Wp

WORKING PAPERS

INTERNET AND LABOR INCOME: PLACES AND ACTIVITIES IN COLOMBIA
21/12/2016
N° 2016/16

INTERNET AND LABOR INCOME: PLACES AND ACTIVITIES IN COLOMBIA

Martin, Diego A.

DEVELOPMENT BANK OF LATIN AMERICA

WORKING PAPERS

INTERNET AND LABOR INCOME: PLACES AND ACTIVITIES IN COLOMBIA

Martin, Diego A.

CAF – Working paper N° 2016/16
21/12/2016

ABSTRACT

Despite the literature in the richest countries about the positive correlation between Internet and income, there is still an open question about the utility of this technology the developing world. This paper uses Propensity Score Matching and includes demographic and industry characteristics, which is the best possible way to control for the productivity of each worker. The matching model meets a broad common support, close gaps between control and treated individual after matching and holding the results for different matching techniques. The results present a positive and statistically significant correlation between Internet and income in Colombia. The key contribution of this paper is to show that the lowest wage premium for using Internet is in the middle of the skill distribution, as in the developed world. Nonetheless, Colombia differs to these countries in the tales of the skill distribution because the highest wage premium is for the lowest skill workers.

Wp

WORKING PAPERS

INTERNET E INGRESO LABORAL

Martin, Diego A.

CAF - Documento de trabajo N° 2016/16
21/12/2016

RESUMEN

A pesar de la literatura en países ricos acerca de la correlación positiva entre Internet e Ingreso, la pregunta sigue abierta en el mundo en desarrollo. Este paper usa el método de emparejamiento e incluye características demográficas y de la industria, la cuales son la mejor aproximación posible de productividad individual. El método de emparejamiento cumple con un amplio soporte común, cierra las brechas entre controles y tratados después del emparejamiento y mantiene los resultados para diferentes técnicas de emparejamiento. Los resultados presentan una correlación positiva y significativa entre Internet e ingreso en Colombia. La contribución principal es mostrar que la prima salarial por usar Internet es más baja en el medio de la distribución de habilidades, al igual que en los países desarrollados. Sin embargo, Colombia difiere con estos países en las colas de la distribución porque la prima salarial más alta es para los trabajadores con menores habilidades.

Internet and Labor Income

Places and Activities in Colombia *

Diego A. Martin†

December 21, 2016

Abstract

Despite the literature in the richest countries about the positive correlation between Internet and income, there is still an open question about the utility of this technology the developing world. This paper uses Propensity Score Matching and includes demographic and industry characteristics, which is the best possible way to control for the productivity of each worker. The matching model meets a broad common support, close gaps between control and treated individual after matching and holding the results for different matching techniques. The results present a positive and statistically significant correlation between Internet and income in Colombia. The key contribution of this paper is to show that the lowest wage premium for using Internet is in the middle of the skill distribution, as in the developed world. Nonetheless, Colombia differs to these countries in the tales of the skill distribution because the highest wage premium is for the lowest skill workers.

JEL codes: C14, J24, J31
Key words: Internet, Propensity Score Matching, Colombia

*I thank Universidad del Rosario for useful comments and discussion

†Development Bank of Latin America - CAF. Department of Researches in Economics (DIS), E-mail: dmartin@caf.com

1 Introduction

The impact of new technologies on productivity is one of the main questions around the world because they can simplify routine tasks, which leaves more free time for thinking and solving problems. In particular, Internet is the key to a universe of communication and education tools, where having a computer without Internet is almost useless because there is not possibility of accessing to these tools. Indeed, developing countries are very interesting in the effects of using Internet given the open questions about how people can improve their quality of life and access to Internet can be part of the solution. The Internet also is a mechanism through which the public sector builds policies of huge impact around all the country and particularly for the farthest places from the urban centers. Finally, policy makers want to know whether the expensive made in Internet is cost-efficient and how Internet can increase the competitiveness of the country.

In the world, the access to Internet is highly heterogeneous across regions. In 2015, Europe has on average 82 Internet user per 100 people, North America has 86 and South America has 51. For this regions, the latter has the highest variation between countries. Bolivia, for example, is in the bottom of the distribution with 35 Internet user per 100 people, Argentina in the top with 65 and Colombia is just on the average. This rank is very similar, looking for the per capita GDP in South America because Colombia is also neither the richest country, nor the poorest (World-Bank, 2015). This makes Colombia an interesting study case for all the region, where there are questions about how people can improve their wealth and there is an effort from the government on increasing the use of Internet.[1] The Colombian government invest 0.8% of the total budget in communication projects, this expenditure is in 15 out of 30 sectors, ordering government spending upwards (MinHacienda, 2014). Under this context, the research question is: what is the impact of the Internet on the labor income in Colombia? Furthermore, the available information in this country allows us to analyze four complementing hypothesis: first, the general impact of using Internet, second a heterogeneity effect depending on the

[1] In 2013, the percentage of households with access to Internet in developed countries was 76%, in America 51% and in Colombia 36%. In terms of individuals, the situation is similar, 75% in developed countries, 61% in America and 52% in Colombia (MTIC, 2014).

place of access, third the impact of the activities done on the Internet and forth who earns more for using Internet the high skill workers or the low skill. Therefore, this paper fills the gap in the literature about the benefits from the Internet in developing countries and provides an empirical support for future investments from the public and private sector in new technologies involving Internet use. Even this study contributes for countries with similar income and challenges in other regions such as Asia and Africa, where the Internet can help to face poverty. In fact, this is a unique study in the country and a benchmark to compare with other studies in other countries in the region, such as Navarro (2010).

There are empirical and theoretical studies about the impact of new technologies on earnings for developed countries (Acemoglu, 1998; Autor, Levy, & Murnane, 2003; Bartel, Ichniowski, & Shaw, 2007). The main idea of these studies is that workers become more productive when they manage new technologies because they can do either more complex or faster tasks than workers without this skill. The labor market notices the rise of abilities, which causes a wage premium for using technology. Although the first studies focus on the appearance of the computer, having this machine is highly related to using Internet in this time.[2] Hence, using Internet has many advantages that only this tool brings such as accessing to information, job vacancies, belonging to the network of professional contacts and obtaining training in education programs. Empirical studies have used parametric and non-parametric models to estimate the effects of the Internet on earnings (Forman, Goldfarb, & Greenstein, 2012; DiMaggio & Bonikowski, 2008). The parametric models use the ordinary least square (OLS), applying a mincer equation with a dummy equal to 1, when an individual uses Internet. These models estimated bias coefficients because there is a problem of omitted variable, and then the coefficient associated with the impact of the Internet is related to other unobservable variables. Furthermore, there is not a natural experiment in Colombia where the treatment is randomly assigned and then the estimators suffer from selection bias, because people choose using Internet according to their own characteristics. The non-parametric models use Propensity Score Matching (PMS) to deal with the omitted variables and bias selection. The method al-

[2]In Colombia, for example, Only 6% of the households with computer does not have Internet.

lows comparing the close people in characteristics, whose only difference in the outcome is being in the treated or controlled group. In this paper, non-paremtric estimations are robust under all the requirements, such that satisfying extensive common support, the gap between treat and control groups are closed in observed and unobserved variables, the results are maintained in sign and statistically significant when controls are included and the methodology of matching is changed.

Using a household surveys from 2009 to 2011 (Colombian Great Integrated Household Survey, GEIH), this study finds a positive and significant correlation of using Internet on labor income in OLS and PSM models. On average, Internet users earn 6.5% of one standard deviation more than not users. For places of access, Internet users at work earn 28.4% and user at home 17.4% of one standard deviation more than not users. Workers using Internet for education and bank service earn 7.1% of one standard deviation more than people using Internet for downloading music and search for videos. Finally, the wage premium depends on the occupation (skill) of each worker, being the middle of the skill distribution the less benefited, as in the developed world. Nonetheless, the higher wage premium is in the lowest tale of the distribution, where farmers and miners using Internet gain 20.7% of one standard deviation more than not users in the same occupation. In general, the OLS model overestimated the results because this empirical strategy cannot eliminated the bias coming for the not random assign of the treatment. Whereas, matching model compare the individuals as similar as possible, even in productivity.

The structure of the paper is as follows: Section 2 presents the literature review about the mechanism of the impact of new technologies and earnings. Section 3 provides a discussion about how the impact of the new technologies is estimated and some results in other countries. Section 4 builds up four hypothesis to develop the research question. Section 5 presents the empirical strategy, using the parametric model (OLS) and non-parametric (PSM) model, including industry controls, which work such as the best proxy of productivity. Section 6 explains the database and the descriptive statistics depending on the treatments: Internet use, Internet use at work, Internet use at home and Productive users. Section 7 shows the results for OLS and PSM, as well as robustness check from the daily Internet use. Finally, Section 8 presents the conclusion and final remarks.

