9. HAPPY E-INCLUSION? THE CASE OF ROMANIA

Geomina ȚURLEA*
Constantin CIUPAGEA**

Abstract

This paper investigates the determinants of adoption of ICT technology by households in Romania, using a probit model based on a time-series cross-section dataset. A particular attention is given to a several psycho-social factors in addition to the recognised role of usual socio-economic determinants, such as income, age, employment status, educational level or gender. The particular findings are that, together with an expected impact of the occupational status and of the educational level, the perceived wellbeing of individuals is one of the most important factors influencing the decision to acquire and use a PC at home. Gender does not seem to have the same importance as in other regions of the world and an opposite influence than elsewhere, whereas income influences the decision, but with a weaker effect.

Keywords: information and communications technology; e-inclusion; Probit model; Romania; determinants of PC use

JEL Classification: O33, O52, L86

1. Introduction

Modelling the use of ICT in various (often overlapping) groups of population and deriving policy implications based on quantitative methods remain serious challenges for e-Inclusion. Recent literature on e-Inclusion swings between behavioural models (individuals belong to groups defined according to attitudes, perceptions and reactions) and social models (individuals belong to broad groups based on socio-economic factors as age, education, revenue or area of residence). However, little has been done so far in combining the two streams of literature and in offering an image sufficiently complex to allow for identification of different rates of growth of ICT use. In

* *Institute for Prospective Technological Studies, DG-JRC, European Commission, Seville, Spain.*

** *Institute of World Economy, Romanian Academy, Bucharest, Romania, e-mail: cciupagea@iem.ro .* *The research is part of a grant/project financed from the Romanian PNCDI II (National Programme for R&D and Innovation II) budget, Programme IDEI 2007. The name of the project is "Dynamic Model for e-Inclusion Policies".*

particular, there is limited knowledge on the set of attitudes, perceptions and reactions associated with a given level of ICT use within each of the broad social groups and on how such a set influences the actual variance within the group. Moreover, little is known about how ICT use of groups and within groups changes over time, or how these changes could be represented or modelled.

Quite often, the significant reason for this knowledge gap is that the empirical investigations are restricted by the insufficient availability of data. In order to perform a study on the determinants of ICT use in a given region or country, one either needs long time series at an aggregate level or large panel data at individual level (households and firms). Many studies look at cross-country data due to scarcity of national-level registered data. Following the same line of reasoning, but adding to it the difficulty of quantifying the propensity to use ICT by households or firms, the variables taken as proxies for such propensity may not give always the exact picture of the real phenomenon (owning a PC doesn't necessarily mean using a PC, paying for an internet link is not directly related to the amount of information transferred).

We use Romanian individual household data taken from a large database, representative at national level and covering several consecutive years in trying to estimate the relevant determinants of computer use in Romanian households. We look not only at socio-economic factors, but also at a category of psycho-social factors, such as attitude towards society and democracy or perception of future well-being, which factors are more and more proven to be important for the decision of an individual in adopting ICT.

2. The literature on determinants of ICT use

A large amount of the literature referring to determinants of ICT use considers access to ICT equipment and infrastructure, such as access to PC, access to internet, access to various types of telecommunication, as an adequate proxy for the propensity to use ICT. The usual determinants that are taken into account in modelling indicators of ICT use are a) socio-economic factors (income, level of education, occupational status); b) demographic factors (gender, age, family status, racial division in social strata); c) psycho-social factors (attitudes towards society, towards information society and technology, perception of economic evolution, perception of democracy).

A major part of the studies and articles refer to a specific country or region (most of the studies were conducted in the US), but there are many analyses dedicated to cross-country comparisons (developed economies, Asian countries, EU member states). The availability of data is the main restriction to having more detailed or more globalised results concerning the identification of determinants in explaining the ICT use in modern society. Nonetheless, regardless of the depth of the analyses - whether they are based on national micro-data or cross-regional and cross-country macro-data -, the main findings are largely similar, pointing to at the validity of the determining factors mentioned in the previous paragraph.

Investigating the ICT diffusion in selected Asian and Pacific countries, including the US, Ghatak (2007) finds that the major determinants of digital divide are income, level of education, size and type of households, age, gender, the urban/rural divide, ethnicity, infrastructure and cost of accessibility, legal framework and institutional setting.

Another regional study, this time based on econometric testing results departing from cross-country panel data for EU member states (Vincente & Lopez, 2006), shows that inside the EU income and access to university education are the main factors in explaining PC adoption and internet use, while the price of ICT seems to be less important. There is also a perceived risk of digital exclusion for unemployed, women and elderly people within the European Union. One particular finding of the authors for the EU is the importance of R&D expenditures in determining a higher rate of ICT usage in a country.

Most of the studies refer to the situation in the United States and use a large data base available for this country, at micro and macro level. A NSF report (2001, M. Papadakis) "shows that there is a consistent socioeconomic (income, education, occupation) early IT adoption bias by individuals who are affluent, more highly educated, and from higher status occupations compared to society as a whole". The findings are equally supported by data and analyses looking at Internet use as well as at PC ownership or use. Early studies (covering the last two decades of the 20^{th} Century) support a strong correlation and consistency between the rate of early adoption of home computers and the income of households (as the main determinant), or the level of formal education among the members of a household. Demographic and institutional factors are a second large category which is shown to have an impact on ICT early adoption by population.

McNeill *et al*. (2007) investigate in more detail the social determinants of PC use and adoption within a selected group of the American society (urban low-income public housing residents). Although the testing methodology could be challenged (being based on bilateral relationships testing leading to multivariate logit model), the authors found that income, ethnicity, gender, family size, level of education, social environment (neighbourhood) for the household and professional/occupational status are all important even in the case of a specific group of population.

