

# Costs or Gross Benefits? - What Mainly Drives Cross-Sectional Variance in Internet Adoption

Joeffrey Drouard\*

1 September 2010

## Abstract

In this paper, we propose an empirical model of Internet adoption which takes into account the household's desire to adopt the Internet. Our research supports three central findings. First, we determine the main factors that explain the cross-sectional variance in gross benefits. Second, we show that while the cross-sectional variance in the predicted probability that a household desires to adopt the Internet and does not because its adoption costs are higher than its gross benefits is relatively low (except for the age factor), the cross-sectional variance in the predicted probability that a household does not desire to adopt the Internet is high. Third, we compute the predicted adoption probability assuming that the adoption costs are homogeneous across households. We show that, for a given dimension (except for the age factor), the adoption rate will be only slightly modified if the adoption costs are homogeneous across households. Our results support the argument that the digital divide is mainly due to differences in gross benefits of adoption.

*Keywords:* Digital Divide; Internet Adoption; Gross Benefits; Adoption Costs.

*JEL Codes:* D12; D63; L86; L96.

---

\*Telecom ParisTech, Paris, France and European University Institute, Florence, Italy. E-mail: [joeffrey.drouard@eui.eu](mailto:joeffrey.drouard@eui.eu).

# 1 Introduction

While the development of information and communication technologies in the late 20th century created hope for a reduction of inequalities by promoting a greater and more egalitarian access to information throughout the population, research on the diffusion of information and communication technologies has shown on the contrary a widening of inequalities with the creation of a digital illiterate class. These disparities in the diffusion and the appropriation of new technologies have often been referred to as the "digital divide" since the mid-1990s.

Numerous papers have analyzed the digital divide in Internet connection. A consensus has emerged concerning the central role of certain factors - including residential area, income, education, age, presence of children - to explain the households' choice to adopt the Internet.<sup>1</sup> Then, in order to understand the ethnic digital divide better, new variables assumed correlated with the variable "ethnic group" were introduced, such as lack of skills (see, e.g., Krueger, 2000), lack of friends and relatives using the new technologies (see, e.g., Goolsbee and Klenow, 2002)<sup>2</sup> and finally the quality of service available<sup>3</sup> and the degree of competition<sup>4</sup> in the households' residential location (see, e.g., Prieger and Hu, 2008). In these papers only the net benefits of Internet adoption are identified. Thus, they do not allow identifying if cross-sectional variance in Internet adoption is mainly due to differences in costs or in gross benefits.

Identifying if the digital divide in Internet connection is mainly due to differences in costs or in gross benefits has implications for policy makers. Indeed, if cross-sectional variance in Internet adoption is mainly due to differences in costs, then stimulating the reduction of Internet service costs and promoting the abundance of free (public) Internet access locations will effectively help to reduce the digital gap. However, if cross-sectional variance is driven by gross benefits (or perceived gross benefits when, for example, potential adopters cannot accurately forecast the gross benefits of Internet adoption) it would be more efficient to inform the non-users about the opportunities offered by the Internet, and also to promote access to the digital skills which are required for efficient use of the Internet in order to encourage Internet adoption.

---

<sup>1</sup>A non-exhaustive list of work using cross-sectional data includes: U.S. Department of Commerce (1995, 1999), Chaudhuri et al. (2005), Chaudhuri and Flamm (2007); while Prieger and Hu (2009) provide a dynamic approach of the digital divide.

<sup>2</sup>Notice that the purpose of their analysis is not Internet adoption, but computer adoption.

<sup>3</sup>The quality of service is measured by the distance of households from the central office - the smaller the distance, the higher the speed offered by ADSL.

<sup>4</sup>The main idea is that competition can increase Internet adoption because it leads to price cuts and quality improvements.

In this paper, we propose a model in which the household’s decision to adopt the Internet is analyzed sequentially. Prior to subscribing to an Internet service provider, the household considers the need for an Internet connection at home. Subsequently, conditional to a need for an Internet connection at home, the household chooses whether to adopt the Internet or not by comparing its adoption costs to its gross benefits of adoption. Therefore, contrary to the papers cited above, the digitally excluded population can be separated into two groups. In the first group, we find the households that are not interested in adopting the Internet; in the second group, we find the households that desire to adopt the Internet but do not do so because their adoption costs are higher than their gross benefits of adoption.

Our research supports three central findings. First, we determine the main factors that explain the cross-sectional variance in gross benefits. While it has been expressed in the literature that many of the people on the wrong side of the divide simply do not desire to adopt the Internet,<sup>5</sup> there is not, to our knowledge, any paper which quantifies the effect of several factors on the gross benefits of adoption. Our paper fills in this gap. In particular, we show that, on average, a wealthy urban household has higher gross benefits than a destitute and rural household. These results suggest that degree of urbanization and income capture more than just availability and affordability constraints, respectively. Second, we estimate the predicted non adoption rate and break it down in two parts: on the one hand, the predicted probability that the household does not desire to adopt the Internet, on the other hand, the predicted probability that the household does desire to adopt the Internet and does not adopt it because its adoption costs are higher than its gross benefits of adoption. While the cross-sectional variance in the predicted probability that a household desires to adopt the Internet and does not because its adoption costs are higher than its gross benefits is relatively low (except for the age factor), we show that the cross-sectional variance in the predicted probability that a household does not desire to adopt the Internet is high. Third, we compute the predicted adoption probability assuming that the adoption costs are homogeneous across households. We show that, for a given dimension (except for the age factor), the adoption rate will be only slightly modified if the adoption costs are homogeneous across households. These results support the argument that the digital divide is mainly due to differences in gross benefits of adoption. Therefore, before taking into consideration inequalities in infrastructure deployment

---

<sup>5</sup>For example Lenhart et al. (2003), Fox (2005) and Van Dijk (2005) notice that many of those who remain on the “wrong side” of the digital divide are just not interested in adopting the Internet. While these papers provide a descriptive approach, in our paper we quantify the effect of several socio-economic factors on the gross benefits of adoption.

or differences in adoption costs, policy makers which attempt to reduce the digital gap should try to increase the (perceived) gross benefits of adoption for the most socially disadvantaged people.

The rest of the paper is organized as follows. In the next section, we outline the conceptual framework and the estimation method. In Section 3, we describe the data and the procedure we follow to restrict our sample. In Section 4, we present the results and also analyze two possible explanations that might affect the cross-sectional variance in gross benefits. In Section 5, we discuss some policy implications of our results. In Section 6, we provide a short conclusion to this paper.

## 2 Empirical framework

In this section, we develop a model of Internet adoption which takes into account the household's desire to adopt the Internet. First, we present the conceptual framework and then we present the estimation method.

### 2.1 Conceptual framework

We note  $B$  the household's gross benefits of Internet adoption and  $C$  the household's costs of adoption. The adoption costs include the financial costs of an Internet connection at home (i.e., the monetary cost of signing up with an Internet service provider and the access fee) but also non-monetary costs such as the transportation costs to subscribe to an Internet offer or the time taken to install the Internet.

Neither the gross benefits of adoption nor the adoption costs are usually observed. For example, previous papers analyzing Internet adoption only observe whether or not the households have adopted the Internet. In other words, they only observe if their net benefits of adoption (i.e.,  $B - C$ ) are positive or negative.<sup>6</sup> Therefore, households which are not interested in adopting the Internet and the ones which benefit from an Internet connection at home but not enough to compensate their adoption costs are grouped into the same class: the excluded ones. In our data, aside from observing if the household has adopted the Internet or not, we also observe, for each household without an Internet connection at home, whether or not it desires to adopt the Internet.<sup>7</sup> This supplementary information regarding the household's desire to adopt the Internet or not, allow

---

<sup>6</sup>In a paper analyzing business Internet adoption, Forman et al. (2005) identify the effect of location size on gross benefits of Internet adoption. Contrary to our paper, where the household's desire to adopt the Internet is used to identify gross benefits of adoption, in their paper, identification is obtained by observing the marginal effect of location for various form of Internet adoption decisions.

<sup>7</sup>The way the variable related to the desire has been built is detailed in Subsection 3.1.

us, contrary to the previous papers in the literature, to explicitly separate the digitally excluded population into two groups: on the one hand, the households which do not desire to adopt the Internet, on the other hand, the households which do desire to adopt the Internet but do not do so because their adoption costs are higher than their gross benefits of adoption.

We have created two binary variables  $D_i$  and  $A_i$ , with  $D_i = 1$  if the household  $i$  desires to have an Internet connection at home and 0 otherwise, and  $A_i = 1$  if the household  $i$  has adopted the Internet and 0 otherwise. Since, we do not observe the household's desire to have an internet connection at home if it has one, we assume that, each household with an Internet connection at home has a desire for it. Therefore, in our model, the occurrence  $D_i = 0$  and  $A_i = 1$  is not possible.