2 New Technologies and Earnings

There are many studies about the relationship between technology and wages, mainly for United States. Since the start of the computation era, there is a rise in the demand for college graduates, who can manage this new technology. Thus, there is a wage premium in favor of technology users in the long term. This is also known as the canonical model, which has explained empirically the evolution of skill premium under the increment of computers (Acemoglu & Autor, 2011). After 1980, computer capital can replace workers in manual tasks, which follow a explicit rule, and complement workers in non-routine tasks such as communication and solving problems (Card & DiNardo, 2002; Autor et al., 2003). This causes a polarization of the labor market in United States, divided into well-paid (skilled) and not well-paid (unskilled) workers (Autor, 2015). In addition, the change in wage is not strikingly monotone across skills because the lowest premium is in the middle of the distribution, instead of being in the bottom, then there is a U-shape in the relationship wage and skills. Such that the automation coming from technology reduces more the demand for middle skill worker than the demand for lower skill workers (Autor & Dorn, 2013). This same pattern is also found in 16 countries in Europe (Goos, Manning, & Salomons, 2009) and United Kingdom (Acemoglu, 1999; Goos & Manning, 2007).

Nowadays, labor market requires mainly the use of Internet to access a quality job because online skills help people to perform their work with higher efficiency. Worker with relatively high ability are more likely to move to non-routine cognitive occupation, while people with low ability relatively stay in routine jobs with less growth pay (Cortes et al., 2016). The Internet also makes easy that workers move to a better jobs because they can find the labor demand concentrate in one place (McDonald & Crew Jr, 2006). In United States, workers who use Internet earn 13.5% more than not users. This wage premium is even higher in industries less intensive in technology, where the only one who uses Internet is the person who can perform the most difficult tasks (Goss & Phillips, 2002). Using panel data, DiMaggio and Bonikowski (2008) also found an increase of 20% in hourly wage of workers for using Internet, not just computers. This shows that this tool is useful for finding information and makes simpler communication. In fact, Bartel et al. (2007) present that companies incentive workers to obtain human capital, when they invest in Internet and other technologies. Gust and Marquez (2004), for example, show that the more flexible regulation in information technology in United State explains in part the fact that this country has higher productivity than 13 industrial countries in the OECD. In developing countries, Paunov and Rollo (2016) emphasize that using Internet also impacts the productivity positively, where firms with the most sophisticated production have the highest gain from this technology.

Internet is important in the labor market and in the daily life. The massive media moving to online channels make the web a focus of culture, education, politics and personal relationships. Indeed, Internet can reduce the isolation, the social exclusion and the physical disabilities, which also affect the productivity (Foley, 2004; Chigona, Beukes, Vally, & Tanner, 2009; Dobransky & Hargittai, 2016). In fact, a biggest change in technology is not a requirement to have high effects in quality life. Blanco and Vargas (2014), for instance, show that sending message to vulnerable population, saying what rights they have and where they can ask for them, increase the access to public aids and their welfare Therefore, there is a open debate about the impact of the Internet in different variables, but the majority of the discussion takes place in richest countries. For this reason, this study is an important way to see whether the technology is useful in the developing world, making more worthy expenses and effort in Internet projects.

3 How to Estimate the Effect

The literature is divided in two ways of estimating the relationship between new technologies and earnings: parametric models using ordinary least square (OLS) and non-parametric model using Propensity Score Matching (PSM). First, Krueger (1993) is the pioneer in using an equation of Mincer (1958), including a dummy variable equal to one when the individual uses computers, to see the relationship of this technology with labor income. This model is also used in different countries and they find a wage premium for using technology in some countries such as United States, United Kingdom and Australia (Goss & Phillips, 2002; Arabsheibani, Emami, & Marin, 2004; Chiswick & Miller, 2007). Second, Rosenbaum and Rubin (1983) presents the method of PSM for doing causal inference, and there are many impact evaluation policies that base on this methodology (Heckman, Ichimura, & Todd, 1998; Heckman, LaLonde, & Smith, 1999). In topics of new technologies and labor income, this model is widely used when there is information from household surveys (DiMaggio & Bonikowski, 2008; Navarro, 2010; Benavente, Bravo, & Montero, 2011).

Parametric models estimate bias coefficient because they do not deal with two issues: omitted variables and reflection problem. The associate coefficient to use Internet can reflect the differential of variables such as ability and there are workers that were more productive even before using Internet (Entorf, Gollac, & Kramarz, 1999). In is also possible that the Internet connection is expensive and only well paid workers can afford Internet access, which implies the reverse causality. Indeed, DiNardo and Pischke (1997) doubt the empirical results in favor of using computer. Although they find a similar wage premium for using computer in Germany and United States in 1997, they argue that there is also a large wage premium for using calculators, telephones, pens or pencil. They claim that those results show the problem of selection bias, where the characteristics of workers decide whether they have access to these tools. According to these authors, given the unlikely of an experiment where Internet is randomly assign, the best way to estimate the effect of technology is including education and fixed effects in the model. In fact, there is not evidence of computer wage premium in some countries such as Ecuador and in Great Britain the computer skills are not very important relatively to math and reading skills (Oosterbeek & Ponce, 2011; Borghans & Ter Weel, 2004). Particularly in using Internet, Lee and Kim (2004) only find a premium wage for Internet users in United State for 1997, but they do not find any result statistically significant for 1998 and 2000.

The non-parametric model using matching techniques deals with the problem of biased estimates, comparing workers as similar as possible in demographic, social and firm characteristics, where the only difference is either using the Internet or not. Using matching techniques in United States, DiMaggio and Bonikowski (2008) find that on average Internet users earn 15% more income and there is heterogeneity in the magnitude of the effect depending of the place where people access. The highest coefficient is using Internet everywhere, following using Internet only in work, and the smallest effect is only using at home. Navarro (2010) also uses matching models to find that the Internet increase the labor income in five out of six countries in Latin America. The magnitude varies between 30% in Honduras until 18% in Mexico.[3]

In developing world, the relationship between Internet, income, education and other variables is a research agenda. In the best of my know, there is not studies for Colombia including the use of household surveys, non-parametric models and fixed effect of size firm, occupation and economic sector, which analyze the impact of the Internet on earnings. In addition, other developing countries can also learn for these results, because the Internet disappears geographical barriers and it makes the world smaller. For this reason, the contribution of this research is to present new evidence and methodology for this country and the corresponding technical support to justify investments on Internet.

4 Hypothesis

This paper proposes four hypotheses to analyze how productivity is the main mechanism between the correlation of Internet and income. First, Internet helps people to develop their jobs faster and more efficiently, causing positive and significant effect on the earnings. A huge development in job activities is not necessary, whenever workers can either communicate easier or access to information faster such as bank loans, education or new technologies, they achieve high salaries. Second, Internet is used in different places, in particular at the workplace and at home. Considering that it is more likelihood that peo-

[3]The other countries are Brazil with 29%, Chile 26% and Costa Rica 24%. Paraguay is the only country where the effect is not statistically significant.

ple use Internet at work in productive activities, this places has a higher wage premium that Internet users at home. Third, Internet offers different kinds of applications for education, communication and entertainment. However, the productive activities have more impact on earnings because the labor market values the effort made by the workers. Thus, employees using Internet for activities such as education have a higher reward than the workers using Internet for entertainment. Finally, there is a polarized labor market in developing world (Autor, 2015), but it is probably that this polarization does not occur in developed world such as Colombia because these two regions look very different in education, industrialization and other variables.

5 Empirical Strategy

The first way to analyze correlation between income and Internet is estimating the equation 1 using ordinary least square (OLS):

$$Y_{ifjt} = \alpha_0 + \theta Internet_{ifjt} + X_{ifjt} + \kappa_{ijt} + \delta_j + \gamma_t + \mu_{ifjt} \quad (1)$$

Y_{ijtf} is the monthly labor income standardized of the individual i, working in the firm, the economic activity and occupation f, in the municipality j in year t. $Internet_{ifjt}$ is the treatment variable, equal to 1 whether the worker uses Internet and 0 otherwise. X_{ifjt} are socio-demographic variables such as age and education, and experience with technology such as computer and telephone holding. κ_{ijt}[4] is a set of fixed effects, combining firms size, economic sector and occupation. These variables build a cell (as small as possible) where Internet user and not user are compared. One of the most relevant omitted variables is the workers ability. In this paper, this omitted variable is controlled using the industries code and the sector code of the place where an employee works. This is a good proxy because shows where a person works and what she or he does in that place. δ_j is the municipality fixed effect, which controls for aggregate shock at muniicpal level such as slope and rain. γ_t is the year fixed effect, which control for aggregate shock in time

[4] κ_{ijt} can be also written such as: κ_{ijt}=$\kappa_{ijts} + \kappa_{ijte} + \kappa_{ijto}$. Where κ_{ijts} is firm size fixed effect, κ_{ijte} is economic sector fixed effect and κ_{ijto} is occupation fixed effect

such as inflation or macroeconomic condition for all the country. μ_{ifjt} is the error term. The OLS model estimates bias coefficients because the model cannot control for omitted variables, such as Internet users are highly skilled workers, and reflection problem, when wealthy firms provide Internet) (DiNardo & Pischke, 1997). [5]

The ideal database to estimate the impact of the Internet on labor income is a experiment framework, where the Internet is randomly assign to one group and not to other group. However, that database is very unlikely available in developing countries and in case it exists, it can suffer with the common experiment's problems with high probability because it is strongly difficult to realize what people use technology and people who are not assigned to the treatment (using Internet) can easily access to Internet for ways such as cell phones, even in this context (this is called not compliners (Angrist & Pischke, 2009). Barrera-Osorio and Linden (2009), for example, assign randomly computes in Colombian schools to analyze the impact of technology on children performance. The authors find a very little effect on students, explaining for the low real inclusion of computer in classroom, teacher have the computer but they do not want to use for teaching. In this context, Propensity Score Matching (PSM) is useful to find unbiased estimators when the treatment does not come from a random assignation. Intuitively, this non-parametric model allows us to find for each Internet user (treated individual), another one being exactly the same, except for not using Internet (controlled individual). This model assumes that the treatment is determined exclusively by observable variables of the individuals and gaps in unobservable variables are closed at the same time as the gaps in observable variables, which is called conditional Independence (Angrist & Pischke, 2009).

