Digital Divide in Internet Usage: Evidence from Five EU Countries[1]

Bertin Martens
Information Society Unit
DG JRC-IPTS European Commission
Edificio Expo, C/Inca Garcilaso, 3
41092 Seville, Spain
Tel. +34 954 48 366
Fax. +34 954 48 802
Email: bertin.martens@ec.europe.eu

Smaranda Pantea
Information Society Unit
DG JRC-IPTS European Commission
Edificio Expo, C/Inca Garcilaso, 3
41092 Seville, Spain
Tel. +34 954 48 7185
Fax. +34 954 48 802
Email: Smaranda.Pantea@ec.europa.eu

February 2013
Draft – work in progress
Please do not quote or circulate

Abstract

This paper examines the digital divide in internet usage in general and for specific purposes (leisure, improving human capital and obtaining goods and services). It uses a unique dataset which covers the entire clickstream of almost 20000 internet users in five largest EU countries during 2011. Our main finding is that, conditional on internet adoption, there is no digital divide in internet usage between internet users with different levels of income, but there is some evidence that might a divide in the ability use websites related to career, education and health or websites related to buying and obtaining goods and services. We find that low income internet users spend more time online overall and on websites related to leisure, and that there is no digital gap in the usage of other types of websites. Tertiary education has no effect on the overall time online, a negative effect on time spent on leisure websites and a positive effect on time spent on websites related to human capital and websites related to obtaining services and goods, suggesting that is a divide in the ability to use these websites. The effect of income on internet usage is higher for intensive users of internet, than for lighter users. The effect of education on usage of websites related to human capital and obtaining goods and services is higher for intensive users of these websites, than for lighter users.

JEL codes: *L86, D12, D13*

Keywords: *Internet Usage, Time allocation, Leisure.*

---

[1] Disclaimer: The views expressed are those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission. Email addresses: bertin.martens@ec.europe.eu (Bertin Martens), smaranda.pantea@ec.europa.eu (Smaranda Pantea).

1. Introduction

The digital divide in access and usage of internet between individuals with different socio-economic characteristics, especially income and education, has been an important concern of economic policy and literature. The digital gap with regard to access to internet has been extensively documented in US (Golfarb and Prince, 2008) and in EU (Orviska and Hudson, 2009). However, less is known about the digital gap in internet usage, especially in Europe. This paper aims to contribute to this area by studying several aspects of digital divide in internet usage in five largest countries in the EU. More precisely, is studies how, conditional on internet adoption, income and education affect internet usage in general and for specific purposes (leisure, improving human capital and obtaining goods and services) for a cross section of almost 20000 internet users in five largest EU countries. It uses a unique dataset that covers the entire clickstream of these internet users during the year 2011.

This paper builds on Goldfarb and Prince (2008) who study the role of income and education levels in internet usage patterns in the US. We extend this study in several ways. First, we study the determinants of time spent online in five largest EU countries using objective data on their online behaviour (their entire clickstream), while previous studies focused on US and used survey data. Second, we study the determinants of time spent on specific types of online activities. We distinguish three types: (a) human capital investment activities that affect future income, such as work, education and health related sites, (b) activities related to the nonmarket work (activities related to obtaining goods and services) and (c) leisure activities. Third, we study whether the relationship between income and education and time spent online differs between employed and not employed internet users (who differ in their access to internet and their opportunity costs of time), and between light and intensitve internet users.

Our main finding is that, conditional on internet adoption, there is no digital divide in internet usage between internet users with different level of income and education. Moreover, we find that low income internet users spend more time online overall and on websites related to leisure. The relationship between income and time spent on websites related to human capital improvement and websites related to buying and obtaining goods and services is more ambiguous but it never indicates a digital divide in usage of these websites based on income. However, we find that education has a positive effect on time spent on websites related to human capital and websites related to obtaining services and goods, suggesting that ability affects the usage of these websites and there might be digital divide in the usage of these types of websites. We find that the effect of income on internet usage is not affected by the employment status and that it higher for intensive users of internet than for light users. The effect of education varies on time spent on websites related to human capital and nonmarket work is higher for intensive users of these websites than for light users.

The paper is organised as it follows. Section 2 reviews related literature. Section 3 describes the data used and presents some preliminary evidence on the relationship between time spent online and income and other demographic characteristics. Section 4 describes the empirical methodology. Section 5 discusses the results of the estimation and Section 6 concludes.

2. Related Literature

Internet usage has been studied from several angles and a review of all the literature on this topic is beyond the purpose of this paper. The paper is related to three strands of literature: studies on the welfare effects and value of internet usage, studies related to usage of internet for a specific purpose and sociological studies on the effect internet use on other activities.

Our study is mostly related to studies that examine internet usage measured as time spent online and its welfare effects (Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Brynjolfsson and Oh, 2012). In these studies, line with Becker (1965), consumer utility depends on his consumption, which requires income obtained partly through labour, and his leisure. Individuals choose the time spent online leisure, offline leisure and work in order to maximise their utility subject to budget and time constraints (the sum of the amount of time spent on each of these activities cannot be higher than 24 hours). Internet pricing consists of fixed cost of adoption and zero usage fees. Then, conditional on internet adoption, the marginal cost of internet use is given only by the opportunity cost of time, which is given by the income that the internet user could earn on the labour market. Therefore, the opportunity cost of spending time online is higher for high income earners than for low income ones. If both low and high income users benefit equally from internet, than given the higher opportunity costs for high income users, they will spend less time on internet. Goolsbee and Klenow (2006), Goldfarb and Prince (2008) and Brynjolfsson and Oh (2012) test empirically variants of these model on a sample of US internet users. All these studies find that income has a negative effect on time spent online, which indicates that conditional on internet access, there is no digital divide in internet usage in US. Most of these studies suggest that higher opportunity costs of spending time online of the high income internet users explain this negative relationship.

There are several studies that examine the use of the internet for specific purposes such as: e-commerce, job search, entertainment etc. (Goldfarb and Prince, 2008; Orviska and Hudson, 2009; Pérez-Hernández and Sánchez-Mangas, 2011). Due to differences in data sources and in the dependent and explanatory variables they are not directly comparable. However, they show some common patterns. The most important and most relevant for our study is that income, education, and other demographic characteristics have different effects on participation in different online activities. Goldfarb and Prince (2008) find that for US internet users, income and university/college education is negatively associated with using internet for activities related to leisure (chat, online games) and e-health, but positively associated with using it for activities related to buying (research purchases and ecommerce). Pérez-Hernández and Sánchez-Mangas (2011) found similar effects of education on online shopping for Spain. Although we draw on these studies, we differ from them in that we our study does not examine the determinants of using internet for a specific purpose, but of the time spent online on each online activity considered.