The adoption decision is modeled as a two stage process. In the first stage, the household wonders if it needs an Internet connection at home; in the second stage, if there is a need for it, the household subscribes to an Internet offer if its gross benefits of adoption exceed its adoption costs. Our model is based on the two following equivalence relations:<sup>8</sup>

$$D_i = 1 \text{ if and only if } B_i > 0, \text{ and } 0 \text{ otherwise,}$$

and

$$A_i = 1 \text{ if and only if } B_i - C_i \geq 0, \text{ and } 0 \text{ otherwise.}$$

Given that the household's decision is sequential (prior to choosing whether to adopt the Internet or not, the household wonders if there is a need for it), we have that:

$$A_i \text{ is observed if and only if } D_i = 1,$$

and

$$D_i \text{ is observed for the overall sample.}$$

---

<sup>8</sup>In our model, we assume that wanting to adopt the Internet indicates that the gross benefits of adoption are greater than zero. More generally, we could have assumed that wanting to adopt the Internet indicates that the gross benefits of adoption are higher than a certain parameter  $\tau$  (which is normalized to 0 in our model). The results would not have been modified: only the interpretation of the constant term in the gross benefits equation would have changed. The important assumption is that the households are homogeneous to characterize their desire to adopt the Internet. I.e., there is a unique and identical parameter  $\tau$  for all of the households, such that a household does not desire (resp., desires) to adopt the Internet if its gross benefits of adoption are smaller than  $\tau$  (resp., higher than  $\tau$ ).

Thus, the household  $i$  can fall into one of the three following categories: it does not benefit from an Internet connection at home ( $D_i = 0$ ), it benefits from an Internet connection at home and its gross benefits are smaller than its adoption costs ( $D_i = 1, A_i = 0$ ), it benefits from an Internet connection at home and its gross benefits are higher than its adoption costs ( $D_i = 1, A_i = 1$ ).

## 2.2 Estimation methodology

We assume that the household  $i$ 's gross benefits of Internet adoption ( $B_i$ ) and the household  $i$ 's costs of Internet adoption ( $C_i$ ) are modeled in the following way:

$$B_i = X_i\beta_1 + u_{1,i}, \quad (1)$$

and

$$C_i = X_i\beta_2 + u_{2,i}, \quad (2)$$

where  $X_i$  is the vector of explanatory variables<sup>9</sup> and  $u_{1,i}$  and  $u_{2,i}$  are the error terms. The error terms  $u_{1,i}$  and  $u_{2,i}$  are assumed to be normally distributed with mean  $(0, 0)$  and variance-covariance matrix  $\Sigma$ , where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

and where  $\sigma_1^2$  and  $\sigma_2^2$  correspond to the variance of unobservable variables in each equation and  $\rho\sigma_1\sigma_2$  corresponds to the covariance between the unobservable variables. Thus, we have:

$$D_i = 1 \text{ if and only if } B_i = X_i\beta_1 + u_{1,i} > 0, \text{ and } 0 \text{ otherwise,} \quad (3)$$

and

$$A_i = 1 \text{ if and only if } B_i - C_i = X_i\delta + v_i \geq 0, \text{ and } 0 \text{ otherwise,} \quad (4)$$

where  $\beta_1$  and  $\delta = \beta_1 - \beta_2$  are vectors of parameters to be estimated and  $u_{1,i}$  and  $v_i = u_{1,i} - u_{2,i}$  are normally distributed with mean  $(0, 0)$  and variance-covariance matrix

$$\tilde{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_1^2 - \rho\sigma_1\sigma_2 \\ \sigma_1^2 - \rho\sigma_1\sigma_2 & \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2 \end{pmatrix}.$$

---

<sup>9</sup>A constant term is added to the explanatory variables.

Notice that only  $\frac{\beta_1}{\sigma_1}$ ,  $\frac{\delta}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}$  and  $\frac{\sigma_1 - \rho\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}$  can be identified. We note these vectors of parameters as  $\tilde{\beta}_1$ ,  $\tilde{\delta}$  and  $\tilde{\rho}$  respectively.<sup>10</sup> The observations being independent and identically distributed, the log likelihood L is then given by:

$$L(\tilde{\beta}_1, \tilde{\delta}, \tilde{\rho} / A, D, X) = \sum_{i/D_i=0} \text{Log} \left\{ \int_{-\infty}^{-X_i \tilde{\beta}_1} f(\varepsilon_1) d\varepsilon_1 \right\} + \sum_{i/D_i=1, A_i=0} \text{Log} \left\{ \int_{-X_i \tilde{\beta}_1}^{\infty} \int_{-\infty}^{-X_i \tilde{\delta}} h(\varepsilon_1, \varepsilon_2) d\varepsilon_2 d\varepsilon_1 \right\} \\ + \sum_{i/D_i=1, A_i=1} \text{Log} \left\{ \int_{-X_i \tilde{\beta}_1}^{\infty} \int_{-X_i \tilde{\delta}}^{\infty} h(\varepsilon_1, \varepsilon_2) d\varepsilon_2 d\varepsilon_1 \right\},$$

where  $f(\cdot)$  denotes the density function of the standard normal distribution and  $h(\cdot, \cdot)$  the density function of the bivariate standard normal distribution. The maximum likelihood method is used to estimate the model.

### 3 The data

This paper was made possible by the availability of a unique database, collected in the form of surveys of French households by INSEE in October 2005.<sup>11</sup> In particular, the database contains information on demographic variables and on ICT<sup>12</sup> equipment of French households.

#### 3.1 The dependent variables

In our database, neither the gross benefits of adoption nor the adoption costs are observed. Nevertheless, for each household, we observe whether or not it has an Internet connection at home. In case, it does not have one and that the household is composed of at least two people, we observe whether or not each individual in this household wishes to have an Internet connection at home. Indeed, the question "Do you wish to have an Internet connection at home?" was asked to all individuals living in a household without an Internet connection at home and composed of at least two people (this question was not asked to single person households). In order to determine whether

---

<sup>10</sup>Notice that the vector of parameters  $\tilde{\beta}_2 = \beta_2/\sigma_2$  is identifiable if we assume that  $\sigma_1 = \sigma_2 = \sigma$  or if we assume that  $\rho = 0$ . In this paper, we do not go deeper into this identification point for two reasons. First, none of these assumptions are testable. Second, we have to use the estimation of the residual correlation to identify this vector of parameters; its standard deviation being not negligible, the identification of the vector of parameters  $\tilde{\beta}_2$  (obtained under one of the two untestable previously mentioned assumptions) will not be robust.

<sup>11</sup>The name of the survey is "technologies de l'information et de la communication."

<sup>12</sup>ICT refers to Information and Communication Technologies.

a household desires to adopt the Internet or not, we assume that household  $i$  desires to have an Internet connection at home if it has one<sup>13</sup> or if it does not have one but at least one member of its reference group has answer "yes" to the former question. The reference group being defined as the group of people that make decisions affecting the household (e.g., regarding the household's consumption).<sup>14</sup> Thus,  $D_i = 0$  if the household  $i$  does not have an Internet connection at home and all the members belonging to its reference group have answer "no" to the question "Do you wish to have an Internet connection at home?".  $D_i = 0$  and  $A_i = 1$  if the household  $i$  does not have an Internet connection at home and at least one of the members of its reference group has answered "yes" to the former question. Finally,  $D_i = 1$  and  $A_i = 1$  if the household  $i$  has an Internet connection at home.

**Remarks regarding our way of classifying the households into the three previous categories.** First, taking into account only the members of the reference group to describe the household's desire to adopt the Internet can be viewed as reductive. For example, one can argue that all the members of the household which are over a certain age should be taken into consideration in order to properly describe the household's desire to adopt the Internet.<sup>15</sup> We have chosen to take into consideration only the desire (to adopt the Internet) of the members of the reference group in order to reduce the effect of the household's size. To ensure the robustness of our results, we have also used other alternatives to define the household's desire to adopt the Internet (i.e., where all the household members which are over a certain age are taken into consideration). Our

---

<sup>13</sup>Therefore, as noticed previously, the occurrence  $D_i = 0$  and  $A_i = 1$  is not possible in our model.

<sup>14</sup>In our database, households specify themselves which members belong to the reference group. The children do not usually belong to the reference group.