The NTIA Report (2002), which is covering one of the largest database in the US and was produced jointly with the ESA and the US Bureau of Census, identifies the same determinants in leading to an increase in the Internet and computer use rate: income, employment status, age, gender, educational attainment, urban or rural location, ethnicity.

A very comprehensive survey conducted by Assael (2005) in the US among Internet users investigated the role of age, gender, marital status, education level, income and number of working hours on the use of the internet, as well as the role of perceptions (self-, social, future) and attitudes (towards telecom and media, towards leisure, liberalism). He found that income and working environment (including occupation) are responsible for a high proportion of the determining factors, as well as the attitude towards technology and towards liberal professions and way of thinking.

A German survey also confirms that income and level of education are the main factors determining the use of computers and the Internet in households, in both Eastern and Western parts of Germany (2000, Springer Link). The study based on

1998 and 1999 data collected from a large representative sample also shows that age, family status and gender contribute also to the decision of computer adoption.

Other studies look at macro-data over a specific time span (Mohan, 2007) or at micro-samples in smaller countries (Jin & Cheong, 2008), concluding with similar results. Despite the lack of individual data (households surveys) in some cases or of short time series in other cases, there are two important recent trends that have to be further investigated: a) the rate of penetration of ICT equipment and infrastructure is continuously increasing and the determinants are practically the same everywhere (socio-economic, demographic and psycho-social); b) the role of psycho-social factors (perceptions, attitudes, norms) is increasingly observed in studies and surveys related to ICT adoption and use.

3. The determinants of computer technology adoption

We use a utility model to describe the decision to endow the household with a PC, following Vincente and Lopez (2006) model of Internet use.

An individual (i) will decide to acquire a PC according to its perceived utility, in turn described by a linear function of the individual's characteristics (vector X). Let Y =1 if the individual has a PC at home. The probit model that we propose simply assumes that:

$$P(Y = 1 \mid X = x) = \Phi(x'\beta) \tag{1}$$

where: Φ is the cumulative distribution function of the standard normal distribution.

We tried to select as many potentially influencing variables (determinants of PC adoption/ownership) as possible and available.

Following the existing literature, we include first the five categories of fundamental determinants: demographic (age, gender, family characteristics), residence, education level, income and occupational status. We assume that the young generation is inherently more prone to adopting the ICT technology. Urban centres are typically considered to concentrate higher levels of ICT use. Income affects the budgetary constraints and in this respect is expected to play an important role, especially in the first stages of adoption, where Romania can be still retrieved. Education is important mainly as it is positively correlated with openness to technology, but also as it is likely to be the main channel through which the ICT skills are acquired. However, especially for the older generations in Romania, it is often the employment status that puts individuals in direct contact with technology and quite often provides training for specific skills.

A sixth group of determinants, the psycho-social variables, includes proxies for perceptions and attitudes. The need to stay in contact with the family abroad and the activism associated with the participation to NGOs is typically associated with the need for communication and/or processing of important amount of information. However, the true innovation of this paper is that, after controlling all these determinants, it tries to approach the link that might still exist between the perceived wellbeing, or "happiness" and the decision to adopt technology. Is there any surplus of genuine welfare that comes with technology, or perhaps is a happier person more willing to take on the challenge of adapting technology? This paper is trying to provide a first set of tentative insights into this relationship using the specific case of Romania.

4. The dataset

The period of availability for our cross-section dataset covers the years 2003 to 2007, with data extracted from the Public Opinion Barometer, the Spring Survey conducted every year by the *Fundatia pentru o Societate Deschisa* (FSD – The Foundation for Open Society Romania, www.osf.ro). The basic data were processed and harmonised by the authors in order to be used for the current experiment. The number of relevant observations for the model we used is 7425, disaggregated by year as follows: 1453 in 2003, 1612 in 2004, 1434 in 2005, 1528 in 2006 and 1398 in 2007.

The dependent variable (ACCESS_PC) takes the value 1 if the interviewed person has a PC in its own household and is a proxy for the decision of technology adoption.

The group of *demographic variables* includes age (AGE); gender (GENDER); and the presence of children in the households (CHILD). Two types of variables are considered to measuring age; a) a continuous variable will simply measure the age of the respondent in years and b) a set of dummies will include the respondent in one of the main age groups (18-24y, 25-34y, 35-55y and 55y and over). By design, only people over 18 years of age are included in the survey. We consider the presence of children in the household as the main determinant regarding family characteristics. This is particularly relevant for the multigenerational Romanian type of family nucleus, where grandparents might have computers in the households as a result of living with the grandchildren. We also looked at the number of education years of the parents of the interviewed person as the main family push factor, but we later excluded this variable from the modelling exercise on the grounds that it was highly correlated with other standard determinants.

Two dimensions are measured to account for the influence of the *residence* place. Firstly, we control for the capital centric model of regional development of Romania using a dummy for residence in the capital city, Bucharest (LOC_BUC). Secondly and most important we took into account the urban-rural division, but allowing a more nuanced transition between the two; rather than separating the respondents into urban and rural residents, we consider among the determinants the distance in km from the residence place to the nearest larger urban centre (LOC_DIST).

The *education* level is measured by three dummies allocated to the level of completed education: up to secondary education or 8 years (PGEDUC_w); high-school education or from 9 to 12 years of schooling (LEDUC_w); and finally higher education or over 12 years of completed education (HEDUC_w). Knowledge of at least another foreign language (LANGUAGES) is included as a proxy to cultural openness and ability to have access to a wider variety of computer content. Because of the delay in software translating in the case of Romania, knowledge of foreign languages was almost a precondition for using a computer during most of the analysed period.