The matching process requires two steps: First, estimating the propensity score (PS) with a Probit model, which is a conditional probability of using Internet following the equation 2:

$$Internet_{ifjt} = \alpha_0 + X_{ifjt} + \kappa_{ijt} + \delta_j + \gamma_t + \mu_{ifjt} : \tag{2}$$

[5]The estimations control for a dummy equal to one for salaried workers an zero for self-employed. Table A-1 shows that difference in variable between these two types of workers is statistically different from zero

$Internet_{ifjt}$ is a dichotomy variable equal to 1 when a worker uses Internet. The independent variables include demographic characteristics, firm size, economic sector, occupation and fixed effects of municipality and time.[6] Second, using the propensity score, each treated individual is matched with a control individual, reducing the gaps in observable variable between Internet users and not users. The PSM calculates an average effect of the treatment in the treated (θ_{ATT}), which means that the estimated parameter is only for a sub-sample, in this case workers, instead of being of the full sample (Angrist & Pischke, 2009). The equation 3 show the general coefficient estimated by PSM.

$$\theta_{ATT} = \sum_{i=1}^{I} mean[(Y_i|Internet_i = 1) - Y_{c(i)}|Internet_i = 0)] \tag{3}$$

θ_{ATT} is the difference in the outcome between the most similar treated and control workers. I is the number of workers in the sample, $Internet_i$ is the treatment, equal to 1 when an individual i uses Internet. c_i is the set of control workers, who do not use Internet but they are very similar in all the characteristics. Even, this method compares workers with similar productivity because equation 3 controls for firm size, economic sector and occupation. In addition, the simplest and more intuitive way to match workers is using the nearest neighbor, which compares a treated individual with the closet control individuals in propensity score, following an Euclidean distance. However, this method can be useless to close gaps in observable variables when there are many covariates and fixed effects, falling in the risk of no control bias of omitted variables because it is not a comparison between the most similar individuals. To solve this problem, the Mahalanobis distance offers an advantage to close gaps in most variables because the distance includes a matrix of variance and covariance of characteristics, which is more informative in order to match treatments and controls.

The main challenge in calculating the correlation between Internet and income is to control for the unobservable productivity of each worker (omitted variable) and the possibility that only the richest companies access to Internet (reverse causality). This study is an improvement of the other matching models because it is built a cell as small as

[6]All the variables including in equation 1 are also including in equation 2.

possible, where the treated and control individuals are compared. This cell includes socio-demographic characteristic and most important, information about size firm, occupation and economic sector. These variables are the better way to create a proxy of productivity because firm size gives information about the environment where the employees work, occupation shows what employee do at work and economic sector presents the market conditions of the place where employees perform their tasks.

6 Data and Descriptive Statistics

This study uses the Colombian Great Integrated Household Survey (GEIH) for 2009 to 2011. This data is a nationally representative cross-section survey and it is the main tool for official statistics because it contains information at individual level in housing, education, income and labor characteristics. The dependent variable is the standardized monthly labor income and the sample is restricted to workers between 18 to 65 years old in urban areas.[7] Only this period includes the following question about Internet use: Have you used Internet in the last 12 months? and Where do you use the Internet?[8] To analyze the productive activities, the questions are: Do you use Internet for education or financial finance?[9] For robustness checks, the question is: Do you use Internet every day?[10] This survey is the best way to know whether an individual uses this service or not in Colombia.

Controlling for economic activities, occupation and firm size is on the most important aspects of this study because this information is a proxy of productivity per worker. This is an important improve comparing with other studies, which reduces the possible bias coming from omitted variable and revers causality. The economic activity is divided in

[7]This sub sample is also used by DiMaggio and Bonikowski (2008) and Benavente et al. (2011). The results are robust to use the full sample, where rural areas are included.

[8]The options are accessing to Internet at home, at work, in educational institutions, in free public access centers, in payed access centers, In the house of another person (relative, friend, neighbor), in other place.

[9]The original question is for which of the following services or activities, do you use Internet? The options are obtaining information, communication, electronic banking and other financial services, education and learning, transactions with government agencies, entertainment (games, download music, etc).

[10]The original question is how often do you use Internet? The options are at least once a day, at least once a week but not every day, at least once a month, but not every week and less than once a month.

13 sectors from agriculture to financial service using the Classification of All Economic Activities (ISIC, 2008).[11] The occupation are classified in 10 categories since farming until professional specially, following the Standard Occupational Classification System (SOCPC, 2010).[12] The size firm considers 5 categories depending on the number of workers. Each category of the variables is included as a fixed effect in all the estimations. Therefore, these variables recover the productivity of each worker and they compare an Internet user and not user in the same firms economic activity (sector), in the activities that workers actually perform (occupation) and in the places where they work (firm size).

Table 1 presents the characteristics of the individuals by location of Internet access and productive activities on the web. It is clear that these aspects are not exclusive, because an individual can use Internet in both places or in any place. On average, the Internet users are women, younger, better educated and well paid than not users. By location (work and home), both groups seem very similar in these characteristics. Internet users in productive activities are women with higher educated and older than Internet users in entertainment. Furthermore, Internet users have more experience with technology than not users, because a large proportion of them have telephone, computer and satellite TV at home, comparing with not users (62%, 81% and 62%, respectively). This is also true for Internet users at work, at home and productive users. Particularly, users at home have the highest experiences, 78% have telephone, 90% satellite TV and 92% have PC at home. People using Internet in education are very close in technology experience to Internet users in entertainment. For the financial situation of the house all the categories, this variable seem very similar, between 41% for not users and 48% for users at home, this is an approach to wealth and living condition.

Table 2 shows Internet users distributed by firm size, economic sector, occupation and type of worker. First, there is a positive relation between number of employees in companies and Internet users in all categories. On average, 78% of workers use Internet in big companies (101 and more employees employees), while 55% of worker in small

[11] The sectors are: agriculture, mining, manufacturing, electricity, construction, transport, financial service, housing sector, public administration, education, health and domestic service.

[12] The occupations are: Professional specialty, executive and managerial, service, sales, machine operator, cleaners and laborers, profession products, transportation, farming and mining

companies (2-5 employees). The number of Internet users decreases when the analysis is made by access and activities. In big companies the number of users is almost the same at work (69%) and at home (66%), while in small firms, only 29% of employees use Internet at work and 32% use Internet at home. The number of productive users is always smaller than not productive users in all the firm size. The small companies seem to have a poor control of the employees activities on the web because only 22% of workers use Internet to education and financial and 72% using Internet for entertainment in the smallest firms.

There are clearly economic sector where workers need more Internet for their daily activities. Financial service and public administration are the sectors with more Internet users (87% and 81%, respectively), while construction and mining have the smallest group of users (56% and 49%, respectively). By place of access, Internet users decrease especially in sector with routine task, where the 28% of workers in domestic service use Internet at work or at home, being the sector with the smallest number of Internet users. Activities in the Internet show an interesting behavior. On one hand, financial service is in the top of productive users (59%) compare with not productive users (40%). On the other hand, domestic service is in the bottom with just 26% of people using Internet for education and financial service comparing with 73% of people using Internet for entertainment.

Although occupation seems very similar to economic sector in terms of the distribution of Internet users, this variable is very important to compare people doing the same activities. To summarize, 83% of white-collar (not routine activities) use Internet, comparing with 60% of blue-collar (routine activities). White-collar workers also use more Internet at work, at home and in productive activities.

7 Results

There is a positive correlation between Internet users and income in all the world. Figure A-1 shows that an increment in the number of Internet users is positively correlated with the logarithm of per capita GDP. This figure also presents a high heterogeneity between regions because the developing world is on the right-top of the figure, where there is high levels of income and Internet and Latin America is in the middle of the figure, being under 50 Internet users per 100 people and around 9 in logarithm per capita GDP.

