challenge-Response the token is unique for each software installation and cannot be reused to illegally install the same application software on different computers when only one license is available. The measure 'Dongle', also referred to as 'Hardware-Key', uses a small stick as a token which can either be plugged into the USB, serial or printer port. The token has to be present in the appropriate port during the application software runtime. The measure 'Expansion Card' uses a PCI-Card inside the computer as a token. The token in this case is also used for additional

functionality (i.e. sound processing, audio interface) within the application software. Table 1 summarizes the different measures studied in this paper and the sequencer software they protect.

| Measure | Protection Level | Sequencer Software |
|---|---|---|
| Serial Number | Software | Cakewalk Sonar, Synapse Orion |
| CD-Check | Software | MOTU Digital Performer |
| Multiple CD-Checks | Software | Propellerheads Reason |
| Challenge-Response | Software | ImageLine Fruityloops |
| Dongle | Hardware | Steinberg Cubase, Steinberg Nuendo, Apple Logic |
| Expansion Card | Hardware | Digidesign ProTools |

*Table 1. Technical Copy Protection Measures Studied*

Hardware-based protection measures are generally harder to circumvent than software-based ones. As hardware tokens cannot be easily duplicated or generated, they provide stronger copy protection than software-based measures.

The above mentioned technical copy protections used to protect the studied sequencer software are similar to those found in other application software like operating systems (i.e., Windows XP's Challenge-Response) or graphics software (e.g. Dongle-protected CAD- and 3D-software).

### 2.2 Costs of technical copy protection

The implementation of technical copy protection measures causes direct and indirect costs for the publisher. The direct costs for the publisher depend on the (1) level of protection and (2) origin of the protection.

(1) Overall hardware-based copy protections are more expensive than software-based ones. Software-based protection only causes fix costs during the implementation of the copy protection methods in the application software source code. Hardware-based copy protection causes additional variable costs during software production. For each application software copy, one hardware token needs to be provided.

(2) When implementing technical copy protection measures, the publisher has to decide if he wants to rely on 3rd party copy protection methods (e.g. SafeDisc, WiBU Key) or develop his own proprietary methods. In the first case, the publisher has to pay license fees to the copy protection producer. For example, one WiBU Key license costs 50 USD including the hardware token. In the second case, the publisher faces additional development costs.

The indirect costs are independent of the origin or level of the copy protection and result from potential bugs and incompatibilities of the copy protection due to the user's computer configuration (Microsoft 2004). These lead to additional support and development costs for the publisher (e.g. support hotline, software updates).

## 3 RESEARCH FRAMEWORK: FACTORS OF SOFTWARE PIRACY

Literature on technical copy protection either focuses on the technical vulnerability of single protection measures (Devanbu et al. 2000, Kingpin 2000, Anckaert et al. 2004) or on the strategic implications of technical copy protection use e.g. network effects (Conner et al. 1991, Shy & Thisse 1999). Empirical research on software piracy factors includes studies on ethical attitude (Lending & Slaughter 2001, Wagner & Sanders 2001, Hinduja 2003, Kini & Ramakrishna & Vijayaraman 2004), cultural aspects (Husted 2000, Marron & Steel 2000) or behavior (Christensen & Eining 1991, Cheng

& Sims & Teegen 1997, Peace & Galletta & Thong 2003). The role of pricing on piracy has also been investigated (Slive & Bernhardt 1998, Gopal & Sanders 2000).

As noted before, technical copy protections can be cracked (Devanbu et al. 2000, Kingpin 2000, Anckaert et al. 2004). Also, different Internet services, such as P2P-networks and public websites, along with broadband connections have simplified the access to illegal software copies (Hinduja 2001, Business Software Alliance 2003b). The use of these Internet services as a distribution network for cracked software copies may diminish the distinction between different levels and measures of technical copy protection.