The respondents are distributed by *income* deciles according to the national level income brackets in which its household revenue falls. This allows pooling data for different years together without concerns on year to year monetary comparability. For the purposes of modelling, the deciles are aggregated into: low incomes or deciles 1 to 4 (D1234); average income or deciles 5 to 7 (D567); and high income or deciles 8 to 10 (D8910).

We include four dummies corresponding to the following *occupational* status: white collar employed (WCEMPL), blue collar employed (BCEMPL), student or pupil (STUDENT), unemployed or inactive, but not a student (INACTIVE). The dummies are built from the status declared by the respondents according to the basic principles of occupational taxonomies.

The *psycho-social variables* are again dummy variables which take value 1 if the respondent has members of the family abroad (NETW_abroad) or if the respondent is engaging in NGOs or other private organisations (NETW_asoc). Our variable of perceived wellbeing or happiness (SOCIS_CONTENT) takes value 1 if the respondent declares himself as more or less happy/content, happy or very happy with his current quality of life. It can be argued that the perceived wellbeing is highly correlated with the socio-economic status. In order to counteract this effect, we instrumented the "happiness" variable. We used as instruments the following dummy variables which proxy various aspects of subjective wellbeing: the perception of the current improvement against past periods (ABS_IMPROVE); expectations of improvement for the future (PERC_FUTURE); trust in institutions and democratic representation (TRUST_INST); and trust in other people (TRUST_PEERS). Trust in institutions is a composite indicator from the trust in seven different institutions: central government, local government, political parties, presidency, parliament, army and police. Indeed, we assume that the perceived wellbeing in general is seen as based on feelings of trust, both in a positive trend of the personal evolution as in the social environment of the person.

Table 1

Basic descriptive statistics of the variables used in the modelling exercise[1]

Variable	Description/codification	Mean (Std. dev)	Correla-tion*
ACCESS_PC	1 – owns a PC; 0 – otherwise	0.241	1
Demographic variables			
AGE Age in years	Min: 18y, Max:96	48 (18.26)	-0.27
AGE18_24	1 – age 18y to 24y; 0 – otherwise	0.135	0.13
AGE25_34	1 – age 25y to 34y; 0 – otherwise	0.145	0.08
AGE35_55	1 – age 35y to 54y; 0 – otherwise	0.347	0.12
AGE55_	1 – age 55y or older; 0 – otherwise	0.373	-0.27
GENDER	1 – male; 0 – female	0.468	0.02
CHILD	1 - children; 0 – otherwise	0.659	0.22

[1] *See the online version of the paper on www.ipe.ro/rjef_princ.html for the correlation matrix of the independent variables.*

Variable	Description/codification	Mean (Std. dev)	Correla-tion*
	Residence		
LOC_BUC	1 – residence in Bucharest; 0 – otherwise	0.089	0.14
LOC_DIST	km between the residence and the nearest urban center (Min:0km, Max: 99km)	10.05 (14.80)	-0.23
	Level of completed education		
PGEDUC_w	1 – <= 8y of education and no longer a student; 0 – otherwise	0.328	-0.293
LEDUC_w	1 – 9 to 12 y of education and no longer a student; 0 – otherwise	0.438	0.005
HEDUC_w	1 – > 12 years of education and no longer a student; 0 – otherwise	0.165	0.289
LANGUAGES	1 – knowledge of one or more foreign languages; 0 – otherwise	0.276	0.248
	Income		
D1234	1 – revenue in deciles 1 to 4; 0 – otherwise	0.589	-0.312
D567	1 – revenue in deciles 5 to 7; 0 – otherwise	0.247	0.067
D8910	1 – revenue in deciles 8 to 10; 0 – otherwise	0.164	0.255
	Occupation status		
STUDENT	1 – student or pupil; 0 – otherwise	0.061	0.219
WCEMPL	1 – employed white collar (manager, clerk, army officer, professional); 0 - otherwise	0.108	0.351
BCEMPL	1 – employed blue collar (skilled/unskilled worker); 0 – otherwise	0.283	0.022
INACTIVE	1 – unemployed or otherwise inactive, but not student or pupil; 0 – otherwise	0.548	-0.336
	Psycho-social variables		
NETW_abroad	1 – family or friends abroad; 0 – otherwise	0.101	0.035
NETW_asoc	1 – activity in an NGO or other civic association; 0 – otherwise	0.116	0.143
SOCIS_CONTENT	1 – happy of very happy with his quality of life; 0 – otherwise	0.785	0.116
	Instruments		
ABS_improve	perceived improvement of the quality of life compared with previous year: 0 – much worse; 1 – worse; 2 – the same; 3 – better; 4 – much better	1.90 (0.862)	0.172 (0.257)
PERC_future	expected improvement of the quality of life in the next year: 0 – much worse; 1 – worse; 2 – the same; 3 – better; 4 – much better	2.06 (0.838)	0.167 (0.253)
TRUST_peers	1 – general trust in other people; 0 – otherwise	0.397	0.015 (0.07)
TRUST_INST	Score function, takes values from 0 – the respon-dent doesn't trust and of the seven institutions to 7 – the respondent trusts all the seven institutions	2.70 (2.131)	-0.08 (0.09)

Notes: *Variables in bold refer to the household of the respondent, the rest to the respondent only*
*** correlation with the dependent variable (correlation with the instrumented variable)*

For a binary variable, the mean value is in fact the share of the observations for which the variable takes value 1. Almost a quarter of the respondents have a computer at home, and almost 78% declares themselves as content with the quality of their lives. However, 86% of those who have a computer at home consider themselves as happy, but they cover only 26% of the happy people. This would remotely suggest rather that happy people are open to technology than that the technology would be a factor of happiness.

