RETURNS TO COMPUTER USAGE:
THE CASE OF UKRAINE

by

Natalia Piskounova

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Arts in Economics

National University "Kyiv-Mohyla Academy"
Economics Education and Research Consortium
Master's Program in Economics

2007

Approved by Mr. Serhiy Korablin (Head of the State Examination Committee)

____

____

____

Program Authorized
to Offer Degree Master's Program in Economics, NaUKMA

Date ____

National University "Kyiv-Mohyla Academy"

Abstract

RETURNS TO COMPUTER
USAGE: THE CASE OF UKRAINE

by Natalia Piskounova

Head of the State Examination Committee: Mr. Serhiy Korablin,
Economist, National Bank of Ukraine

During the last two decades two tendencies have been observed: the increase of wage dispersion among high and low skilled labor and vast computerization among most industries. This paper analyses possible a wage premium associated with computer usage during 2003 in Ukraine based on ULMS dataset. Usual OLS procedure is used. As well as more advanced techniques are implemented: heckman selection model to remove the selectivity bias and instrumental variables to remove endogeneity problem. Estimates suggest that there exists a positive return on computer usage in OLS and heckman procedure and the returns are not statistically different from each other. Instruments used don't give appropriate results because of their weakness. Hence, there exists a wage premium associated with computer usage which is equal to 15.8% for men and 19.1%.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, professor Mark E. Schaffer, for all the guidance, invaluable help and comments during my work on thesis, professor Tom Coupe for all useful comments, professor Olena Nizalova for helping me in choosing the topic. Also, I would like to thank my mom and dad for all their love and support during all my life, my friends in Minsk and Kiev for sharing my good and bad times. Last but not least, I would like to thank Lado for all his care and patience with me.

GLOSSARY

Returns to computer usage – annual yield for the worker for using computer in comparison with the one which does not use it.

Returns to Internet usage - annual yield for the worker for using Internet in comparison with the one which does not use it.

Chapter 1

INTRODUCTION

"A pair of strong hands are not what they used to be. Now those hands have to be able to use a keyboard."

Senator Bill Bradley

October 7, 1999

Computers and information technologies now play a crucial role in the labor market when workers are searching for a job because possessing a computer literacy is one of the main requirements demanded by employers. The knowledge of using word, spreadsheets, electronic mail and databases is asserted to increase worker's productivity. It also gives a chance "to access instant information, and reach new markets" [Dolton and Makepeace 2002]. Consequently, in order to survive in modern and highly competitive labor market, the workers have to possess at least basic computer skills.

Wage differentials have been growing in last couple of decades all around the world. There has been proposed several reasons for that. It is important to distinguish them for developed countries and for developing ones. For countries like the USA, Germany, the UK, France, Canada etc the causes might lie in rising of globalization pressures [Wood 1994, 1998] and in greater pace of skill-biased technological change [Bound and Johnson 1992]. The third explanation relates to the fact that during the period of 1970s and 1980s, there has been observed a "substantial slowdown in the rate of expansion of the relative supply of more-educated workers" [Autor, Katz and Krueger 1998].

Considering ones with transition economy like post-soviet countries, it is worth mentioning that during the soviet times almost no wage inequalities could be observed due to severe regulations from above [Arabsheibani, Mussurov, 2006]. But as soon as the transition period started, the wage differentials started to rise very rapidly due to observed heterogeneity of workers which reflected the differences in their education, experience, knowledge. Putting it differently, differences in human capital.

There have been a lot of discussions in the last decade whether such factor as "Computer Usage at Work" really influences the worker's earnings. Some investigations in this field were done. Krueger in his pioneering work (1993) asserts that this factor raises the earnings by as much as 15% in the US labor market. The support for such conclusions can be found in the further investigation made by Arabsheibani and Marin (2000) who found that there is a return to computer usage of 19.1%, which is in agreement with previous researches made by the same authors for the UK labor market that state that return was a little bit higher than 20% in 1985 and 1990 though it was based on the different sample and other explanatory variables. Another empirical study for the UK labor market was made by Dolton and Makepeace (2002) who used instrumental variable method, the treatment effects model with selection and again they found a positive and significant returns to computer usage of at least 10%, but this returns might be a little less for women. Another supporting view is presented by the Daldy and Gibson (2002) who investigated the New Zealand labor market. They found that there exist a positive and significant relationship between computer-related training and earnings of 12.1%. Oosterbeek (1996) finds for the Netherlands that using computer at work, brings a wage premium of 12,3%. Dostie, Jayaraman and Trepanier (2006) using North American data and using mixed effects model in order to clear away the worker heterogeneity and workplace heterogeneity state that returns to computer anyway are almost 4%.

The main counter argument to these investigations was made by DiNardo and Pischke (1997) who argued that this large number for the returns on computer usage only reflects the omitted variable bias. Krashinsky (2005) also finds that using more advanced econometric techniques like GLS, fixed effects estimation but not simple OLS regression which was used by this author in 2000 makes these returns insignificant, consequently he asserts that this returns only reflect the ability bias. Franzen (2001) analyzing the labor market in Switzerland finds no supporting evidence for the Krueger's statements.

What concerns the empirical study for the transition countries, there has been made one by Kuku, Orazem and Singh (2006) who investigated 9 countries in transition. They found that there is a 25% premium for computer usage at work but if they include the correction mechanism with instruments this returns to computer disappear.

As no research has been done for transition country alone, only a combined research for 9 post-soviet countries including Ukraine, I will investigate such a relationship for the Ukrainian labor market alone but not in combination with other countries. It is important to have such an investigation in order to see whether a wage differential for computer literacy exists here and whether it is worth spending time on obtaining such skills. In my work I will use the ULMS dataset.

Also, this research will be a useful one because if the returns on computer usage will be positive and significant, it may be a signal for the government to develop telephone, cable and satellite services because people who are equipped with necessary skills will learn more, this will in turn lead to increased total welfare of the economy as companies will be more productive in terms of advanced technologies, workers will be paid higher wages, hence consumption of the economy will increase. Also with this respect, the adequate legislation needs to be elaborated in order to attract more investments into Ukraine. This in turn may help to develop infrastructure of home IT market, and help firms to penetrate into this market with less costs and difficulties, and thus, this policy will lead to more developed IT infrastructure, which in turn will lead to a more development in general.

This work is organizes as follows. In section 2 the previous works will be examined, in section 3 the methodology is given, section 4 will present the data description, section 5 – empirical results of this investigation. And, finally, in section 6 the conclusion will follow.

Chapter 2

LITERATURE REVIEW

Many researches argue that plenty of changes in the wage premium occurred in the past few decades in the developed countries. One of the main reasons named is the returns to education. For example, in the United States college graduates wage premium in comparison with high school graduates increased from 34 percent in 1979 to 56 percent in 1991 (Mishel and Berstein 1992, Table B1).

Another important explanation is related to the development of informational technologies and its affect on the wage premium of more technically skilled workers in comparison with lower-skilled labor.

The general question which must be raised before mentioning all supporting and contradicting views concerns the perspective of computerization by Borghans and Weel (2000). They highlight that it is not that reasonable to put up huge sums of money into computer skills but "to concentrate on policies and regulations relating to the infrastructure for information and communication technologies (ICT) (the digital highway), the availability and development of software, et cetera" [Borghans and Weel 2000]. Their reason for this lies in the fact that some obstacles can be met in the growth potential due to "underdeveloped or too expensive cable network, or a too protected software market"[Borghans and Weel 2000]. Also, it is very important to follow the tendency of the labor market in order to understand in what direction it's better to conduct the research and policy-making.

In this literature review I will examine the pioneering work by Krueger (1993) done in this field which predicts positive and significant returns on computer usage, the supporting views for positive returns on computer usage as well as the investigations which indicate the conflicting inferences for developed countries. Then, mention the analysis made in countries with transition economies. Also, the separate investigation for women with respect to computer returns might be considered.

The pioneering work written about the influence of computer usage on wage premium is of Alan Krueger's. He determines the computer usage as probability of using programming, word processing, e-mail, computer-aided design, etc at work as a source of increase in worker productivity and hence increase in the wage premium. He asserts that there are some groups of people who have higher probability of computer usage at work than others. The former includes women, Caucasians, and highly educated people. The latter group consists of men, African Americans, and less-educated workers. On basis of CPS data he runs the simple OLS regression model:

$$lnW_i = X_i\beta + C_i\alpha + \epsilon_i$$

where X_i - vector of observed characteristics;

C_i - a dummy variable that equals to one if the *i*th individual uses a computer at work, and zero otherwise;

W_i - wage rate of *i*th individual;

$\alpha \text{ and } \beta$ - are parameters to be estimated.

The possible bias can occur because C_i can be correlated to ϵ_i. The author tries to deal with it using four different specifications.

First of all, he estimates the equation, where the explanatory variables are only computer usage at work and at home. Here he asserts, that variable "Computer use at home" can explain some unobserved heterogeneity. The results are the following: if an individual used computer only at work, he earned 18% more than otherwise. If an individual used computer only at home, he earned 7% more than otherwise, and, finally, if an individual, used computer at work and at home, he earned only 9% more, then otherwise. The results are very similar for 1984 and for 1989.

The second approach – estimates for narrow occupations. He takes a group of secretaries. The results show that the wage bonus for secretaries in 1984 is 6% and in 1989 – 9%. Both are significant. If to consider secretaries with high school education, these numbers are the following: 9.2% in 1984 and 8.6% in 1989.

The third approach – estimates based on the High School and Beyond Survey. The interesting finding can be traced here: the higher were the grades of the worker, the higher the probability that he will use computer at workplace.

The last approach – occupational level. Here the author, using WLS, estimates the relationship between the growth of wage premium and the growth of computer usage at the occupational level. The conclusion here is: if the worker changed his job from the one with no computer to the one with computer, his wage bonus will be 10.5%.

Then in the analysis, he estimates two specifications, one – where he includes only the dummy variable, and the other – where he includes such characteristics as education, potential experience and its square, gender, and the union status. In the former case the author found that those who use computer at work earn 31.8% in 1984 and 38.4% in 1989 higher than those who do not. In the latter case, this difference fell to 18.5% in 1984 and 20.6% in the 1989. Both coefficients are significant at conventional significance levels.

Also, Krueger runs a regression where he uses dummies for specific computer tasks and he discloses that using an e-mail creates a positive wage premium of 14.5% while playing games – negative wage premium of -10.9%.

An interesting inference was made by Krueger in the sense that around 40% of the increase in the returns to education is due to the use of computer at work.

But according to the assumption that more high-skilled workers will enter the labor force in the future and the market will not need it in the same huge amount as in the past, it is reasonable to assume that the wage premium for computer use will fall with time. This might not be the case for Ukraine, which is a transition country. Because to develop and to grow, it has to attract more well-educated workers for a long period of time. Hence, it can be said that the market for computer-educated workers is far from being satiated. Hence, the demand for such workers will grow in the future.

Let's start with investigations that support Krueger's findings.

Concerning the developed countries, it is necessary to mention the analysis made for the North American labor market by Dostie, Jayaraman and Trepanier (2006). The dataset which was utilized is Workplace and Employer Survey dated 1999-2002 conducted by Statistics Canada. This data is beneficial in the sense that workers and workplaces are observed during the whole period of investigation. Also, it contains a huge bunch of other characteristics which help to avoid the most vital problem in this field – the "ability" bias. Using pooled OLS regression, the authors find a very similar result to returns to computer use in France made by Entorf, Gollac and Kramarz in 1999 – 9.7%. Though it doesn't control for the worker unobserved heterogeneity.

Then utilizing the mixed effects model, which allows to control for worker heterogeneity and workplace heterogeneity separately as well as simultaneously. The mixed effects model represents the mixture of random and fixed effects. In such model it is necessary to distinguish factors that are fixed and that are random. The former are factors in which levels are chosen in such a way that their levels can be compared to one another. The latter are those that are randomly chosen from the population in order to study the variation in that population. And still there is a wage bonus associated with computer usage at work though it is almost two times less – 4.6% in case of controlling for worker unobserved heterogeneity, if controlling for workplace heterogeneity, the result is a little bit higher – 6.1%. If we control for both – still the wage premium exists and amounts to 3.8%. Though the opponents to these results suggest that computer industries are willing to employ more high-skilled labor and actually they do so.