Our paper is also related to sociology studies that examine which activities are displaced by time spent on internet. The findings of this literature are mixed and overall suggest that spending time on internet is not consistently associated with notable changes in media use or social activities and other daily activities (for a review see Martin and Robinson (2010)). A more nuanced view is proposed by Nie and Hillygus (2002). They suggest that only intensive internet use (more than 60 minutes per day) has a large effect on other activities, especially on leisure and to a lesser extent on work, childcare, housework and sleeping, while light internet use has a small and often insignificant effect. These findings suggest that the opportunity costs of time spent online is higher for intensive internet users, than for light users. We will test this hypothesis in the empirical part of the paper.

In conclusion, there is a large and very heterogeneous literature related to the topic of internet usage, including a few studies on the relationship between income and/or education and internet usage. However, most of these empirical studies are based on US survey data and most of them do not take into account several aspects of this relationship documented in other strands of literature (different types of online activities, intensity of use, and differences across different demographic groups). In this paper, we study different aspects of the effect of income and education on internet usage using objective clickstream data from five largest EU countries.

3. Data Description

The data used in this paper have been collected by Nielsen NetRatings through voluntary online consumer panels. The dataset contains information on all web pages clicked on from their home computers by 25,000 internet users in the five largest EU economies (France, Germany, Italy, Spain and United Kingdom) during the entire year 2011. According to Nielsen, the sample of internet users is representative of the online population in these countries in terms of gender and age[2]. For each click it contains information on the URL, the time and date the website is accessed and time spent on the website. The data on the online activity is collected through a piece of software that internet users in the online panel voluntarily install on their PC. The data collection procedure uses information in the computer about the webpage that is in focus (the page to which the keyboard and mouse activity is directed to). This helps correct for errors in measurement of the time spent on websites due to minimising tabs, tabbed browsing and periods of inactivity. For most websites the dataset contains their brand names, which are classified into subcategories and categories based on the content of the websites using a methodology developed by Nielsen. For each user[3] the dataset contains information on basic social and economic characteristics, gathered. through a questionnaire when the user installs the Nielsen software.

[2] Nielsen provides incentives to participate and to remain in the panel in the form of vouchers and points which can be redeemed from their reward website or used in online games and sweepstakes(prize drawing), which might bias our sample towards people who are more likely to value these activities. In the empirical part we will discuss potential implications for our estimated effects and as a robustness check we will repeat the estimations excluding time spent on online games and gambling websites to make sure that our results are not driven by time spent on these websites.

[3] There are households in which more users are registered with Nielsen. In these households the meter prompts the internet user to log in; however the match between user profile and his online activity is likely to be imperfect. To ensure that our results are affected by this problem, we will estimate our model also on the sample of single households.

The sample that we use in the empirical analysis excludes records with missing information on the website category and on the demographic characteristics of the internet user, and records of unlikely young and old[4] internet users and outliers[5] and records of self-employed internet users[6]. Table 1 shows how excluding these observations affects the sample used for the empirical analysis. The remaining samples are still large covering close to 4,000 users in each country, which represent close to 80% of the initial user sample and more than 70% of the initial clickstream sample.

We examine how much time users spend online and what type of activities they carry out (leisure, investment in human capital and non-market work buying/obtaining goods or services). Table 2 present the classification of websites into these groups of online activities, which is based on how each activity contributes to consumer utility.

Table 3 presents the average time spent on each of the three groups of online activities. This table shows that the average person in each of the five observed countries allocates time on these three groups of activities in a similar way. Average weekly time spent online varies between 3.5 (France) and 6 (UK) hours. The average person spends most time on leisure websites - between 2 and 3 hours per week, more than one hour per week on nonmarket work online and between 8 and 10 minutes per week on websites related to work, education and health. The most popular leisure websites are, in order, social networks, online games, videos/movies and adult websites and the most popular nonmarket work websites are general portals, search and e-commerce websites.

[4] The dataset includes internet users between 2 and 99 years old. It is likely that very young and very old consumers did not answer the questionnaire themselves and that they did not use the internet themselves. To ensure comparability we will focus on internet users aged between 16 and 74 in line with Eurostat for Information Society Indicators and previous empirical studies (Orviska and Hudson, 2009, Perez-Hernandez and Sanchez-Mangas, 2011).

[5] We exclude internet users who spend an implausible large or small amount of time online (internet users in the highest and the lowest 1% of average weekly time spent online). The main results are not affected by this exclusion.

[6] In the case of self employed we do not know what part of their time online is related to their work and which part is leisure. They represent 9% of our sample, but the results are robust to including these observations in the sample.

We examine the social and economic characteristics of the internet users in the sample. The definition of all variables used and their summary statistics for the internet users in the sample aged 16-74 - excluding outliers and those with missing information - are given in Table 4. The summary statistics of the demographic characteristics of the internet users show that the sample used in the empirical analysis includes a large variety of internet users in terms of education, occupation, income and other demographic characteristics.

Figure 1 presents the distribution of time spent online (minutes per week) in the pooled 5-country sample on all types of websites and on the leisure, human capital related website and obtaining goods and services websites. It shows the percentage of internet users in the sample on the vertical axis that spent a specific number of minutes (a multiple of 60 minutes for all activities except human capital for which we use a multiple of 10 minutes) online per week on the horizontal axis. The figure reveals large heterogeneity in the intensity of using internet. Many internet users use spent little time online and at there is long tail that spends many hours online.

In Figure 2, we present some patterns that show how time spent online is linked with income. It shows that total time spent online and time spent on all online activities considered decreases with income. This relationship is strongest for all time spent online and for time spent on leisure websites, for the other two types of websites it is weaker. These patterns show clearly that there is no digital divide based on income in internet usage, on the contrary lower income people spend more time online types of websites. These patterns are consistent with the hypothesis that high income users have a higher opportunity cost of time and therefore spend less time on these online activities.

Figure 3 presents how average time spent online varies with educational attainment. Internet users with tertiary education spent less time online than users with lower educational attainment. This pattern might indicate higher opportunity cost of time for internet users with tertiary education. There is a clear positive relationship between human capital websites and times spent online, which indicates that there might be a digital divide in the ability to use these websites between people with different education levels. Finally, there is not relationship between time spent online on nonmarket work websites and education.

In summary, our descriptive analysis shows that there is large heterogeneity regarding the intensity of internet usage and that there is a negative relationship between income and time spent online and mixed relationship between education and time spent online.

4. Methodology

Following Goldfarb and Prince (2008), we assume time spent online is a function of total leisure time, total income, price of internet and other individual characteristics. We include controls for occupational and demographic characteristics related to life stage (being married/cohabitating and having children) to control for leisure time. Household income is our proxy for total money available. We include country and region dummies to control for the fixed internet connection price. We also include several demographic characteristics which previous studies have shown to have an effect on time spent online. We estimate the following regression:

$$y_i = \alpha + \beta_I' Income_i + \beta_x' x_i + \beta_o' o_i + \beta_c' c_i + \beta_{cr}' r_{ci} + u_i \qquad (1)$$

y_i is the average time spent online per week by internet user i, measured in minutes per week. Since we do not have continuous income variable but only income groups, $Income_i$ is measured as a set of dummies for household income in a given interval. o_i are dummy variable for different occupations, c_i and r_{ci} are dummy variables that control for country and regions within each country where the internet user resides. x_i are other social and economic characteristics of the internet user. One such characteristic is education. Its effect in this regression shows the effect of ability of using internet in different activities. In addition we include the following control variables: gender, age, being single, children in the household.