<sup>15</sup>Notice that the definition used to state whether or not a household desires to adopt the Internet raises two different potential errors. First, if we take into consideration only the members of the reference group to describe the household's desire to adopt the Internet, then it is possible that the desire of some individuals relevant to describe the household's desire to adopt the Internet (e.g., a child which is at the university) were erroneously not taken into consideration. Thus, it is possible that Internet adoption is observed only because a third person (who does not belong to the reference group) wanted to have an Internet connection at home and not because at least one of the members of the reference group did it. This potential issue will be captured in the errors terms of the gross benefits and the net benefits equations. In these cases, the error terms will be large in both equations (and thus positively correlated). Second, in the opposite way, if we take into consideration too many people to describe the household's desire to adopt the Internet, then it is possible that the desire of some individuals (e.g., a young child), which do not affect the household's adoption choice, were erroneously taken into consideration. Thus, some households will fall into the category "desire to adopt the Internet but do not do so because its adoption costs are higher than its gross benefits" while they should fall into the category "do not desire to adopt the Internet." In these cases, while the non-adoption choice is actually due to too low gross benefits of adoption, we will erroneously consider that it is due to an affordability issue (i.e., due to too high adoption costs). Thus, the error terms will be large in the gross benefits equation and small in the net benefits equation (and thus negatively correlated). Therefore, the smaller the number of people taken into consideration to describe the household's desire to adopt the Internet, the higher the estimated residual correlation (i.e.,  $\tilde{\rho}$ ) will be.



main results are unchanged.<sup>16</sup>

Second, it could be inaccurate to conclude that a household desires to adopt the Internet when at least one member of its reference group has answered "yes" to the question "Do you wish to have an Internet connection at home?". For example, if the household's gross benefits are defined as the sum of members' gross benefits, then it could be possible that a household does not desire to adopt the Internet even if one member of its reference group desires to do so.<sup>17</sup> If the potential benefits from an Internet connection at home are considered as higher than potential disadvantages, the likelihood of encountering such cases is low. Furthermore, notice that the adoption-making mechanism could also be something different of "maximizing the sum of members' gross benefits." For example, a household could choose whether to adopt the Internet or not under the principle of maximizing the member's utility with the higher value among the members of its reference group. Rationality of such extreme group decision-making rule is supported by group decision theory (see, Zhang et al., 2009). In this case, observing that at least one member of the reference group desires to adopt the Internet will be a sufficient condition to ensure that the household desires to do so. There is no clear a priori criterion to select which type of adoption-making mechanism is more suitable. This could probably be tested. However, this question is beyond the scope of this paper.

### 3.2 Sample restriction

The database contains initially 5603 observations and is composed of two parts: one at the household level and the other at the individual level. In each household, one individual (belonging or not to the reference group) was randomly drawn to participate to a detailed (individual) investigation.

All the households in which the individual drawn to participate to this detailed investigation did not belong to the reference group have been excluded from the sample.<sup>18</sup> After the exclusion of missing data, the "whole" sample contains 4491 observations.

Since the question "Do you wish to have an Internet connection at home?" was not asked

---

<sup>16</sup>In appendix I, we show the results of the estimation obtained when all the household members which are over 15 years old are taken into consideration to define the household's desire to adopt the Internet (i.e., we assume that household  $i$  desires to have an Internet connection at home if it has one or if it does not have one but at least one of its member over 15 years old has answer "yes" to the question "Do you wish to have an Internet connection at home?"). The results obtained when all the household members are taken into consideration to define the household's desire to adopt the Internet are available from the author upon request.

<sup>17</sup>For example, one could be better off without an Internet connection at home because she/he is worried about online pornography, credit card theft or virus attack. Thus, the positive gross benefits of one individual could be too small to compensate the disutility of a home Internet connection of the other one.

<sup>18</sup>Excluding these households is similar as excluding missing at random data and therefore does not raise any issue regarding the robustness of our results.

to single person households, it is not possible to define whether or not households with a single person desire to adopt the Internet. Therefore, single person households must be excluded from the sample.<sup>19</sup> Once the single person households have been excluded, 3114 observations remain. Among these 3114 observations, there are 2872 households in which the reference group is composed of two people. To the extent that a household will be more likely to desire to have an Internet connection at home that the number of people belonging to its reference group is high,<sup>20</sup> only households in which the reference group is composed of two people have been kept.<sup>21</sup> These selection rules leave us with a total of 2872 cases for the empirical analysis. Table 1 shows the sharing out of the households between the three possible categories:  $D_i = 0, D_i = 1$  and  $A_i = 0, D_i = 1$  and  $A_i = 1$ .

Table 1: Descriptive statistics of the dependent variables

	Desire to have an Internet connection at home	
	No	Yes
No	1153 (0.402)	0 (0.000)
Yes	423 (0.147)	1296 (0.451)

Percentage in parentheses. Number of observations= 2872.

From Table 1 we can see that 40.2% of the households do not desire to adopt the Internet, 14.7% of the households desire to adopt the Internet but did not adopt it because their gross benefits of adoption are smaller than their adoption costs and finally 45.1% of the households have adopted the Internet.

<sup>19</sup>The fact that we can not use the single person households to estimate the model reduce by more than 30% the whole sample. The sample is not anymore representative of the french population and those for two reasons. First, we do not observe single persons households. Second, the size of the household is correlated to some of the explanatory variables. For example, since the proportion of single person households is higher at both ends of the age distribution than at the middle of the age distribution, the dispersion of the age distribution will be lower in the selected sample than in the overall population. Not observing the single person households raises two issues. First, the estimated coefficients could be biased because in our selected sample, the elderly (i.e., older than 70 years old), the young (younger than 30 years old), the destitute and the urban households are less represented than in the population. However, to the extent that these variables are (only) slightly correlated with the size of the household (except for the age) and that we still observe high variability in these variables in our selected sample, this is not a fundamental issue. Second, if living alone has a significant effect on the gross benefits of adoption and/or on the adoption costs then the predicted probabilities calculated in Subsection 4.2 could overestimate or underestimate the true rates in the population. Therefore, these rates have to be interpreted with caution: they are valid for the subsample of households with more than one person.

<sup>20</sup>Higher is the number of people belonging to the reference group higher is the probability that at least one member of the reference group has answer "yes" to the question "Do you wish to have an Internet connection at home?."

<sup>21</sup>The results obtained when all the households with more than one person (i.e., with 3114 households) are used to estimate the model are similar. They are available from the author upon request

### 3.3 The explanatory variables

The explanatory variables<sup>22</sup> used for my econometric studies can be separated into two sets. One set is the data at the individual level, which controls for age, education, opinion on new technologies, individual's technophilia<sup>23</sup> and sociability<sup>24</sup> of one individual of the reference group. The second set consists of the data at the household level, which controls for adjusted income,<sup>25</sup> characteristics of the metropolitan area (average income, degree of urbanization), number of persons, household's technophilia<sup>26</sup>, presence or absence of children and the age of the older child in the household. A detailed description and descriptive statistics<sup>27</sup> of these variables are summarized in Table 2. Columns (2) and (3) of Table 2 show the means of the explanatory variables on the selected sample (i.e., the one used to estimated the model) and on the whole sample, respectively.<sup>28</sup>

---

<sup>22</sup>A constant term is added to these explanatory variables.

<sup>23</sup>We have created one variable related to the individual's technophilia. This variable is labeled *computer* and takes two values, 1 if she/he has already used a computer, 0 otherwise.

<sup>24</sup>Similar to Suire (2007), we have created two variables related to the social environment of the individual. The first variable is labeled *association*. This variable takes two values, 1 if she/he is an active member of an association, 0 otherwise. This variable concerns the capital within the meaning of Putnam (1993). According to Putnam, being a member of an association strengthens, among other things, the trust between individuals. The second is named *friends*. This variable takes two values, 1 if she/he meets with her/his friends at least once a week, 0 otherwise. This variable refers, stricto sensu, to the social network (i.e., the number of acquaintances).

<sup>25</sup>See Guillemin and Roux (2001) for further investigation on the weighting of income.

<sup>26</sup>We have created one variable related to the household's technophilia. This variable is labeled *mobile* and takes three values, 0 if the household does not have any mobile phone, 1 if it has one mobile phone and 2 otherwise.

<sup>27</sup>Apart from the variable *number of persons* and *mobile*, all other variables are binary variables. Therefore we do not defer in this table the minimum, the maximum and the variance of each variable. Concerning the variable *number of persons*, its minimum value is 1 (resp., 2), its maximum value is 9, and its standard deviation is 1.23 (resp., 1.08) in the whole sample (resp., in the selected sample). Concerning the variable *mobile*, its minimum value is 0, its maximum value is 2 and its standard deviation is 0.49 (resp. 0.48) in the whole sample (resp., the selected sample).

<sup>28</sup>In the whole sample, the households are, on average, more destitute and living in a more urbanized area than the ones in the selected sample. Moreover, the variability of the age distribution is higher in the whole sample than in the selected sample.