Our research framework (Figure 1) includes the dependent variable *Type of Software installation (SI)*, which is a binary variable with the value '1' for a legal software installation and '0' for an illegal, pirated installation. We introduce eleven independent variables: The first one is *Level of Copy Protection (CP)*, which represents the two levels of technical copy protection (Copy Protection Context). The value '0' represents software-based technical copy protection used in the sequencer software installation; and the value '1' represents implementation of hardware-based technical copy protection. The next 7 independent variables refer to the Internet context. We build on Hinduja (2003), who differentiates between five 'software piracy' media as ways to obtain illegal software copies. We add P2P-networks as a sixth one and replace instant messaging programs with eMail. In detail, the six variables are *World Wide Web Usage (WWW)*, *eMail Usage (EM)*, *Newsgroups Usage (NG)*, *FTP Usage (FTP)*, *P2P Usage (P2P)* and *Chat Usage (CHAT)*. These independent 'Internet Context' variables measure the usage of each software piracy medium by the user who owns the sequencer software installation on a relative four-point scale ranging: 'Never', 'Rarely', 'Occasionally' and 'Often'.

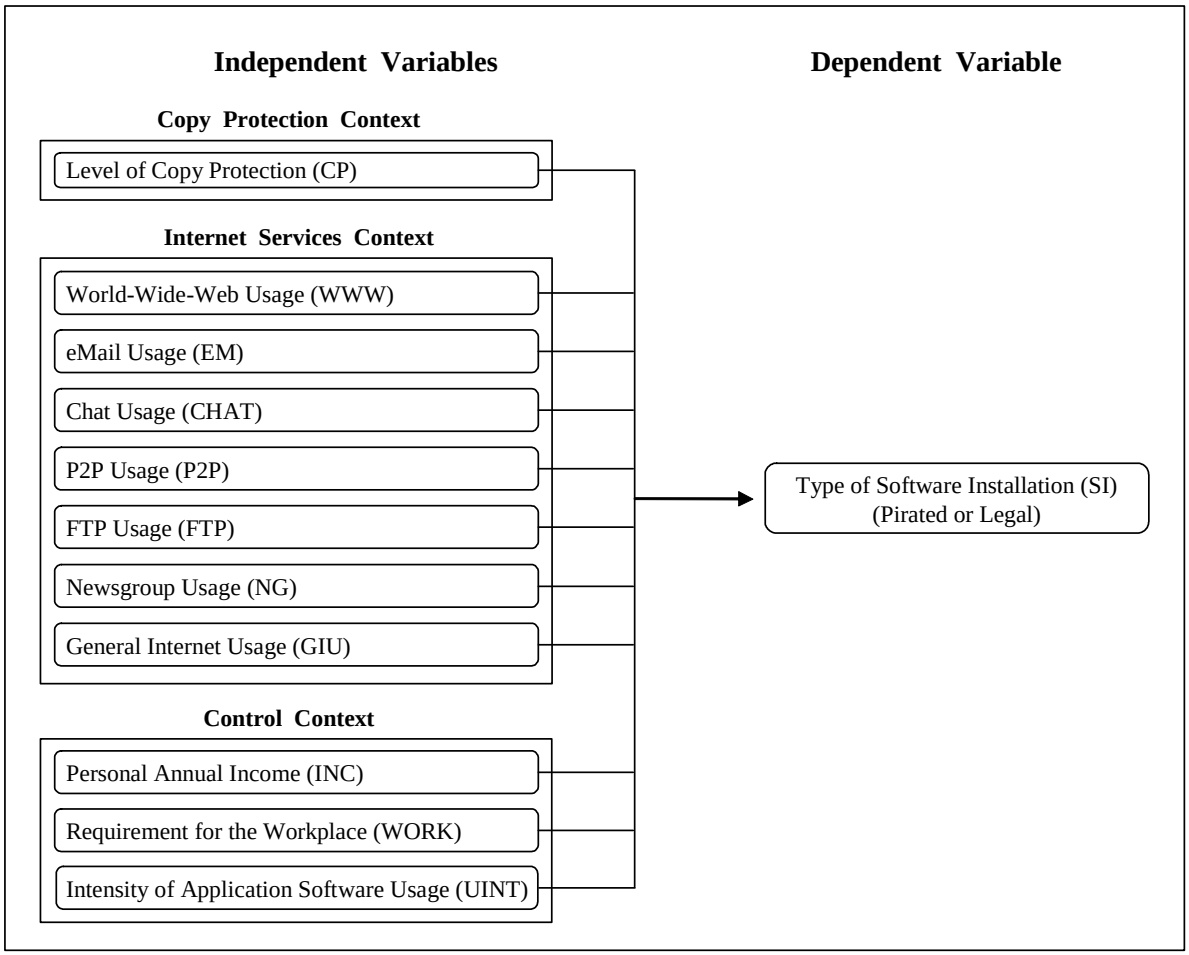


*Figure 1. Research Framework: Factors of Software Piracy*

Further, we add the independent variable *General Internet Usage (GIU)* to measure the effect of overall Internet usage on software piracy. In addition, describing the 'Control Context', *Personal Income (INC), Intensity of Application Software Usage (UINT)* and *Requirement for the Workplace (WORK)* are included as independent variables. The impact of these variables on software piracy has been investigated in earlier studies (Cheng et al. 1997, Chiang & Assane 2002, Hinduja 2003, Business Software Alliance 2003b). They are included to assess the importance of technical copy protection with regard to these factors of piracy. The *Personal Annual Income* is measured in this study as the user's annual income in US Dollar. *Intensity of Application Software Usage* indicates how much time is spent using the sequencer software in hours per week. Finally, *Requirement for the Workplace (WORK)* measures the requirement of the sequencer software for the workplace. This variable is proxied via the share of the income earned from working with sequencer software, i.e. the percentage of music-specific income over total personal income.

# 4 RESEARCH METHODOLOGY

In the descriptive part we first contrast the 'Regional Piracy Rates' surveyed in this study with the rates published by the Business Software Alliance. A Piracy Rate is the percentage of pirated application software installations (i.e. software without license) over all application software installations. The Business Software Alliance publishes Piracy Rates for six world regions in their 2002 annual piracy report (Business Software Alliance 2003a). These Regional Piracy Rates are calculated for standard applications like spreadsheets or word processors without technical copy protection. One should expect the Piracy Rates of 'our' copy-protected software to be below the ones of the unprotected standard applications. In addition, we measure and compare the Piracy Rates of each technical copy protection and sequencer software.