We test several hypotheses. First, we test the opportunity cost of time hypothesis[7], which predicts, the opportunity cost of spending time online is higher for high income earners and therefore they will spend less time online overall and on different types of internet use. If this hypothesis were true the coefficients on the income dummies ($Income_i$) should be negative and significant. Such results would also suggest that there is no digital divide in internet usage at least not based on income. We will test a similar hypothesis regarding education. Highly educated internet users also experience higher opportunity cost of time because they earn higher wages. However, the effect of education on time spent on different online activities measure also the ability of using internet for different purposes. If an online activity requires certain skills/abilities we would expect a positive coefficient of education. We would expect this to be true for more sophisticated internet usages, especially human capital related and nonmarket work related, which includes ecommerce and use different services online such as online banking or government websites.

We test whether the effects of income and education on time spent online vary for intensive and light internet users. Testing this hypothesis is relevant given the positively skewed distribution of time spent online (Figure 1). Theory also suggests that these effects might differ. Nie and Hillygus (2002) found that spending a considerable part of time on the internet

[7] We assume fixed monthly internet usage fees and consequently a zero marginal financial cost of internet usage.

crowds out other daily activities, while light use of internet does not. If hypothesis were true, then we would observe a larger negative effect of income on time spent online for more intensive internet users than for less intensive ones. Ability (education) might also have a different effect on light or intensive internet users. It could be that light users only use basic digital content which does not require much ability, while intensive users use more diverse and sophisticated features of the websites and digital content, which requires higher ability or digital skills.

Equation (1) is estimated using OLS. In our sample all individuals have access to internet at home and were active users of internet during the period studied. Therefore, we cannot examine the determinants of selection into internet usage or control for it. Moreover, the descriptive statistics presented in the previous section show that all users spent positive amounts of time on consumption and leisure websites and 98.6% do so on human capital websites. Given that our dependent variable is not censored, or in the case of human capital websites would be very little affected by censoring, we conclude that OLS is the appropriate estimation method[8]. In addition, we will carry out several robustness checks which will be discussed in the results section.

To test whether the effects of income and education vary with the intensity of internet use we will use quantile regressions. This method provides a more complete characterisation of the conditional distribution of time spent online by allowing the effect of income and education and other explanatory variable that to vary and it is more appropriate given the positively skewed distribution of our dependent variables (see Figure 1). It is not sensitive to outliers and it is more robust and efficient than OLS when the error is non normal (Buchinsky, 1998).

[8] We have estimated the equation (1) using tobit and the results are very similar to the ones obtained using OLS.

5. Estimation Results

Table 5 reports the estimation of equation (1). The results in the first column of Table 5 confirm that all income coefficients are negative and statistically significant. The household income group 0-18.000 Euros is taken as the reference group. *Ceteris paribus*, internet users with a household income between 18.000 and 27.000 Euros spend on average 52 minutes less time online than internet users with a household income of less than 18.000. The results suggest that time spent online decreases almost monotonically with the household income. The differences between the coefficients of income intervals are statistically significant (See bottom part of Table 5). Thus, the results suggest a negative relationship between income and time spent online. These results suggest that there is not digital divide in terms on interne usage and they are consistent with the opportunity cost hypothesis and with previous studies (Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Brynjolfsson and Oh, 2012).

Columns 3, 4 and 5 in Table 5 show that there is a negative relationship between income and time spent on each type of website considered. These results confirm Goldfarb and Prince (2008) results that lowest income internet users spend more time on websites related to leisure and websites related to health. For leisure, the differences between coefficients of income intervals are statistically significant, suggesting a negative and monotonic relationship between time spent on leisure online and household income. This is consistent with the findings for the overall time spent online, with the opportunity cost hypothesis, with previous empirical studies on leisure online (Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Brynjolfsson and Oh, 2012). It is also similar to results on TV watching (Frey *et al.*, 2007). The differences in the coefficients of income intervals on time spent on human capital and nonmarket work online (reported in the bottom part of Table 5) are not statistically significant. They suggest that for internet users above lowest income interval there is no relationship between time spent online and income.

Secondary education has a positive effect on time spent on all types of websites, although it is insignificant for leisure websites. Tertiary education has an insignificant effect on overall time spent online, but a negative effect on time spent on leisure websites, and an unambiguous positive effect on time spent on human capital and nonmarket work websites. These results are consistent with Goldfarb and Price (2008) result that education has a negative effect on time spent on leisure, but a positive effect on ecommerce and research purchases. The results for education on time spent on human capital and nonmarket work websites suggests there might be a divide in terms of ability to use these types of websites.

We carry out several robustness checks. To address possible problems with the measurement of duration of time spent online, we re-estimate equation (1) replacing time spent with the number of clicks per week as the dependant variable for all categories. To check the robustness of our findings to possible mismatch between online activity and user profile we estimate the baseline models on the sample of single household sample. We also estimate equation (1) excluding time spent on online games and gambling/sweepstakes to check that our results are not driven by time spent on these websites. To confirm that our results are not driven by one country or a group of countries, we estimate country specific results (reported in Table 6). The results of these robustness checks confirm our baseline results. We also test whether the effects of income and education on time spent online varies by employment status of internet users. These estimations reported in provides a test of whether the negative relationship between income and internet is driven by access to internet at work for higher income internet users. The results (reported in Table 7)show that income has a negative effect on overall time use and on time spent on specific websites, both for working and non-working user and the differences in coefficients of two types of internet users are not statistically significant. These results indicate that the negative relationship between times spent online and income are not driven by the fact that high income internet users a more likely to have access to internet at work than low income ones. The findings for employed internet users confirm the opportunity cost hypothesis. The findings for not working users are more difficult to interpret. The opportunity cost of time of internet users who do not earn a working time-related income is not given by income earned in the labour market, but by other possible uses of their time. However, these opportunity costs may be correlated with household income. These results could also be interpreted as lending supports to the hypothesis that low income internet users benefit more from internet than high income internet users who may have better alternatives or different preferences as suggested by Goldfarb and Prince (2008).

We check whether the wide heterogeneity and long-tail distribution of the observations affects the findings regarding income and other explanatory variables. For this purpose we use quantile regressions. The estimation results for the 10th, 25th, 50th, 75th and 90th quantiles, for the four categories of websites, are reported in Table 8 -Table 11. The main result of the quantile regressions estimations is that effect of income has a larger impact at the top than at the bottom of the distribution of time spent online. This is true for overall time spent as well as for time spent on each of the three categories of websites. We also tested and confirmed that the differences in the income coefficient for different quantiles are statistically significant. This result can be interpreted as consistent with the hypothesis that opportunity cost of time spent online increases with the time spent online because heavy use of internet crowds out other activities. These results are in line Nie and Hillygus (2002) results that light use of internet has a small, often insignificant effect, on other daily activities, but heavy use crowds out activities daily activities.