The UK labor market was investigated a lot of times and the supporting views were obtained in the field whether we have a wage premium associated with the use of computer or the presence of necessary computer skills. Arabsheibani and Marin (2000) found that computer use coefficient is positive and highly significant and amounts to 19.1%. This is very similar to their previous study, where they found that this estimate amounted to around 20% in 1985 and 1990 in the UK (in spite the fact that other dataset was used). Also they augmented their base regression by including the variable, which indicates whether the user became better in using computer (increase his/her productivity). The analysis shows that the coefficient is positive and highly significant. This indicates that if the worker acquires more computer skills his productivity increases, which in turn give a rise to the wage premium.

According to Bell (1996), who used panel data (which is usually used to control for unobserved individual heterogeneity) from the British National Child Development Survey Study, 1981-1991, the wage premium is 11-13% higher for those who make use of their computer skills and there is almost no indication that such a correlation might be explained by unobserved characteristics.

Also, for the UK labor market Dolton and Makepeace (2002) found that the wage premium for computer use is as high as 10% but it can be a little less for women. The utilized data: the National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS70) are of a great value due to fact that the interested variable "computer usage" is traced over a period of time when information technologies got its highest development. Another positive peculiarity of these data is that age of worker is the same, thus it helps to avoid the problems with compositional effects. In investigating the impact of computer usage at work as well as at home, these authors included five different categories of control variables: human capital, employment, employer, socio-demographic and other variables. Using the specification procedure proposed by Krueger, the following results were obtained: the returns to computer use are 18-20%. The next used specification was augmentation of previous specification by including a term which represents unobserved heterogeneity of worker. The returns to computer usage fall to 13-16%. The next specification – inclusion of the dummy variable for using computer at home. This almost didn't change the results obtained in the first specification. Next they tried to figure out whether such variable as "future computer usage" could be a proxy for unobserved heterogeneity, they explain it by the fact that using computer in the future is an indicator whether a worker is able and willing to learn anew. The returns are 8% from men and 2% for women. Instrumental variables approach and treatment effects model are also implemented in this research. The endogenous variable is "computer use determination".

The model in this case can be estimated by two methods: maximum likelihood estimation and Heckman Two Step method. In this case the necessary thing to be is that Z is not correlated with the error term in (1). In order to proceed with this estimation, the authors checked whether the "computer usage at work" is endogenous and whether the instruments are valid and qualitative. The Hausman *t*-statistic is 6.12 (for NCDS) and 4.57 (BSC), hence the hypothesis of "computer usage" endogeneity is confirmed. The quality of instruments is checked by joint significance of instrumental variable by chi-square test. The results are 1261.55 (for BSC) and 1707.8 (for NCDS). Hence, the quality of instruments is confirmed. Sargan test is used to check for validity of instruments. Uncentered R^2 must exceed the critical value in order to be able to assert that instruments are not valid. The values are 0.0002 (for BCS) and 0.0022 (for NCDS), which implies that instrumental variables are valid. The returns to computer usage vary from 17.9% from women and 34.9% for men. Using random effects and fixed effects estimation gives the following results: the former methods suggests that the wage premium from men is 17% and for women – 13%, the latter approach lowers the returns to 6.3% for men and 2.1% for women. Consequently, it can be stated that using different econometric tools in order to investigate the relationship between the "computer usage at work and earnings", the authors came to conclusion that the wage bonus associated with using information technologies increases both for men and for women in the UK labor market.

Daldy and Gibson (2002) investigate New Zealand labor market using Education and Training Survey (ETS). Their research is focused on the training programs, they used eight subjects for training among which there was "Computing". From analysis it can be inferred that computer training created 12.1% increase in the wage premium, which is rather similar to what was found in other countries with the only distinction that instead of training program, "computer use at work" dummy was used.

They extended the research by choosing the sample of more skilled (educated) workers in order to put away the sample selection bias. They divided all the training programs into two groups: "Computing" and "Other Programs". Using panel data they got the same findings as before: the wage premium for the first group is higher than in the other (11.6% versus 2.7%).

Bruinshoofd and Weel (1998) investigate the Netherlands labor market, and in particularly how changes in the computer technologies relate to the fact that there exists a change toward the high-skilled labor force during the 90's. They came to conclusion that high-skilled labor can take the benefits from technological changes while low-skilled labor is badly hurt through such changes. They got their results using cross-sectional data from OECD's STAN Database. Using simple OLS procedure, they assert that in 1992 there existed a large gender wage inequalities (33.1% more for men), though such wage differential disappeared almost entirely in 1996 (2%). Comparing the wage inequalities between white-collar and blue-collar workers, they find that in 1992 wage premiums were 1.1% for white-collar, and in 1996 – 2.7%, but in both years they are insignificant. Hence, they find no evidence to support a wage differential between white-collar and blue-collar workers. Also, they investigate the wage premiums in different sectors of economy, and determine that white-collar workers in R&D sector in 1992 had a wage bonus of 6.8%, but it is insignificant, while this index doubled in 1996 – 14% and became highly significant. As a conclusive statement here, it can be stressed that in 1992 in Netherlands there was no wage premium associated with technological changes though the situation changed completely in 1996, and high-skilled labor and white-collar workers face a significant increase in their wages.

Now let's proceed with investigations that are in conflict with Krueger's findings.

In 1997 DiNardo and Pischke presented their counter-evidence on the Krueger's paper. They used three large cross-sectional datasets for German workers and tried to evaluate the returns to the calculators, telephones, pens or pencils on the wage premium in order to shed the doubt on Krueger's conclusions. They argue that this positive wage premium might be associated with omitted variables bias or worker's unobserved heterogeneity. The author criticizes that while Kruger uses only cross-sectional data, he is not able in his research to control for individual fixed effects. First, it was found that wage premiums are similar for two countries. For Germany the wage premium associated with computer use amount to 0.139 in 1979, 0.239 in 1985-1986, 0.288 in 1991-1992.

As a criticism DiNardo and Pischke run a regression where they included variables as calculators (the use of which increased a lot recently), telephones, pens or pencils and whether the employee is sitting during his work. It was found that sitting on the job and use of telephone creates a wage premium of 5-7%.

Krueger (1999) partially agrees to the results found by DiNardo and Pischke (1997) that "a causal interpretation of the effect of possessing computer skills on pay is based on circumstantial evidence, and has not been proven beyond a shadow of doubt" [Krueger 1999]. Though he criticizes their arguments in a way that, it is more likely that exactly the computer literacy but not the usage of pens or pencils determines the cognitive ability.

Another evidence for German labor market that supports DiNardo and Pishke's views was made by Anger and Schwarze (2002). They relate the wage premium not to the development of new information technologies but to the worker's abilities. The longitudinal micro-database German Socio-Economic Panel Study (GSOEP) dated 1985-1999 has a lot of useful information related to that kind of study. The regressions are run separately for men and women in order to see whether wage premiums are different due to gender gap. 14 time periods, 14 industries, standard set of control variables such as education, experience, age, marital status, computer usage are included. Three different models are investigated: simple cross-sectional study, panel data to run pooled regression using fixed and random effects in order to control for heterogeneity bias and regression which includes the dummy variable for future computer usage at work. The results for men indicate that under simple cross-sectional study, the returns to computer use are as much as 7%, but they are much smaller in pooled OLS – 2.5%, in random effects which controls for heterogeneity bias – 1.1%, and in fixed effects regression they are negative but statistically insignificant. These results relate to the first two models mentioned above. The third model is estimated by random effects procedure because dummy variable for future computer usage at work is constant over time. The results show that men and women will receive a wage bonus if they will use computer at work in their future careers but not at the present moment. For men the wage premium is about 3% in case of pooled OLS regression and 5% in case of using the random effects procedure. For women this estimate is 4% in both cases. Consequently, the authors came to the following conclusion that the wage premium is not associated with the computer usage at work but relates to the worker's abilities, and the future computer usage may be used as a dummy for worker's abilities.

Another contradiction to Krueger's views comes from the Switzerland labor market investigated by Franzen (2001). The analysis is based on two cross-sectional datasets which were conducted for Switzerland in 1998. The first one is Swiss Labor Market Survey (SAMS). The second one – Swiss Labor Force Survey (SAKE). The distinction feature of this research in comparison with other investigations consists of the fact that dummy variable indicating "on-the-job use of Internet" was also included in the regression. Using SAMS data set, the author runs simple OLS regression which includes 9 occupational dummies of the International Standard Classification of Occupations (ISCO-88) and 13 dummies controlling for the workplace sector. The results in this case show that the wage premium associated with computer usage falls to 1% and becomes insignificant. Later considering the SAKE dataset it is found that there exists 8% wage premium for those workers who use computer at their workplaces. And this result is highly significant. But the usage of Internet in that particular case has a negative effect on earnings. Then including the dummy variables controlling for computer use at home as well as at work and dummies of the International Standard Classification of Occupations (ISCO88), decreases the wage premium from 8% to 3%. The weakness of this dataset is that it cannot give the researcher a chance to control for workers and workplaces unobserved heterogeneity. Consequently, the author asserts that in case if controlling for this heterogeneity the wage premium is likely to disappear or to become insignificant. Hence, workers in Switzerland do not observe higher wages which might be related to computer usage at work and at home.

The conflicting inference with the Krueger's work can be found in the researches done by Krashinsky in different years, 2000 and 2005.

In the research (2000) he uses the data on identical twins in order to find out whether the computer usage influences the wage premium.

Equations for identical twins are:

$$y_{1j} = \beta'_{1j} X_{1j} + \alpha' Z_j + A_j + \varepsilon_{1j}$$
$$y_{2j} = \beta'_{2j} X_{2j} + \alpha' Z_j + A_j + \varepsilon_{2j}$$

where X_{ij} - vector of individual characteristics for twin *i* from family *j*;

Z_j - common characteristics for family *j*;

A_j – family-specific ability term;

ε_{ij} - an individual-specific error term.

He runs simple OLS regression, fixed effects model and correlated random-effect models: with and without a measurement-error correction. The correlated random effects model comes from the assumption that "the returns to individual characteristics X_{ij} are the same for both twins, and the ability is correlated between twins" [Krashinsky 2000]. Hence, $A_j = \frac{\gamma(X_{1j}+X_{2j})}{2} + \vartheta_j$. Consequently, we can write down the reduced-form correlated random effects (Chamberlain 1982):

$$Y_{1j} = \beta X_{1j} + \alpha Z_j + \frac{\gamma(X_{1j} + X_{2j})}{2} + \vartheta_j + \varepsilon_{1j}$$
$$Y_{2j} = \beta X_{2j} + \alpha Z_j + \frac{\gamma(X_{1j} + X_{2j})}{2} + \vartheta_j + \varepsilon_{2j}$$

Where γ - the correlation between a family's ability level and each twin's individual characteristics.

The conclusion was made that the presence of computer doesn't create the wage premium at all, only that more computer educated labor force usually finds and gets a job where computer skills are a requirement.

Then in his research of 2005, he uses data from the CPS, NLSY and identical twins. It is beneficial to mention the main advantages and disadvantages of these data sets because non of them constitutes an ideal sample and some problems can be due to drawbacks in the data itself.

The advantage of CPS is that we can trace the wage of individual before he got married and after thus find out whether the wage differs in two states (not married and married). The disadvantage of CPS is that it is not possible to trace the individual in case if he moves out of his house to other place.

The advantage of NLSY lies in the fact that it is a long panel data set that has information about young people of 14-22 in 1979 and that it is optimal for making research on individuals who just got married. Also in NLSY one can find Armed Forces Qualifying Test (AFQT), which can be used as a source of the ability for each individual.

The data on twins analysis is very much alike the CPS, with the only slight differences in wages and education. This data set can possible suffer the measurement error, which in turn alleviates the estimator's coefficients. Consequently, this measurement error should be accounted for.

He used the generalized least squares procedure, correlated random effects and fixed effects: both with and without the measurement-error correction.