We then perform a binary logistic regression analysis on the variables presented in our research framework (Figure 1). The binary logistic regression is used as the dependent variable is dichotomous. It is applied to predict the probability of an event under specific circumstances. For other applications of the logistic regression in the context of IS-related topics, see for instance Zhu, Kraemer and Xu (2002). In our setting, the logistic regression is applied to predict whether a given installation of sequencer software is pirated or legal.

# 5 EMPIRICAL STUDY

## 5.1 Survey Design, Data Collection and Quality

The online survey questionnaire was available for 53 days from November 8 to December 31, 2003. We chose an online survey to collect the data for two reasons: Firstly, a 'traditional', paper-based survey would have been too costly on an international level. Secondly, participants 'perceive' electronic surveys as more anonymous than traditional ones (Kiesler & Sproull 1986). As the survey deals with copyright infringement, anonymity has been essential for the participants.

In October 2003, the questionnaire was pre-tested with eleven users of sequencer software. The afterwards revised questionnaire was offered in English, French and German on the public website 'www.musicsurvey.uni-koeln.de'.

Participants were invited to the survey either by eMail, the Mail-Group, or by posting on public message boards dealing with music software, the Board-Group. As the composition of both samples is different, the data of each group is analyzed separately.

The required eMail addresses for the Mail-Group where collected from musicians' Internet yellow pages. Of the 2,742 invitations sent, 219 resulted in a valid, completed questionnaire (7.99 %). Email feedback showed that the low response rate in this group was partially due to (1) the cover letter being

regarded as spam and (2) outdated eMail addresses in the musicians' Internet yellow pages. Nobody in this group could fill out multiple questionnaires as each person needed an individual key provided with the cover letter to access the survey website.

2,159 persons from thirteen message boards were invited to the Board-Group. Due to the participants' self-selection, we could not define the Board-Group sample in advance. In this group, 575 participants filled out a questionnaire (26.63%). Multiple questionnaires were avoided (1) by placing a cookie with the current survey status and (2) by saving the IP-address of each participant. 3% of Board-Group participants completed more than one questionnaire; the multiple questionnaires were excluded from the analysis.

Two indicators measure the quality of our survey data:

(1) The country specified in the questionnaire and that revealed by the IP-address matches in 97% of the cases in the Mail-Group and 93% in the Board-Group.