Education has different effect for different online activities. For overall time and leisure, education has a positive effect on the lower quantiles of these distributions but an insignificant and even negative effect on the higher quantiles of these distributions. For time spent on human capital and nonmarket work websites, education has a positive and significant effect on all quantiles of the distribution and its effect is higher for higher quantiles of these distributions. These results could be interpreted as evidence of the need for skills to use these types of websites intensively; leisure sites apparently require less skills or only at very low levels of usage.

6. Conclusions

This paper aims to contribute to the debate regarding digital divide in access and usage of internet between individuals with different socio- economic characteristics, especially income and education. While there is a large literature on the digital divide in access, less is known about the digital divide in usage and the evidence that exist is based on survey data and most of is for US.

In this paper we examine digital divide in internet usage in general and internet usages for specific activities (leisure, human capital improvement and obtaining goods and services) in five largest EU countries. It studies how, conditional on internet adoption, income and education affect internet usage in general and for specific purposes for a cross section of close to 20.000 internet users during 2001. We use a unique dataset which covers their entire clickstream of these internet users during 2011.

We build on Goldfarb and Prince (2008) model who study the role of income and education levels on internet usage patterns in the US, which we extent to study the determinants of the specific online activities mentioned above. In addition, we study whether the relationship between income and education and time spent online differs between employed and not employed internet users (who differ in their access to internet and their opportunity costs of time), and between light and heavy internet users.

Our main finding is that, conditional on internet adoption, there is no digital divide in internet usage between internet users with different level of income and education. On the contrary, we find that low income internet users spend more time online overall and on websites related to leisure. The relationship between income and time spent on websites related to human capital improvement and websites related to buying and obtaining goods and services is more ambiguous but it never indicates a digital divide in usage of these websites based on income. However, we find that tertiary education has a positive effect on time spent on websites related to human capital and websites related to obtaining services and goods, suggesting there might be a divide in the ability of using websites related to career, education and health or websites related to buying and obtaining goods and services.

These results are robust to several robustness checks and the effects of income and education on internet usage are not affected by the employment status, but are affected by the intensity of internet usage. The effect of income is higher for more intensive users of internet than for less intensive ones. The effect of education on time spent on websites related to human capital and nonmarket work is higher for more for more intensive users of lighter users of these websites.

Taken together our results are consistent with the hypothesis that the opportunity costs of time explains the negative relationship between time spent online and income, although there is also some evidence that can interpreted that low income internet users may benefit more from internet than high income ones. These findings show that there is no digital divide in internet usage with regard to income, however there is some evidence that education affect the ability to use websites related to career, education and health or websites related to buying and obtaining goods and services.

References

Aguiar, M., Hurst, E. and Karabarbounis, L. (2012). "Recent Developments in the Economics of Time Use," *Annual Review of Economics*, vol. 4(1), pp. 373-397.

Brynjolfsson, E. and Oh, J. H. (2012)"The Attention Economy: Measuring the Value of Free Goods on the Internet", mimeo.

Becker, G. S. (1965) "A Theory of the Allocation of Time," *Economic Journal,* vol. 75 (299), pp. 493-517.

Buchinsky, M (1998) "Recent Advances in Quantile Regression Models: A Practical Guide for Empirical Research," *Journal of Human Resources*, vol. 33(1), pp.88 -126.

Frey, B. S., Benesch C., and Stutzer, A. (2007) "Does Watching TV Makes Us Happy?", *Journal of Economic Psychology*, vol. 28(3), pp.283-313.

Goldfarb, A. and 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.

Goolsbee, G and Klenow, P. J. (2006) "Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet," *American Economic Review*, vol. 96(2), pp 108-113.

Gronau, R. (1977) "Leisure, Home Production, and Work-The Theory of the Allocation of Time Revisited," *Journal of Political Economy*, vol. 85(6), pp. 1099-1123.

Kuhn, P. J. and Mansour, H. (2011) "Is Internet Job Search Still Ineffective?," IZA Discussion Papers 5955, Institute for the Study of Labor (IZA).

Kuhn, P. And Skuterud, M. (2004) "Internet Job Search and Unemployment Durations," *American Economic Review*, vol. 94(1), pages 218-232.

Nie, N. D. and Hillygus, D. S. (2002) "The Impact of Internet Use on Sociability: Time-Diary Findings," IT & Society, vol. 1(1), pp. 1-20.

OECD (2012) Internet Economy Outlook 2012.

Orviska, M. and Hudson, J. (2009) "Dividing or Uniting Europe? Internet Usage in the EU," *Information Economics and Policy*, vol. 21(4), pp. 279-290.

Pérez-Hernández, J. and Sánchez-Mangas, R. (2011) " To Have or Not to Have Internet at Home: Implications for Online Shopping," *Information Economics and Policy*, vol. 23(3), pages 213-226.

Robinson, S. M and Marin, J. P. (2010) "IT Use and Declining Social Capital? More Cold Water from the General Social Survey (GSS) and the American Time Use Survey (ATUS)," *Social Science Computer Review*, vol. 28(1), pp. 45-63.

Annexes

Tables

Table 1 Summary of available information about clicks.

Country	Total Clicks		Categorised clicks		Categorised clicks with complete demographic data		Categorised clicks of consumers aged 16-74 years, with complete demographic data and not considered outliers	
1	2	3	4	5	6	7	8	9
	users	clicks	share users	share clicks	share users	share clicks	share users	share clicks
FR	5000	147.601.904	100%	89%	96%	85%	83%	74%
DE	5000	246.568.640	100%	87%	93%	81%	86%	73%
IT	5000	211.011.296	100%	91%	90%	81%	82%	73%
ES	5000	222.590.768	100%	90%	91%	83%	83%	75%
UK	5000	199.006.384	100%	90%	84%	78%	76%	70%

Source: Nielsen Click stream

Table 2 Online Activity Definition

Activity	Nielsen Category
All Time Online	All categories
Leisure	Entertainment, Family and Lifestyle (except subcategory Health, Nutrition and Fitness), News & Information, subcategories Member Communities and Targeted Member Communities from Portals & Communities category and Internet Services.
Human Capital	Education& Careers, Corporate[9], subcategory Health, Nutrition and Safety from Family and Lifestyle).
Nonmarket work (including obtaining goods and services)	Home & Fashion, Ecommerce, Travel, Government & Nonprofit, Finance, Search Engines, General Portals & Search (subcategories General Portals and Search from Search Engines, Portals & Communities category), Special Occasions, Automotive, Computers & Electronics.