	SOCIS_CONTENT		
ACCESS_PC	0	1	Total
0	1,788	5,570	7,358
1	305	2,039	2,344
Total	2,093	7,609	9,702

The age and incomes distribution is implicitly simulated by the corresponding sets of dummies. Hence, 28% of the adult population is under 35 years, almost 35% between 35 and 55 years, and around 37% 55 years or older. We broke down the lower segment into 18-24 years and 25-34 years, because of the well-known higher dynamics of technology adoption specific to younger age brackets. More than half of our respondents have low revenues (deciles 1 to 4) and this is the segment that is most likely to be sensitive to the cost of technology. For further steps of this research we plan to increase the granulation of the income classification. A final note regards the students or pupils. From the occupational point of view, students make up a subgroup of inactive persons. We treat them separately simply because students have a very different model of technology adoption than the unemployed of other inactive people. But students are also persons in education, unlike people that are inactive on the labour market. So, from the point of view of educational status, we also treat them separately from the persons with some completed form of education (primary, secondary or tertiary). From the respondents in our sample 93.1% have some form of completed education (most of them, 43.8% high-school education), while over 6% are still in schooling.

Regarding our selection of instruments it is interesting to note that the presence of the positive trend in the personal evolution has the strongest correlation with the perception of wellbeing, while, quite the opposite, there is only some mixed evidence to support the trust in institutions as a factor. The distribution of the TRUST_INST variable clearly shows a bias of the interviewed population towards the sceptical positions when it comes to democracy and its institutions. In fact almost 20% of the respondents would not trust any of the seven institutions considered here (central government, local government, political parties, presidency, parliament, army and police), while half of the respondents won't trust more than two of those institutions The correlation between the TRUST_INST and our ICT adoption variable is actually negative. This seems difficult to explain, but it is probably related with a specific tension between the political turmoil during the transition period and the attitude of forefront technology adopters.

5. The estimation procedure

We use a standard specification of a binary response model, the probit model, in order to estimate the parameters in our probability function. Two features specific to our case needed to be considered when designing further the overall estimation procedure.

Firstly, the problem with our specification (and indeed with most of the similar specifications) is that one or more repressors are very likely endogenously determined. Fitting the limited dependent variables with endogenous repressors has received considerable attention in the economic literature and methods like the one proposed by Newey (1987) with maximum likelihood estimation became standard and are already embedded in statistical packages. We use a STATA function that employs instruments of control for the correlation of one or more repressors with the error term. As explained in Chapter 4, we instrumented the perceived wellbeing variable, using four other variables, namely perceived improvement of the quality of life, hope for the future, trust in other people and trust in institutions and democratic representation.

Secondly, the statistical description in Table 1 is shown for the entire data sample, for the sake of simplicity. In fact, we observe a lot of dynamics in some of the independent variables from one year to another and expectably, in their influence on the dependent variable. Starting with the share of PC owners and PC users in total population (proxy for our unobservable dependent variable, probability of acquiring a PC) with itself increased at a high rate from one year to another during the period of estimation, as it is shown by official statistics and also analysed in various studies looking at the Romanian ICT market and information society (Ciupagea *et al*., 2008; Țurlea & Gheorghiu, 2006, 2007).

We deal in two alternative ways with the year by year variation within our time-series cross-section dataset. The first option is to introduce year dummies that would cumulate various sources of year to year variability. This is the standard in a pooled regression. The second option is to construct time interactions with the selected variables in an attempt to identify the sources of those variations. We choose interaction with the time of the main explanatory variables (income, age, education, occupational status, gender and distance), also looking at the yearly stability of the correlations between the explained variable (ACCESS_PC) and the selected explanatory variables (see Appendix 1 in on-line version on www.ipe.ro/rjef_princ.html). We did not explicitly build other type of interactions than with time, but applied alternative strategies for minimising cross correlations between dependent variables (see the case of variable STUDENT from occupational and educational perspective as explained in Chapter 4).

Finally we decided to test the two alternative approaches by including the age variable (as continuous variable or as age groups), as explained in Chapter 4.

Departing from these considerations, in the next stage we built four different specifications of the same probit model described above, which we tested separately, using a maximum-likelihood econometric estimation that takes into account the potential endogeneity of some repressors. We also include the option of robust estimates in the estimation procedure. We further compare the results in search for relevant and significant differences.

The set of explicative variables (vector X in eq.1) is different in each of these four versions, depending on two selection criteria: a) one type of model version uses age as a continuous variable, the other uses 3 different age groups (categorical variable); b) one type of model version takes into account different time interaction for each variable separately, while the other type considers year dummies for the entire sample (the year influence is global and not variable-specific). A formalisation of these different approaches is described hereunder [eq. 2]:

$$X_1 = f(Age, Age*t, W, W*t, Z)$$
$$X_2 = f(Age, W, Z, YearDummies)$$
$$X_3 = f(Age18\_24, Age18\_24*t, Age25\_34, Age\ 25\_34*t, Age35\_54, Age35\_54*t, Age55, Age55*t, W, W*t, Z)$$
$$X_4 = f(Age18\_24, Age25\_34, Age35\_54, Age55, W, Z, YearDummies)$$

where: X_i represents the vector of descriptive variables (repressors) for model i and the list of repressors contains, except for the age variables, the following items:

W – a vector of descriptive variables that are considered to show potential time interaction;

Z – a vector of the rest of descriptive variables which were not considered for time interaction testing;

YearDummies – dummies for each year in the sample (2003 to 2007).