In the Colombian case, the OLS model in table 3 shows the positive correlation between Internet and income controlling for socio-demographic characteristics, firm size, economic sector, occupation and fixed effects of municipality and time. The correlation is statistically significant adding each set of controls. Workers gain 6.9% of one standard deviation more when they use Internet. Although the OLS model includes control for industry characteristics (the best proxy of productivity), the matching model reduces the bias at maximum because within each small cell of firm size, economic sector and occupation compares the most similar workers according to the propensity score.

The first step for Propensity Score Matching is running a Probit model using equation 2. Table 4 presents the marginal effects of this model for different treatment: Internet users, workplace users, worker using Internet at home and productive activities on the Internet,[13] including in all the estimation include municipality and time fixed effects. The first set of fixed effects controls relevant variables related to cities characteristics. The second set of fixed effects, among other things, control for weather and geographic conditions. Hence, the probabilities of receiving the treatments take into account, for example, that big cities have higher probability of accessing to Internet because they enjoy more available services than those with difficult geographic and weather conditions. The results in table 4 show that schooling and technology experience increase the probability of using Internet in all the treatments, while age reduces the probability. At the same time, working in a big firm (more than 101 employees) in the financial service and performing the occupancy of professional increase the probability of using Internet at both places and in productive activities. This table also presents that the independence conditional assumption holds because all the variables are statistically significant, showing that the treatment can be caused by observable variables.

Matching models have to overcome three proofs to find unbiased estimators (Smith

[13]In the treatment of using Internet, using at work and using at home the control group is not users. In the treatment of productive users, the control group is Internet user in not productive activities

& Todd, 2005; Arceneaux, Gerber, & Green, 2006; Porto, 2016). The first requirement is a common support in the probability of using Internet between treated and control group. Figure A-2 shows the propensity score for Internet users and not users, where the 90% of the sample is in the intersection between the probabilities of both groups. In this common support, the matching method looks for treated and control individuals as similar as possible in all the covariates. Second, the matching method has to close the majority of the gaps in observable and unobservable characteristics between Internet users and not users. Table 5 presents the difference in average between treated and controlled individuals before and after the matching process, using two different methods: the nearest neighbor and the Mahalanobis distance. In the unmatched sample, all the differences are statistical different from zero. After using nearest neighbor, there are still many differences between Internet users and not users, while Mahalanobis distance closes the gaps for all the variables. Thereby, this matching technique builds a small cell where the individuals are very similar and the unique difference is the treatment, using Internet or not. In this way, the problem of reverse causality is solved, because the study compares a clone in socio-demographic characteristics, in the same firm, performing the same occupation. Third, estimated coefficients have to be robust to changes in the variables and the matching method. Table 6 presents the results using two matching methods and adding sets of controls. There is a positive correlation statistically significant between the Internet and the labor Income, under different matching methods and controls. Column 5 using the Mahalanobis distance and all the covariates shows that Internet users earn 6.5% of one standard deviation more than not users. Table A-2 estimates the matching model for each year in the sample and it presents a reduction in the effect during time. In 2009, Internet users earn 7.8% of one standard deviation more than not users, while in 2011 the correlation is 6.1%. These results evidence that as the technology expands the wage premium reduces because there more labor supply who can use the new tool. Autor et al. (2003) find the same pattern for United States, the increase in salary for using new technology decrease over time.

Is this coefficient big or small in magnitude? Considering that 54% of the population in Colombia live with less than one monthly minimum salary, which is defined as the minimum money to have a decent life (MinTrabajo, 2014). 6.5% of one standard deviation in income is equivalent to 11% of one minimum salary, which can definitely help worker to overcome this threshold. Is investment in technology cost-efficient? Comparing the per capita expenditure in technology projects and wage premium for using Internet, for each dollar invests in Internet, there is an increase of 3.4 in the salary of the workers.[14]

There is a difference wage premium depending on the places where people uses Internet, regardless using OLS or matching models. On one hand, Internet user access to this tool from different places and they do any task.[15] On the other hand, Internet access at work is linked with productivity activities because it is more likely that workers use Internet for company issues than personal purposes when there is some control of either from colleges or employers. Similarly, Internet access at home is related to productive activities because the sample is restricted only for workers, who can even work from their houses using telecommuting. Table 7 show the results for running the OLS and PSM model including all the controls in table 1 and 2. Looking the matching model because it reduces the possible problem of omitted variable and self-selection bias, Internet user at work earn 28.4% of one standard deviation more than not users, and Internet users at home gain 17.4% more than not users. In addition, the correlation between Internet and income has a high heterogeneity between places of access. People using Internet anywhere obtain only 22.8% of the wage premium of users at workplace and individuals using Internet at home receive 61.2% of the group mentioned before. The difference in the wage premium is explained mainly by the difference in productivity activities.[16] For this reason, the last column in table 7 shows that workers using Internet for education and financial services earn 7.1% of one standard deviation more than Internet users in

[14]This relationship is calculated in for steps: First, dividing the budget spending in technology and communication projects (0.8% of the total budget) over total population in Colombia. The results is \$5.92. Second, 6.5% of one standard deviation of income is \$17.09. Third, dividing \$17.09/\$5.92. In all the steps the prices are constant to 2011 and converted to dollar using PPP.

[15]Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, In the house of another person (relative, friend, neighbor). Internet use is either a bigger set than Internet at work or at home.

[16]Productive users earn 27.6% of one standard deviation more than not users, using Mahalanobis distance and including all the set of controls.

less productive tasks, such as searching for music and videos. The last variable to take into account is the intensity of using Internet, for this reason table A-3 run the four treatments using daily Internet and the results show a positive correlation between Internet and income, with a slightly increase in magnitude for all the treatments.[17] Recalling the three hypothesis in section 4, there is a positive correlation between Internet and income, the wage premium depends on place of access and the online activities really matter for the final increase in the salary.[18]

There is evidence that new technology causes a polarized labor market in developed countries such as United States (Autor, 2015), where the workers doing the most sophisticated task have the higher wage premium for using these tools, the lower wage premium is for workers in middle distribution task and worker in routine tasks have a wage premium between these two groups. In Colombia, table 8 evidence the matching model of equation 3, including all the set of of covariates and dividing the sample depending on the occupation of workers. The evidence for developing world holds partially in a developed country such as Colombia. There is a polarized market where workers doing middle-skills task have the lowest wage premium. Indeed, the figure A-3 seems very similar to the findings for United States (Autor & Dorn, 2013). However, the higher wage premium is in the task requiring the lowest skills. Internet users working either in farming or in mining earn 20% of one standard deviation more than not users in this same occupation. The following wage premium is for the task more intensive in solving problems and not routine task. Internet users being professional specialty gain 18% of one standard deviation more than not users in the same occupation.

[17]On average, Daily Internet users earn 8.1% of one standard deviation more than not users.

[18]Most of salaried workers use Internet only at job, while self-employed workers also use Internet at home. Furthermore, the decision of buying Internet is not made by workers, it is taken by companies. On the other hand, self-employed workers take the decision and the cost of accessing Internet by their own. Finally, according to the literature self-employed workers are widely different in several characteristics, which are related to education, expected income, risk and preferences for autonomy (Guataqui, Martin, & Porto, 2016). Table A-4 shows that self-employed workers earn around the double of salaried workers when they use Internet

8 Conclusion

There is an open question about the relationship between Internet and income around the world. This study analyze this correlation for the Colombian case, a developing country investing in technology projects and with a lot of inequality in income, education and other issues. The empirical strategy builds a novel framework for evaluating the effect of Internet when there is not a random experiment, as in the case of developing world. Besides that, even field experiment can suffer heavily from not complainers, people can access to Internet when they are in the control group, because it almost impossible to realize who use Internet or not all the time. The methodology creates a proxy of productivity including firm size, economic sector, occupation and other variables, which controls for omitted variable and self-selection bias. The estimators from the matching specification are robust to different proof, which makes these parameters reliable. The results show a positive correlation between Internet and income in the Colombian case, using parametric and non-parametric models. The wage premium depends on the places of access to Internet, the activities doing in the web, the intensity of the use and the occupation of each worker. The highest increase in income is for Internet users at work in the lowest skill occupation (farming and mining). Whereas, the lowest wage premium is for Internet users in a place in the middle of the skill distribution.

This paper presents a full set of results following the hypothesis that the Internet increases productivity, which causes a rise in the salary. People learn how to do something faster and more efficiently when they can access to Internet. In addition, investments in technology are worthy and cost efficient because the total amount of budget for this projects is just 0.8% and the results are always positive statistically significant, between 6% and 20% of one standard deviation in income for Internet users. As the best of my know, this is the first time that the hypothesis of polarization labor market is tested in a developing world and the finding are much more important because the main beneficiaries are the most poor people in Colombia, farmers and miners.This makes the Internet a strong tool to face poverty and lack of opportunities. Developing world need for this type of policies, faster to implement and with a high effect on key variables such as income.