Table 3 Average time allocated on different online activities (minutes/week)

	All Time Online		Leisure		Human Capital		Nonmarket Work	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
France	220.28	[244.60]	142.87	[194.53]	8.88	[19.77]	68.53	[75.32]
Germany	345.23	[355.74]	239.96	[306.40]	7.52	[15.04]	97.75	[110.07]
Italy	311.48	[309.01]	224.92	[261.54]	7.71	[15.13]	78.85	[89.09]
Spain	317.97	[311.86]	231.90	[262.14]	9.78	[18.83]	76.29	[85.54]
UK	356.95	[351.31]	242.25	[295.09]	10.19	[22.40]	104.51	[108.86]

Source: Calculations based on Nielsen Click stream

[9] We assume that people searched this category mainly for finding information about job vacancies. However, classifying it as a residual category or as nonmarket work does not change the results.

Table 4 Variable definitions and summary statistics

Variable	Definition	Obs.	Mean	Std. Dev.	Min	Max
All Time Online	Average time spent online per week (in minutes)	18680	306.02	320.30	1.78	1973.38
Leisure	Average time spent online per week on websites related to entertainment, communication and	18680	214.05	270.18	0.02	1838.78
Human Capital	Average time spent online per week on websites related to career, education and health (in minutes)	18680	8.79	18.53	0.00	556.86
Nonmarket Work	Average time spent online per week on websites related to buying/obtaining/ finding information	18680	83.18	93.62	0.04	1479.06
All Clicks	Average number of views per week	18680	697.15	889.64	1.37	11161.60
Leisure Clicks	Average number of views per week on leisure pages	18680	482.83	758.66	0.08	11028.79
Human Capital Clicks	Average number of views per week on work and HC pages	18680	20.01	51.60	0.00	2630.19
Nonmarket Work Clikcs	Average number of views per week on websites related to buying/obtaining/ finding information	18680	194.31	242.47	0.17	3562.83
Female	A dummy variable that indicates whether the internet user is female.	18680	0.51	0.50	0	1
Single	A dummy variable that indicates whether the internet user is single (not married or cohabitating).	18680	0.25	0.43	0	1
Age	Age of the internet users (in years)	18680	41.68	13.57	16	74
Children	Dummy variable that takes value 1 if there are children present in the household and 0 otherwise.	18680	0.31	0.46	0	1
Household Income ≤18000	A dummy variable that indicates whether the household income of the internet user is in this interval.	18680	0.21	0.41	0	1
Household Income 18001-27000	A dummy variable that indicates whether the household income of the internet user is in this interval.	18680	0.23	0.42	0	1
Household Income 27001-36001	A dummy variable that indicates whether the household income of the internet user is in this interval.	18680	0.17	0.38	0	1
Household Income 36001-54000	A dummy variable that indicates whether the household income of the internet user is in this interval.	18680	0.22	0.42	0	1
Household Income 54001-72000	A dummy variable that indicates whether the household income of the internet user is in this interval.	18680	0.10	0.30	0	1
Household Income >72000	A dummy variable that indicates whether the household income of the internet user is in this interval.	18680	0.06	0.23	0	1
Below Secondary Education	A dummy variable that indicates whether this is the educational attainment of the internet user	18680	0.26	0.44	0	1
Secondary Education	A dummy variable that indicates whether this is the educational attainment of the internet user	18680	0.26	0.44	0	1
Tertiary Education	A dummy variable that indicates whether this is the educational attainment of the internet user	18680	0.48	0.50	0	1
Employed	A dummy variable that indicates whether the internet user is employed.	18680	0.66	0.47	0	1
Unemployed	A dummy variable that indicates whether the internet user is unemployed.	18680	0.09	0.29	0	1
Student	A dummy variable that indicates whether the internet user is a student.	18680	0.09	0.29	0	1
Retired	A dummy variable that indicates whether the internet user is retired.	18680	0.10	0.29	0	1
Homemaker	A dummy variable that indicates whether the internet user is a homemaker or carer.	18680	0.07	0.25	0	1

Source: Calculations based on Nielsen Click stream

Table 5 Baseline model

	All Time Online	Leisure	Human capital	Nonmarket Work
Income 18001-27000	-52.09	-44.20	-1.65	-6.24
	[7.81]***	[6.75]***	[0.42]***	[2.15]***
Income 27001-36001	-67.86	-61.43	-1.68	-4.75
	[8.21]***	[7.04]***	[0.46]***	[2.32]**
Income 36001-54000	-103.02	-91.11	-2.20	-9.71
	[7.86]***	[6.67]***	[0.46]***	[2.31]***
Income 54001-72000	-127.26	-110.95	-2.60	-13.71
	[9.06]***	[7.50]***	[0.60]***	[2.85]***
Income >72000	-153.05	-127.00	-3.90	-22.16
	[10.36]***	[8.52]***	[0.63]***	[3.13]***
Secondary education	18.29	6.76	2.34	9.19
	[7.23]**	[6.18]	[0.37]***	[2.02]***
Tertiary education	-0.75	-15.86	3.40	11.71
	[6.75]	[5.74]***	[0.35]***	[1.94]***
Constant	384.22	324.03	5.57	54.62
	[21.01]***	[17.89]***	[1.16]***	[5.58]***
N	18680	18680	18680	18680
R^2	0.10	0.10	0.03	0.05
F tests of differences in income coefficients (p values)				
$\beta_{Inc.\ 18\text{-}27000} = \beta_{Inc.\ 27\text{-}36000}$	0.03	0.00	0.94	0.47
$\beta_{Inc.\ 27\text{-}36000} = \beta_{Inc.\ 36\text{-}54000}$	0.00	0.00	0.21	0.02
$\beta_{Inc.\ 36\text{-}54000} = \beta_{Inc.\ 54\text{-}72000}$	0.00	0.00	0.46	0.11
$\beta_{Inc.\ 54\text{-}72000} = \beta_{Inc.\ >72000}$	0.00	0.02	0.04	0.01

Notes: Dependent variable is average time spent online total or on specified category of websites, measured in minutes. All equations include occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 6 Country specific results for all time online

	DE	FR	IT	ES	UK
Income 18-27000	-86.36	-42.64	-44.84	-39.51	-54.01
	[18.76]***	[18.13]**	[16.79]***	[14.84]***	[19.58]***
Income 27-36001	-86.05	-69.78	-79.48	-37.84	-77.6
	[19.47]***	[17.80]***	[17.36]***	[16.82]**	[23.41]***
Income 36-54000	-139.9	-100.64	-116.54	-75.77	-99.16
	[18.60]***	[17.29]***	[16.77]***	[16.19]***	[20.75]***
Income 54-72000	-152.54	-119.03	-93.19	-87.78	-165.26
	[21.22]***	[18.38]***	[22.42]***	[21.39]***	[22.54]***
Income >72000	-200.56	-134.81	-136.81	-134.89	-146.12
	[22.93]***	[19.51]***	[25.32]***	[24.20]***	[33.67]***
Secondary ed.	-19.18	17.14	33.99	69.47	35.60
	[16.02]	[14.31]	[16.39]**	[16.37]***	[28.86]
Tertiary ed.	-47.70	0.82	31.43	41.01	15.89
	[14.88]***	[11.24]	[18.14]*	[15.07]***	[27.00]
Constant	416.86	248.71	354.60	356.52	475.16
	[43.28]***	[30.85]***	[33.56]***	[34.18]***	[45.85]***
N	3928	4028	3535	3767	3422
R^2	0.10	0.11	0.08	0.08	0.09

Notes: Dependent variable is average time spent online, measured in minutes. All equations include controls for other demographic characteristics and occupation, and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 7 Baseline model: comparison of working and not working internet users.