Table 2: Description and summary statistics of the explanatory variables

		(1)	(2)	(3)
			Mean in our sample	Mean in the whole sample
<b>HOUSEHOLD</b>				
NUMBER OF PERSONS	= number of people in the household		2.879	2.287
<b>CHILD</b>				
no child	=1 if there is not any child in the household, 0 otherwise		0.533	0.649
from 0 to 9 years	=1 if the older child in the household is less than 10 years old, 0 otherwise		0.209	0.149
from 10 to 15 years	=1 if the older child in the household is between 10 and 15 years old, 0 otherwise		0.117	0.091
16 years and more	=1 if the older child in the household is older than 15 years old, 0 otherwise		0.140	0.111
<b>INCOME (adjusted )</b>				
less than 899 euros	=1 if the monthly income is inferior than 899 euros, 0 otherwise		0.157	0.191
from 900 to 1149 euros	=1 if the monthly income is between 900 and 1149 euros, 0 otherwise		0.173	0.184
from 1150 to 1499 euros	=1 if the monthly income is between 1150 and 1499 euros, 0 otherwise		0.205	0.202
from 1500 to 1999 euros	=1 if the monthly income is between 1500 and 1999 euros, 0 otherwise		0.222	0.206
more than 2000 euros	=1 if the monthly income is superior than 2000 euros, 0 otherwise		0.243	0.218
<b>MUNICIPALITY</b>				
low average income	=1 if the annual average income in the area is smaller than 20000 euros, 0 otherwise		0.291	0.323
<b>URBANIZATION</b>				
low	=1 if the population density in the metropolitan area is low, 0 otherwise		0.195	0.168
medium	=1 if the population density in the metropolitan area is intermediate, 0 otherwise		0.383	0.347
high	=1 if the population density in the metropolitan area is high, 0 otherwise		0.423	0.485
MOBILE	=number of mobile phones in the household (takes the value 2 if the household has more than 2 mobile phones)		1.607	1.534
<b>INDIVIDUAL</b>				
<b>EDUCATION</b>				
no diploma	=1 if she/he has got no diploma, 0 otherwise		0.141	0.144
school certificate	=1 if her/his highest degree is a school certificate, 0 otherwise		0.215	0.225
high school prof. degree	=1 if her/his highest degree is a high school professional degree, 0 otherwise		0.264	0.249
high school acad. degree	=1 if her/his highest degree is a high school academic degree, 0 otherwise		0.118	0.126
2 years university degree	=1 if her/his highest degree is a 2 years university or technical degree, 0 otherwise		0.100	0.092
4 years university degree	=1 if her/his highest degree is a 4 years university degree, 0 otherwise		0.161	0.163
<b>AGE</b>				
71 years and more	=1 if she/he is older than 71 years old, 0 otherwise		0.129	0.168
from 61 to 70 years	=1 if she/he is older than 61 years old and younger than 70 years old, 0 otherwise		0.143	0.144
from 51 to 60 years	=1 if she/he is older than 51 years old and younger than 60 years old, 0 otherwise		0.196	0.177
from 41 to 50 years	=1 if she/he is older than 41 years old and younger than 50 years old, 0 otherwise		0.187	0.174
from 31 to 40 years	=1 if she/he is older than 31 years old and younger than 40 years old, 0 otherwise		0.221	0.196
30 years and less	=1 if she/he is younger than 30 years old, 0 otherwise		0.124	0.141
<b>COMPUTER</b>				
SOCIABILITY	=1 if she/he has already used a computer, 0 otherwise		0.541	0.520
association	=1 if she/he is an active member of an association, 0 otherwise		0.505	0.547
friends	=1 if she/he meets with her/his friends at least once a week, 0 otherwise		0.334	0.329
<b>OPINION</b>				
Oa	=1 if she/he totally agrees that new technologies make life easier, 0 otherwise		0.415	0.409
Ob	=1 if she/he rather agrees that new technologies make life easier, 0 otherwise		0.463	0.460
Oc	=1 if she/he does not agree that new technologies make life easier, 0 otherwise		0.122	0.130
Number of observations			2872	4491

## 4 The results

### 4.1 The estimation results

Table 3 provides the estimation results.<sup>29</sup> All results in Table 3 are marginal effects.<sup>30</sup> Columns (1), (2) and (3) of Table 3 presents the marginal effects for the probabilities that the household desires to adopt the Internet, that the household has adopted the Internet given that it has a desire to do so and that the household has adopted the Internet, respectively. We carried out a Wald test of the null hypothesis that residual correlation  $\tilde{\rho}$  is equal to zero against the alternative that it is different from zero. A Wald test rejects at the 10% level the null hypothesis that the residual correlation  $\tilde{\rho}$  is equal to zero.<sup>31</sup>

---

<sup>29</sup>When the same independent variables appear in both of equations (i.e., equations (3) and (4)) the model is technically identified, but identification only occurs on the basis of distributional assumptions about the residuals alone and is not due to variation in the explanatory variables. In order to increase the precision of the estimation, three explanatory variables (friends, Ob and Oc) were excluded from the net benefits equation (i.e., equation (4)). When they are added to the set of explanatory variables used in the net benefits equation, these three variables do not have significant effect. Therefore, they are good candidates for valid exclusion.

<sup>30</sup>These marginal effects are based on the average level of the covariates. For example, for the income, everything else equal to the average level of the covariates, Column (1) of Table 3 shows that the difference between the probability that a household with an income higher than 2000 euros desires to adopt the Internet and the probability that a household with an income lower than 899 euros desires to adopt the Internet is equal to 0.230.

<sup>31</sup>Notice that in the model where all the household members which are over 15 years old are taken into consideration to define the household's desire to adopt the Internet, a Wald test does not reject at the 10% level the null hypothesis that the residual correlation  $\tilde{\rho}$  is equal to zero (see, Table 7 in Appendix I for the estimation results). The reason of this difference is discussed in Footnote 15. Besides, the only significant differences between the results in Table 3 and in Table 7 are those regarding the marginal effects of the variables related to the age of the older child in the household. This is a basic automatic effect attributable to the way used to describe the household's desire to adopt the Internet

Table 3: Estimation results

	(1)		(2)		(3)	
	The marginal effects for p(D=1/X)		The marginal effects for p(A=1/D=1, X)		The marginal effects for p(A=1/X)	
<b>HOUSEHOLD</b>						
NUMBER OF PERSONS	0.021	( 0.019 )	0.031	( 0.023 )	0.034 *	( 0.019 )
CHILD						
no child	-0.117 ***	( 0.045 )	-0.173 ***	( 0.049 )	-0.203 ***	( 0.048 )
from 0 to 9 years	-0.076 *	( 0.045 )	-0.218 ***	( 0.047 )	-0.204 ***	( 0.045 )
from 10 to 15 years	-0.041	( 0.044 )	-0.173 ***	( 0.049 )	-0.151 ***	( 0.047 )
16 years and more		ref.		ref.		ref.
INCOME (adjusted )						
less than 899 euros	-0.230 ***	( 0.039 )	-0.174 ***	( 0.051 )	-0.267 ***	( 0.039 )
from 900 to 1149 euros	-0.161 ***	( 0.036 )	-0.150 ***	( 0.045 )	-0.211 ***	( 0.037 )
from 1150 to 1499 euros	-0.115 ***	( 0.034 )	-0.080 **	( 0.039 )	-0.139 ***	( 0.035 )
from 1500 to 1999 euros	-0.071 **	( 0.032 )	-0.049	( 0.034 )	-0.088 ***	( 0.033 )
more than 2000 euros		ref.		ref.		ref.
MUNICIPALITY						
low average income	-0.060 **	( 0.024 )	-0.063 **	( 0.030 )	-0.080 ***	( 0.024 )
URBANIZATION						
Low	-0.076 ***	( 0.029 )	-0.028	( 0.037 )	-0.071 **	( 0.030 )
Medium	-0.047 *	( 0.024 )	-0.032	( 0.029 )	-0.054 **	( 0.024 )
High		ref.		ref.		ref.
MOBILE	-0.005	( 0.022 )	0.064 **	( 0.028 )	0.037	( 0.023 )
<b>INDIVIDUAL</b>						
<b>EDUCATION</b>						
no diploma	-0.314 ***	( 0.043 )	-0.195 ***	( 0.056 )	-0.349 ***	( 0.043 )
school certificate	-0.209 ***	( 0.038 )	-0.184 ***	( 0.044 )	-0.277 ***	( 0.040 )
high school prof. degree	-0.181 ***	( 0.035 )	-0.169 ***	( 0.037 )	-0.250 ***	( 0.036 )
high school acad. degree	-0.140 ***	( 0.042 )	-0.086 **	( 0.041 )	-0.170 ***	( 0.040 )
2 years university degree	-0.051	( 0.044 )	0.012	( 0.038 )	-0.033	( 0.044 )
4 years university degree		ref.		ref.		ref.
AGE						
71 years and more	-0.491 ***	( 0.044 )	0.227 ***	( 0.071 )	-0.185 ***	( 0.048 )
from 61 to 70 years	-0.314 ***	( 0.042 )	0.116 *	( 0.062 )	-0.108 **	( 0.048 )
from 51 to 60 years	-0.242 ***	( 0.038 )	0.239 ***	( 0.050 )	0.011	( 0.045 )
from 41 to 50 years	-0.155 ***	( 0.036 )	0.253 ***	( 0.047 )	0.088 **	( 0.044 )
from 31 to 40 years	-0.106 ***	( 0.034 )	0.141 ***	( 0.043 )	0.046	( 0.039 )
30 years and less		ref.		ref.		ref.
COMPUTER	0.276 ***	( 0.024 )	0.165 ***	( 0.032 )	0.293 ***	( 0.025 )
SOCIABILITY						
Association	0.050 **	( 0.023 )	0.063 **	( 0.027 )	0.076 ***	( 0.023 )
Friends	-0.042 **	( 0.021 )	0.018	( 0.015 )	-0.018 *	( 0.011 )
OPINION						
Oa		ref.		ref.		ref.
Ob	-0.111 ***	( 0.024 )	0.049 **	( 0.025 )	-0.046 *	( 0.025 )
Oc	-0.214 ***	( 0.035 )	0.093 *	( 0.049 )	-0.098 **	( 0.041 )

athrho = 0.787\* (0.417)

rho (residual correlation) = 0.657 (0.237)

Wald test of rho = 0: chi2(1) = 3.56 Prob > chi2 = 0.0590

Log Likelihood of the basic model=-2094.186.