(2) The percentage of female and male participants in both groups (Table 2) compares to the one of the Electronic Musician subscribers (Electronic Musician 2004).

| Data | Male | Female |
|---|---|---|
| Board-Group | 98 | 2 |
| Mail-Group | 92 | 8 |
| Electronic Musician subscribers | 89 | 11 |

*Table 2. Gender Percentages*

### 5.2 Descriptive Statistics

All Piracy Rates were calculated as percentage of sequencer software installations without license over all installations of sequencer software. For each subtype of Piracy Rate, the data was either split by: the region of the sequencer software installation, the copy protection measure applied to the sequencer software or the name of the sequencer software.

$$PR = \frac{\text{Number of sequencer installations without license}}{\text{Total number of sequencer installations}}$$

| Region | Mail-Group | Board-Group | BSA |
|---|---|---|---|
| World Piracy Rate | 19 | 22 | 39 |
| North-America | 10 | 17 | 24 |
| Western Europe | 24 | 25 | 35 |
| Middle East/Africa | No data | 27 | 49 |
| Latin America | No data | 67 | 55 |
| Asia/Pacific | 21 | 19 | 55 |
| Eastern Europe | 100 | 63 | 71 |

*Table 3. Regional Piracy Rates (in %)*

Table 3 shows the Regional Piracy Rates resulting from the survey data compared to the rates from the Business Software Alliance (BSA) study (Business Software Alliance 2003a). The 'World Piracy Rate' shows the Piracy Rate over all regions.

| Copy Protection | Mail-Group | Board-Group |
|---|---|---|
| Serial Number | 25 | 16 |
| CD-Check | 24 | 18 |
| Multiple CD-Checks | 33 | 35 |
| Challenge-Response | 50 | 29 |
| Dongle | 19 | 20 |
| Expansion Card | 7 | 9 |

*Table 4. Piracy Rates per Technical Copy Protection (in %)*

Table 4 shows the Piracy Rates for each technical copy protection, while Table 5 shows the Piracy Rates for the different kinds of sequencer software presented in Table 1.

| Sequencer Software | Mail-Group | Board-Group |
|---|---|---|
| Apple Logic | 14 | 12 |
| Cakewalk Sonar | 20 | 9 |
| Digidesign Protools | 10 | 10 |
| ImageLine Fruityloops | 50 | 32 |
| MOTU Digital Performer | 16 | 16 |
| Propellerheads Reason | 33 | 35 |
| Steinberg Cubase | 27 | 27 |
| Steinberg Nuendo | 21 | 26 |
| Synapse Orion | 100 | 18 |

*Table 5. Piracy Rates per Sequencer Software (in %)*

Table 6 presents mean, median, and standard deviation for the independent variables of our research framework model. This table demonstrates some differences in the characteristics between the Mail-Group and the Board-Group. The *Personal Annual Income (INC)* is 20% and the *Requirement for the Workplace (WORK)* is 80% higher in the Mail-Group than in the Board-Group. Obviously, Mail-Group participants use their sequencer software more for business purposes than Board-Group members do. The *Intensity of Application Software Usage (UINT)* is 16% higher in the Mail-Group, which supports the previous statement.

| Variable | Mail-Group | | | Board-Group | | |
|---|---|---|---|---|---|---|
| | Mean | Median | Std. Dev. | Mean | Median | Std. Dev. |
| INC ($) | 30.480 | 35.000 | 13.884 | 25.243 | 25000 | 15334 |
| WORK (%) | 63 | 88 | 31% | 22 | 13 | 32 |
| UINT (h) | 18,0 | 8 | 18,5 | 14 | 8,0 | 14,4 |
| GIU (h) | 15,0 | 13,0 | 10,8 | 19,9 | 18 | 9,8 |
| CP | - | Hardware | - | - | Hardware | - |
| WWW | - | Often | - | - | Often | - |
| EM | - | Often | - | - | Often | - |
| FTP | - | Occasionally | - | - | Occasionally | - |
| NG | - | Rarely | - | - | Rarely | - |
| CHAT | - | Never | - | - | Rarely | - |
| P2P | - | Never | - | - | Rarely | - |

*Table 6. Mean, Median, and Standard Deviation Values for Independent Variables*

The *General Internet Usage (GIU)* and *P2P Usage (P2P)* are higher in the Board-Group. The average usage of the other Internet services (WWW, FTP, CHAT, NG and EM) and the average *Level of Copy Protection (CP)* are the same for both groups. Finally, the usage of P2P- and Chat-networks is more intense in Board-Group. Additional calculations show that the *P2P Usage* and *Chat Usage* frequency of 'Often' is 114% and 90% respectively higher in the Board-Group than in the Mail-Group.