	All Time Online All		Leisure Leisure		Human Capital Human Capital		Nonmarket Work Nonmarket Work	
	Empl.	Not Empl.	Empl.	Not Empl.	Empl.	Not Empl.	Empl.	Not Empl.
Income 18001-27000	-52.69	-46.76	-46.74	-38.10	-0.99	-2.37	-4.96	-6.28
	[10.06]***	[12.84]***	[8.73]***	[11.03]***	[0.54]*	[0.64]***	[2.78]*	[3.51]*
Income 27001-36001	-57.69	-85.76	-55.75	-73.95	-0.71	-2.98	-1.23	-8.83
	[10.55]***	[13.71]***	[9.09]***	[11.68]***	[0.58]	[0.70]***	[2.98]	[3.85]**
Income 36001-54000	-101.27	-105.37	-93.72	-88.19	-1.26	-3.30	-6.30	-13.89
	[9.95]***	[13.99]***	[8.48]***	[11.83]***	[0.54]**	[0.83]***	[2.90]**	[4.11]***
Income 54001-72000	-127.41	-123.02	-114.90	-103.82	-1.43	-4.17	-11.07	-15.04
	[11.35]***	[16.60]***	[9.47]***	[13.61]***	[0.74]*	[1.02]***	[3.56]***	[5.09]***
Income >72000	-160.42	-126.97	-138.36	-98.56	-3.01	-4.29	-19.05	-24.12
	[12.26]***	[21.77]***	[10.14]***	[17.99]***	[0.67]***	[1.54]***	[3.86]***	[5.78]***
Secondary education	-0.56	37.34	-8.15	22.91	1.84	2.58	5.75	11.85
	[9.36]	[11.77]***	[7.96]	[10.12]**	[0.49]***	[0.59]***	[2.69]**	[3.16]***
Tertiary education	-20.40	18.39	-29.51	-2.17	2.26	4.62	6.85	15.94
	[8.94]**	[10.79]*	[7.61]***	[9.18]	[0.45]***	[0.58]***	[2.57]***	[3.19]***
Constant	399.05	465.73	331.61	403.05	5.67	8.85	61.78	53.83
	[25.19]***	[35.20]***	[21.11]***	[30.66]***	[1.43]***	[1.79]***	[6.99]***	[8.69]***
Mean DV	284.73	347.18	195.69	249.53	8.25	9.83	80.79	87.82
Std. dev.	[299.00]	[354.30]	[250.54]	[301.43]	[18.42]	[18.70]	[89.68]	[100.65]
N	12311	6369	12311	6369	12311	6369	12311	6369
R^2	0.10	0.13	0.11	0.13	0.07	0.10	0.06	0.15

Notes: Dependent variable is average time spent online total or on specified category of websites, measured in minutes. All equations ncluded occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 8 Quantile regressions: All Time Online

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18001-27000	-8.59	-26.76	-66.66	-85.49	-84.89	-52.09
	[2.28]***	[5.26]***	[9.29]***	[15.98]***	[26.56]***	[7.81]***
Income 27001-36001	-11.46	-36.39	-77.01	-102.69	-140.25	-67.86
	[2.40]***	[5.57]***	[9.15]***	[15.43]***	[27.15]***	[8.21]***
Income 36001-54000	-15.28	-46.11	-110.34	-143.33	-200.02	-103.02
	[2.30]***	[5.45]***	[8.96]***	[15.18]***	[25.95]***	[7.86]***
Income 54001-72000	-16.96	-54.99	-126.32	-178.20	-266.33	-127.26
	[2.52]***	[5.49]***	[10.04]***	[16.14]***	[29.43]***	[9.06]***
Income >72000	-20.22	-66.31	-141.43	-210.39	-289.59	-153.05
	[2.56]***	[5.71]***	[11.08]***	[16.72]***	[33.02]***	[10.36]***
Secondary education	6.13	14.40	23.25	29.86	23.11	18.29
	[1.49]***	[3.33]***	[6.87]***	[12.17]**	[22.84]	[7.23]**
Tertiary education	5.49	17.44	25.98	0.64	-48.78	-0.75
	[1.48]***	[3.17]***	[6.40]***	[11.70]	[23.25]**	[6.75]
Constant	39.52	134.18	418.25	686.20	757.41	384.22
	[27.13]	[55.70]**	[146.12]***	[91.41]***	[85.33]***	[21.01]***
Pseudo R^2	0.02	0.04	0.06	0.07	0.10	0.09

Notes: Dependent variable is average time spent online total or on specified category of websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets.*, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 9 Quantile regressions: Leisure

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18001-27000	-4.63	-17.87	-51.22	-80.36	-100.88	-44.20
	[1.48]***	[3.12]***	[6.79]***	[13.55]***	[24.83]***	[6.75]***
Income 27001-36001	-6.43	-24.16	-65.22	-112.93	-142.19	-61.43
	[1.44]***	[2.99]***	[7.39]***	[13.15]***	[25.97]***	[7.04]***
Income 36001-54000	-7.61	-29.59	-82.56	-152.70	-213.09	-91.11
	[1.41]***	[2.95]***	[7.03]***	[12.88]***	[25.71]***	[6.67]***
Income 54001-72000	-8.73	-34.94	-94.80	-173.08	-263.18	-110.95
	[1.48]***	[3.11]***	[7.51]***	[13.28]***	[26.11]***	[7.50]***
Income >72000	-10.25	-39.68	-102.60	-183.92	-291.26	-127.00
	[1.51]***	[3.17]***	[7.58]***	[13.43]***	[29.83]***	[8.52]***
Secondary education	2.39	5.38	15.56	13.89	-2.65	6.76
	[0.78]***	[1.97]***	[4.70]***	[10.84]	[24.71]	[6.18]
Tertiary education	2.18	6.99	10.02	-11.18	-65.37	-15.86
	[0.68]***	[1.70]***	[4.25]**	[9.29]	[21.21]***	[5.74]***
Constant	20.21	80.31	300.13	496.40	807.60	324.03
	[13.96]	[30.26]***	[83.74]***	[108.73]***	[112.89]***	[17.89]***
Pseudo R^2	0.01	0.03	0.06	0.08	0.11	0.10