Without correction for measurement error, under GLS, the parameter "Computer at work" gives a wage premium of 20.5% and is significant, while corrected random effect and fixed-effect states that this wage premium becomes insignificant.

Even when correcting for measurement error, the results are robust.

The contradictory results with Bruinshoofd and Weel (1998) were found by Osterbeek (1996) for the Netherlands labor market shows no supporting evidence for Krueger's conclusions. This research is unique in the sense that it has some peculiarities in comparison with other countries and studies. They are the following: the author utilizes the first dataset related to this particular topic, the fixed effects estimation procedure is employed and such variable as "the intensity of computer usage" is used. In order to examine this issue, the author uses two distinct methods which measure "computer usage". The first one is dichotomous index, where all the respondents are divided into two groups: users and non-users, the former consists of people who use computer daily, weekly, monthly and almost never, and the latter includes people who do not use it at all under all the circumstances. The second technique is to include five different dummies, each reflecting the groups mentioned earlier. The conclusions of the author are the following: the technological change has nothing to do with the growing wage differentials among workers. And that such factor as "the intensity of computer usage" does not influence the wage premium in case the worker uses computer at work and it doesn't matter whether he uses on the daily, monthly or other basis.

Now it's worth considering what returns to computer usage are observed in the countries with transition economy.

The recent research made in this field is of Kuku, Orazem and Singh (2006) who used data from nine countries of Eastern Europe and Central Asia where transition process is in its progress: Armenia, Belarus, Bulgaria, Georgia, Moldova, Romania, Russia, Ukraine, and Uzbekistan. The data set was gathered by the InterMedia Survey Institute based in Washington D.C. First they run a probit estimation procedure where computer usage at work stands as endogenous variable. And among factors which influence are the following: age negatively affects the usage, every additional year of educational attainment increases the probability of computer usage at work by 2%, those workers who live in the cities are 5% likely to use computer at work than those who live in the rural areas, those who are interested in political situation of their country also are 4% more like to use computer.

Then using the results of the probit estimation, the authors utilize the Mincerian function, in order to investigate whether the computer usage at work or future computer usage leads to wage inequalities. In this case, computer usage is treated as exogenous variable.

Concerning the returns to computer use at work, they draw a conclusion that the wage premium is about 25% when using computer at work but if using the correction mechanism with instruments this wage premium connected with computer use disappears. And they make two conclusions based on these results: first of all, the wage premium is associated not with the computer usage at work but to unobserved characteristics of worker, and, secondly, no wage premium is associated with computer usage at work in transition countries. But the authors do not report any tests for controlling the exogeneity and quality of instruments.

They assert that as long as transition countries don't have a highly developed IT, and Internet access is constraint due to its huge costs, developing countries have much less wage premium associated with computer usage at work. Another analysis made by Kuku, Orazem, Singh has been made using Russian Longitudinal Monitoring Survey (RLMS) during 2000-2004. The result is that if we run simple OLS regression, computer use has a positive (and significant) effect of 17% on wage premium.

Another aspect of this field relates to the fact that recently, relative wages of women to men has risen as well as the relative women labor participation. Concerning the supply of women labor, it increased from 35% in 1975 to 42% in 1993. Thus, these changes stimulated researches to conduct some analysis in this direction.

Weinberg examines this area. He argues that women are more prone to use computers at the their workplaces than men as well as changes in the workplace are taken place in such a way that women are more likely to get merits of it. In support of this assertion, there exists the research of Goldin (1987) who uses the data set from 1890 til 1980 where he argues that women are more favored from the technological changes.

Weinberg using the cross-industry-occupation analysis found out that demand for labor increases by one-half if women use computers at their workplaces. Also he clarifies the fact that there is no reverse causality (that increases in women labor participation can lead to the increase in the use of computer at the workplace) between computer use and demand for women labor. He checks this issue utilizing the individual level dataset on computer use which helps to explore what are the changes in men and women computer usage rates in occupation for which data is available. This is done in two ways. The first consists of estimating the regression where the dependent variable is change in computer use over all and independent variable – change in computer usage among men. The results show that not only increased computer usage of women influences the overall change, but also changes among men are responsible for that change. Another way of looking at this problem is to run a regression where dependent variable is women's employment share and independent variable – changes in computer usage among men. The results tell that there is no relationship between these two factors, hence, the relationship between women's employment and computer usage is not explained by reverse causality. According to his investigation, if there occurred advances in women's technical skills; it can lead to two-thirds increase in their wage premium. Also, he claims that women are inclined to use computer at the workplace by 45% and 33% more in 1984 and 1993 respectively and this computer use increased their hours worked by 3.7 percents. This is rather convincing because the model explanatory power is 55%.

Concerning the case of Ukraine, there has been done a relevant research in this particular area only using pooled data, where Ukraine was one of the 9 countries included. Hence, my contribution lies in the fact that I will use ULMS dataset in order to figure out whether the computer usage has any influence on wage among Ukrainian workers.

Chapter 3

METHODOLOGY

A basic framework for investigating the returns to computer usage is the Mincerian function (1974) which is a standard human capital earnings equation. It states that (log) annual wages depend on years of schooling, age and age squared.

In 1993 Krueger expanded this earnings equation to the following form:

$$lnY_i = \beta_0 + \beta_1 C_i + \beta_2 H_i + \beta_3 Z_i + \varepsilon_i \quad (1),$$

where β_1 - returns on computer use;

C_i – dummy variable if individual *i* uses computer at work;

H_i - vector of individual characteristics;

Z_i - vector of demographic characteristics.

$\varepsilon_i = \alpha_i \xi_i + \phi_i$ – error term which characterizes the worker's unobserved heterogeneity and a pure random term.

In case of zero correlation between ξ_i and regressors, β_1 estimated by OLS is unbiased. Otherwise, the bias will be present.

According to the previous studies conducted, the following specification was usually used to estimate the returns to computer use:

$$lnY = X'\delta + \beta C + u \quad (2)$$

where *lnY* stands for the logarithm of wages;

X – the vector of characteristics;

C – the dummy variable which is equal to 1 in case of using computer at work and zero otherwise.

The vector of characteristics can be further divided into subgroups [Dolton, Makepeace, 2002]:

1) Human capital variables: education, years of schooling, institutions attended, field of studies, number of languages spoken, job tenure
2) Employment variables: occupation, training, part-time or full-time work
3) Employer variables: location, industry, size variables
4) Socio-demographic variables: marital status, age, number of children, nationality, university attended.
5) Other variables: family educational and occupational background

Computer tasks can be classified as "preparation of texts and charts, the design of three-dimensional objects, calculations, the transfer of data, the classification of data and the maintenance and manipulation of data sets" [Kuku, Orazem, Singh, 2006].

Measuring unobserved heterogeneity of workers

All individuals possess different skills, knowledge and work experience, thus individuals are not homogeneous in their characteristics. Due to such distinctions, it's necessary to control for unobserved heterogeneity of a particular worker, where some measures of ability can be used to control for such heterogeneity [[Dostie, Jayaraman, Trenier, 2006].

Consequently, the following specification can be applied:

$$lnY = X'\delta + \beta C + A + u \quad (3)$$

where A represents the ability of individual.

This ability shows the competence of the individual which can be measured by general tests on individual level of development. But as mentioned before in case if this ability measure is correlated with the computer use, omitted variable bias will appear.

As a measure of ability such variable as "using personal computer at home" proposed by Krueger (1993) to capture the unobserved heterogeneity or just using a PC for all other purposes rather than work can be used. The argument lies in the fact that in case of using computer at home, the worker can be more productive at work as he/she might be familiar with much more computer-related tasks. And this is very much appreciated by employers. In case of not including this, the omitted variable bias can occur.

Consequently, the following equation can be used:

$$lnY = X'\delta + \beta_1 C_h + \beta_2 C_w + \beta_3 C_h C_w + u \quad (4)$$

where C_h - the dummy variable which is equal one if the individual uses computer at home and zero otherwise;

C_w - the dummy variable which is equal to 1 if the individual uses computer at work and zero otherwise;

$C_h C_w$ - the interaction term between computer use at home and at work, which states that an individual uses computer at home and at work. Unfortunately, that most popular technique cannot be applied due to the fact that ULMS does not have information about whether computer was used at home or at work. But it allows to control for the Internet usage. In this case we have:

$$lnY = X'\delta + \beta_1 C_i + \beta_2 C_w + u \quad (5)$$

Where C_i stands for the use of Internet.

There exist different tasks which can be done in Internet: sending emails, searching for necessary information, reading news, buying necessary things, etc.

According to Krueger investigation results, sending e-mails has the highest wage premium in case if considering the computer usage in a broad way. That is why, it is appropriate to use such variable as "Internet usage" for any purposes, hence, whether an individual is familiar with working in the web. From the common sense logic, it can be stated that it takes a little more knowledge and experience for the individual to use Internet and to work in this global network. Hence, possessing such skills can be very much appreciated by employers.

Using computer for different purposes other than for work might be interpreted as an ability of individuals to learn as the computer technologies develop very fast and it's necessary to adjust to such progress. Also it indicates that individuals are able to learn fast, which can be a signal that their innate abilities are quite high. In case if worker is adopting new computer skills, this might in turn mean that he/she has more inclination to perform certain tasks which others cannot.

But here might be another problem related to selection bias. Employers usually choose their workers according to their knowledge of computer, and controlling for computer usage at home will not give an appropriate measure for computer usage. Because the worker might use computer at work and possess only skills which are necessary at work but not at home. Hence, he can still get a wage premium for computer usage not using it at home.

In some works other methods are proposed to deal with unobserved heterogeneity.

The first one is using tools, such as pencils, pens, sitting on a chair during the working hours, etc. In all of the investigations these tools captured the unobserved heterogeneity of workers. Thus, the usage of such tools reduced the returns to computer usage to a very low level and made those returns insignificant. But unfortunately, the ULMS data does not allow to control for such tools, hence this hypothesis cannot be verified for Ukraine, at least at this stage.

Another way to capture the unobserved heterogeneity of workers is to use "future computer use" as a proxy for ability. This might reflect that such a fact can be attributed not to the possibility of being taught in the future and as an innate ability of the individual. Unfortunately, due to the absence of such variable in ULMS, this cannot be checked for Ukraine.

Controlling for Occupational, Industry and Firm Size Heterogeneity

Here it is appropriate to include variables which characterize employer status. These variables might help to capture the unobserved heterogeneity which in this case will not be attributed to the error term and will not lead to the omitted variable bias. Including such variables as occupation and industry will lower the returns to computer usage. But still the upward bias can be present in case of existence of "unmeasured ability differences across occupations and industries" [Daldy, Gibson, 2002]. The importance of firm size results from the fact that there may be present human capital differences between different sized enterprises [Reilly, 1995]. Thus, the equation will be of the following form:

$$lnY = X'\delta + \beta C + \sum_{j=1}^{n} \eta_j Occupation_{ij} + \sum_{k=1}^{m} \phi_k Fsize_{ik} + \sum_{s=1}^{l} \tau_s Industry_{is} + \varepsilon_i \qquad (6)$$

Where *Fsize* – the set of dummies which describe the size of the enterprise;

Occupation – the set of dummies describing the individual's occupation;

Industry – the set of dummies that describe the industry where the individual works.

In case of controlling for individual's occupation, the returns to computer use can be lowered by as much as 30-50%[1]. According to Reilly investigation results, if the firm size is controlled in the regression analysis, then it is supposed that the coefficient on "Computer Usage" will be insignificant.

Not controlling for all of these variables – occupation, size and industry – the returns to computer use will capture their influence on wages. ULMS data allows to control for these three variables.

IV estimation

In case of trying to avoid the endogeneity problem the IV approach can be applied

$$lnY = X'\delta + \beta C + u \quad (7)$$

$$C^* = \alpha Z + \eta \quad (8)$$

Where C=1 if $C^* > 0$ and C=0 if $C^* \leq 0$

Z – the vector of explanatory variables which influence the computer usage.