Robust standard errors in parentheses. Number of observations= 2872.

\*  $p < 0.10$ . \*\*  $p < 0.05$ . \*\*\*  $p < 0.01$ .

The effects of socio-economic factors on the gross benefits of adoption are displayed in Column (1) of Table 3. Interestingly, the income and the degree of urbanization positively affect the household's gross benefits of adoption. Those results suggest that the budgetary constraint is not the only reason why, on average, destitute households are less connected to the Internet than wealthy households and also that higher infrastructure deployment in the urban areas than in the rural areas cannot entirely explain the "rural digital divide." Recalling that the diffusion of the Internet heavily relies upon intensive interactions between people,<sup>32</sup> spatial proximity as well as belonging to a group where Internet usage is widespread might increase the value of Internet adoption. Therefore, network effects could partly explain why an urban, wealthy and educated<sup>33</sup> household has, on average, higher gross benefits of adoption than a rural, less wealthy and less educated household.

Column (2) of Table 3 shows that the main determinants cited in the literature to explain the digital divide in Internet connection - the income, the residential area,<sup>34</sup> education - are still significant when the sample is restricted to households which desire to adopt the Internet. Indeed, conditionally to the desire for it, wealthy, educated households located in high income area are more likely to have adopted the Internet than destitute, uneducated households located in low income area.<sup>35</sup>

Columns (1) and (2) of Table 3 show that the income, the residential area and the education affect the gross benefits of adoption and the net benefits of adoption provided that the household desires to adopt the Internet in the same way (i.e., the sign of the marginal effects are similar). This similarity is not verified for every socio-economic factors. Indeed, whereas the gross benefits are negatively affected by the age, we find that provided that they desire to adopt the Internet, young households (i.e., younger than 30 years old) are not more willing to adopt the Internet than others. This result is even reversed. Provided that they desire to adopt the Internet, households with residents in their twenties are less likely to have adopted the Internet than the others. The

---

<sup>32</sup>For example, Goldfarb (2006) shows that people living with students in the Mid-1990s are more likely to have adopted the Internet in 2001. He notices that one of the explanation could be that the Internet may exhibit network externalities and thus the benefits of adoption will increase if a household member is online (because of university attendance in the mid-1990s).

<sup>33</sup>To the extent that they communicate mostly with others that have identical educational background and income, people with high salary and high education degree are more likely to know someone that has adopted the Internet.

<sup>34</sup>Notice that only the average income in the area has a significant effect. Provided that the household desires to adopt the Internet, the population density does not significantly affect the probability of adoption.

<sup>35</sup>This result must be interpreted with caution: it does not imply that these variables (the income, the residential area, education) have an effect on the adoption costs. Indeed, the cross-sectional variance in Internet adoption provided that the household desires to adopt the Internet might be due to differences in adoption costs as well as differences in gross benefits of adoption. For example, if the adoption costs were independent of the variable  $x$  and the gross benefits were increasing with the variable  $x$ , then the variable  $x$  will have a positive effect on the probability that the household has adopted the Internet provided that it desires to do so.

fact that the households in the twenties age bracket have smaller net benefits of adoption provided that they desire to adopt the Internet and also higher gross benefits of adoption indicates that, in the sub-population of households desiring to adopt the Internet, the adoption costs are higher for people in their twenties than for the others.

## 4.2 Divided the digitally excluded group into two parts

In addition to determine the main factors that affect the gross benefits of adoption and the net benefits of adoption provided that the household desires to adopt the Internet, our second contribution is to quantify (for various socio-economic characteristics) what percentage of the divide is due to differences in interest levels and what percentage is due to differences in affordability levels. Thereafter, we can analyze whether the cross-sectional variance in Internet adoption is mainly due to differences in costs or in gross benefits. To this end, we have calculated the predicted adoption and non-adoption rates. The predicted non-adoption rate is divided into two parts: on the one hand, the predicted probability that a household does not desire to adopt the Internet (i.e.,  $p(D_i = 0/X_i)$ ), on the other hand, the predicted probability that a household desires to adopt the Internet and does not because its adoption costs are higher than its gross benefits (i.e.,  $p(D_i = 1, A_i = 0/X_i)$ ).

In order to gain understanding of the variability of the adoption costs, these predicted rates are also computed assuming that the adoption costs were homogeneous across households. The value of the homogeneous adoption costs is calibrated such that the predicted probability that an "average" household (i.e., with the average level of the covariates) has adopted the Internet is unchanged whether the adoption costs are homogeneous or not.<sup>36</sup>

The predicted probabilities that a household does not desire to adopt the Internet, that a household desires to adopt the Internet and does not because its adoption costs are higher than its gross benefits and that a household has an Internet connection at home are displayed in columns

---

<sup>36</sup>Denoting by  $K$  the value of the homogeneous adoption costs,  $K$  is such that  $p(\bar{X}\hat{\beta}_1 - K + w_1 > 0/\bar{X}) = p(\bar{X}\hat{\beta}_1 + w_1 > 0, X\hat{\delta} + w_2 > 0/\bar{X})$ , with  $w_1$  and  $w_2$  the error terms assumed to be normally distributed with mean  $(0, 0)$  and variance-covariance matrix

$$\hat{\Sigma} = \begin{pmatrix} 1 & \hat{\rho} \\ \hat{\rho} & 1 \end{pmatrix},$$

$\bar{X}$  the average level of the covariates,  $\hat{\beta}_1$ ,  $\hat{\delta}$  and  $\hat{\rho}$  the estimated values of  $\tilde{\beta}_1$ ,  $\tilde{\delta}$  and  $\tilde{\rho}$ , respectively. Thus, if the adoption costs are homogeneous across households, the predicted probability that a household  $i$  with the covariates  $X_i$  adopt the Internet is equal to  $p(X_i\hat{\beta}_1 + w_1 \geq K/X_i)$  and the predicted probability that a household  $i$  with the covariates  $X_i$  is interested in adopting the Internet and does not do so because its gross benefits of adoption are lower than the homogeneous adoption costs is equal to  $p(0 < X_i\hat{\beta}_1 + w_1 < K/X_i)$ .



(1), (2) and (3) of Table 4, respectively. The similar predicted probabilities when the adoption costs are assumed to be homogeneous across households are displayed in columns (4), (5) and (6) of Table 4, respectively.

Table 4: Predicted non-adoption and adoption rates

	(1)	(2)	(3)	(4)	(5)	(6)
	Standard			Homogeneous adoption costs		
	p(D=0/X)	p(D=1,A=0/X)	p(D=1,A=1/X)	p(D=0/X)	p(D=1,A=0/X)	p(D=1,A=1/X)
INCOME (adjusted )						
less than 899 euros	0,498	0,197	0,305	0,498	0,185	0,317
from 900 to 1149 euros	0,429	0,210	0,361	0,429	0,189	0,382
from 1150 to 1499 euros	0,383	0,183	0,434	0,383	0,189	0,428
from 1500 to 1999 euros	0,340	0,176	0,484	0,340	0,187	0,474
more than 2000 euros	0,268	0,159	0,572	0,268	0,177	0,555
URBANIZATION						
Low	0,413	0,180	0,408	0,413	0,190	0,397
Medium	0,384	0,191	0,425	0,384	0,189	0,427
High	0,337	0,184	0,479	0,337	0,187	0,477
EDUCATION						
no diploma	0,530	0,177	0,294	0,530	0,181	0,290
school certificate	0,424	0,210	0,366	0,424	0,189	0,387
high school prof. degree	0,396	0,211	0,393	0,396	0,189	0,415
high school acad. degree	0,355	0,171	0,473	0,355	0,188	0,457
2 years university degree	0,267	0,123	0,610	0,267	0,176	0,557
4 years university degree	0,215	0,141	0,643	0,215	0,164	0,621
AGE						
71 years and more	0,659	0,081	0,259	0,659	0,154	0,187
from 61 to 70 years	0,483	0,181	0,337	0,483	0,186	0,331
from 51 to 60 years	0,411	0,134	0,456	0,411	0,190	0,399
from 41 to 50 years	0,324	0,144	0,532	0,324	0,185	0,491
from 31 to 40 years	0,275	0,235	0,490	0,275	0,178	0,547
30 years and less	0,169	0,387	0,444	0,169	0,147	0,684

These predicted rates are based on the average level of the covariates.

Column (1) of Table 4 shows that a significant proportion of households, and particularly the most socially disadvantaged ones, simply does not desire to adopt the Internet. For example, 49.8% of the most destitute households and 53% of the less educated ones do not desire to adopt the Internet. Furthermore, it shows that, for a given dimension, the predicted probability that a household does not desire to adopt the Internet varies significantly according to which category the household belongs (i.e., most destitute vs. wealthiest, less educated vs. most educated...). For instance, the most destitute households are approximately two times more likely to not be interested in adopting the Internet than the wealthiest ones.