When a set of independent dummy variables covers the entire range of population in the sample, one of them has to be excluded from the list of independent variables to be tested in order to avoid colinearity, which become a reference group. We have considered as reference the groups of higher educated white-collar workers, workers with revenues between the 5th and the 7th income deciles, and persons with completed higher education. When the age groups are used (instead of the continuous variable), we consider the youngest as the reference group. This selection is not meant to draw a reference profile. The exclusion of some variables is made independently based on statistical criteria, as minimising the correlation between the variables in the vector X (for instance, the correlation between white collar employment and completed higher education or between the age group 18-24 years and the status of student).

We report the results of the probit estimations in Appendix 1, but only for Model 1. The results for all other three models are reported within the on-line version of this article, see www.ipe.ro/rjef_princ.html. The Pseudo-R^2 is not computable by the econometric method that we use hence we rely on an analysis of the correct prediction rates to assess the goodness of fit of the model. The table in Appendix 2 presents the classification prediction probabilities for Model 1 (see on-line version on www.ipe.ro/rjef_princ.html for each of the other three probit models) and shows that over 80% of the observations are correctly classified by the model. The results of the probit model show that most of the repressors discussed above are significant at 5%. In general their coefficients have the expected sign. The Wald test and the p-value on the top of the estimation reject the null that all the coefficients would be jointly zero.

The significant Wald test for the exogeneity of the instrumented variables reported at the bottom of the results rejects the hypothesis of no endogeneity and supports our choice for the specific econometric procedure. This is confirmed by the significant *rho* value also reported by the testing algorithm.

However, it is not useful to look at the coefficients of independent variables resulting from the probit model because in the probit model the derivative of the probability with respect to X varies with the level of X and with the levels of the other repressors. The solution is to compute the values of the derivatives at the mean values of the repressors (the "slopes"), considering that each of them represents a "typical" observation (reported as y - probability of positive outcome in first row of Table 2 below).

6. The results and directions for further research

The estimated marginal effects are defined as the variation of the estimated probability with respect to a marginal variation in the repressors (or a discrete change in case of binary variables and dummies).

Therefore, we look at a change of one unit in the categorical or continuous variable or to a step from 0 to 1 in the case of a binary variable when searching for the marginal effects on the probability of a positive outcome (prediction probability) for the predicted endogenous variable. Table 2 below presents these marginal effects (except for the year dummies).

Table 2

Marginal effects for the probability of ICT technology adoption model for Romania in each of the four versions (model 1 to 4), period of observations 2003-2007

	Model 1	Model 2	Model 4	Model 3
Decision to adopt ICT technology (ACCESS_PC)	0.151	0.152	0.152	0.152
Demographic variables				
Age (AGE)	-0.0029922***	-0.0010844***		
Age*t	0.0006258***			
Age25_34			0.0262708	-0.0817517***
Age25_34*t				0.0411856***
Age35_54			0.0476311*	-0.0460836***
Age35_54*t				0.0319559*
Age55_			-0.0250756	-0.1139298***
Age55_*t				0.0342071***
Gender (GENDER)	-0.0255577**	-0.024175***	-0.0208273*	-0.021778**
children in the household (CHILD)	0.184823***	0.1856456***	0.1812609***	0.1815041***
Residence				
residence in Bucharest	0.105721***	0.1121771***	0.1130684***	0.1071375***

	Model 1	Model 2	Model 4	Model 3
(LOC_BUC)				
distance to the nearest city (LOC_DIST)	-0.0101085***	-0.0078531***	-0.0076843***	-0.0097873***
(distance to the nearest city)^2	0.0000624***	0.0000747**	0.000727***	0.0000609***
distance*t	0.0007914***			0.0007309**
Level of completed education				
up to 8y (PGEDUC_w)	-0.1431921***	-0.1492205***	-0.146668***	-0.1420807***
12y (LEDUC_w)	-0.156402***	-0.0693436***	-0.0708117***	-0.1373272***
12y*t	0.0304486***			0.0232619***
knowledge of foreign languages (LANGUAGES)	0.0573001***	0.0569203***	0.0611327***	0.059945***
Income				
low income (D1234)	-0.096622***	-0.0963487***	-0.0982397***	-0.0993***
Average income (D567)	0.0960089***	0.0924805***	0.0920538***	0.0943341***
Occupation status				
Student of pupil (STUDENT)	-0.105906***	-0.0122471	0.0414439	-0.0943684***
Student*t	0.0437323***			0.0549945***
blue collar employed (BCEMPL)	-0.1237054***	-0.1245431***	-0.121162***	-0.1208849***
Inactive (INACTIVE)	-0.1253105***	-0.1720381***	-0.1543978***	-0.1267995***
Inactive*t	-0.0139717*			-0.0085108
Psycho-social variables				
Family/friends abroad (NETW_abroad)	0.0433135***	0.0428206**	0.0466708***	0.0470192***
participation to civil associations (NETW_asoc)	0.0652804***	0.0617288***	0.0608956***	0.0637684***
perceived wellbeing (SOCIS_CONTENT)	0.110683***	0.1099119***	0.1168584***	0.1157777***
Year dummies				
2004		0.0486266***	0.048758***	
2005		0.1190472***	0.1186861***	
2006		0.1660083***	0.1653314***	
2007		0.2490352***	0.2496051***	

(*) dy/dx is for discrete change of variable from 0 to 1

With the exception of the variables that have attached a time interaction, the coefficients for the rest of the variables are similar in all four models: this means that we generally succeeded in breaking down the residual year variation captured by the time dummies and allocate it to several specific variables. Not all the time interaction tested proved to be statistically significant, which means that some of those variables actually have constant impact over the period under consideration. The most interesting example is the income, for which the interaction with time proved clearly insignificant. In our opinion, this comes somehow in support of our assumption according to which our income groups are too aggregated, especially towards the lower end (deciles 1 to 4), where the elasticity to relative price variations is expected to be the highest. However, it is still expected that the probability of acquiring a PC increases with the income level (the bigger the rank of the deciles the higher the probability) and that income remains an important factor.