References

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *Quarterly journal of economics*, 1055–1089.

Acemoglu, D. (1999). Changes in unemployment and wage inequality: An alternative theory and some evidence. *The American Economic Review*, *89*(5), 1259-1278. Retrieved from `http://www.jstor.org/stable/117057`

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics*, *4*, 1043–1171.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricist's companion* (Vol. 1). Princeton university press Princeton.

Arabsheibani, G. R., Emami, J., & Marin, A. (2004). The impact of computer use on earnings in the uk. *Scottish Journal of Political Economy*, *51*(1), 82–94.

Arceneaux, K., Gerber, A. S., & Green, D. P. (2006). Comparing experimental and matching methods using a large-scale voter mobilization experiment. *Political Analysis*, *14*(1), 37–62.

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *The Journal of Economic Perspectives*, *29*(3), 3-30. Retrieved from `http://www.jstor.org/stable/43550118`

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *The American Economic Review*, *103*(5), 1553-1597. Retrieved from `http://www.jstor.org/stable/42920623`

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*(4), 1279-1333. Retrieved from `http://www.jstor.org/stable/25053940`

Barrera-Osorio, F., & Linden, L. L. (2009). The use and misuse of computers in education: evidence from a randomized experiment in colombia. *World Bank Policy Research Working Paper Series, Vol.*

Bartel, A., Ichniowski, C., & Shaw, K. (2007). How does information technology affect productivity? plant-level comparisons of product innovation, process improvement, and worker skills. *The Quarterly Journal of Economics*, *122*(4), 1721-1758. Retrieved from `http://www.jstor.org/stable/25098887`

Benavente, J. M., Bravo, D., & Montero, R. (2011). Wages and workplace computer use in chile. *The Developing Economies*, *49*(4), 382–403.

Blanco, M., & Vargas, J. F. (2014). Can sms technology improve low take-up of social benefits? *Peace Economics, Peace Science and Public Policy*, *20*(1), 61–81.

Borghans, L., & Ter Weel, B. (2004). Are computer skills the new basic skills? the returns to computer, writing and math skills in britain. *Labour Economics*, *11*(1), 85–98.

Card, D., & DiNardo, J. (2002). Skillbiased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, *20*(4), 733-783. Retrieved from `http://www.jstor.org/stable/10.1086/342055`

Chigona, W., Beukes, D., Vally, J., & Tanner, M. (2009). Can mobile internet help alleviate social exclusion in developing countries? *The Electronic Journal of Information Systems in Developing Countries*, *36*.

Chiswick, B. R., & Miller, P. W. (2007). Computer usage, destination language proficiency and the earnings of natives and immigrants. *Review of Economics of the Household*, *5*(2), 129–157.

Cortes, G. M., et al. (2016). Where have the middle-wage workers gone? a study of polarization using panel data. *Journal of Labor Economics*, *34*(1), 63–105.
DiMaggio, P., & Bonikowski, B. (2008). Make money surfing the web? the impact of internet use on the earnings of us workers. *American Sociological Review*, *73*(2), 227–250.
DiNardo, J. E., & Pischke, J.-S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *The Quarterly Journal of Economics*, *112*(1), 291-303. Retrieved from `http://www.jstor.org/stable/2951283`
Dobransky, K., & Hargittai, E. (2016). Unrealized potential: Exploring the digital disability divide. *Poetics*, *58*, 18–28.
Entorf, H., Gollac, M., & Kramarz, F. (1999). New technologies, wages, and worker selection. *Journal of Labor Economics*, *17*(3), 464–491.
Foley, P. (2004). Does the internet help to overcome social exclusion. *Electronic Journal of e-government*, *2*(2), 139–146.
Forman, C., Goldfarb, A., & Greenstein, S. (2012). The internet and local wages: A puzzle. *The American Economic Review*, *102*(1), 556–575.
Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The review of economics and statistics*, *89*(1), 118–133.
Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in europe. *The American Economic Review*, *99*(2), 58–63.
Goss, E. P., & Phillips, J. M. (2002). How information technology affects wages: Evidence using internet usage as a proxy for it skills. *Journal of labor Research*, *23*(3), 463–474.
Guataqui, C., Martin, D., & Porto, I. (2016). El raro caso de los trabajadores cuenta propia en colombia. *Documment not publish*.
Gust, C., & Marquez, J. (2004). International comparisons of productivity growth: the role of information technology and regulatory practices. *Labour economics*, *11*(1), 33–58.
Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261–294.
Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of labor economics*, *3*, 1865–2097.
ISIC. (2008). International standard industrial classification of all economic activities. revision 4. *United Nations Publications*. Retrieved from `http://unstats.un.org/unsd/publication/seriesM/seriesm_4rev4e.pdf`
Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata. *The Quarterly Journal of Economics*, *108*(1), 33–60.
Lee, S.-H., & Kim, J. (2004). Has the internet changed the wage structure too? *Labour Economics*, *11*(1), 119–127.
McDonald, S., & Crew Jr, R. E. (2006). Welfare to web to work: Internet job searching among former welfare clients in florida. *J. Soc. & Soc. Welfare*, *33*, 239.
Mincer, J. (1958). Investment in human capital and personal income distribution. *The journal of political economy*, 281–302.
MinHacienda. (2014). Presupuesto de 2015 por 216, 2 billones tendr prioridad en lo social: Ministerio de hacienda. *Ministerio de Hacienda. Colombia.*
MinTrabajo. (2014). Definicin de salario mnimo. *Ministerio de Trabajo. Colombia.*
MTIC. (2014). Reporte de la industria tic. *Ministerio de tecnologas de la informacin y las comunicaciones. Colombia.*

Navarro, L. (2010). *The impact of internet use on individual earnings in latin america* (Tech. Rep.). Institute for Advanced Development Studies.

Oosterbeek, H., & Ponce, J. (2011). The impact of computer use on earnings in a developing country: Evidence from ecuador. *Labour Economics*, *18*(4), 434–440.

Paunov, C., & Rollo, V. (2016). Has the internet fostered inclusive innovation in the developing world? *World Development*, *78*, 587–609.

Porto, I. (2016). Impacto del programa de alimentación escolar en el trabajo infantil: Una aproximación desde la toma de decisiones familiares.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Smith, J. A., & Todd, P. E. (2005). Does matching overcome lalonde's critique of nonexperimental estimators? *Journal of econometrics*, *125*(1), 305–353.

SOCPC. (2010). Standard occupational classification. *Bureau of Labor Statistics On behalf of the Standard Occupational Classification Policy Committee*. Retrieved from `https://www.bls.gov/soc/soc_structure_2010.pdf`

World-Bank. (2015). World bank indicators. *The World Bank Group*.

Table 1: Socio-demographic statistics by place and use Internet

	Not		User		Productive	
	users	everywhere	at work	at home	not users	users
Logarithm of income	13.276	13.632	13.839	13.836	13.588	13.699
Standardized income	-0.326	0.211	0.590	0.519	-0.152	0.301
	Panel A: Population					
Schooling	11.812	13.797	14.545	14.513	13.614	14.468
Age	35.757	33.073	34.871	35.178	31.004	35.074
Women	0.479	0.525	0.562	0.524	0.481	0.603
	Panel B: Technology experience					
Telephone at home	0.419	0.618	0.650	0.779	0.630	0.618
Satellite TV at home	0.653	0.808	0.837	0.907	0.814	0.799
PC at home	0.263	0.615	0.669	0.925	0.640	0.614
	Panel C: Proxy of wealth					
Property owner	0.415	0.423	0.440	0.478	0.407	0.463

Notes. The average of each variable is presented in the table. The samples is restricted to employed population between 18 to 65 years old. The standard deviation of the labor income is $262 and the mean is $295 for this sample. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor.
Source: Colombian Great Integrated Household Survey (GEIH) between 2009 to 2011