We can estimate (7) by simple OLS procedure, but the coefficient β will be unbiased only if C is exogenous. Hence, it's appropriate to use IV approach which is widely used and in particular Krueger (1991), Harmon and Walker (1995) used it.

In this case we need to use in (8) variables that are not included in (7). In other words, we need to have a variable which determines the computer usage at work but which is not present in the earnings equation. Some are proposed with this

[1] DiNardo and Pischke, 1997

respect: DiNardo and Pischke propose such instruments as parental background, achievement scores and school grades. The other instrument proposed by Kuku, Orazem, Singh is region dummy variables. It is so, because the usage of computers and Internet at home and at work is highly dependent on the regional location of the city or village. So wage also can be highly dispersed in different regions of the country. Hence, as an instrument in my research I will use settlement size for computer usage.

The correlations between computer usage and settlement size are present in the appendix 3, as well as the correlations between wages and settlement size.

Dolton and Makepeace asked several questions in their survey in order to see the attitudes of the respondent to the computer usage, whether an individual agrees or disagrees to the proposition. Answers to these questions might be potential instrumental variables. Among them are "computers are destroying people's skills", "computers enrich the lives of users", "every family should have a computer", "learning to use a computer is more trouble than it is worth". Also answers to such questions might serve as instrumental variables – whether an individual previously used computer at work, whether he/she has computer at home and whether individuals used computer at home more or less than once a week. Unfortunately, ULMS data set is very restrictive in terms of computer literacy of the workers. If new data set would be available containing all that information, it could be the topic for deeper research.

Sample Selection

In case if the sample is not randomly chosen from the population, the sample selection problem occurs.

$$d_i^* = z_i'\gamma + v_i$$

$$y_i^* = x_i'\beta + \varepsilon_i$$

v_i and ε_i are independent and jointly normally distributed error terms.

We can observe the value y_i only:

$$d_i = \begin{cases} 1 \; if \; d_i^* > 0 \\ 0 \; otherwise \end{cases}$$

$$y_i = \begin{cases} y_i^* \; if \; d_i = 1 \\ n.a. \; otherwise \end{cases}$$

According to the first equation – we determine the presence of the observation in the sample, to the second one – we evaluate y_i

y_i represents the conditional expectation of y_i^* in case of being observed ($d_i = 1$):

$$E(y_i \parallel x_i, z_i) = E(y_i^* \parallel d_i = 1, x_i, z_i) = x_i'\beta + \rho\sigma_\varepsilon \frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)} = x_i'\beta + \rho\sigma_\varepsilon\lambda(z_i'\gamma)$$

where $\lambda(\alpha) \equiv \phi(\alpha)/\Phi(\alpha)$ – *inverse Mills ratio.*

The selection bias occurs if we just run OLS regression:

$$y_i = x_i'\beta + u_i$$

β will be biased because $\rho\sigma_\varepsilon \frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)}$ is omitted and belongs to the error term.

Hence, Heckman's Two-Step Procedure can be applied to deal with sample selection. This done by:

1) running a standard probit regression: estimate γ by maximum likelihood estimation

$$d_i^* = z_i^{'}\gamma + \upsilon_i$$

$$d_i = 1 \text{ if } d_i^* > 0, \text{ 0 otherwise}$$

Also, the Mills ratio can be estimated as

$$\hat{\lambda}_i = \frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)}$$

for all observations.

2) running a linear regression model (with inverse Mills ratio):

$$y_i = x_i^{'} + \beta_\lambda \hat{\lambda}_i + u_i$$

The coefficient obtained by this procedure is consistent but not efficient.

Chapter 4

DATA DESCRIPTION

The dataset to be used in this research is ULMS for 2003. The data is also available for 2004 but the variation in interested variables: computer usage, internet usage, wage and education is very small and results are expected to be very similar. Hence, regressions only for 2003 are run. The data describes the working-age population aged 15-72 and the whole sample consists of 8641 individuals.

The variables are constructed as follows: minimum wage in 2003 was 180 hryvnas, but a lot of people reported lower than that, hence I take the lower bound for fulltime workers to be 50% of the minimum wage, that is, 90 hryvnas, and 45 hryvnas for part-time workers, wage consists of wage from the main job, second job and in-kind payments in hryvnas; the dummy variable for female is equal to one if the person is female and zero if male; those who are in a non-registered or registered marriage are labeled as married and the dummy variable for them is equal to one and zero if a person is single, widowed, divorced or separated; if the person is employed, the dummy is equal to one and zero otherwise; the dummy variable for being in the union is equal to one and zero otherwise; the dummy variable for computer usage is equal to one if person ever used a computer and zero otherwise; the dummy variable is equal to one if the person used Internet during the last year and zero otherwise, occupations analyzed are: manager, professional, technician, clerk, service, skilled, craft, machinery, unskilled and armed; industries are the following: agriculture, industry, electricity, construction, sales, transport, financial, public administration, education and health, municipal and other. The region dummy variable are consolidated into several groups due to the fact that each region represents a very small part of the sample, only Kiev and Crimea are taken as separate groups. Settlement size is constructed as follows: village, urban settlement, small town (up to 20 thousands), town (20-99 thousands), city (100 thousands – 499 thousands), and large city (more than 500 thousands). Concerning education, I grouped all workers having bachelor, specialist, master and PhD into “higher education” because each group separately represents a very small fraction. Also, those who have 1-6 and 7-9 grades into “under 10 grades” because of the same reason.

A drawback of the ULMS data is that it has information only about whether a person used computer during his/her life and whether he/she used Internet during the last year. It cannot control whether an individual used computer at work or at home; what particular tasks were performed: word processing, internet, email, games, data analysis, programming; or what is the individual’s experience is in using a computer and Internet. Also, it cannot control for how many hours a person worked with the computer. Also, it doesn’t reveal information how an individual evaluates his own knowledge of computer, for example, on a scale – good, fair, poor, not skilled – in different fields of computer usage.

After dropping observations of people who do not report whether they used computer or internet, people who do not report their wages, people who are unemployed or who are employed but report zero wages, the sample reduces to 2598 individuals.

Table 1. Computer and Internet usage for full sample of workers.

	Computer Usage, %	Internet Usage, %
Yes	29.42	7.29
No	70.58	92.71
Total	100	100

As can be seen from 2003 sample, people who ever used computer represent 29.42%, and those who used Internet this year – 7.29%.

Table 2. Computer usage among men, women, single and married.

Computer Usage, %	Men	Women	Single	Married
Yes	29.78	29.15	43.88	21.48
No	70.22	70.85	56.12	78.52
Total	100	100	100	100

In contradiction with findings from developed countries, where women are more likely to use computer, women in Ukraine are equally likely to use computer with men – 29.15% of women from the sample used computer while slightly higher percentage 29.78 of men did that. Also, single people who ever used computers represent higher percentage – 43.88% in comparison with their married counterparts – 21.48%.

Table 3. Internet usage among men, women, single and married.

Internet usage, %	Men	Women	Single	Married
Yes	8.56	6.37	12.81	4.27
No	91.44	93.63	87.19	95.73
Total	100	100	100	100

Men were using Internet more often than women during the last year – 8.56% for men in contrast to 6.37% for women. Single people were using Internet more than three times often – 12.81% against 4.27%.

Table 4. Differences in wages between worker's who use computer and Internet and those who do not.

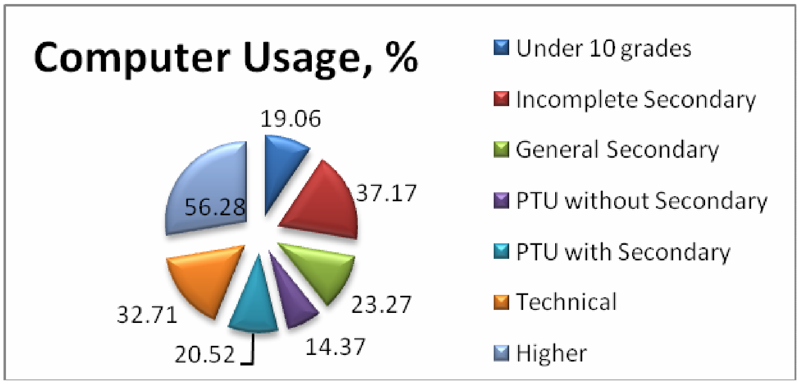
	Computer Usage		Internet Usage	
	Yes	No	Yes	No
Wage, hr	362.4	276	431.1	295.2

As can be seen there is a substantial wage differential between those who used computer and those who didn't. Those who ever used computer received 362.4 hryvnas and those who didn't – 276 hryvnas.

What concerns Internet usage, there is also a substantial wage differential, much higher than in computer usage. For those who used Internet during the preceding year – 431.1 hryvnas against those who didn't – 295.2 hryvnas.

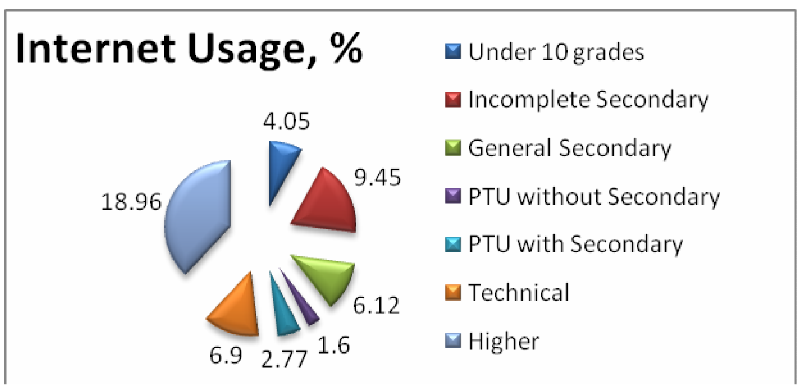
Different authors use different variables for wage – some use logarithm of hourly wages and some – logarithm of monthly wages. The main criticism about using the former goes from the fact that a lot of people work long hours like managers, hence, if hourly wage is used, we will get the underestimation of wages. In this research the logarithm of monthly wage is used. For Ukraine it is more appropriate to use logarithm of monthly wages, because all the workers initially are paid a monthly salary and not an hourly wage.

Figure 1. Computer usage according to educational attainment



As can be seen from the table, among workers who hold degrees – bachelor, specialist, master or doctor – the percentage of people using computer is the highest – 56.28%. The smallest percentage of people who ever used computer relates to the category "PTU without having a secondary education" – 14.37%.

Figure 2. Internet usage according to educational attainment



Similar picture can be observed for the usage of Internet. Among people who have a degree the highest percentage of people who use Internet during the last year is observed – 18.96 %.The smallest fraction of people who used Internet during the last year is observed among people with "PTU without secondary" education – 1.6%.

Computer usage as well as the Internet usage is dependent on the location of a particular worker because some settlements can be situated in an area where the cable connection is not connected, it is important to see where do the sample workers have their jobs located.

The division between settlements is made on the following scale: village, urban settlement, small town (up to 20 thousand people), town (20-99 thousand people), city (100 -4999 thousands people), Large city (more than 500 thousands people).

Table 5. Computer and Internet usage in different types of settlements

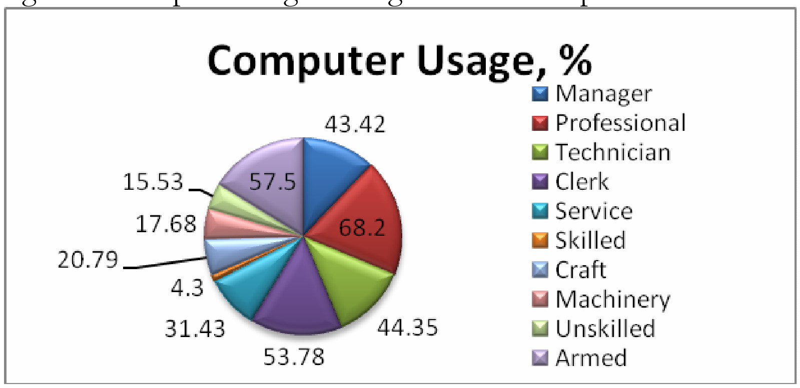
	Computer Usage, %	Internet Usage, %
village	18.4	2.6
urban	27	4.2
small town	20.4	2.6
medium town	27.4	6.1
city	36.6	10.9
large city	45.7	15.3

According to ULMS data, the highest fraction of those living in large city ever used a computer and Internet during the last year – 45.7% and 10.9% accordingly. The interesting fact that this data disclose is that there is a lower fraction of those living in the small town who ever used computer and Internet than among those who live in the urban settlement – 20.4% and 2.6% accordingly.