Notes: Dependent variable is average time spent online total or on specified category of websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets. *, ** and *** indicate significance at10%, 5% and 1%, respectively

Table 10 Quantile regressions: Human Capital

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18001-27000	-0.08	-0.32	-0.80	-1.72	-4.21	-1.65
	[0.03]**	[0.08]***	[0.19]***	[0.43]***	[1.15]***	[0.42]***
Income 27001-36001	-0.11	-0.39	-0.90	-2.17	-4.06	-1.68
	[0.03]***	[0.08]***	[0.20]***	[0.47]***	[1.39]***	[0.46]***
Income 36001-54000	-0.18	-0.49	-1.15	-2.61	-5.28	-2.20
	[0.03]***	[0.08]***	[0.19]***	[0.44]***	[1.20]***	[0.46]***
Income 54001-72000	-0.18	-0.57	-1.62	-3.82	-8.13	-2.60
	[0.04]***	[0.09]***	[0.19]***	[0.48]***	[1.33]***	[0.60]***
Income >72000	-0.24	-0.78	-2.11	-4.88	-9.45	-3.90
	[0.04]***	[0.09]***	[0.21]***	[0.57]***	[1.36]***	[0.63]***
Secondary education	0.11	0.31	0.89	2.14	4.26	2.34
	[0.02]***	[0.05]***	[0.11]***	[0.30]***	[0.81]***	[0.37]***
Tertiary education	0.16	0.45	1.47	3.89	7.53	3.40
	[0.02]***	[0.05]***	[0.12]***	[0.33]***	[0.83]***	[0.35]***
Constant	0.54	0.40	1.44	1.85	17.75	5.57
	[0.25]**	[0.53]	[1.64]	[5.07]	[7.48]**	[1.16]***
Pseudo R^2	0.01	0.01	0.03	0.04	0.05	0.03

Notes: Dependent variable is average time spent online total or on specified category of websites, measured in minutes. All equationsinclude controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets. *, ** and *** indicate significance at10%, 5% and 1%, respectively

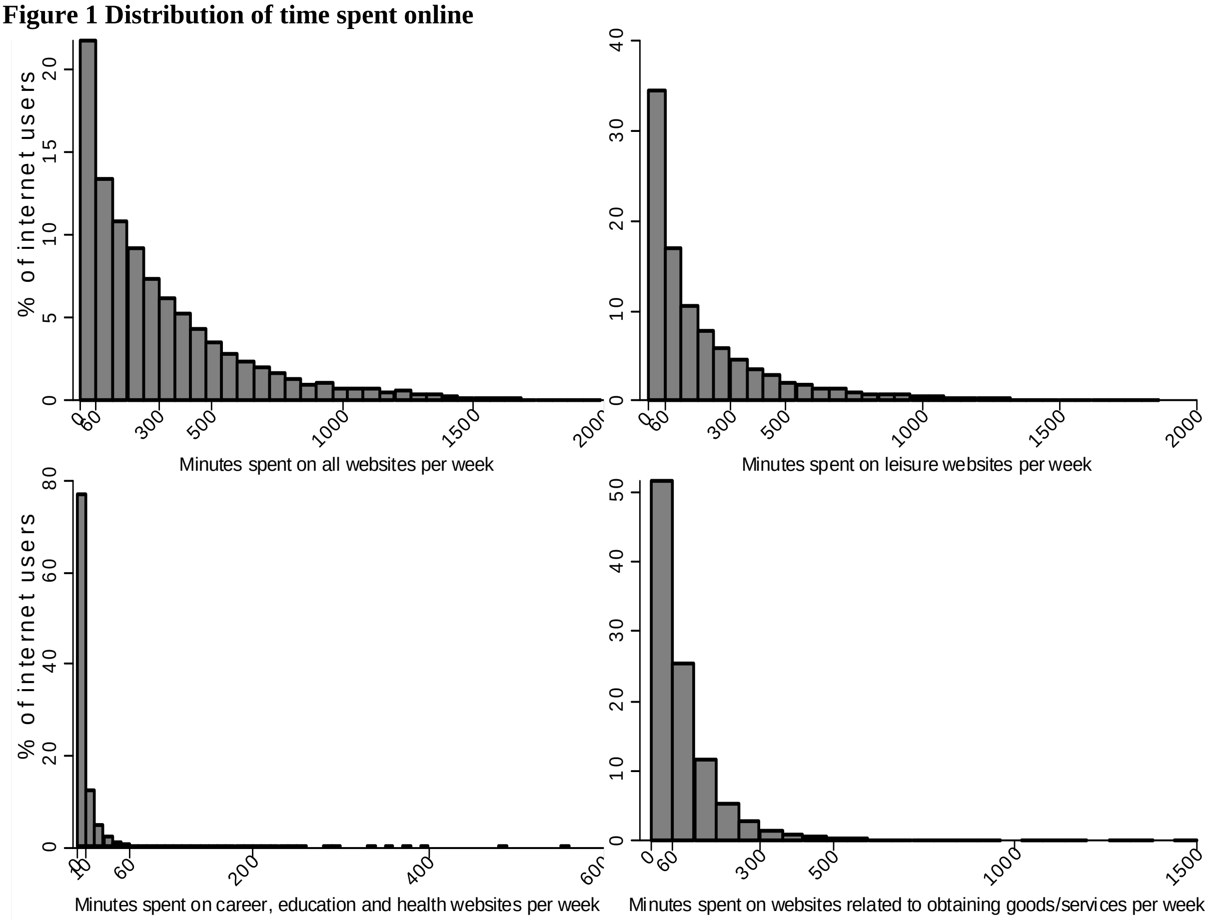
Table 11 Quantile regressions: Nonmarket work

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18001-27000	-0.24	-2.49	-3.83	-5.75	-12.62	-6.24
	[0.57]	[1.16]**	[1.97]*	[3.10]*	[6.71]*	[2.15]***
Income 27001-36001	-0.92	-2.47	-3.74	-1.98	-6.50	-4.75
	[0.63]	[1.29]*	[2.01]*	[3.62]	[7.32]	[2.32]**
Income 36001-54000	-1.73	-5.52	-9.03	-12.07	-14.54	-9.71
	[0.55]***	[1.14]***	[1.88]***	[3.46]***	[7.19]**	[2.31]***
Income 54001-72000	-2.26	-6.97	-11.77	-21.20	-28.55	-13.71
	[0.63]***	[1.36]***	[2.27]***	[4.04]***	[8.57]***	[2.85]***
Income >72000	-3.14	-11.05	-18.18	-25.53	-42.91	-22.16
	[0.65]***	[1.37]***	[2.69]***	[4.67]***	[9.32]***	[3.13]***
Secondary education	2.14	5.22	10.41	12.89	19.45	9.19
	[0.45]***	[0.88]***	[1.62]***	[3.36]***	[6.21]***	[2.02]***
Tertiary education	2.45	6.62	14.62	15.96	25.71	11.71
	[0.44]***	[0.99]***	[1.69]***	[3.08]***	[5.98]***	[1.94]***
Constant	6.29	17.86	57.08	81.58	94.45	54.62
	[6.21]	[12.40]	[24.49]**	[37.18]**	[79.38]	[5.58]***
Pseudo R^2	0.03	0.04	0.04	0.04	0.05	0.05