### 5.3 Logistic Regression

We use the following binary logistic regression model (for legend see Figure 1):

$$\ln[SI/1-SI] = \alpha + \beta_1 \cdot CP + \beta_2 \cdot INC + \beta_3 \cdot WORK + \beta_4 \cdot UINT + \beta_5 \cdot GIU + \beta_6 \cdot WWW + \beta_7 \cdot EM + \beta_8 \cdot NG + \beta_9 \cdot FTP + \beta_{10} \cdot CHAT + \beta_{11} \cdot P2P$$

The binary logistic regression model is estimated with the 'Binary Logistic' method provided in SPSS. The outcomes of the binary logistic regression for each model variable in the Mail- and Board-Group are displayed in Table 7.

Testing our model of software piracy with the binary logistic regression means testing if the values of the $\beta_{1-11}$-coefficients are non-zero: Positive and significant coefficients raise the probability of legal application software installations; negative and significant coefficients lower the probability. The $\beta$-coefficient value indicates the direction of the influence of a variable; the Exp ($\beta$)-value points to the strength of the influence.

| Variable | Mail-Group | | | | Board-Group | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta$ | Std.Err. | Sig. | Exp (b) | $\beta$ | Std.Err. | Sig. | Exp (b) |
| CP | 0,846 | 0,322 | 0,009 | 2,329 | 0,382 | 0,210 | 0,069 | 1,465 |
| INC | 0,475 | 0,124 | 0,000 | 1,608 | 0,447 | 0,076 | 0,000 | 1,564 |
| WORK | 0,294 | 0,120 | 0,015 | 1,341 | 0,299 | 0,084 | 0,000 | 1,348 |
| UINT | 0,032 | 0,012 | 0,006 | 1,033 | 0,046 | 0,010 | 0,000 | 1,047 |
| GIU | -0,048 | 0,084 | 0,563 | 0,953 | 0,044 | 0,058 | 0,452 | 1,045 |
| WWW | 0,476 | 0,184 | 0,010 | 1,610 | 0,576 | 0,165 | 0,000 | 1,780 |
| NG | -0,060 | 0,174 | 0,730 | 0,942 | 0,057 | 0,096 | 0,551 | 1,059 |
| EM | -0,321 | 1,100 | 0,771 | 0,726 | 0,513 | 0,218 | 0,019 | 1,670 |
| CHAT | -0,143 | 0,175 | 0,415 | 0,867 | -0,206 | 0,100 | 0,039 | 0,814 |
| P2P | -0,097 | 0,170 | 0,568 | 0,908 | -0,889 | 0,103 | 0,000 | 0,411 |
| FTP | 0,270 | 0,174 | 0,120 | 1,311 | 0,070 | 0,113 | 0,535 | 1,073 |

*Table 7. Values for Independent Model Variables*

We assess the overall Model Fit with three Goodness-of-Fit tests (Table 8):

(1) The Likelihood Ratio-Test (Menard 1995): It analyzes if the independent variables have an explanatory power. Significance of this test indicates a good fit of the model to the data, which is the case in both groups.

(2) The Hosmer-Lemeshow-Chi (Hosmer & Lemeshow 2000): It compares our model to a model with a perfect fit. Non- significance of the test indicates a good fit of the model to the data, which is the case in both groups.

(3) Two Pseudo-$R^2$ values - McFadden (1974) and Nagelkerke (1991): The interpretation of both Pseudo-$R^2$ values is similar to the $R^2$ in a linear regression. Values above 0.2 indicate a good fit of the model to the data. Again, in both groups the Pseudo-$R^2$ are above 0.2, indicating good model fit.

| Goodness-of-Fit Test | Mail-Group | Board-Group |
|---|---|---|
| Likelihood Ratio-Test | 262.670 (Sig. 0.0) | 604.221 (Sig. 0.0) |
| Hosmer-Lemeshow-Chi | 12.823 (Sig. 0.118) | 9.186 (Sig. 0.327) |
| McFadden Pseudo-$R^2$ | 0.202 | 0.333 |
| Nagelkerke Pseudo-$R^2$ | 0.287 | 0.454 |

*Table 8. Model Fit*

# 6 STUDY RESULTS

In brief, our work finds that

(1) Neither hardware- nor software based protection fully avoids piracy.