In all models, the most important determinants of technology adoption are: the presence of children in the household, completed higher education, white-collar employment, perceived wellbeing and residence in the capital city, followed by the rest of the factors. It is not entirely expectable, nor in line with the existing literature on ICT adoption, particularly within the EU (see, for instance, Vincente and Lopez, 2006), that all these factors mentioned above would be more important than the income in inducing a propensity to acquire a PC. It is probably the case that the deployment of the new technologies really took off in Romania at times of falling prices, hence revenues, although important, were not ranking at the very top of the potential drivers of ICT consumption.

Other result that needs further explanation is the negative coefficient associated with the status of student or pupil. Being a student and or a white collar employee have both a high impact on ICT technology adoption. At the level of the overall sample, the probability of having a computer when on white-collar employment (proxies by the share of the white collars employees who own a computer at home), is 64.3%. The same probability when a student is 58.1%. With the white collars as a reference group, comparatively, the probability that a student has a computer is lower, hence the negative sign for the STUDENT variable in models 1 and 3.

Nevertheless, year by year this probabilities change as follows:

	2003	2004	2005	2006	2007
When a white collar	44%	54.9%	60.5%	75.4%	78.9%
When a student	38.1%	44.3%	58.7%	81.48%	80.8%

By the fourth year of our sample, being enrolled in an education program means a higher probability to acquire a computer than being hired as a white collar. As the group of reference in the case of occupational status is formed by white collars, the positive sign for the coefficient of the complex variable student*t (time trended variable) is a proof for convergence between the two groups presented in the above table, although the model predicts a faster convergence than observed.

Both model 1 and model 3 results show that the gap between the probability of acquiring a PC for a highly educated person and the same probability for a person with medium education (up to 12 years) has decreased over the period of reference (2003-2007). This is confirmed by the positive sign of the coefficient corresponding to the time-trended variable. Nevertheless, it would be a hazard to conclude that such an outcome is the result of a recent better PC-endowment of high-schools in Romania. Most of the people in the group of those with high-school completed higher education belong to age groups above 35 years. Moreover, there are small differences by age groups in the proportion of people owning a PC in total people having medium education level (9 to 12 years of education) who are not currently students.

Becoming inactive lowers dramatically the probability of buying a computer. But when the time interaction is taken into consideration, a worrying trend is revealed: a negative coefficient for the time interaction means an increasing gap over time between the white collars and the inactive persons which instead means that the inactive persons show an increasing risk of social exclusion (and e-exclusion).

This confirms the observed dynamics of the probabilities:

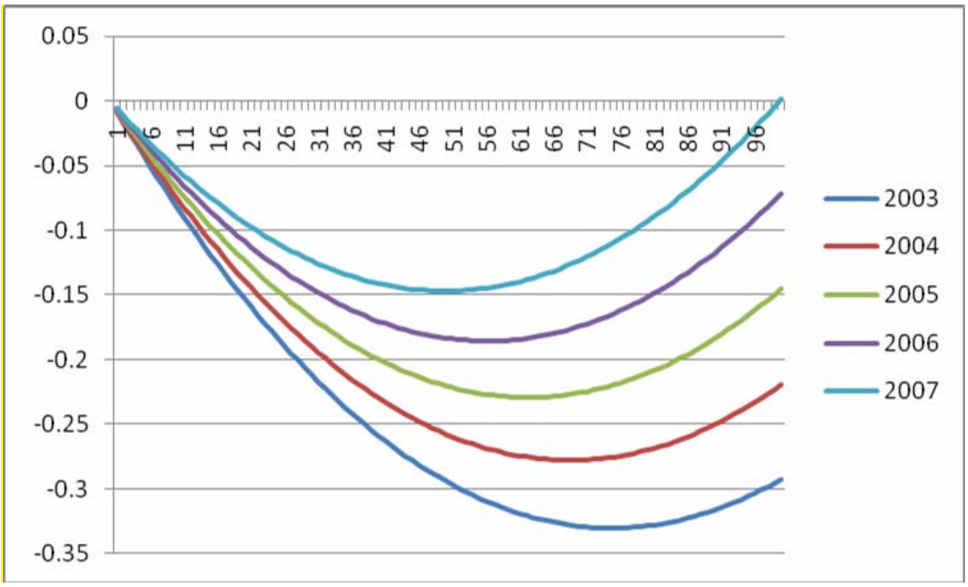
	2003	2004	2005	2006	2007
When a white collar	44%	54.9%	60.5%	75.4%	78.9%
When inactive (but not a student)	5.6%	7.7%	10.2%	11.66%	15.5%%

Considering the strong impact age has according to our model, the gap between the share of 18-year-old persons who own a computer and the same share for people over 55 years have increased significantly with the global increase in the early share of PC owners in total population. Again, it must be mentioned that this result holds at this stage of the technology adoption, when less than a quarter of respondents are actually owning a computer.

One very surprising result is the low impact of gender difference in the decision of acquiring a PC, as well as the sign of this estimated impact (the negative sign for the slope denotes that women are most likely to acquire a PC than men in Romania). This is an opposite finding from the usual results in the literature, particularly in the studies referring to the US or Asian markets. The explanation may partially reside in the heritage of the previous socialist system, characterised by an artificially imposed equality between women and men. However, the gender factor is also the less statistically significant among all the factors took into account. Time trend also had little impact when tested, regardless of the model version used.