But still the result is the following: the larger the city, the higher percentage of people is using computer and Internet, which is logical because larger cities are well equipped in terms of telephone, cable and satellite facilities.

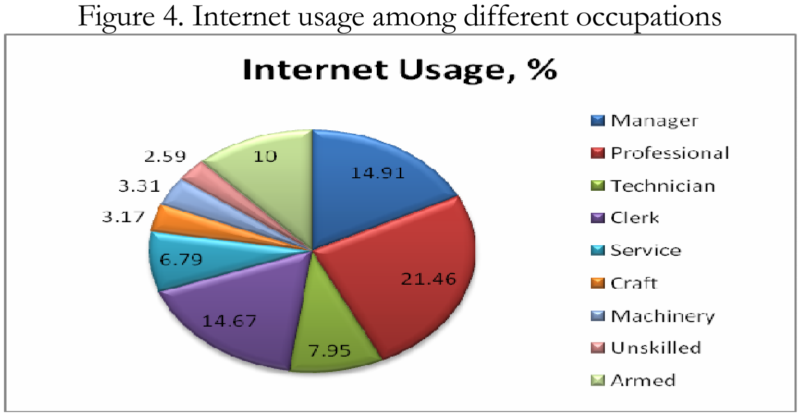
Also it would be beneficial to see how many people use computer and Internet in different occupations, regions and industries.

Figure 3. Computer usage among different occupations



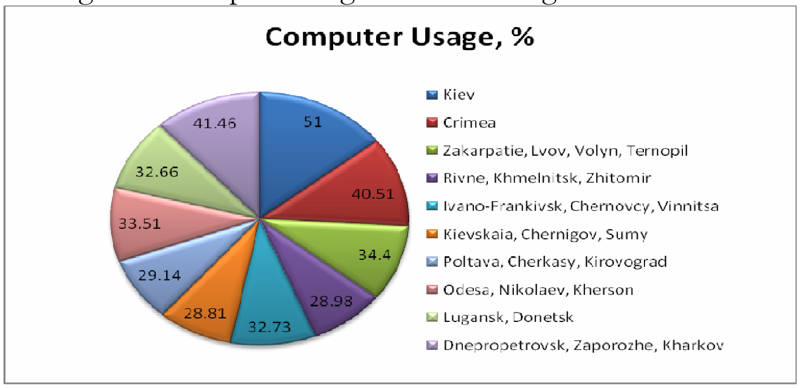
From the graph it can be inferred that the highest proportion of people who used computers was among professionals and armed (68.2% and 57.2% accordingly), and the lowest – among skilled workers (4.3%).

Figure 4. Internet usage among different occupations



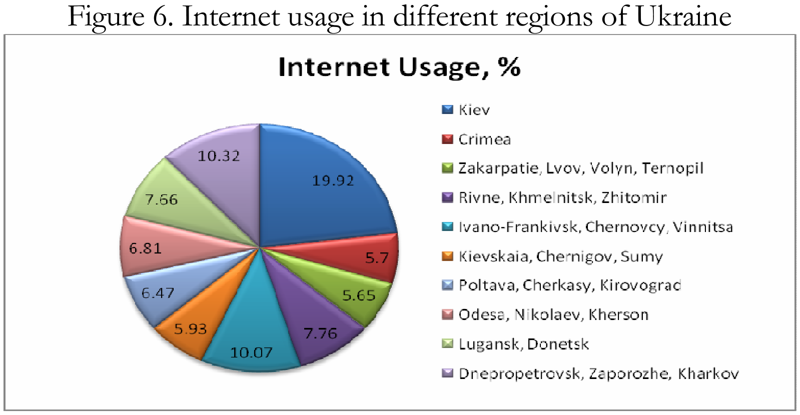
Among professionals there exists the highest percentage of workers who used Internet during the last year – 21.46%. The striking result is that among skilled workers no worker ever used Internet during the last year.

Figure 5. Computer usage in different regions of Ukraine



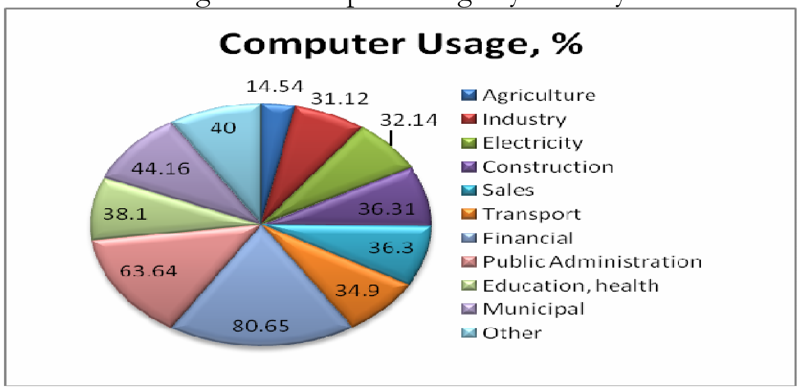
As expected, the highest percentage of those who live in Kiev ever used a computer – 51%, which can be explained by the developed IT infrastructure in the capital of Ukraine in comparison with other regions, then goes – Dnepropetrovsk, Zaporozhe, Kharkov (41.46%) and Crimea – 40.51%. Interesting result is that in Kievskaia, Chernigovskaia, Sumskaia region the smallest fraction of workers ever used computers – 28.81%.

Figure 6. Internet usage in different regions of Ukraine



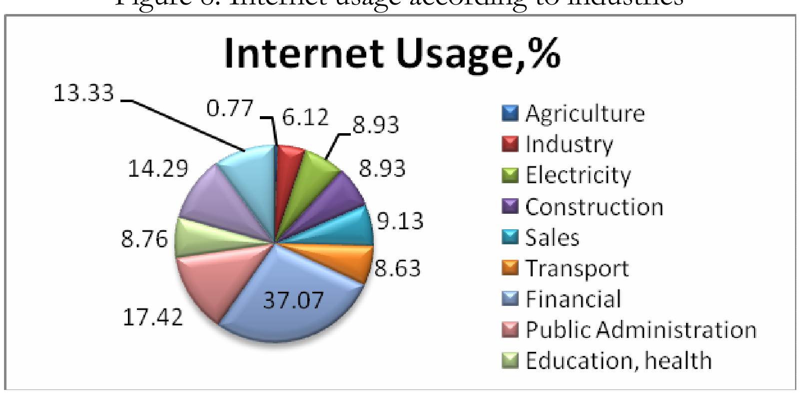
Similar situation is present in Internet usage: in Kiev the highest proportion of people used Internet during the last year – 19.92%, then again goes – Dnepropetrovsk, Zaporozhe, Kharkov (10.32%) and then – Ivano-Frankivsk, Chernovcy, Vinnitsa region – 10.07%. The smallest proportion of people who used Internet during the last year is in Kiev, Chernigiv, Sumy region – 5.93%. Which is similar to the computer usage in this region.

Figure 7. Computer usage by industry



A very large fraction who work in the financial sector ever used computer – 80.65%, which is in accordance with common sense logic. The lowest portion of workers who used computer is present in agriculture – 14.54% which is also natural to expect.

Figure 8. Internet usage according to industries



Similar results are for Internet usage during the last year – the largest fraction is present for those who work in the financial sector – 37.07%, and the lowest – in agriculture (0.77%).

Chapter 5

EMPIRICAL FINDINGS

Running the OLS regression of the standard Mincerian function augmented by gender, occupation, the size of the enterprise, industry, region, whether individual works full-time, computer usage and Internet usage gives the following results (the whole output of this regression is in the Appendix 1).

	Full sample	Men	Women
COEFFICIENT	ln_wage	ln_wage	ln_wage
e_secondary	0.0312	0.0372	0.0106
	-0.038	-0.057	-0.045
e_ptu_without	0.0186	-0.00947	0.0608
	-0.048	-0.066	-0.06
e_ptu_with	0.0679*	0.076	0.0212
	-0.038	-0.057	-0.047
e_technical	0.0595	0.0828	0.0206
	-0.037	-0.057	-0.042
e_higher	0.142***	0.133*	0.0903*
	-0.045	-0.071	-0.053
age	0.0351***	0.0452***	0.0184*
	-0.0072	-0.01	-0.0097
fulltime	0.498***	0.636***	0.436***
	-0.044	-0.095	-0.047
comp_use	0.135***	0.147***	0.152***
	-0.023	-0.04	-0.028
e_usage	0.127***	0.1	0.148***
	-0.04	-0.061	-0.054
Constant	3.847***	3.286***	4.364***
	-0.25	-0.37	-0.32
Observations	2598	1203	1395
R-squared	0.38	0.34	0.36
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

As can be seen from the table, coefficients on computer usage and Internet usage are highly significant. Returns to computer usage are 0.135 for full sample, 0.147 for men and 0.152 for women. The returns are calculated as follows: $100(e^{0.135}-1)$, which is equal to 14.5%, 15.8% and 16.4% accordingly. The coefficients on Internet usage are 0.127 for full sample, 0.1 for men and 0.148 for women. The returns are 13.5%, 10.5% and 16% accordingly. The returns to the Internet usage for men are insignificant. This can be due to the fact that a lot of women work as secretaries or connected with jobs where one of the major tasks is sending emails, while men usually use computer for working with spreadsheets, databases and making different estimation procedures, that is why it does not matter whether they will use Internet or not.

For the full sample only higher education is significant and is equal to 0.142, hence, returns are 15.3%. While for men and women OLS estimates give the result that education give the returns of 14.2% for men having higher education (with t=1.9) and 9.5% for women having higher education (t=1.7).

Working full time gives a premium of 64.5% for full sample, 88.9% for males, and 54.7% for female. Age also positively relates to earnings: 3,6% premium for full sample, 4.6% for men, and 1.9% for women.

The next technique to be performed – Heckman selection procedure. The results are the following (The whole output of this regression is in the appendix 2).

	Log pseudolikelihood=-3754,049 Wald chi2(53)=1746,83 Prob>chi2=0,0000	Log pseudolikelihood=-1951,049 Wald chi2(53)=794,32 Prob>chi2=0,0000	Log pseudolikelihood=-1658,539 Wald chi2(53)=916,87 Prob>chi2=0,0000
	Full Sample	**Men**	**Women**

	ln_wage	Robust SE	ln_wage	Robust SE	ln_wage	Robust SE
e_secondary	0.051166	0.0397981	0.0427296	0.0564334	0.0221913	0.0459537
e_ptu_with~t	0.0438285	0.0496339	-0.007239	0.0654968	0.0790869	0.0616533
e_ptu_with	0.1093059	0.0432699	0.0815802	0.061275	0.0385673	0.048365
e_technical	0.0939743	0.0406621	0.080052	0.0583057	0.0496816	0.0445675
e_higher	0.1737275	0.0488538	0.1186266	0.0724433	0.1139593	0.0564868
age	0.0518294	0.0086809	0.048301	0.0114818	0.0404942	0.011843
fulltime	0.502062	0.0437746	0.6425063	0.0908093	0.4369583	0.0467739
comp_use	0.1424682	0.0244267	0.1315761	0.0384088	0.1745286	0.0301983
e_usage	0.1210039	0.0407137	0.088332	0.0596094	0.1480964	0.0548667
cons	2.816096	0.4024714	3.077025	0.5538274	3.239978	0.794664
/athrho	0.5445462	0.2050572	-0.0299416	0.2600422	0.646501	0.2464554
/lnsigma	-0.7748684	0.0432723	-0.7482176	0.0219152	-0.9076021	0.0495396
rho	0.4964215	0.1545241	-0.0299327	0.2598092	0.5693098	0.1665758
sigma	0.4607644	0.0199383	0.4732092	0.0103705	0.4034906	0.0199888
lambda	0.2287334	0.0805101	-0.0141644	0.12297	0.2297111	0.077488
	Wald test (rho=0): chi2(1)= 7.05 prob>chi2=0.0079		Wald test (rho=0): chi2(1)= 0.01 prob>chi2=0.9083		Wald test (rho=0): chi2(1)= 6.88 prob>chi2=0.0087	

Coefficient on "rho" is the correlation between two equations. "athrho" is the transformation of "rho". Wald test shows the significance of the "athrho": if it is significant, then there is sample selection. Hence, from the regression results, it can be stated that there is sample selection for the full sample and for female workers (p-value is equal to 0.0079 and 0.0087 accordingly), while there is no sample selection in men sample, hence, it's appropriate to report OLS results.