On the contrary, the cross-sectional variance in the predicted probability that a household desires to adopt the Internet and does not do so is relatively low (except for the age factor). Furthermore, this variability is partly attributable to differences in gross benefits.

The small differences (except for the age factor) between the results in Column (3) and in Column (6) of Table 4 indicate that the cross-sectional variance in Internet adoption is mainly due to differences in gross benefits of adoption. For example, while the wealthiest households are 88% more likely to adopt the Internet than the destitute households, Column (6) of Table 4 shows that the wealthiest households will still be 75% more likely to adopt the Internet than the destitute ones if the adoption costs are homogeneous across households.

Some comments are in order with regard to the uneducated category. Our result suggest that the adoption costs of the uneducated households are similar to the ones of an "average" household. Indeed, the adoption rate of the uneducated households is (almost) unchanged whether the adoption costs are homogeneous across households or not. This could partly be due to the presence of unobservable variables (correlated with the uneducated category) which affect positively the gross benefits of adoption and negatively the adoption costs. From this perspective, the uneducated households could desire to adopt the Internet because they have a high value of an unobservable variable (i.e., because they have large error terms). As a consequence, in the sub-population of households desiring to adopt the Internet, while other households will have a more normal range of value of this unobservable variable, the less educated ones will have a high value of this unobservable variable. Therefore, the uneducated households will be observed as having small adoption costs and a relatively small probability of desiring to adopt the Internet and not doing so (see, Column 2 of Table 4).

Finally, as noticed in Subsection 4.1, we can see that the age factor has an opposite effect on gross benefits and on adoption costs. Indeed, the young households (with residents in their

twenties or thirties) are more likely to be interested in adopting the Internet than the older ones but they are also more likely to encounter an affordability issue (because of higher adoption costs). Interestingly, while the adoption rate of the households in their twenties is similar to the one of the households in their fifties, Column (6) of Table 4 shows that households in their twenties will be approximately 70% more likely to adopt the Internet than the ones in their fifties if the adoption costs are homogeneous across households.

### 4.3 Analyzing different patterns to explain the cross-sectional variance in gross benefits

In this subsection, we identify two possible explanations that might affect the gross benefits of adoption. The two explanations we consider are (i) the level of computer skills and (ii) the expectations regarding the potential benefits of using the Internet.

#### 4.3.1 Two explanations

Below, we discuss each of these explanations and describe the devices used to capture this information.

##### (i) Level of computer skills

In contrast with former innovations in the information and communication markets, such as the phone or the radio, specific knowledge is required in order to benefit entirely from the advantages offered by the Internet. Thus, the lack of computer skills may prevent the appropriation of those benefits and discourage Internet adoption. In our data set, we observe for one individual of the reference group whether she/he has already used a computer or not. In case she/he did it, we also observe her/his computer skills.<sup>37</sup> We have created one variable related to the computer skills. This variable is labeled *computerskills* and takes 7 values according to the number of computer tasks that she/he can accomplish.<sup>38</sup>

##### (ii) Expectations regarding the potential benefits of using the Internet

---

<sup>37</sup>For each individual who has already used a computer, we observe if she/he knows how to use the tool copy/paste to move information in a document, how to install a new hardware, how to install a software, how to compress or uncompress files (Winzip, Winrar, Winace...), how to use basic arithmetic formulas in a spreadsheet (Excel, Quattro, Lotus...), how to write a computer program using a specific computer language (Visual Basic, Fortran, Java, C++...).

<sup>38</sup>The minimum value and the maximum value being 0 and 6, respectively. Descriptive statistics of the variable *computerskills* are displayed in Table 8 in Appendix II.

Many people on the wrong side of the divide might not perceive the potential usefulness of an Internet connection at home and thus cannot accurately forecast the gross benefits of Internet. In our survey, it was asked to one individual of the reference group if she/he totally agrees, rather agrees, disagrees or totally disagrees with the statement "people who do not use the Internet are disadvantaged." We have created one variable related to the opinion regarding the potential disadvantage of not using the Internet.<sup>39</sup> This variable is labeled *expectation* and takes 4 values.<sup>40</sup>

It is not a debate about what is plausible, since it is clear that uninformed and unskilled people have lower gross benefits of adoption than well-informed and skilled ones.<sup>41</sup> The empirical question is how do these different patterns vary with the socio-economic characteristics.

#### 4.3.2 How do these different patterns vary with the socio-economic characteristics?

To determine how the computer skills vary with the socio-economic characteristics, we estimate two models. In the first one (Model A), the dependent variable is *computer*<sup>42</sup> (we use a probit model). In the second one (Model B), the dependent variable is *computerskills* (we use an ordered probit model). Only the individuals who have already used a computer are kept to estimate Model B. To determine how the opinion regarding the potential disadvantage of not using the Internet varies with the socio-economic characteristics, we estimate an ordered probit model (Model C) in which the dependent variable is *expectation*.<sup>43</sup> The explanatory variables used to estimate these models are similar to those used in the main model (see, Table 2). The estimation results of models A, B, C are displayed in Table 5.

---

<sup>39</sup>Notice that 94 individuals did not answer to this question. Therefore, we have only 2778 observations.

<sup>40</sup>The variable *expectation* is equal to 3 if she/he totally agrees, 2 if she/he rather agrees, 1 if she/he disagrees and 0 if she/he totally disagrees with the statement "people who do not use the Internet are disadvantaged." Descriptive statistics of the variable *expectation* are displayed in Table 9 in Appendix II.

<sup>41</sup>For example, while 88.7% of the individuals with the highest level of computer skills have adopted the Internet or desire to do so, only 30.9% of the individuals who have never used a computer have adopted the Internet or desire to do so. In the same manner, while 80.5% of the individuals that totally agree with the statement "people who do not use the Internet are disadvantaged" have adopted the Internet or desire to do so, only 40.6% of the individuals that totally disagree with this statement have adopted the Internet or desire to do so. See Appendix II for further descriptive statistics.

<sup>42</sup>Description and summary statistics of the variable *computer* are displayed in Table 2.

<sup>43</sup>This model has been estimated twice: with and without the explanatory variables Ob and Oc.

Table 5: Estimation results of models A, B, C.

	Model A		Model B		Model C			
	(1)		(2)		(3)	(4)		
	computer		computerskills		expectation	expectation		
<b>HOUSEHOLD</b>								
NUMBER OF PERSONS	-0.110**	(0.048)	-0.045	(0.042)	0.043	(0.037)	0.046	(0.038)
CHILD								
no child	-0.075	(0.123)	-0.095	(0.126)	0.018	(0.099)	0.030	(0.099)
from 0 to 9 years	0.285**	(0.120)	-0.209*	(0.108)	-0.125	(0.088)	-0.115	(0.088)
from 10 to 15 years	0.394***	(0.118)	-0.250**	(0.103)	0.129	(0.090)	0.151*	(0.090)
16 years and more	ref.		ref.		ref.		ref.	
INCOME (adjusted )								
less than 899 euros	-0.856***	(0.108)	0.019	(0.110)	-0.264***	(0.083)	-0.234***	(0.082)
from 900 to 1149 euros	-0.580***	(0.103)	-0.029	(0.088)	-0.218***	(0.074)	-0.196***	(0.074)
from 1150 to 1499 euros	-0.650***	(0.094)	-0.044	(0.087)	-0.164**	(0.068)	-0.162**	(0.068)
from 1500 to 1999 euros	-0.298***	(0.091)	0.020	(0.076)	-0.072	(0.062)	-0.060	(0.062)
more than 2000 euros	ref.		ref.		ref.		ref.	
MUNICIPALITY								
low average income	-0.091	(0.067)	-0.070	(0.064)	0.012	(0.049)	0.013	(0.049)
URBANIZATION								
Low	-0.176**	(0.081)	-0.046	(0.076)	-0.210***	(0.060)	-0.191***	(0.060)
Medium	-0.070	(0.067)	-0.099*	(0.060)	-0.113**	(0.048)	-0.116**	(0.048)
High	ref.		ref.		ref.		ref.	
MOBILE	0.173***	(0.061)	0.157***	(0.059)	-0.059	(0.044)	-0.058	(0.044)
<b>INDIVIDUAL</b>								
EDUCATION								
no diploma	-1.561***	(0.126)	-0.861***	(0.142)	-0.482***	(0.094)	-0.443***	(0.095)
school certificate	-0.991***	(0.112)	-0.504***	(0.105)	-0.336***	(0.079)	-0.309***	(0.079)
high school prof. degree	-0.821***	(0.104)	-0.523***	(0.080)	-0.331***	(0.070)	-0.309***	(0.070)
high school acad. degree	-0.158	(0.126)	-0.446***	(0.083)	-0.201**	(0.078)	-0.180**	(0.078)
2 years university degree	-0.010	(0.137)	-0.103	(0.089)	-0.040	(0.075)	-0.029	(0.075)
4 years university degree	ref.		ref.		ref.		ref.	
AGE								
71 years and more	-2.348***	(0.157)	-0.642***	(0.176)	-0.112	(0.103)	-0.069	(0.105)
from 61 to 70 years	-1.548***	(0.133)	-0.524***	(0.133)	-0.116	(0.100)	-0.086	(0.100)
from 51 to 60 years	-1.073***	(0.124)	-0.286***	(0.100)	-0.077	(0.087)	-0.062	(0.088)
from 41 to 50 years	-0.661***	(0.125)	-0.001	(0.098)	0.014	(0.086)	0.034	(0.086)
from 31 to 40 years	-0.268**	(0.119)	0.183**	(0.079)	-0.087	(0.077)	-0.089	(0.077)
30 years and less	ref.		ref.		ref.		ref.	
COMPUTER	dropped		dropped		0.286***	(0.057)	0.251***	(0.057)
SOCIABILITY								
Association	0.199***	(0.065)	0.069	(0.057)	0.073	(0.045)	0.077*	(0.045)
Friends	-0.006	(0.061)	-0.087	(0.056)	-0.044	(0.043)	-0.065	(0.043)
OPINION								
Oa	ref.		ref.		dropped		ref.	
Ob	-0.225***	(0.063)	-0.244***	(0.056)	dropped		-0.196***	(0.044)
Oc	-0.463***	(0.100)	-0.462***	(0.106)	dropped		-0.511***	(0.079)
Log Likelihood	-1191.250		-2762.740		-3610.376		-3583.406	
Number of observations	2872		1553		2778		2778	