(2) Stronger technical copy protection does not necessarily result in lower piracy rates.

(3) A low level of protection does not necessarily lead to high Piracy Rates.

(4) The use of legal or pirated software mainly depends on personal income, the use of the software for business purposes and the intensity of the software usage.

(5) Intense use of P2P- and Chat-networks, but not the Internet in general, fosters piracy.

(6) P2P- and Chat-networks are used to access 'cracked' software copies.

**Ad (1)** Almost 80% of our piracy rates (see Table 3) are below those of the Business Software Alliance. Except for the regions 'Eastern Europe' and 'Latin America', our 'copy-protected' software has lower piracy rates than an unprotected one. However, in no region the piracy rate is 0%. In fact, despite technical copy protection the piracy rates go up to 100% in 'Eastern Europe'. According to Table 4, the protection measure Expansion Card shows the overall lowest piracy rates, Challenge-Response the overall highest. Further, the impact of a specific copy protection on piracy can vary for the groups. For example, the Piracy Rate of the protection measure Challenge-Response is 21% higher in the Mail-Group than in the Board-Group. In the logistic regression, the level of copy protection is significant at the 5% level only in the Mail-Group; the variable has the second highest standard error for both groups.

**Ad (2)** While the hardware-based protection Dongle has a piracy rate of 20% in the Board-Group, the software-based protection CD-Check has only 18% in the same group. The copy protection Multiple CD-Checks is harder to circumvent by the user than the protection CD-Check as it requires multiple CD-ROMs for the authorization process instead of only one. Yet the piracy rates of Multiple CD-Checks are higher than CD-Check in both groups (see Table 4).

**Ad (3)** The Piracy Rate for the protection measure Serial Number offers a support for this finding. While this measure is comparatively easy to circumvent for users and almost costless to producers, the according piracy rates are not as high as one might expect. Serial Number has also lower Piracy Rates than technically superior copy protections like Challenge-Response or Multiple CD-Checks. This seems to indicate that more factors influence the 'pirate-or-buy' decisions than just the ease of copying software.

**Ad (4)** According to Table 7, *Personal Annual Income*, *Requirement for the Workplace* and *Intensity of Application Software Usage* are significant at the 5% level in both groups. The higher the value of any of these variables, the greater is the probability of a sequencer software installation being legal. Unlike the variable *Level of Copy Protection* these variables are significant in both groups stressing their importance as factors of software piracy.

**Ad (5)** The higher the value of *P2P Usage* or *Chat Usage*, the greater is the probability of pirated software installations. Both variables are significant at the 5% level only in the Board-Group; yet their

direction of influence is the same in both samples. In contrast, the higher the value of *WWW Usage*, the greater is the probability of legal sequencer software installations. The impact of the other Internet context variables remains unclear, as the signs of their $\beta$-coefficient differ between both groups (Table 7).

**Ad (6)** P2P-software and Chat-software provide access not only to pirated software copies, but also to cracked copies. In the Mail-Group, where the usage of P2P- and Chat-networks is lower than in the Board-Group, the variable *Level of Copy Protection* is significant at the 5% level. In contrast, in the Board-Group, where the usage of such services is more intensive, the *Level of Copy Protection* does not have any influence on the installations of legal or pirated software copies.

# 7 CRITICAL STUDY ASSESSMENT

While the survey presents some interesting insights into the relationship between technical copy protections, Internet services usage and software piracy, a few points should be noted:

(1) Due to the characteristics of online surveys and the focus on users of sequencer software the results cannot be generalized and need to be checked carefully. Still, the findings of this study may apply to software with similar user characteristics (e.g. Internet affinity) and technical copy protections (e.g. CAD- or 3D-software).

(2) Due to the anonymity of participants, it is difficult to verify the validity of answers. Inclusion of partially filled-out questionnaires resulted in even higher Regional Piracy Rates, which indicates underreporting to some extent. Nevertheless the three data quality indicators (see above) promise good quality of the completely filled-out questionnaires.