Table 2: Industrial statistics by place and use Internet

	Not		User		Productive	
	users	everywhere	at work	at home	not users	users
	Panel A: Firm size					
101 and over employee	0.223	0.777	0.694	0.669	0.532	0.468
51-100 employee	0.278	0.722	0.625	0.583	0.574	0.426
11-50 employee	0.285	0.715	0.607	0.558	0.600	0.400
6-10 employee	0.342	0.658	0.489	0.460	0.654	0.346
2-5 employee	0.448	0.552	0.288	0.316	0.725	0.275
	Panel B: Economic sector					
Financial service	0.132	0.868	0.840	0.799	0.409	0.591
Public administration	0.191	0.809	0.769	0.705	0.567	0.433
Education	0.210	0.790	0.705	0.711	0.445	0.555
Housing sector	0.255	0.745	0.676	0.608	0.564	0.436
Health	0.272	0.728	0.636	0.593	0.598	0.402
Electricity	0.311	0.689	0.610	0.552	0.611	0.389
Transport	0.321	0.679	0.576	0.503	0.647	0.353
Manufacturing	0.383	0.617	0.412	0.445	0.687	0.313
Commerce	0.397	0.603	0.392	0.393	0.698	0.302
Agriculture	0.422	0.578	0.394	0.382	0.708	0.292
Mining	0.429	0.571	0.408	0.481	0.663	0.337
Construction	0.441	0.559	0.360	0.353	0.718	0.282
Domestic Service	0.512	0.488	0.281	0.282	0.735	0.265
	Panel C: Occupation					
Professional specialty	0.101	0.899	0.872	0.861	0.347	0.653
Administrative support	0.114	0.886	0.864	0.835	0.438	0.562
Executive and managerial	0.150	0.850	0.811	0.744	0.497	0.503
Sales	0.360	0.640	0.399	0.433	0.698	0.302
Service	0.371	0.629	0.437	0.431	0.707	0.293
Machine operator	0.426	0.574	0.230	0.374	0.756	0.244
Precision production	0.487	0.513	0.174	0.290	0.775	0.225
Transportation	0.488	0.512	0.176	0.313	0.786	0.214
Cleaners and gardeners	0.499	0.501	0.208	0.267	0.750	0.250
Farming and mining	0.589	0.411	0.144	0.175	0.802	0.198
White-collar	0.118	0.882	0.853	0.824	0.415	0.585
Blue-collar	0.400	0.600	0.378	0.392	0.708	0.292
	Panel D: Type worker					
Salaried worker	0.288	0.712	0.592	0.565	0.591	0.409

Notes. The average of each variable is presented in the table. The samples is restricted to employed population between 18 to 65 years old. Economic sector is built using International Standard Industrial Classification of all Economic Activities (ISIC, 2008). Occupation is built using Standard Occupational Classification System (SOCPC, 2010). The standard deviation of the labor income is $262 and the mean is $295 for this sample. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor.

Source: Colombian Great Integrated Household Survey (GEIH) between 2009 to 2011

Table 3: Paremetric model: Ordinary least square

	Dep. Var: Standardized labor income				
	(1)	(2)	(3)	(4)	(5)
Internet use	0.169***	0.125***	0.116***	0.095***	0.069***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)
Controls					
Socio-demographic	✓	✓	✓	✓	✓
Firm size		✓	✓	✓	✓
Economic sector			✓	✓	✓
Occupation				✓	✓
Municipality and time FE					✓
R-squared	0.367	0.386	0.392	0.419	0.429
No. of observations	468,166	468,166	465,331	462,014	462,014

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. FE means municipality and time fixed effects. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor)

Table 4: Propensity score: Probit marginal effects model

Dep. Var:	Internet use	Internet at work	Internet at home	Productive users
	(1)	(2)	(3)	(4)
Schooling	0.058***	0.086***	0.086***	0.055***
	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.026***	-0.015***	-0.025***	-0.029***
	(0.000)	(0.000)	(0.000)	(0.000)
Age squared	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Women	0.028***	0.027***	0.064***	0.017***
	(0.000)	(0.000)	(0.000)	(0.000)
Telephone at home	0.065***	0.091***	0.171***	0.058***
	(0.000)	(0.000)	(0.000)	(0.000)
Satellite TV at home	0.059***	0.082***	0.129***	0.028***
	(0.000)	(0.000)	(0.000)	(0.000)
PC at home	0.191***	0.237***	0.624***	0.158***
	(0.000)	(0.000)	(0.000)	(0.000)
Property owner	-0.039***	-0.047***	-0.036***	-0.018***
	(0.000)	(0.000)	(0.000)	(0.000)
Agriculture	-0.002***	-0.019***	0.015***	-0.037***
	(0.000)	(0.001)	(0.001)	(0.000)
Mining	-0.002**	0.015***	0.159***	0.014***
	(0.001)	(0.002)	(0.002)	(0.001)
Manufacturing	-0.007***	-0.068***	0.017***	-0.019***
	(0.000)	(0.000)	(0.000)	(0.000)
Electricity	-0.058***	-0.077***	-0.029***	-0.034***
	(0.000)	(0.000)	(0.001)	(0.000)
Construction	-0.021***	-0.131***	0.011***	-0.055***
	(0.000)	(0.000)	(0.000)	(0.000)
Commerce	-0.021***	-0.084***	0.001***	-0.031***
	(0.000)	(0.000)	(0.000)	(0.000)
Transport	0.024***	0.074***	0.043***	-0.007***
	(0.000)	(0.000)	(0.000)	(0.000)
Financial service	0.043***	0.062***	0.084***	0.084***
	(0.000)	(0.000)	(0.000)	(0.000)
Housing sector	0.040***	0.078***	0.091***	0.030***
	(0.000)	(0.000)	(0.000)	(0.000)
2-5 employee	-0.087***	-0.197***	-0.133***	-0.062***
	(0.000)	(0.000)	(0.000)	(0.000)
6-10 employee	-0.051***	-0.089***	-0.091***	-0.047***
	(0.000)	(0.000)	(0.000)	(0.000)
11-50 employee	-0.031***	-0.031***	-0.049***	-0.033***
	(0.000)	(0.000)	(0.000)	(0.000)
51-100 employee	-0.032***	-0.032***	-0.040***	-0.012***
	(0.000)	(0.000)	(0.000)	(0.000)
Professional specialty	0.143***	0.316***	0.238***	0.179***
	(0.000)	(0.000)	(0.001)	(0.001)
Executive and managerial	0.192***	0.426***	0.290***	0.153***
	(0.000)	(0.000)	(0.001)	(0.001)
Administrative support	0.175***	0.368***	0.301***	0.153***
	(0.000)	(0.000)	(0.001)	(0.001)
Service	0.072***	0.202***	0.133***	0.008***
	(0.000)	(0.001)	(0.001)	(0.001)
Sales	0.101***	0.256***	0.174***	0.039***
	(0.000)	(0.000)	(0.001)	(0.001)
Machine operator	0.042***	0.103***	0.074***	-0.030***
	(0.001)	(0.001)	(0.001)	(0.001)
Equip. cleaners and laborers	0.036***	0.106***	0.058***	0.012***
	(0.000)	(0.001)	(0.001)	(0.001)
Precision production	0.026***	0.049***	0.033***	-0.027***
	(0.000)	(0.001)	(0.001)	(0.001)
Transportation	0.015***	0.019***	0.054***	-0.045***
	(0.000)	(0.001)	(0.001)	(0.001)
Municipality and time FE	✓	✓	✓	✓

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. FE means municipality and time fixed effects.

Table 5: Difference in means between treated and controls before and after matching selected variables for salaried workers

Variable	Unmatched (U) Matched (M)	Mean Treated	Mean Control	Difference
Schooling more than 12	U	0.603	0.190	0.412***
	M-NN	0.503	0.504	-0.001
	M-MD	0.503	0.503	0.000
Age	U	32.989	35.543	-2.554***
	M-NN	33.856	33.565	0.291***
	M-MD	33.847	33.871	-0.024
Age squared	U	1194.20	1374.30	-180.1***
	M-NN	1259.70	1242.90	16.80***
	M-MD	1259.10	1258.90	0.200
Woman	U	0.493	0.525	-6.800***
	M-NN	0.492	0.509	-0.017**
	M-ND	0.492	0.492	0.000
Telephone at home	U	0.603	0.399	0.205***
	M-NN	0.567	0.563	0.003***
	M-MD	0.566	0.566	0.000
Satellite TV at home	U	0.812	0.651	0.161***
	M-NN	0.792	0.787	0.006**
	M-MD	0.792	0.792	0.000
PC at home	U	0.608	0.239	0.369***
	M-NN	0.528	0.517	0.011***
	M-MD	0.529	0.528	0.001
Property owner	U	0.406	0.400	0.006**
	M-NN	0.409	0.409	-0.000
	M-MD	0.409	0.409	0.000
White-Collar	U	0.430	0.139	0.291***
	M-NN	0.337	0.347	-0.010***
	M-MD	0.338	0.337	0.000

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. The row M-NN shows the result of matching treated and control individual using the Nearest Neighborhood. The row M-MD presents the results of matching individuals using the Mahalanobis Distance.