Robust standard errors in parentheses. \*  $p < 0.10$ . \*\*  $p < 0.05$ . \*\*\*  $p < 0.01$ .

Columns (1) and (2) of Table 5 show that the computer skills are unequally distributed in the population. Indeed, the socially disadvantaged (uneducated, low income) and elderly people are less likely to have already used a computer. However, provided that they have already used it, destitute people do not have fewer computer skills than wealthy ones. Thus, income is identified as a barrier to computer use, but once this barrier is overcome, there is not any significant difference in computer knowledge between destitute and wealthy people.

Column (3) of Table 5 shows that destitute people are less likely than wealthy people to agree with the statement that "people who do not use the Internet are disadvantaged." This result suggests that the negative relation between the income and the gross benefits of adoption might partly be due to the fact that destitute people have, on average, lower expectations concerning the usefulness of the Internet. On the contrary, age does not affect the expectations concerning the usefulness of the Internet. Thus, age has mainly a negative effect on the gross benefits of adoption, not because older people are not aware of the opportunities offered by the Internet, but because their abilities to appropriate the benefits of the Internet are too low (partly because they are not as good as others to use a computer).

Finally, expectations regarding the potential benefits of using the Internet are negatively affected by the degree of urbanization, reinforcing the thesis that rural populations are more resistant to new technology which could affect their daily lives.

## 5 Policy implications

The fact that the number of Internet applications that policymakers consider as socially beneficial - particularly in relation to education, health and government services - is increasing and that people on the wrong side of the digital divide are likely to engage in many of these online activities if given access (see, Goldfarb and Prince, 2008), are strong arguments in favor of a public intervention to facilitate Internet adoption.<sup>44</sup> If a public intervention is socially desirable (i.e., its social benefits surpass its social costs), a central question remains regarding the form of this intervention.

We have shown previously that prior to having an affordability issue, a large part of the excluded population has a motivational problem and also that the cross-sectional variance in Internet adoption is mainly due to differences in gross benefits of adoption. This is why, if desirable, a policy

---

<sup>44</sup>For example, Goldfarb and Prince (2008) predict that 46% and 58% of low-income non adopter would use (if given access) the Internet for e-government and health information, respectively.

that aims to reduce the digital divide should, first, try to increase the (perceived) gross benefits of adoption, before taking into consideration differences in adoption costs.

Notice that a public intervention considering only the affordability issue (i.e., an intervention that aims only at reducing the adoption costs through Internet access subsidies or technical assistance to install the Internet) will have two main drawbacks. First, only excluded households that desire to adopt the Internet will eventually be encouraged to do so. Therefore, such intervention will have an incentive effect only on a limited part of the digitally excluded population: in our sample, the households without an Internet connection at home and desiring to adopt the Internet represent less than 27% of the total excluded population. Second, reducing the adoption costs could have a perverse effect by increasing, in particular, the "generational digital divide." Indeed, contrary to the young households who encounter mainly an affordability issue, the non-adoption choice of the elderly households is mainly due to a lack of gross benefits. Thus, reducing the adoption costs will more effectively encourage the young households than the elderly ones to adopt the Internet.

There are several reasons that might explain why a large part of the population has too low gross benefits of adoption. As shown in Subsection 4.3, many people on the wrong side of the divide might not perceive the usefulness of having an Internet connection at home (particularly, the destitute and less educated ones). Therefore, it is necessary to inform non-users about the opportunities and advantages offered by the Internet. This need seems to be taken into consideration by the authorities. For example, in France, the "Numérique 2012" report (see, France Numérique 2012, 2008) suggests to emphasize the opportunities offered by the Internet through a national TV advertising campaign (see, proposition "Action 23"). However, even fully informed regarding the potential benefits of having an Internet connection, the lack of computer skills to use effectively the Internet may prevent the appropriation of those benefits. This is why, public interventions should also aim to provide universal digital literacy, so as to allow individuals to autonomously and effectively use the Internet and so fully benefit from its advantages. For example, the creation of global offers for the elderly people including the Internet connection and an educational tutoring is likely to have a positive effect on reducing the "generational digital divide."

## 6 Conclusion

The aim of this paper was to analyze whether the cross-sectional variance in Internet adoption is mainly due to differences in adoption costs or in gross benefits of adoption. To this end, we have

proposed a model where the digitally excluded population can be separated into two groups: on the one hand, the households which do not desire to adopt the Internet, on the other hand, the households which do desire to adopt the Internet but do not do so because their adoption costs are higher than their gross benefits of adoption. Our results show that a large part of the population is not interested in adopting the Internet and that the cross-sectional variance in Internet adoption is mainly due to differences in gross benefits.

Nowadays, in the developed countries, universal broadband accessibility at a basic speed turns out to be a secondary issue. Indeed, broadband is accessible through various electronic devices (e.g., computers, mobile phones), in various places (e.g., at home, in public places), for attractive price and almost universally.<sup>45</sup> Therefore, before wondering how to guarantee broadband access to every resident for a reasonable price within the "broadband for all" debate,<sup>46</sup> policy makers should formulate effective action plans to increase the gross benefits of the excluded ones.

---

<sup>45</sup>Indeed, broadband coverage is higher than 95% in several OECD countries (see, OECD website: [www.oecd.org/statistics](http://www.oecd.org/statistics)).

<sup>46</sup>Many countries have included broadband as part of their definition of universal service/access. Some of them plan to go (or already did) one step further by recognizing broadband as a universal legal right (for example, in Finland, access to 1 Mbit/s connection has become a legal right for all its citizens in July 2010).



## References

- [1] Chaudhuri, A., & Flamm, K., 2007. An Analysis of the Determinants of Broadband Access. *Telecommunications Policy*, Vol. 31 (6-7), pp. 312-326.
- [2] Chaudhuri, A., Flamm, K., & Horrigan, J., 2005. An Analysis of the Determinants of Internet Access. *Telecommunications Policy*, Vol. 29 (9-10), pp. 731-755.
- [3] Fox, S., 2005. Digital Divisions. Working Paper, Pew Internet and American Life Project.
- [4] France Numérique 2012, Plan de Développement de l'Economie Numérique, 2008. Electronic copy available at [http://francenumerique2012.fr/pdf/081020\\_FRANCE\\_NUMERIQUE\\_2012.pdf](http://francenumerique2012.fr/pdf/081020_FRANCE_NUMERIQUE_2012.pdf).
- [5] Forman, C., Goldfarb, A., & Greenstein, S., 2005. How Did Location Affect Adoption of the Commercial Internet: Global Village vs. Urban Leadership. *Journal of Urban Economics*, Vol. 58 (3), pp. 389-420.
- [6] Goldfarb, A., 2006. The (Teaching) Role of Universities in the Diffusion of the Internet. *International Journal of Industrial Organisation*, Vol. 24 (2), pp. 203-225.
- [7] Goldfarb, A., & Prince, J., 2008. Internet Adoption and Usage Patterns are Different: Implications for the Digital Divide. *Information Economics and Policy*, Vol. 20 (1), pp. 2-15.
- [8] Goolsbee, A., & Klenow, P. J., 2002. Evidence on Learning and Network Externalities in the Diffusion of Home Computers. *The Journal of Law and Economics*, Vol. 45 (2), pp. 317-343.
- [9] Guillemin, O., & Roux, V., Juillet 2001. Comment se Détermine le Niveau de vie d'une Personne?. Division revenu et patrimoine des Menage, Insee. N°798.
- [10] Krueger, A. B., 2000. The Digital Divide in Educating African-American Students and Workers, Working Papers 434. Industrial Relations Section, Princeton University.
- [11] Lenhart, A., Horrigan, J., Rainie, L., Allen, K., Boyce, A., Madden, M., & O'Grady, E., 2003. The Ever-Shifting Internet Population: A new Look at Internet Access and the Digital Divide. Washington, DC: Pew Internet and American Life Project. Electronic Copy available at [http://pewinternet.org/pdfs/PIP\\_Shifting\\_Net\\_Pop\\_Report.pdf](http://pewinternet.org/pdfs/PIP_Shifting_Net_Pop_Report.pdf).