The returns to computer usage didn't change much in case of correcting for selectivity bias. They are 15.3% for the full sample, 19.1% for women. They are all significant. Concerning Internet usage, it is significant for the full sample, as well as for females. The returns are 12.9% for the full sample and 16% for female workers.

Education is significant for the full sample for "ptu with secondary education" and gives the returns of 11.5%, "technical education" with returns 9.9%, and "higher education" with returns of 19%. For women premium for higher education is equal to 12.1% and it's significant.

Working fulltime give a premium for full sample and for women and represents 65.2% and 54.8% accordingly and it's significant.

Testing heckman coefficients for higher education, age, working fulltime, computer usage, internet usage against OLS coefficients tells that they are insignificantly different from each other. For women and full sample p-values are 0.6753 and 0.5161 accordingly for higher education, 0.0621 and 0.0540 for age, 0.9837 and 0.9261 for working fulltime, 0.4557 and 0.7598 for computer usage, 0.9986 and 0.8829 for internet usage.

Next technique to be used – IV to deal with endogeneity problem. In order to be a good instrument, it has to satisfy two conditions: relevance (correlated with endogenous regressors), and exogeneity (uncorrelated with error terms).

As an instrument variable I used "settlement size". I chose it because it is correlated with computer usage, because the higher the settlement size, the higher the probability that it will be well-equipped with telephone, cable and satellite facilities, and vice versa. The first problem which arises here is that settlement size can be correlated with wages: the larger the settlement size, the higher will be the wages. In Appendix 3 there presented the correlations between computer usage and settlement size, as well as correlations between wages and settlement size.

After obtaining the regression results (which are in Appendix 4), it can be stated that instruments doest not satisfy the relevance condition. It can be seen from the following statistics obtained: robust F-statistics is equal to 3.18 for full sample, 1.77 for males and 1.93 for females, from which it can be stated that the model is weakly identified. Comparing this to critical values of Cragg-Donald F statistic, it also can be stated that the null about the weakly identified model definitely cannot be rejected. Also, Staiger-Stock "rule of thumb" tells that F-statistics should be more than 10 in order to satisfy the relevance condition which is definitely not the case of our IV regression.

Hence, instruments are not sufficiently correlated with endogenous regressor to avoid "weak identification" problem such as large finite-sample bias.

Chapter 6

CONCLUSIONS

A lot of discussion has been present in the past decade among researches about whether the higher wage premium can be associated with skill-biased technological change.

Hence, the purpose of my research was to investigate the Ukrainian labor market with respect to whether computer literacy of the worker influences his wage. My results are in accordance with the main findings of Krueger (1993) that there exists a wage premium for male and female workers that is equal to 15.8% and 16.4 % accordingly using simple OLS and 19.1% for women using heckman selection procedure. Also, the returns are a little higher for women which is also the case for all developed countries. These results are high and significant even after controlling for individual and job characteristics. Such human capital variables were controlled: education, age, age squared, gender, marital status. Also such job characteristics were controlled: the activity of the enterprise, occupation, the region where the worker is employed, whether the worker is familiar working with Internet, which are suggested by researches in this field and results still show the positive premium associated with computer usage, which is significant.

Hence, it can be stated that there exists a premium for computer usage, which is used as a proxy for skill-biased technological change, for the Ukrainian labor market. This will lead to the increased productivity of workers, which in turn will positively influence the activity of the enterprise. Thus, this should be a signal for the government to develop the telephone, cable and satellite facilities. The developed IT infrastructure will lead to more development in the whole economy, where the welfare of the individuals will be increased.

Though the results should be treated with caution because of the drawbacks of the dataset used: it has only the information about whether the individual ever used a computer, there does not exist variable which could control for whether the computer is used at work or at home, or for what purposes it could be used – for programming, modeling, spreadsheet analysis, word processing, or games, also it would be beneficial to figure out whether the frequency of computer usage somehow influences the wage structure of the worker. If such information were present, the results might be significantly different. In this perspective the coefficients I obtained can be overestimated. Hence, it can the basis for further investigation in this field in order to find out whether the skill-biased technological change is present in the Ukrainian labor market or all wage inequalities can be attributed to unobserved heterogeneity of workers and workplaces.

BIBLIOGRAPHY

Anger and Schwarze. 2002. *Does Future PC Use Determine Our Wages Today? Evidence from German Panel Data*. Discussion Paper No. 429, February.

Arabsheibani and Marin. 2000. *If Not Computers Then What? Returns to Computer Use in the UK Revisited.* Research Paper 2000-4.

Arabsheibani and Mussurov. 2006. *Returns to Schooling in Kazakhstan; OLS and Instrumental Variables Approach.* IZA Discussion Paper No 2462.

Autor, Katz and Krueger. *Computing Inequality: Have Computers Changed the Labor Market?* The Quarterly Journal of Economics. Vol.113. No.4, November, 1169-1213.

Bell. 1996. *Skill-biased Technical Change and Wages: Evidence from a Longitudinal Dataset.* Economic Papers with number W25.

Borghans and ter Weel. 2000. *How Computerization Changes the UK Labor Market: The Facts Viewed from a New Perspective.* Provided by Maastricht University. Research Memoranda with number 025.

Bruinshoofd and Ter Weel. 1998. *Skill-biased Technical Change: On Technology and Wages in the Netherlands.* Provided by Maastricht University. Research Memoranda with number 021.

Daldy and Gibson. 2002. *Have Computers Changed the New Zealand Wage Structure? Evidence from Data on Training.* New Zealand Journal of Industrial Relations Online. Vol. 28. No.1

DiNardo and Pischke. 1997. *The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?* The Quarterly Journal of Economics. Vol. 112, No.1, February, 291-303.

Dolton and Makepeace. 2002. *The Returns to Computer Use: An Empirical Analysis for the UK.* The Economic Journal. Vol.114. Issue 494

Dolton, Makepeace and Robinson, 2007. *Use IT or Lose IT? The Impact of Computers on Earnings.* IZA Discussion paper No. 2588.

Dostie, Jayaraman and Trepanier. 2006. *The Returns to Computer Use Revisited, Again.* IZA Discussion paper No. 2080.

Entrof, Gollac and Kramarz. 1999. *New Technologies, Wages, and Worker Selection*. Journal of Labor Economics, 17, 3, 464-491

Franzen and Axel. 2001. *Wages and the Use of New Technologies: An Empirical Analysis of the Swiss Labor Market.* Swiss Journal of Economics and Statistics*,* December, 505-523.

Goldin. 1987. *Women's Employment and Technological Change: A Historical Perspective.* In Computer Chips and Paper Chips: Technology and Women's Employment, Heidi I Hartmann, Editor. Washington, D.C.: National Academy Press.

Krashinsky and Harry. 2000. *Do Marital Status and Computer Usage Really Change the Wage Structure? Evidence from a Sample of Twins*. Working Paper No. 439.

Krashinsky and Harry. 2005. *Do Marital Status and Computer Usage Really Change the Wage Structure?* The Journal of Human Resources*,* 774-791.

Krueger. *How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989*. The Quarterly Journal of Economics. Vol.108, No.1, February, 33-60.

Krueger. 2000. *The Digital Divide in Educating African-American Students and Workers.* Working paper No. 434. Princeton University.

Kuku, Orazem, and Singh. 2006. *Computer Adoption and Returns in Transition.* IZA Discussion Paper No.1360.

Lichtenberg. 1993. *The Output Contributions of Computer Equipment and Personnel: A Firm-level Analysis.* National Bureau of Economic Research. Working paper 4540.

Morrison and Berndt. *Assessing the Productivity of Information Technology Equipment in U.S. Manufacturing Industries.* National Bureau of Economic Research, Working Paper 3582.

Oliner and Sichel. 1994. *Computers and Output Growth Revisited: How Big is a Puzzle?* Brookings Papers on Economic Activity, 2.

Oosterbeek. 1996. *Returns to Computer Use: A Simple Test on the Productivity Interpretation*. Economic Letters. Vol.55. 273-277

Schmidheiny Kurt. 2004. *Limited Dependent Variable Models*. Applied Econometrics II.

Siegel and Donald. *The Impact of Computers on Manufacturing Productivity Growth: A Multiple-Indicators, Multiple-Cases Approach.* The Review of Economics and Statistics. Vol. 79, No. 1, February, 68-78

Weinberg. 2000. *Computer Use and the Demand for Women Workers.* Ohio State University. Industrial and Labor Relations Review. Vol.53, No.2, p. 290-308.

Appendix 1. OLS estimation procedure.

Table 1. OLS regression

	Full sample	Men	Women
COEFFICIENT	ln_wage	ln_wage	ln_wage
e_secondary	0.0312	0.0372	0.0106
	-0.038	-0.057	-0.045
e_ptu_without	0.0186	-0.00947	0.0608
	-0.048	-0.066	-0.06
e_ptu_with	0.0679*	0.076	0.0212
	-0.038	-0.057	-0.047
e_technical	0.0595	0.0828	0.0206
	-0.037	-0.057	-0.042
e_higher	0.142***	0.133*	0.0903*
	-0.045	-0.071	-0.053
age	0.0351***	0.0452***	0.0184*
	-0.0072	-0.01	-0.0097
age2	-0.000283***	-0.000369***	-0.000154**
	-0.000058	-0.000085	-0.000078
fulltime	0.498***	0.636***	0.436***
	-0.044	-0.095	-0.047
female	-0.217***		
	-0.021		
comp_use	0.135***	0.147***	0.152***
	-0.023	-0.04	-0.028
e_usage	0.127***	0.1	0.148***
	-0.04	-0.061	-0.054
o_manager	0.303***	0.323**	0.318***
	-0.099	-0.15	-0.11
o_proffesional	0.181**	0.198*	0.173*
	-0.082	-0.11	-0.092
o_technician	0.0365	0.0846	-0.0156
	-0.077	-0.11	-0.086
o_clerk	-0.0262	-0.064	-0.0258
	-0.081	-0.13	-0.088
o_service	-0.0133	-0.0187	-0.0651
	-0.082	-0.13	-0.091
o_craft	0.051	0.0308	0.0612
	-0.078	-0.1	-0.092
o_machinery	0.0767	0.0643	0.143
	-0.084	-0.11	-0.12
o_unskilled	-0.0843	-0.0322	-0.131
	-0.075	-0.1	-0.085
o_armed	0.445***	0.360**	0.515***
	-0.12	-0.15	-0.11
s_2_4	0.0667	0.135	0.0615
	-0.11	-0.2	-0.12
s_5_9	0.106	0.207	0.0707
	-0.11	-0.2	-0.12
s_10_19	0.126	0.232	0.11
	-0.1	-0.19	-0.11
s_20_49	0.13	0.183	0.132
	-0.1	-0.18	-0.11
s_50_99	0.179*	0.205	0.199*
	-0.1	-0.18	-0.11
s_100_249	0.206**	0.27	0.206*
	-0.1	-0.18	-0.12
s_250_499	0.234**	0.322*	0.200*