Distance to the closest urban centre proves to be significant, but the results point out at a non-linear function, having the characteristics of a parabola which turns positive after 100 or more km. This cannot be confronted with the reality for the lack of corresponding observations in our sample, as there are not individuals/households in the sample located at such a distance from the nearest city (given a high density of population and of towns spread rather uniformly on the territory of Romania).

The increase in the slopes corresponding to the year dummies from the earliest to the latest (2004 to 2007) is of no surprise, as it follows the clear trend of increasing share of people that own a PC in total population of Romania, a trend which is perfectly in line with the period of strong economic growth and with the expansion of the ICT market.



Being “happy” (i.e. high perceived well-being) is an important factor. Moreover, it seems to have a higher influence than income and matter as much as living in the capital city. This conclusion is not immediately straightforward and, therefore, an important result of our research. In fact, it supports the hypothesis that after getting acquainted with the new technology through their work/learning environment, Romanians are more susceptible to buy a PC when perceiving that their quality of life is stable and good or has improved.

The profile with the highest probability of acquiring a computer in Romania is the wealthy female student, content with her quality of life, living in Bucharest, with children living in the household and having family members working abroad. She speaks foreign languages and is involved in non-governmental organisations.

At the other extreme we find the elderly man, with poor retirement revenues and little education. He lives alone and is not involved in social work. From the policy-making point of view, there is no or very limited scope of intervention at the level of these extreme profiles.

Yet, there are very limited non-linearities in our model. However, and despite the very few non-liniarities in our model, there are a lot of potential combinations of factors that create profiles with similar probabilities of adopting technology, but susceptible for targeting by different policy measures. The next step of this research will attempt to isolate few contrasting profiles and studying them in more detail, including specific modules of the FSD Public Opinion Barometer, in particular those on media and on happiness and its determinants.

The available data also allow for future analyses of regional differences in the profiles of PC owners and users in Romania (and possibly of internet users).

References

Henry Assael, "A demographic and psychographic profile of heavy internet users and users by type of internet usage", *Journal of Advertising Research*, March 2005, Goliath – Business Knowledge on Demand, http://goliath.ecnext.com.

Avner Bar-Hen, *Generalized Principle Component Analysis of Continuous and Discrete Variables*, September (2002), Ravi Khatree Editor, interstat.statjournals.net/YEAR/2002/abstracts/0209001.php.

Bianchi A., Barrios S., Cabrera M., Cachia R., Compañó R., Malanowski N., Punie Y., Turlea, G., Zinnbauer D., Centeno C., *Revisiting eInclusion: from Vision to Action*, IPTS Report, 22549EN.

Bonin H. and Schneider H., *Analytical Predictions of Transition Probabilities in the Conditional Logit Model*, IZA DP No. 1015.

Caselli, F., Coleman, W.J. (2001), "Cross-country technology diffusion: the case of computers", *American Economic Review,* 91(2):328–335.

Menzie D. Chinn, Robert W. Fairlie - *ICT use in the developing world: An analysis of differences in computer and internet penetration*, NBER Working Paper 12382, July 2006, JEL No. O30, L96.

Newey, W.K. (1987), "Efficient estimation of limited dependent variables models with endogeneous explanatory variables" – *Journal of Econometrics* 36: 231-250.

Constantin Ciupagea, Radu Gheorghiu, (2008), *Dezvoltarea societății informaționale în România şi perspectivele de convergența cu statele Uniunii Europene*, Institutul National de Cercetări Economice, Colecția CIDE, 25/2008.

CpQD (1), *Usage Modeling*, (2006), http://www.jotmi.org/index.php/GT/article/view/art8/20.

CpDQ (2), Mapping Solutions – Telecommunication Solutions for Digital Inclusion Project (in Portuguese). Technical Report, http://www.cpqd.com.br/img/mapeamento_de_solucoes _ab.pdf.

Shambhu Ghatak – Brief Note on ICTs, dgCommunities, Development Gateway Foundation online resources, April 2007. http://topics.develop mentgateway.org .

Radu Gheorghiu, Geomina Ţurlea, (2006), "The need for a new paradigm for policy making for the knowledge based economy: Stimulating the knowledge production in Romania", in *The knowledge based economy in Central and East European Countries*, Edited by Krzysztof Piech and Slavo Radosevic, Palgrave Macmillian Publishing House.

Holanda G.M. and Dall'Antonia C. J, (2006), "An Approach for e-inclusion: Bringing illiterates and disabled people into play", *Journal of Technology, Management and Innovation*.

Jianbin Jin, Angus Weng Hin Cheong, (2008), "Measuring Digital Divide: The Exploration in Macao", *Observatorio (OBS*) Journal*, 6 (2008), pp 259-272, ISSN 1646-5954.

Sangwon Lee, Justin S. Brown, (2008), "Examining Broadband Adoption Factors: An Empirical Analysis between Countries", *The Journal of Policy, Regulation, and Strategy for Telecommunication, Information, and Media*, ISSN: 1463-6697, Volume 10 (1), pp. 25- 39, Emerald Group Publishing Ltd..

Kelly K. Levy, Lee Price, (2002), *A Nation On-line: How Americans are expanding their Use of the Internet*, NTIA, ESA & US Bureau of the Census Joint Report, February 2002.