- [12] Prieger, J. E., & Hu, W-M. , 2008. The Broadband Digital Divide and the Nexus of Race, Competition, and Quality. *Information Economics and Policy*, Vol. 20 (2), pp. 150-167.
- [13] Prieger, J. E., & Hu, W-M. , 2009. The Empirics of the Digital Divide: Can Duration Analysis Help?". *Handbook of Research on Overcoming Digital Divides: Constructing an Equitable and Competitive Information Society*, Enrico Ferro, et al. (eds), Hershey, PA: IGI Global.
- [14] Putnam R., 1993. *Making Democracy Work - civic traditions in modern Italy*. Princeton University Press.
- [15] Suire, R., 2007. Encastrement Social et Usages de l'Internet : une Analyse Jointe du Commerce et de l'Administration Electronique. *Economie et Prévision*, n°180-181.
- [16] U.S Department of Commerce, 1999. *Falling Through the Net : Defining the Digital Divide*. Electronic copy available at: <http://www.ntia.doc.gov/Ntiahome/Fttn99/>.
- [17] U.S Department of Commerce, 1995. *Falling Through the Net: Survey of the "Have-Nots" in Rural and Urban America*. Electronic copy available at: <http://www.ntia.doc.gov/ntiahome/fallingthru.html>.
- [18] Van Dijk, Jan A. G. M., 2005. *From Digital Divide to Social Opportunities*. Paper written for the 2nd International Conference for Bridging the Digital Divide Seoul, December. Electronic copy available at: <http://doc.utwente.nl/59814/>.
- [19] Zhang, J., Kuwano, M., Lee, B., & Fujiwara, A., 2009. Modeling Household Discrete Choice Behavior Incorporating Heterogeneous Group Decision-Making Mechanisms. *Transportation Research Part B: Methodological*, Vol. 43 (2), pp. 230-250.

**Appendix I** Household desires to have an Internet connection at home if it has one or if it does not have one but at least one member over 15 years old has answer "yes" to the question "Do you wish to have an Internet connection at home?".

Table 6 shows the sharing out of the households between the three possible categories:  $D_i = 0$ ,  $D_i = 1$  and  $A_i = 0$ ,  $D_i = 1$  and  $A_i = 1$ .

Table 6: Descriptive statistics of the dependent variables

Desire to have an Internet connection at home	Internet connection at home	
	No	Yes
No	1081 (0.376)	0 (0.000)
Yes	495 (0.172)	1296 (0.451)

Percentage in parentheses. Number of observations= 2872.

Table 7 presents the marginal effects for the probabilities that the household desires to adopt the Internet, that the household has adopted the Internet given that it has a desire to do so and that the household has adopted the Internet, respectively.

Table 7: Estimation results (all the household members which are over 15 years old are taken into consideration to define the household's desire to adopt the Internet)

	(1)		(2)		(3)	
	The marginal effects for p(D=1/X)		The marginal effects for p(A=1/D=1. X)		The marginal effects for p(A=1/X)	
<b>HOUSEHOLD</b>						
NUMBER OF PERSONS	0.032	( 0.020 )	0.022	( 0.021 )	0.037 *	( 0.019 )
<b>CHILD</b>						
no child	-0.282 ***	( 0.034 )	0.024	( 0.054 )	-0.169 ***	( 0.050 )
from 0 to 9 years	-0.297 ***	( 0.039 )	0.020	( 0.048 )	-0.181 ***	( 0.046 )
from 10 to 15 years	-0.262 ***	( 0.042 )	0.070	( 0.048 )	-0.126 ***	( 0.048 )
16 years and more	ref.		ref.		ref.	
<b>INCOME (adjusted )</b>						
less than 899 euros	-0.217 ***	( 0.039 )	-0.190 ***	( 0.049 )	-0.269 ***	( 0.039 )
from 900 to 1149 euros	-0.136 ***	( 0.036 )	-0.169 ***	( 0.042 )	-0.209 ***	( 0.037 )
from 1150 to 1499 euros	-0.113 ***	( 0.034 )	-0.084 **	( 0.038 )	-0.140 ***	( 0.035 )
from 1500 to 1999 euros	-0.064 **	( 0.031 )	-0.061 *	( 0.034 )	-0.091 ***	( 0.033 )
more than 2000 euros	ref.		ref.		ref.	
<b>MUNICIPALITY</b>						
low average income	-0.046 *	( 0.024 )	-0.076 ***	( 0.029 )	-0.081 ***	( 0.024 )
<b>URBANIZATION</b>						
Low	-0.088 ***	( 0.029 )	-0.028	( 0.035 )	-0.078 ***	( 0.030 )
Medium	-0.056 **	( 0.024 )	-0.023	( 0.028 )	-0.053 **	( 0.025 )
High	ref.		ref.		ref.	
MOBILE	-0.008	( 0.022 )	0.069 **	( 0.027 )	0.041 *	( 0.023 )
<b>INDIVIDUAL</b>						
<b>EDUCATION</b>						
no diploma	-0.311 ***	( 0.043 )	-0.224 ***	( 0.055 )	0.038 *	( 0.021 )
school certificate	-0.185 ***	( 0.038 )	-0.203 ***	( 0.043 )	-0.362 ***	( 0.043 )
high school prof. degree	-0.162 ***	( 0.035 )	-0.184 ***	( 0.036 )	-0.276 ***	( 0.040 )
high school acad. degree	-0.096 **	( 0.041 )	-0.130 ***	( 0.040 )	-0.249 ***	( 0.036 )
2 years university degree	-0.055	( 0.043 )	0.014	( 0.037 )	-0.170 ***	( 0.041 )
4 years university degree	ref.		ref.		ref.	
<b>AGE</b>						
71 years and more	-0.498 ***	( 0.044 )	0.227 ***	( 0.072 )	-0.169 ***	( 0.048 )
from 61 to 70 years	-0.318 ***	( 0.041 )	0.125 **	( 0.060 )	-0.093 **	( 0.047 )
from 51 to 60 years	-0.244 ***	( 0.037 )	0.246 ***	( 0.048 )	0.028	( 0.044 )
from 41 to 50 years	-0.123 ***	( 0.034 )	0.232 ***	( 0.046 )	0.109 **	( 0.043 )
from 31 to 40 years	-0.079 **	( 0.031 )	0.125 ***	( 0.042 )	0.057	( 0.039 )
30 years and less	ref.		ref.		ref.	
COMPUTER	0.267 ***	( 0.024 )	0.180 ***	( 0.031 )	0.293 ***	( 0.025 )
<b>SOCIABILITY</b>						
Association	0.056 **	( 0.023 )	0.055 **	( 0.026 )	0.075 ***	( 0.024 )
Friends	-0.037 *	( 0.022 )	0.005	( 0.008 )	-0.021 *	( 0.013 )
<b>OPINION</b>						
Oa	ref.		ref.		ref.	
Ob	-0.117 ***	( 0.022 )	0.016	( 0.021 )	-0.067 ***	( 0.021 )
Oc	-0.207 ***	( 0.035 )	0.028	( 0.038 )	-0.122 ***	( 0.033 )
athrho =0.246 (0.323)						
rho (residual correlation) =0.242 (0.304)						
Wald test of rho = 0: chi2(1) = 0.58 Prob > chi2 = 0.4453						

The results are marginal effects.

Log Likelihood of the basic model=-2117.552.

Robust standard errors in parentheses. Number of observations= 2872.

\*  $p < 0.10$ . \*\*  $p < 0.05$ . \*\*\*  $p < 0.01$ .

**Appendix II** Descriptive statistics of the variables *computerskills* and *expectation* are displayed in Table 8 and Table 9, respectively.

Table 8: Descriptive statistics of the variable "computerskills"

Number of computer tasks that she/he can accomplish	Desire to have an Internet connection at home	
	No	Yes
0	23 (0.015)	69 (0.044)
1	51 (0.033)	127 (0.082)
2	42 (0.027)	162 (0.105)
3	39 (0.025)	190 (0.123)
4	38 (0.025)	234 (0.151)
5	44 (0.028)	387 (0.249)
6	15 (0.01)	132 (0.085)

Percentage in parentheses. Number of observations= 1553.

Table 9: Descriptive statistics of the variable "expectation"

"People who do not use the Internet are disadvantaged"	Desire to have an Internet connection at home	
	No	Yes
Totally agree	79 (0.028)	334 (0.120)
Rather agree	219 (0.078)	569 (0.205)
Disagree	266 (0.096)	436 (0.157)
Totally Disagree	512 (0.184)	363 (0.131)

Percentage in parentheses. Number of observations= 2778.