(3) Another issues arises with the comparison to the Business Software Alliance data. The Business Software Alliance does not fully document the examined software in their annual piracy study. Hence, it is difficult to check whether any copy protection is implemented in the software under investigation by the Business Software Alliance.

(4) The result of a logistic regression model depends on the characteristics of the sample. A few outliers in the data (Menard 1995) or an unbalanced sample (Cramer 1999) can lead to a bad Model Fit, resulting, for example, in a low Pseudo-$R^2$ McFadden value. However, as the fit in our study is sufficient in both groups including outlier datasets and without ex-post balancing, no changes to the collected data were made.

# 8 CONCLUSIONS AND FUTURE RESEARCH

From an economic perspective, the anti-piracy strategy of a publisher should not only focus on technical copy protection measures since their costs may not be accounted for by higher software sells. The amount of illegal copying does not depend on the protection measure or level, more costly protections (i.e. hardware-based) are not always better than cheaper protections (i.e. software-based protections). P2P- and Chat-networks foster access to cracked software copies, diminishing publisher's efforts to protect their software. On the other hand, a low level of protection does not necessarily foster piracy as the software 'Cakewalk Sonar' shows. This software was rated best sequencer software in 2003 by the readers of the British magazine 'Computer Music' and shows a comparatively low Piracy Rate (see Table 5), although it is only protected by a simple Serial Number. It has even lower piracy rates than the sequencer software 'Synapse Orion' with the same type of protection. This may indicate that the user valuation of the software plays an important part in software piracy besides economic factors. The importance of personal income as a factor of software piracy triggers a discussion about the need for adequate pricing strategies. Further, publishers should try to integrate processes previously covered by 3$^{rd}$ party application software into a single product. This would take some pressure off piracy as an intense usage of application software makes a legal copy more likely.

Software publishers should carefully analyze if their target group has a high Internet affinity similar to the sequencer software users in our study. If so, most likely implementations of technical copy protections may not pay off. While sophisticated protections annoy the legal users, the pirate enjoys the cracked and hassle-free software copy obtained on the Internet. These insights may not only be valid for software, but also for similar digital goods like music or video content.

The security of hardware-based copy protection measures could be improved if the hardware token is not only used for copy protection but also provides functionality of the application software itself. Using this method, the hacker would need to rewrite essentials parts of the application, in order to replace the functionality of the token. An example of this method is the protection Expansion Card, which has the lowest piracy rates in both groups of our study (see Table 4). The downsides of this method are increased copy protection costs resulting from additional hardware development, production and support.

Publishers using software-based measures need to improve tamper resistance of their protections and application software. This could be solved either by making each installed copy unique (Anckaert et al. 2004) or by partial 'outsourcing' of the application binary code on an Internet server where it cannot be altered by the hacker. Both approaches increase copy protection costs for the publisher due to additional software development and Internet server expenses.

This research could be extended in scale and scope by refining the independent variables and collecting additional data. Samples among users of graphics and video software could help to increase the validity of our findings. Refining the independent variables would give better insights into the users buy or pirate decision and help publishers in using our logistic regression model to predict the probability of their software products being pirated. For example, independent variables regarding software price, users' software valuation and ethic attitude could be integrated.

# 9 REFERENCES


Anckaert, B., De Sutter, B. and De Bosschere, K. (2004). Software piracy prevention through diversity. In Proceedings of the 4th ACM workshop on Digital rights management, 63-71, Washington DC, USA.

Business Software Alliance (2003a). Annual Business Software Alliance Global Software Piracy Study. www.bsa.org/globalstudy/2003_GSPS.pdf. access on 2004-04-01.

Business Software Alliance (2003b). Internet Piracy on Campus. www.bsa.org/resources/US_Research_Education_0903.ppt. access on 2004-04-01.

Cheng, K., Sims, R. and Teegen, H. (1997). To Purchase or to Pirate Software: An Empirical Study. Journal of Management Information Systems, 13(4), 49-60.

Chiang, E. and Assane, D. (2002). Software copyright infringement among college students. Applied Economics, 34(2), 157-166.

Christensen, A. and Eining, M. (1991). Factors Influencing Software Piracy: Implications for Accountants. Journal of Information Systems, 5(1), 67-80.