	-0.1	-0.19	-0.12
s_500_999	0.233**	0.277	0.264**
	-0.11	-0.19	-0.12
s_1000	0.319***	0.409**	0.270**
	-0.1	-0.18	-0.12
r_krim	-0.112**	-0.217**	-0.0443
	-0.055	-0.099	-0.061
r_zkrp_lv_vl_trn	-0.232***	-0.279***	-0.176***
	-0.042	-0.067	-0.053
r_rvn_xml_jtm	-0.305***	-0.231***	-0.338***
	-0.045	-0.077	-0.054
r_ivn_chrn_vnc	-0.316***	-0.322***	-0.282***
	-0.048	-0.081	-0.059
r_kiv_chrng_cm	-0.314***	-0.379***	-0.239***
	-0.046	-0.075	-0.056
r_pltv_chrks_kir	-0.212***	-0.252***	-0.167***
	-0.046	-0.072	-0.058
r_ods_mklv_xrsn	-0.188***	-0.283***	-0.0967*
	-0.045	-0.071	-0.057
r_lgnsk_dnck	-0.182***	-0.134**	-0.242***
	-0.041	-0.065	-0.052
r_dnprv_zaprj_xrkv	-0.235***	-0.304***	-0.181***
	-0.041	-0.066	-0.05
i_industry	0.258***	0.383***	0.0621
	-0.046	-0.065	-0.055
i_electricity	0.340***	0.443***	0.176*
	-0.062	-0.079	-0.1
i_construction	0.329***	0.439***	0.138*
	-0.063	-0.082	-0.081
i_sales	0.207***	0.229***	0.127**
	-0.052	-0.076	-0.063
i_transport	0.341***	0.510***	0.122*
	-0.052	-0.072	-0.064
i_financial	0.277***	0.412***	0.116
	-0.084	-0.13	-0.1
i_public_adm	0.136**	0.342***	-0.0559
	-0.059	-0.097	-0.068
i_educ_health	-0.0529	-0.0615	-0.152***
	-0.043	-0.069	-0.047
i_other_municipal	0.142***	0.223***	0.0292
	-0.053	-0.085	-0.06
i_other	0.0326	0.0236	0.173
	-0.13	-0.16	-0.19
Constant	3.847***	3.286***	4.364***
	-0.25	-0.37	-0.32
Observations	2598	1203	1395
R-squared	0.38	0.34	0.36

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2. Heckman selection model

	Log pseudolikelihood=-3754,049 Wald chi2(53)=1746,83 Prob>chi2=0,0000		Log pseudolikelihood=-1951,049 Wald chi2(53)=794,32 Prob>chi2=0,0000		Log pseudolikelihood=-1658,539 Wald chi2(53)=916,87 Prob>chi2=0,0000	
	Full Sample		**Men**		**Women**	
	ln_wage	Robust SE	ln_wage	Robust SE	ln_wage	Robust SE
e_secondary	0.051166	0.0397981	0.0427296	0.0564334	0.0221913	0.0459537
e_ptu_with~t	0.0438285	0.0496339	-0.007239	0.0654968	0.0790869	0.0616533
e_ptu_with	0.1039059	0.0432699	0.0815802	0.061275	0.0385673	0.048365
e_technical	0.0939743	0.0406621	0.080052	0.0583057	0.0496816	0.0445675
e_higher	0.1737275	0.0488538	0.1186266	0.0724433	0.1139593	0.0564868
age	0.0518294	0.0086809	0.048301	0.0114818	0.0404942	0.011843
age2	-0.0004217	0.0000712	-0.0003924	0.0000938	-0.0003374	0.0000968
fulltime	0.502062	0.0437746	0.6425063	0.0908093	0.4369583	0.0467739
female	-0.2096702	0.0215234				
comp_use	0.1424682	0.0244267	0.1315761	0.0384088	0.1745286	0.0301983
e_usage	0.1210039	0.0407137	0.088332	0.0596094	0.1480964	0.0548667
o_manager	0.2739446	0.0979164	0.3073445	0.1406617	0.2877925	0.1027824
o_proffesi~l	0.1630332	0.0807251	0.1879047	0.1084544	0.1635342	0.0885618
o_technician	0.0145585	0.0755772	0.0709209	0.1008718	-0.022649	0.0809848
o_clerk	-0.0551298	0.0801305	-0.0823901	0.1234805	-0.038047	0.0841156
o_service	-0.0392439	0.0805216	-0.0314154	0.1266237	-0.0805491	0.0866036
o_craft	0.0328091	0.0763915	0.0190151	0.0962157	0.0518726	0.0865805
o_machinery	0.0612271	0.0823485	0.0523035	0.102407	0.1451441	0.1148371
o_unskilled	-0.0998821	0.073581	-0.0454446	0.0955271	-0.1333654	0.0798076
o_armed	0.4368035	0.1148095	0.3773724	0.1409904	0.515693	0.1088204
s_2_4	0.0799652	0.110621	0.1718828	0.2122198	0.0713743	0.1150573
s_5_9	0.1163307	0.1100964	0.260437	0.2094655	0.0710505	0.1148541
s_10_19	0.1219654	0.1072643	0.265349	0.2004957	0.1016057	0.1127363
s_20_49	0.1416073	0.1063069	0.2302052	0.1964029	0.1294161	0.1111951
s_50_99	0.18052	0.1067308	0.2454942	0.1966517	0.1895212	0.1121142
s_100_249	0.1978313	0.1073129	0.3014726	0.1964283	0.1837472	0.1134539
s_250_499	0.228607	0.1082139	0.358077	0.1979121	0.1797517	0.114747
s_500_999	0.2280256	0.1094135	0.3279695	0.1997465	0.2327912	0.1172754
s_1000	0.3119003	0.10822	0.4402336	0.1982428	0.2469386	0.1148978
r_krim	0.234141	0.1267859	-0.1736403	0.1795456	0.3001326	0.1505327
r_zkrp_lv_~n	0.1810914	0.1353202	-0.2081842	0.1908657	0.2200898	0.1526833
r_rvn_xml_~m	0.0943669	0.1377865	-0.1664498	0.200962	0.0219915	0.1486862
r_ivn_chrn~c	0.0552604	0.1229138	-0.2308179	0.1621014	0.0860752	0.1503561
r_kiv_chrn~m	0.0923088	0.134868	-0.3096494	0.1868678	0.1580836	0.1558474
r_pltv_chr~r	0.1709834	0.1314442	-0.2044687	0.1836425	0.2073176	0.1532402
r_ods_mklv~n	0.2019791	0.1321392	-0.2161711	0.1839075	0.2779054	0.1518166
r_lgnsk_dnck	0.2126175	0.1403787	-0.103017	0.1939521	0.1407728	0.1590257
r_dnprv_za~v	0.0966198	0.1249796	-0.3041142	0.1704636	0.154867	0.1409913
i_industry	0.2122249	0.0461716	0.3186836	0.0632202	0.0179674	0.0571444
i_electric~y	0.3148666	0.0612646	0.3950557	0.0771666	0.1567535	0.1017761
i_construc~n	0.2754966	0.0624333	0.3710354	0.0800211	0.0936168	0.0761923
i_sales	0.1596399	0.0515851	0.1671899	0.0745127	0.0819101	0.0625194
i_transport	0.3071619	0.0502588	0.4619505	0.0695081	0.0933118	0.0629794
i_financial	0.224066	0.0818813	0.3316227	0.1243838	0.0671592	0.0971049
i_public_adm	0.1102216	0.0585692	0.2925011	0.0950335	-0.077342	0.0662195
i_educ_hea~h	-0.0834406	0.0424593	-0.1057569	0.0670885	-0.1803049	0.0464907
i_other_mu~l	0.096109	0.0523587	0.1640151	0.0826909	-0.0143424	0.059578
i_other	-0.0430291	0.1301568	-0.0665885	0.1619321	0.1123028	0.1796382
ss_urban	0.0933717	0.0319453	0.1044653	0.0514697	0.0228485	0.0338564
ss_small_t~n	-0.0233828	0.0546618	-0.0027375	0.0838044	-0.0492622	0.0713099
ss_town	0.0652737	0.0337899	0.0736369	0.0505215	0.0296195	0.0407679
ss_city	0.1209148	0.0301887	0.1228139	0.0450802	0.1073891	0.0356365
ss_large_c~y	0.2318045	0.0443986	0.2105165	0.0685894	0.1722565	0.0516095

cons	2.816096	0.4024714	3.077025	0.5538274	3.239978	0.794664
employed						
married	-0.0088084	0.0482605	0.2777586	0.0903535	-0.1474549	0.0715265
e_secondary	0.1491875	0.0696516	0.1768994	0.0962391	0.1551337	0.104161
e_ptu_with~t	0.1891531	0.0950843	0.2414948	0.1150489	0.1460573	0.1633278
e_ptu_with	0.3701047	0.0754371	0.4639958	0.1010677	0.2590236	0.1120645
e_technical	0.2969994	0.0687447	0.2763432	0.1010837	0.3114509	0.1002256
e_higher	0.3744719	0.0786184	0.3626367	0.1136304	0.3396895	0.1133446
age	0.1260945	0.0132417	0.0577934	0.0174251	0.1958739	0.0224209
age2	-0.0010718	0.0001106	-0.0004582	0.00015	-0.001677	0.0001836
female	0.0555753	0.0484107				
comp_use	0.1756108	0.0588303	0.0194789	0.0882384	0.3382301	0.0802188
e_usage	0.128342	0.0988526	0.0105809	0.1334126	0.3316088	0.1502966
r_krim	2.208415	0.1121358	1.88654	0.150553	2.471745	0.1752878
r_zkrp_lv_~n	2.542957	0.0835873	2.358402	0.1164944	2.773373	0.1262723
r_rvn_xml_~m	2.512531	0.0979581	2.459346	0.1448473	2.527826	0.1330126
r_ivn_chrn~c	2.138936	0.0870757	1.751552	0.1249876	2.486878	0.1238219
r_kiv_chrn~m	2.480395	0.0952088	2.21683	0.13781	2.698216	0.1309665
r_pltv_chr~r	2.452249	0.0929487	2.197204	0.1301213	2.67914	0.1301887
r_ods_mklv~n	2.463905	0.0851288	2.191847	0.1176856	2.702299	0.1219031
r_lgnsk_dnck	2.749755	0.0836834	2.50053	0.108679	2.987871	0.1373119
r_dnprv_za~v	2.261195	0.0782431	2.055526	0.1119646	2.426131	0.1090777
ss_urban	0.2766597	0.063269	0.3679464	0.0894979	0.1614269	0.0941768
ss_small_t~n	0.0571507	0.1173688	0.1556974	0.1563995	-0.0347212	0.1847638
ss_town	0.227845	0.0666487	0.2546188	0.0928744	0.1978939	0.09353
ss_city	0.1932915	0.0565226	0.2039367	0.0812461	0.1971912	0.0813295
ss_large_c~y	0.8817005	0.0707053	0.858804	0.1029749	0.9169771	0.1002332
cons	-5.797057	0.3598538	-4.07316	0.4440398	-7.704791	0.6398005
/athrho	0.5445462	0.2050572	-0.0299416	0.2600422	0.646501	0.2464554
/lnsigma	-0.7748684	0.0432723	-0.7482176	0.0219152	-0.9076021	0.0495396
rho	0.4964215	0.1545241	-0.0299327	0.2598092	0.5693098	0.1665758
sigma	0.4607644	0.0199383	0.4732092	0.0103705	0.4034906	0.0199888
lambda	0.2287334	0.0805101	-0.0141644	0.12297	0.2297111	0.077488
	Wald test (rho=0): chi2(1)= 7.05 prob>chi2=0,0079		Wald test (rho=0): chi2(1)= 0.01 prob>chi2=0,9083		Wald test (rho=0): chi2(1)= 6.88 prob>chi2=0,0087	

Appendix 3.