Table 6: Non-parametric model: Propensity score matching for Internet use

	Dep. Var: Standardized labor income				
	(1)	(2)	(3)	(4)	(5)
Nearest Neighbor	0.166***	0.123***	0.115***	0.094***	0.068***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)
Mahalanobis Distance	0.165***	0.122***	0.114***	0.092***	0.065***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
Controls					
Socio-demographic	✓	✓	✓	✓	✓
Firm size		✓	✓	✓	✓
Economic Sector			✓	✓	✓
Occupation				✓	✓
Municipality and time FE					✓
R-squared	0.369	0.383	0.394	0.419	0.429
No. of observations	421,349	421,349	418,798	415,813	415,813

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. FE means municipality and time fixed effects

Table 7: Places and activities

	Dep. Var: Standardized labor income			
	Treatment: Internet			
	Use	at work	at home	Productive users
	(1)	(2)	(3)	(4)
OLS	0.069***	0.291***	0.183***	0.076***
	(0.005)	(0.006)	(0.006)	(0.005)
Matching Mahalanobis Distance	0.065***	0.284***	0.174***	0.071***
	(0.005)	(0.005)	(0.005)	(0.004)
All controls	✓	✓	✓	✓
R-squared	0.429	0.439	0.432	0.430
No. of observations	451,813	451,813	451,813	451,813

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. All the controls are socio-demographic. firm size, economic sector, occupation, municipality and time fixed effects. The control group for column 1, 2 and 3 is *not* Internet Use and for column 4 is Internet users in not productive activities. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor). The number of observation shows the individuals in the common support for the Propensity Score Matching, using Mahalanobis Distance. The OLS number of observation is 462,014. .

Table 8: Internet use per skill level

	Dep. Var: Standardized labor income									
	Treatment: Internet use									
	Farming and mining	Transportation	Precision production	Cleaners and laborers	Machine operator	Sales	Service	Administrative support	Executive administrative	Professional speciality
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OLS	0.207***	0.049***	0.060***	0.027***	0.074***	0.087***	0.101***	0.096***	0.156***	0.184***
	(0.058)	(0.012)	(0.011)	(0.007)	(0.037)	(0.011)	(0.009)	(0.010)	(0.031)	(0.020)
Matching Mahalanobis Distance	0.196***	0.048***	0.058***	0.023***	0.071***	0.084***	0.098***	0.091***	0.152***	0.178***
	(0.061)	(0.013)	(0.014)	(0.006)	(0.039)	(0.012)	(0.009)	(0.011)	(0.033)	(0.021)
All controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	
R-squared	0.472	0.288	0.285	0.247	0.348	0.241	0.397	0.351	0.336	0.361
No. of observations	3,888	33,010	27,618	63,671	3,029	72,478	45,160	33,937	64,225	68,796

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. All the controls are socio-demographic. firm size, economic sector, occupation, municipality and time fixed effects. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor). The number of observation shows the individuals in the common support for the Propensity Score Matching, using Mahalanobis Distance. The OLS number of observation is around 10% more of the mentioned. Occupation is built using Standard Occupational Classification System (SOCPC, 2010).

A Appendix

Table A-1: Differences in mean by type of worker

Variables	Salaried worker	Self-employed worker	Difference
Logarithm of labor income	13.526	13.111	0.416***
	(0.0013)	(0.0024)	(0.005)
Age	33.385	37.010	-3.625***
	(0.0188)	(0.0243)	(0.054)
Age squared	1,217.343	1,488.661	-271.318***
	(1.392)	(1.900)	(4.112)
Schooling	13.035	12.594	0.442***
	(0.0047)	(0.0053)	(0.013)
12 or more years of education	0.512	0.412	0.100***
	(0.0009)	(0.0010)	(0.003)
Telephone at home	0.622	0.556	0.066***
	(0.0009)	(0.0011)	(0.003)
PC at home	0.770	0.717	0.053***
	(0.0007)	(0.0010)	(0.002)
Satellite TV at home	0.519	0.473	0.047***
	(0.0009)	(0.0011)	(0.003)
House Owner	0.399	0.427	-0.028***
	(0.0009)	(0.0011)	(0.003)
2-5 employee	0.192	0.823	-0.631***
	(0.0007)	(0.0008)	(0.002)
6-10 employee	0.074	0.033	0.040***
	(0.0004)	(0.0003)	(0.001)
11-50 employee	0.159	0.040	0.119***
	(0.0006)	(0.0004)	(0.001)
51-100 employee	0.056	0.009	0.047***
	(0.0004)	(0.0002)	(0.001)
101 and over employee	0.520	0.095	0.425***
	(0.0009)	(0.0006)	(0.002)
Agriculture	0.016	0.021	-0.006***
	(0.0002)	(0.0003)	(0.001)
Mining	0.001	0.001	0.000
	(0.0000)	(0.0000)	(0.000)
Manufacturing	0.171	0.120	0.051***
	(0.0007)	(0.0007)	(0.002)
Electricity	0.013	0.002	0.010***
	(0.0002)	(0.0001)	(0.000)
Construction	0.033	0.048	-0.015***
	(0.0003)	(0.0004)	(0.001)
Commerce	0.235	0.331	-0.095***
	(0.0007)	(0.0010)	(0.002)
Transport	0.077	0.130	-0.053***
	(0.0004)	(0.0007)	(0.002)
Financial service	0.038	0.007	0.032***
	(0.0003)	(0.0001)	(0.001)
Housing sector	0.084	0.111	-0.028***
	(0.0005)	(0.0007)	(0.002)
Public administration	0.078	0.031	0.047***
	(0.0005)	(0.0003)	(0.001)
Education	0.111	0.038	0.073***
	(0.0005)	(0.0004)	(0.001)
Health	0.073	0.069	0.005***
	(0.0004)	(0.0005)	(0.001)
Domestic Service	0.069	0.090	-0.021***
	(0.0004)	(0.0006)	(0.001)
Professional specialty	0.169	0.144	0.026***
	(0.0007)	(0.0007)	(0.002)
Executive and managerial	0.219	0.070	0.149***
	(0.0007)	(0.0005)	(0.002)
Administrative support	0.093	0.080	0.013***
	(0.0005)	(0.0006)	(0.001)
Service	0.113	0.089	0.024***
	(0.0005)	(0.0006)	(0.002)
Sales	0.112	0.249	-0.137***
	(0.0005)	(0.0009)	(0.002)
Machine operator	0.005	0.014	-0.009***
	(0.0001)	(0.0002)	(0.001)
Equip. cleaners and laborers	0.151	0.142	0.008***
	(0.0006)	(0.0007)	(0.002)
Precision production	0.060	0.092	-0.032***
	(0.0004)	(0.0006)	(0.001)
Transportation	0.067	0.099	-0.032***
	(0.0004)	(0.0006)	(0.001)
Farming and mining	0.011	0.022	-0.011***
	(0.0001)	(0.0003)	(0.001)
White-collar	0.398	0.266	0.132***
	(0.0009)	(0.0009)	(0.002)

Notes. Standard error in brackets. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level

Table A-2: Internet use in time

Dep. Var: Standardized labor income				
	Treatment: Internet use			
	2009-2011	2009	2010	2011
	(1)	(2)	(3)	(4)
OLS	0.069***	0.080***	0.072***	0.065***
	(0.005)	(0.008)	(0.008)	(0.008)
Matching Mahalanobis Distance	0.065***	0.078***	0.069***	0.061***
	(0.004)	(0.004)	(0.004)	(0.004)
All controls	✓	✓	✓	✓
R-squared	0.429	0.425	0.443	0.431
No. of observations	451,813	130618	138,221	146,966

Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. All the controls are socio-demographic. firm size, economic sector, occupation, municipality and time fixed effects. The number of observation shows the individuals in the common support for the Propensity Score Matching, using Mahalanobis Distance. The OLS number of observation is 462,014 in column 1, 145,212 in column 2, 153,506 in column 3, 163,296 in column 4.

Table A-3: Places, activities and intensity

Dep. Var: Standardized labor income				
	Treatment: Internet daily			
	Use	at work	at home	Productive users
	(1)	(2)	(3)	(4)
OLS	0.081***	0.295***	0.186***	0.092***
	(0.009)	(0.006)	(0.006)	(0.005)
Matching Mahalanobis Distance	0.077***	0.289***	0.178***	0.080***
	(0.007)	(0.006)	(0.005)	(0.004)
All controls	✓	✓	✓	✓
R-squared	0.429	0.439	0.432	0.430
No. of observations	451,813	451,813	451,813	451,813