Notes: Dependent variable is average time spent online total or on specified category of websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets. *, ** and *** indicate significance at10%, 5% and 1%, respectively

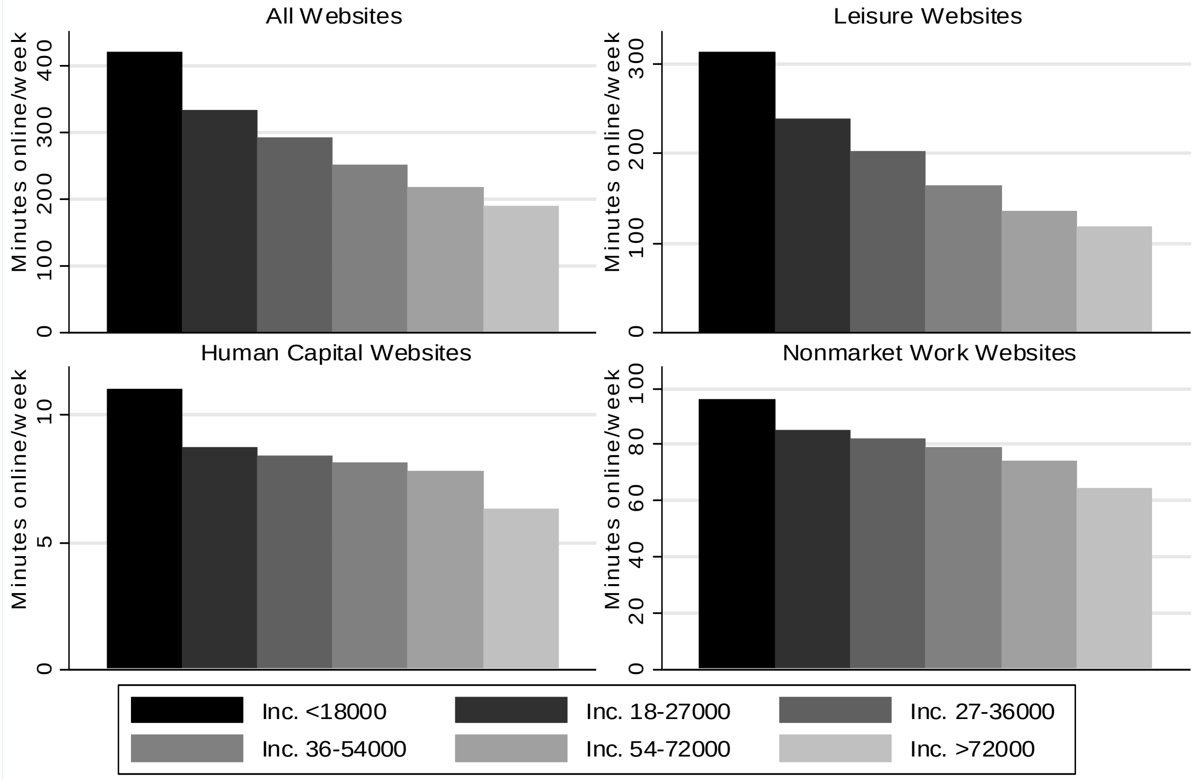
Figures

Figure 1 Distribution of time spent online



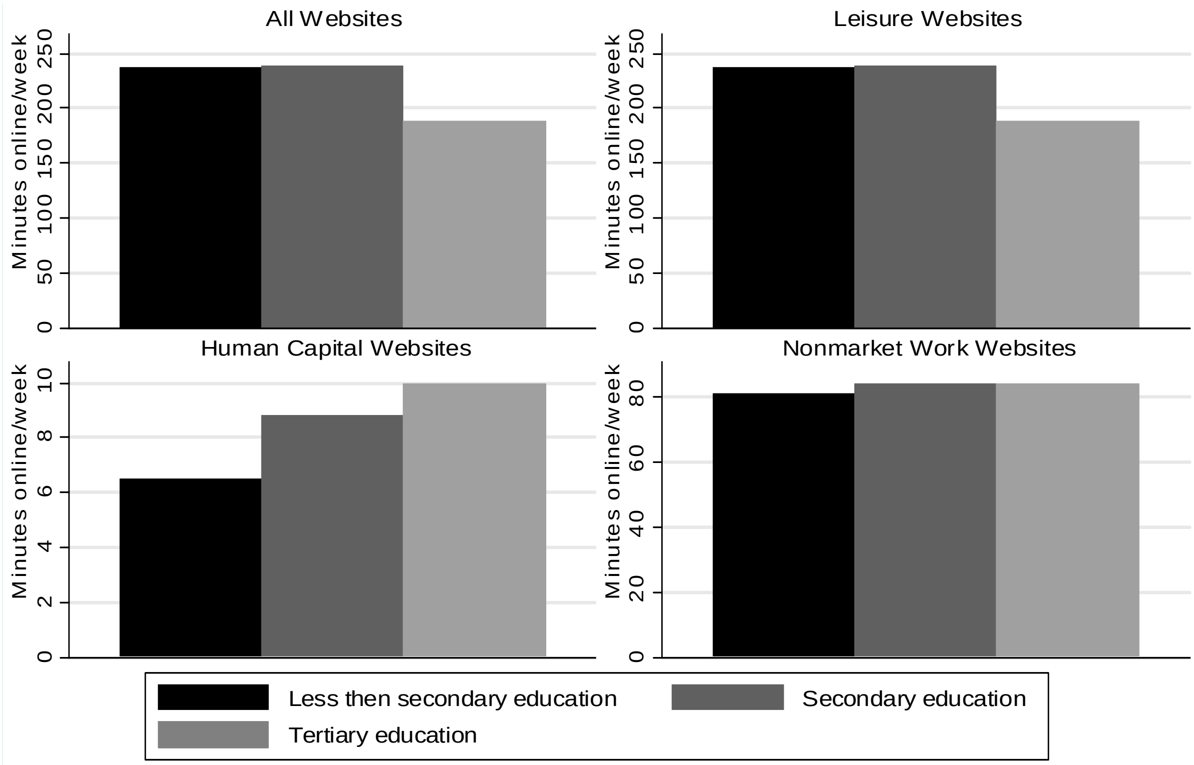
Source: Calculations based on Nielsen Click stream

Figure 2 Time spent on different websites and household income



Source: Calculations based on Nielsen Click stream

Figure 3 Time spent on different websites and education



Source: Calculations based on Nielsen Click stream