Lorna H McNeill, Elaine Puleo, Gary G Bennett, Karen M Emmons, (2007), "Exploring Social Contextual Correlates of Computer Ownership and Frequency of Use Among Urban, Low-Income, Public Housing Adult Residents", *Journal of Medical Internet Research*, 9(4):e35.

Ramesh Mohan – "Determinants of ICT Expenditure Using Logit Transformation for Proportion Data Analysis", *Journal of Information Technology Impact*, Vol. 7, No. 2, pp. 145-158, 2007.

OSF POB database and related papers, Soros Foundation Romania (http://www.osf.ro/en/program_articol.php?articol=18)

Maria C. Papadakis - *The Application and Implications of Information Technologies in the Home: Where Are the Data and What Do They Say,* NSF 01-313, Project Director: Eileen L. Collins, National Science Foundation, Division of Science Resources Studies, Arlington, VA, February 2001.

Springer Link – "Use of Computers and the Internet Depends Heavily on Income and Level of Education", *Economic Bulletin*, Volume 37, Number 11, November 2000, pp. 369-374, ISSN 0343-754X.

Geomina Țurlea, Radu Gheorghiu – "Information society in Romania", in *EU Enlargement. Economic Development and the Information Society*, Eds: Geomina Țurlea and Marc Bogdanowicz, VUBPRESS Brussels University Press, 2007.

Maria Rosalia Vincente, Ana Jesus Lopez – "Patterns of ICT diffusion across the European Union", *Economic Letters* 93 (2006), pp 45-51, www.sciencedirect.com, Elsevier.

Appendix 1

MODEL 1 - AGE CONTINUOUS, WITH TIME INTERACTIONS

Probit model with endogenous regressors Number of obs = 7425
Wald chi2(22) = 1915.18
Log pseudolikelihood = -6181.4916 Prob > chi2 = 0.0000

|		Robust				
|	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
SOCIS_CONT~T	.5604458	.1747536	3.21	0.001	.2179349	.9029566
LOC_BUC	.3881292	.0596553	6.51	0.000	.271207	.5050514
CHILD	.9169628	.0562033	16.32	0.000	.8068064	1.027119
NETW_abroad1	.172394	.0623104	2.77	0.006	.0502679	.29452
LANGUAGES	.2317528	.043101	5.38	0.000	.1472765	.3162292
NETW_asoc	.2533033	.0550688	4.6	0.000	.1453705	.3612361
STUDENT	-.6084335	.1594819	-3.82	0.000	-.9210122	-.2958548
I_STUDENT	.1863362	.0453341	4.11	0.000	.097483	.2751893
BCEMPL	-.6004485	.065696	-9.14	0.000	-.7292103	-.4716867
INACTIVE	-.5257831	.1197571	-4.39	0.000	-.7605027	-.2910634
I_INACTIVE	-.059531	.0320049	-1.86	0.063	-.1222594	.0031975
D1234	-.4112839	.0493555	-8.33	0.000	-.5080189	-.3145489
D8910	.3623193	.0559266	6.48	0.000	.252705	.4719335
PGEDUC_w	-.6983371	.064211	-10.88	0.000	-.8241883	-.5724858
LEDUC_w	-.6917444	.095258	-7.26	0.000	-.8784467	-.5050421
I_LEDUC_W	.1297367	.0259444	5.00	0.000	.0788866	.1805868
AGE	-.0127491	.0022839	-5.58	0.000	-.0172255	-.0082727
I_AGE	.0026663	.0005348	4.99	0.000	.0016181	.0037144
LOC_DIST2	.0002658	.0000595	4.46	0.000	.0001491	.0003824
LOC_DIST	-.0430706	.0050889	-8.46	0.000	-.0530447	-.0330966
I_DIST	.0033718	.001234	2.73	0.006	.0009533	.0057904
GENDER	-.109061	.03913	-2.79	0.005	-.1857545	-.0323676
_cons	-.6479657	.1869412	-3.47	0.001	-1.014364	-.2815677
/athrho	-.1860675	.0726046	-2.56	0.010	-.3283699	-.0437651
/lnsigma	-.9512148	.0084259	-112.89	0.000	-.9677293	-.9347004
rho	-.1839495	.0701478			-.3170553	-.0437371
sigma	.3862715	.0032547			.3799448	.3927035

Instrumented: SOCIS_CONTENT

Instruments:

LOC_BUC CHILD NETW_abroad1 LANGUAGES NETW_asoc STUDENT I_STUDENT BCEMPL INACTIVE I_INACTIVE D1234 D8910 PGEDUC_w LEDUC_w I_LEDUC_W AGE I_AGE LOC_DIST2 LOC_DIST I_DIST GENDER ABS_improve TRUST_PEERS PERC_future TRUST_INST

Wald test of exogeneity (/athrho = 0): chi2(1) = 6.57 Prob > chi2 = 0.0104

Marginal effects after ivprobit

y = Probability of positive outcome (predict, p) = **.15148748**

Appendix 2

Probit model – prediction rates

Model 1

Probit model for ACCESS_PC
-------- True --------

Classified	D	~D	Total
+	1009	402	1411
-	832	5182	6014
Total	1841	5584	7425

Classified + if predicted Pr(D) >= .5
True D defined as ACCESS_PC != 0

Sensitivity	Pr(+| D)	54.81%
Specificity	Pr(-|~D)	92.80%
Positive predictive value	Pr(D| +)	71.51%
Negative predictive value	Pr(~D| -)	86.17%
False + rate for true ~D	Pr(+|~D)	7.20%
False - rate for true D	Pr(-| D)	45.19%
False + rate for classified +	Pr(~D| +)	28.49%
False - rate for classified -	Pr(D| -)	13.83%

Correctly classified 83.38%