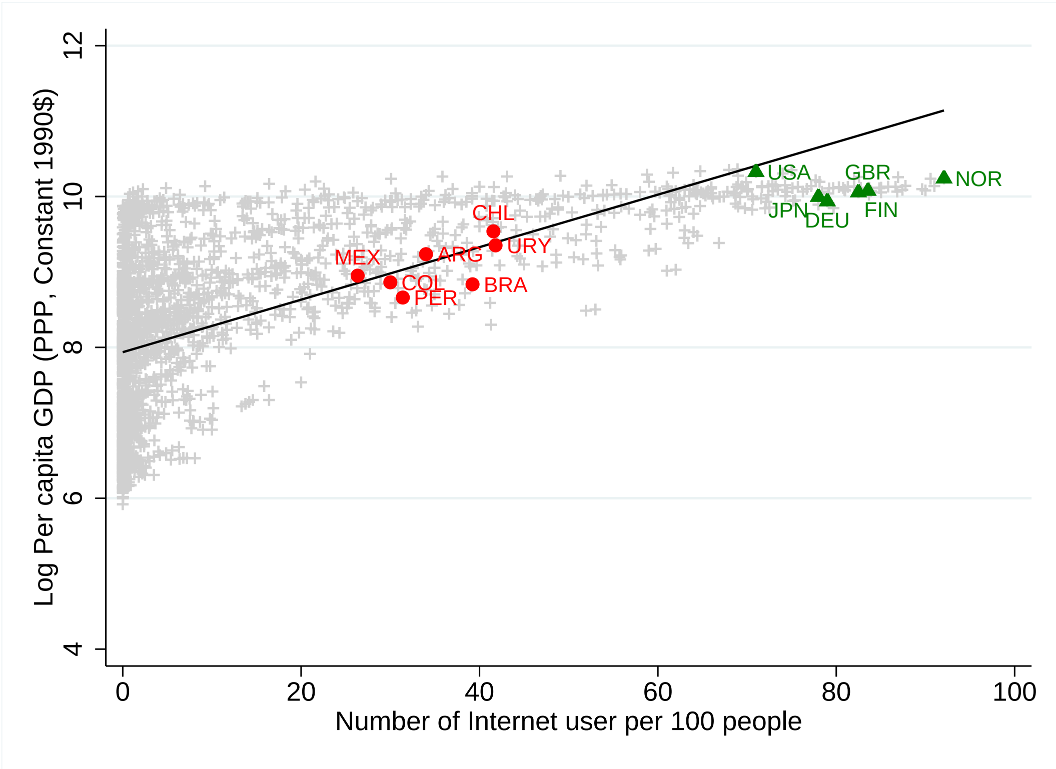
Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. All the controls are socio-demographic. firm size, economic sector, occupation, municipality and time fixed effects. The control group for column 1, 2 and 3 is *not* Internet Use and for column 4 is Internet users in not productive activities. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor). The number of observation shows the individuals in the common support for the Propensity Score Matching, using Mahalanobis Distance. The OLS number of observation is 462,014. .

Table A-4: Places and activities per worker type

Dep. Var: Standardized labor income				
	Treatment: Internet			
	Use	at work	at home	Productive users
	(1)	(2)	(3)	(4)
Panel A: Salaried Worker				
OLS	0.054***	0.206***	0.159***	0.058***
	(0.005)	(0.006)	(0.007)	(0.006)
Matching Mahalanobis distance	0.054***	0.192***	0.145***	0.054***
	(0.004)	(0.005)	(0.003)	(0.004)
All controls	✓	✓	✓	✓
R-squared	0.495	0.501	0.498	0.495
No. of observations	246,359	246,359	246,359	246,359
Panel B: Self-employemnt worker				
OLS	0.098***	0.529***	0.243***	0.112***
	(0.009)	(0.017)	(0.013)	(0.011)
Matching Mahalanobis distance	0.092***	0.511***	0.229***	0.097***
	(0.008)	(0.010)	(0.009)	(0.009)
All controls	✓	✓	✓	✓
R-squared	0.368	0.387	0.372	0.368
No. of observations	169,462	169,462	169,462	169,462

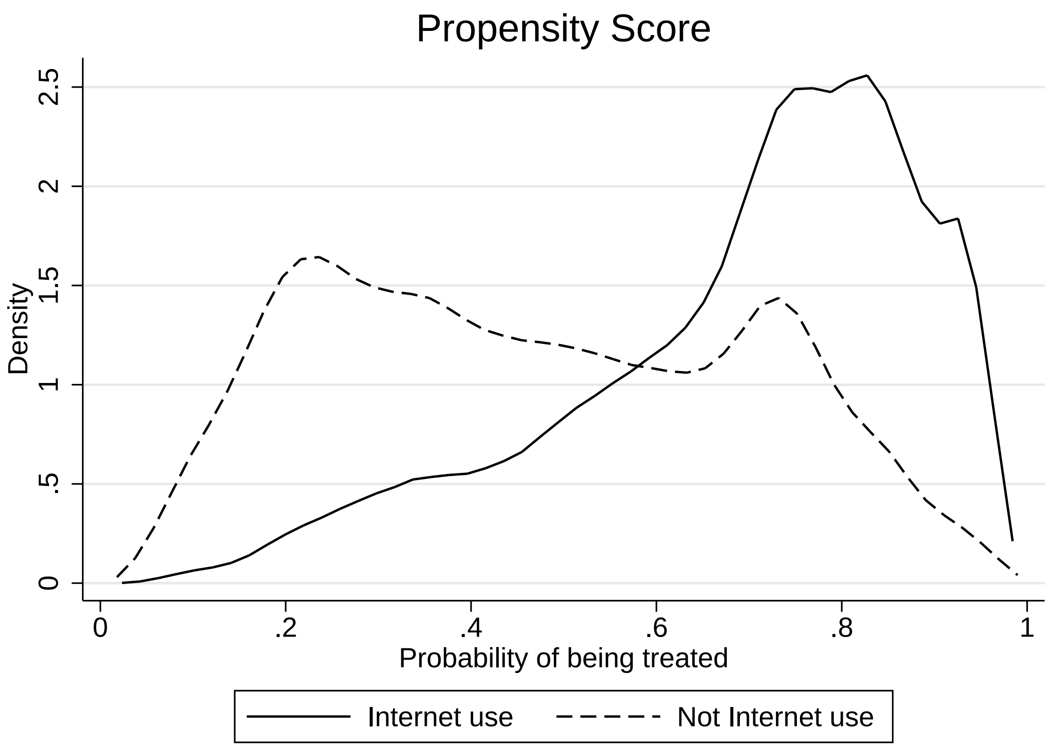
Notes. Robust standard errors are shown in parentheses. *** is significant at the 1% level, ** is significant at the 5% level, * is significant at the 10% level. All the controls are socio-demographic. firm size, economic sector, occupation, municipality and time fixed effects. The control group for column 1, 2 and 3 is *not* Internet Use and for column 4 is Internet users in not productive activities. Internet users are workers accessing from any of the following places: at home, at work, in educational institutions, in free public access centers, in payed access centers, in the house of another person (relative, friend, neighbor). The number of observation shows the individuals in the common support for the Propensity Score Matching, using Mahalanobis Distance. The OLS number of observation is 273,733 for salaried workers and 188,281 for self-employed workers. .

Figure A-1: Internet and Income in the World



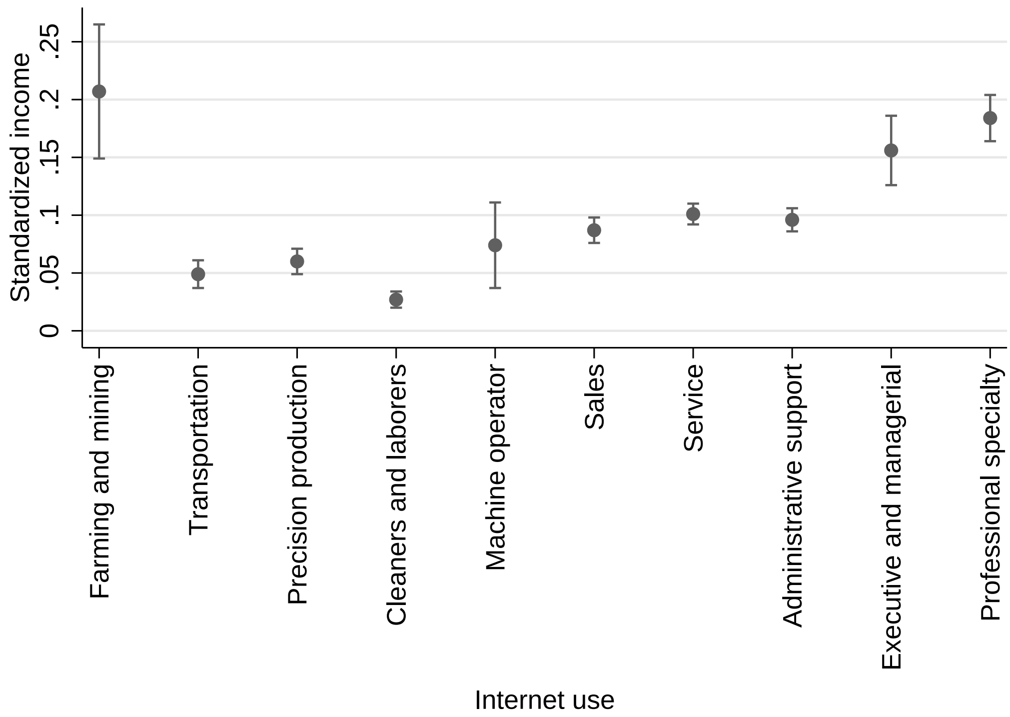
Note: Own calculation using indicators from the World World-Bank (2015)

Figure A-2: Common support between treated and control group



Note: Propensity score built using equation 2 and the indicators from the World-Bank (2015).

Figure A-3: Correlation between Internet use and standardized income



Note: The figure shows the estimators and interval coefficients using Mahalanobis distance and including all the set of covariates from table 1 and 2, for 10 sub-samples depending the occupation. Each category is built using SOCPC (2010).