Cramer, S. (1999). Predictive performance of the binary logit model in unbalanced samples. The Statistican, 48(1), 85-94.

Conner, K. and Rumelt, R. (1991). Software Piracy: An Analysis of Protection Strategies. Management Science, 37(2), 125-139.

Devanbu, P. and Stubblebine, S. (2000). Software Engineering for Security: a Roadmap. In Proceedings of the International Conference on Software Engineering, Limerick (Irland), 227-239.

Electronic Musician (2004). Subscriber Profile. advertisers.emusician.com/market/, 2004, access on 2004-04-15.

Givon, M., Mahajan, V. and Muller, E. (1995). Software piracy - Estimation of Lost Sales and the Impact on Software Diffusion. Journal of Marketing, 59(1), 29-37.

Gopal, R.D. and Sanders, G.L. (1997). Preventive and deterrent controls for software piracy. Journal of Management Information Systems, 13(4), 29-48.

Gopal, R.D. and Sanders, G. L. (2000). You can't get blood out of a turnip. Communications of the ACM, 43(9), 83-89.
Haruvy, E., Mahajan, V. and Prasad, A. (2004). The Effect of Piracy on the Penetration of Subscription Software. Journal of Business, 77(2), 81-107.
Hosmer, W. and Lemeshow, S. (2000). Applied logistic regression, Wiley&Sons, New York (USA).
Hinduja, S. (2001). Correlates of Internet Software Piracy. Journal of Contemporary Criminal Justice, 17(4), 369-382.
Hinduja, S. (2003). Trends and patterns among online pirates. Ethics and Information Technology, 5(1), 49-61.
Husted, W. (2000). The Impact of National Culture on Software Piracy. Journal of Business Ethics, 26(3), 197-211.
Kiesler, S. and Sproull, L. (1986). Response Effects in the Electronic Survey. Public Opinion Quarterly, 50(3), 402-413.
Kingpin, J. (2000). Attacks on and Countermeasures for USB Hardware Token Devices. In Proceedings of the Fifth Nordic Workshop on Secure IT Systems Encouraging Co-operation, Reykjavik, Iceland, 35–57.
Kini, R., Ramakrishna, H. and Vijayaraman, B.S. (2004). Shaping of Moral Intensity Regarding Software Piracy: A Comparison Between Thailand and U.S. Students. Journal of Business Ethics, 49(1), 91-104.
Lending, D. and Slaughter, S. (2001). Research in progress: the effects of ethical climate on attitudes and behaviors toward software piracy. In Proceedings of the ACM SIGCPR Conference on Computer Personnel Research, San Diego (USA), 198-200.
Marron, D. and Steel, D. (2000). Which countries protect intellectual property? The case of software piracy. Economic Inquiry, 38(2), 159-174.
McFadden, D. (1974). The Measurement of Urban Travel Demand. Journal of Public Economics, 3(4), 303-328.
Menard, S. (1995). Applied Logistic Analysis, Quantitative Applications in the Social Sciences, Thousand Oaks, Sage Publications.
Microsoft (2004). SafeDisc Windows XP Fix for Microsoft Games, http://www.microsoft.com/downloads/details.aspx?FamilyID=eae20f0f-c41c-44fe-84ce-1df707d7a2e9&displaylang=en, 2004, access on 2004-11-10.
Nagelkerke, D. (1991). A Note on a General Definition of the Coefficient of Determination. Biometrika, 78(3), 691-693.
Peace, G., Galletta, D. and Thong, L. (2003). Software Piracy in the Workplace: A Model and Empirical Test. Journal of Management Information Systems, 20(1), 153-177.
Shy, O. and Thisse, J.-F. (1999). A strategic approach to software protection. Journal of Economics & Management Strategy, 8(2), 163-190.
Slive, J. and Bernhardt, D. (1998). Pirated for profit. Canadian Journal of Economics, 31(4), 886-899.
Wagner, S C. and Sanders, G. L. (2001). Considerations in Ethical Decision Making. The Journal of Business Ethics, 29(2), 161-167.
Zhu, K. Kraemer, K. and Xu, S. (2002). A Cross-Country Study of Electronic Business Adoption Using the Technology-Organization-Environment Framework. In Proceedings of the 23rd International Conference on Information Systems, Barcelona (Spain), 337-348.