Correlations between computer usage and settlement size:

	comp_use	ss_village	ss_urban	ss_small _town	ss_town	ss_city	ss_large_city
comp_use	1						
ss_village	-0.1734	1					
ss_urban	-0.0192	-0.2629	1				
ss_small_t own	-0.0369	-0.1331	-0.0685	1			
ss_town	-0.0164	-0.2594	-0.1336	-0.0676	1		
ss_city	0.0791	-0.3584	-0.1845	-0.0934	-0.1821	1	
ss_large_ci ty	0.1739	-0.3481	-0.1792	-0.0907	-0.1768	-0.2443	1

Correlations between wages and settlement size:

	ln_wage	ss_village	ss_urban	ss_small_town	ss_town	ss_city	ss_large_city
ln_wage	1						
ss_village	-0.1709	1					
ss_urban	-0.0196	-0.2193	1				
ss_small_town	-0.0585	-0.093	-0.0615	1			
ss_town	-0.0196	-0.2131	-0.1409	-0.0598	1		
ss_city	0.07	-0.3183	-0.2105	-0.0892	-0.2046	1	
ss_large_city	0.1544	-0.3275	-0.2166	-0.0918	-0.2105	-0.3144	1

Appendix 4. IV estimation procedure

COEFFICIENT	Full sample ln_wage	Men ln_wage	Women ln_wage
comp_use	1.582***	2.121***	1.017**
	-0.47	-0.78	-0.44
e_secondary	-0.0485	-0.0445	-0.0477
	-0.061	-0.097	-0.064
e_ptu_without	0.0743	0.0728	0.0956
	-0.067	-0.11	-0.066
e_ptu_with	0.074	0.0565	0.0294
	-0.055	-0.094	-0.057
e_technical	-0.129	-0.15	-0.0968
	-0.084	-0.14	-0.08
e_higher	-0.247*	-0.442	-0.14
	-0.15	-0.27	-0.14
age	0.0636***	0.0780***	0.0356***
	-0.014	-0.022	-0.014
age2	-0.000662***	-0.000824***	-0.000391***
	-0.00015	-0.00024	-0.00015
fulltime	0.504***	0.655***	0.439***
	-0.063	-0.12	-0.058
female	-0.269***		
	-0.036		
e_usage	-0.383**	-0.722**	-0.118
	-0.17	-0.33	-0.15
o_manager	-0.0728	-0.268	0.193
	-0.19	-0.32	-0.17
o_proffesional	-0.227	-0.348	0.00142
	-0.18	-0.29	-0.14
o_technician	-0.217	-0.224	-0.109
	-0.14	-0.21	-0.12
o_clerk	-0.279*	-0.296	-0.101
	-0.14	-0.21	-0.12
o_service	-0.0769	-0.201	-0.0367
	-0.11	-0.21	-0.12
o_craft	0.0349	-0.0869	0.145
	-0.1	-0.15	-0.13
o_machinery	0.0748	-0.0627	0.291*
	-0.11	-0.16	-0.16
o_unskilled	-0.07	-0.139	-0.0197
	-0.098	-0.15	-0.12
o_armed	0.292	-0.11	0.614*
	-0.19	-0.32	-0.32
s_2_4	0.0798	-0.00272	0.145
	-0.17	-0.34	-0.17
s_5_9	0.0222	-0.0274	0.0702

	-0.17	-0.35	-0.17
s_10_19	0.0372	-0.00399	0.113
	-0.17	-0.34	-0.16
s_20_49	-0.0157	-0.0927	0.095
	-0.17	-0.34	-0.16
s_50_99	0.0149	-0.121	0.149
	-0.18	-0.35	-0.17
s_100_249	0.0378	-0.112	0.175
	-0.18	-0.36	-0.17
s_250_499	0.0659	-0.00472	0.145
	-0.18	-0.35	-0.17
s_500_999	0.103	-0.0341	0.25
	-0.18	-0.35	-0.17
s_1000	0.0917	-0.00012	0.172
	-0.18	-0.36	-0.17
r_krim	-0.155*	-0.441**	-0.0253
	-0.087	-0.19	-0.08
r_zkrp_lv_vl_trn	-0.179***	-0.334***	-0.0733
	-0.067	-0.12	-0.085
r_rvn_xml_jtm	-0.161*	-0.0467	-0.230**
	-0.085	-0.15	-0.091
r_ivn_chrn_vnc	-0.209***	-0.203	-0.177**
	-0.08	-0.14	-0.09
r_kiv_chrng_cm	-0.226***	-0.325**	-0.168**
	-0.077	-0.13	-0.083
r_pltv_chrks_kir	-0.0723	-0.227*	-0.0193
	-0.084	-0.13	-0.11
r_ods_mklv_xrsn	-0.131*	-0.254**	-0.0383
	-0.068	-0.12	-0.075
r_lgnsk_dnck	-0.0762	-0.0744	-0.137
	-0.071	-0.11	-0.085
r_dnprv_zaprj_xrkv	-0.183***	-0.241**	-0.132*
	-0.065	-0.12	-0.068
i_industry	0.179**	0.309***	-0.0207
	-0.072	-0.11	-0.084
i_electricity	0.242***	0.429***	0.00366
	-0.088	-0.12	-0.14
i_construction	0.133	0.18	0.00337
	-0.11	-0.17	-0.12
i_sales	0.11	0.00566	0.0838
	-0.081	-0.15	-0.082
i_transport	0.169*	0.407***	-0.0805
	-0.091	-0.13	-0.12
i_financial	0.033	0.426**	-0.146
	-0.14	-0.21	-0.17
i_public_adm	-0.127	0.257	-0.307**
	-0.12	-0.16	-0.15
i_educ_health	0.0211	-0.179	-0.0994

	-0.072	-0.13	-0.07
i_other_municipal	-0.0301	-0.089	-0.0853
	-0.095	-0.19	-0.095
i_other	-0.179	-0.147	-0.0694
	-0.17	-0.24	-0.24
Constant	3.570***	3.268***	4.050***
	-0.37	-0.61	-0.4
Observations	2598	1203	1395
R-squared	-0.55	-1.09	-0.09
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1			
Test of excluded instruments	F(5, 2545)=3.18 Prob >F =0.0073 Robust F-stat=3.18	F(5,1151)=1.77 Prob>F=0.1170 Robust F-stat=1.77	F(5, 1343)=1.93 Prob>F=0.0871 Robust F-stat=1.93
Weak identification tests Ho: equation is weakly identified Weak-instrument-robust inference			
Anderson-Rubin test	F(5,2545)=6.39	F(5,1151)=4.88	F(5, 1343)=3.17
Cragg-Donald F statistic (weak identification test):	3.064	1.77	1.8

Appendix 5. Description of the variables

e_under_10	dummy variable which is equal to 1 if individual's education is less than 10 grades and 0 otherwise
e_secondary	dummy variable which is equal to 1 if individual's education is secondary and 0 otherwise
e_ptu_without	dummy variable which is equal to 1 if individual's education is PTU without secondaryand 0 otherwise
e_ptu_with	dummy variable which is equal to 1 if individual's education is PTU with secondary and 0 otherwise
e_technical	dummy variable which is equal to 1 if individual's education is technical and 0 otherwise
e_higher	dummy variable which is equal to 1 if individual's education is higher and 0 otherwise
married	dummy variable which is equal to 1 if individual is married and 0 otherwise
age	age of the respondent
age2	age square
fulltime	dummy variable which is equal to 1 if individual works fulltime and 0 otherwise
female	dummy variable which is equal to 1 for female and 0 for male
comp_use	dummy variable which is equal to 1 if individual uses computer and 0 otherwise
e_usage	dummy variable which is equal to 1 if individual used Internet during the last year and 0 otherwise
o_skilled	dummy variable which is equal to 1 if individual is skilled and 0 otherwise
o_manager	dummy variable which is equal to 1 if individual works as a manager and 0 otherwise
o_professional	dummy variable which is equal to 1 if individual works as a professional and 0 otherwise
o_technician	dummy variable which is equal to 1 if individual works as a technician and 0 otherwise
o_clerk	dummy variable which is equal to 1 if individual works as a clerk and 0 otherwise
o_service	dummy variable which is equal to 1 if individual works in employed in services and 0 otherwise
o_machinery	dummy variable which is equal to 1 if individual works in employed in machinery and 0 otherwise
o_unskilled	dummy variable which is equal to 1 if individual is unskilled and 0 otherwise
o_armed	dummy variable which is equal to 1 if individual is employed in army and 0 otherwise
s_1	dummy variable which is equal to 1 if firm size consists of 1 worker and 0 otherwise
s_2_4	dummy variable which is equal to 1 if firm size consists of 2-4 workers and 0 otherwise
s_5_9	dummy variable which is equal to 1 if firm size consists of 5-9 workers and 0 otherwise
s_10_19	dummy variable which is equal to 1 if firm size consists of 10-19 workers and 0 otherwise
s_20-49	dummy variable which is equal to 1 if firm size consists of 20-49 workers and 0 otherwise

s_50-99	dummy variable which is equal to 1 if firm size consists of 50-99 workers and 0 otherwise
s_100_249	dummy variable which is equal to 1 if firm size consists of 100-249 workers and 0 otherwise
s_250_499	dummy variable which is equal to 1 if firm size consists of 250-499 workers and 0 otherwise
s_500-999	dummy variable which is equal to 1 if firm size consists of 500-999 workers and 0 otherwise
s_1000	dummy variable which is equal to 1 if firm size consists of more than 1000 workers and 0 otherwise
r_kiev	dummy variable which is equal to 1 if the region is Kiev and zero otherwise
r_krim	dummy variable which is equal to 1 if the region is Krim and zero otherwise
r_zkrp_lv_vl_trn	dummy variable which is equal to 1 if the region is Zakarpatie, Lvovskaia, Volynskaia, Ternopilskaia and zero otherwise
r_rvn_xml_jtm	dummy variable which is equal to 1 if the region is Rovnenskaia, Khmelnickaia, Zhitomirskaia and zero otherwise
r_ivn_chrn_vnc	dummy variable which is equal to 1 if the region is Ivano-Frankivskaia, Cernoveckaia, Vinnickaia amd zero otherwise
r_kiv_chrng_cm	dummy variable which is equal to 1 if the region is Kievskaia, Chernigovskaia, Sumskaia and zero otherwise
r_pltv_chrks_kir	dummy variable which is equal to 1 if the region is Poltavskaia, Cherkasskaia, Kirovogradskaia and zero otherwise
r_ods_mklv_xrsn	dummy variable which is equal to 1 if the region is Odesskaia, Nikolaevskaia, Khersonskaia and zero otherwise
r_lgnsk_dnck	dummy variable which is equal to 1 if the region is Luganskaia, Doneckaia and zero otherwise
r_dnprv_zaprj_xrkv	dummy variable which is equal to 1 if the region is Dnepropetrovskaia, Zaporozhskaia, Kharkovskaia and zero otherwise
i_agriculture	dummy variable which is equal to 1 if individual is employed in agriculture and 0 otherwise
i_industry	dummy variable which is equal to 1 if individual is employed in industry and 0 otherwise
i_electricity	dummy variable which is equal to 1 if individual is employed in elecricity sector and 0 otherwise
i_construction	dummy variable which is equal to 1 if individual is employed in construction sector and 0 otherwise
i_sales	dummy variable which is equal to 1 if individual is employed in sales and 0 otherwise
i_transport	dummy variable which is equal to 1 if individual is employed in transport and 0 otherwise
i_financial	dummy variable which is equal to 1 if individual is employed in financial sector and 0 otherwise
i_public_adm	dummy variable which is equal to 1 if individual is employed in public administration sector and 0 otherwise
i_educ_health	dummy variable which is equal to 1 if individual is employed in education and health sectors and 0 otherwise
i_other_municipal	dummy variable which is equal to 1 if individual is employed in municipal sector and 0 otherwise
i_other	dummy variable which is equal to 1 if individual is employed in other sectors and 0 otherwise

ss_village	dummy variable which is equal to 1 if the settlement size is a village and zero otherwise
ss_urban	dummy variable which is equal to 1 if the settlement size is an urban settlement and zero otherwise
ss_small_town	dummy variable which is equal to 1 if the settlement size is an small town (up to 20 thds) and zero otherwise
ss_town	dummy variable which is equal to 1 if the settlement size is a town (20-99 thds) and zero otherwise
ss_city	dummy variable which is equal to 1 if the settlement size is a city (100-499 thds) and zero otherwise
ss_large_city	dummy variable which is equal to 1 if the settlement size is a large city (more than 499 thds) and zero otherwise