ORIGINAL ARTICLE   **Open Access**

Gender differences in earnings and labor supply in early career: evidence from Kosovo's school-to-work transition survey

Francesco Pastore[1*], Sarosh Sattar[2] and Erwin R Tiongson[2]

* Correspondence:
francesco.pastore@unina2.it
[1]Seconda Università di Napoli, Italy
Full list of author information is
available at the end of the article

Abstract

Very little is known about gender wage disparities in Kosovo and, to date, nothing is known about how such wage disparities evolve over time, particularly during the first few years spent by young workers in the labor market. More generally, not much is known about gender wage gaps in early career worldwide, a period which is perceived to be an important determinant of the overall gender wage disparity. This paper analyzes data from the *School-to-Work Transition* (SWT) survey, an unusual survey conducted by the ILO between 2004 and 2006 in eight countries, including Kosovo, that documents the labor market experiences of the youngest age segment in the labor force (age 15–25 years). The results of the analysis suggest that, on average, women have lower education attainment than men but this educational disparity is masked among the sample of employed men and women who tend to be well-educated. The consequences of this dramatic segmentation of labor market participation are striking. On average, there is little or no gender wage gap. The results of the Juhn et al. (J Polit Econ 101:410–442, 1993) decomposition analysis reveals that gender wage differences are almost entirely driven by differences in characteristics (rather than either the returns to those characteristics or the residual). The greater average educational attainment of employed women, among other characteristics, tends to fully offset the gender wage gap. Not surprisingly, the returns to women's education among employed women are low because there is little variation in educational attainment among the sample of well-educated employed women. When the analysis controls for sample selection bias and heterogeneity, the returns to women's education rise, confirming the lower productivity-related characteristics of non-employed women compared to employed women. The relatively small sample constrains a fuller analysis of the emergence of the gender wage gap, which, according to a small but growing international literature, typically materializes during childbearing years.

JEL codes: I21, J13, J15, J16, J24, J31, J7, P30

Keywords: Gender wage gap and dynamics, Early labor market outcomes, School-to-work transitions, Earnings equations, Decomposition analysis, Balkans area, Kosovo

1. Introduction

In the Western Balkans region, Kosovo has been hit hard by various shocks over the last two decades. As a result, in addition to numerous consequences of previous conflict and crises, the labor market is presently characterized by substantial imbalances along several dimensions, with gender disparities possibly being the most dramatic. As reported in a number of policy documents (e.g., UNFPA 2007), women are often the object of discrimination and violence; they also represent the weakest segment of the working population.

Existing studies[1] of gender differences in Kosovo have noted that in contrast to gender disparities in other developing countries, the women in Kosovo are at a disadvantage compared to men both in terms of their educational attainment and their employment prospects. However, it is not clear how this affects gender differences in earnings, as available databases to date typically do not provide information on wages in Kosovo. As a result, little is known about the gender wage gap in Kosovo. In addition, in Kosovo and in many other countries more generally, little is known about the evolution of gender differences in wages (as well as in labor supply) over the life cycle. An understanding of the first few years in the labor market is particularly important because it is likely that the wage growth during this period is "the primary cause of the overall gender pay gap," as has been argued recently[2].

Data collected by the ILO on School-to-Work Transitions (SWT) in 8 countries, including Kosovo, provide an opportunity to analyze the earnings differentials between men and women, particularly in their early labor market experience. ILO collected the data between 2004 and 2006 covering young workers aged 15–29, and recently made the micro data available to researchers. As the name of the survey suggests, the ILO collected information on the experiences of young workers at the beginning of their career, as they left school and entered the labor market. The ILO conducted the SWT survey in September and October 2004 in 5 out of 7 regions in Kosovo (Pristina, Mitrovica, Gijlan, Gjakova, and Prizren)[3]. In Kosovo's case, available SWT data allow an analysis of early careers only up to age 25, because the final sample includes only individuals aged 15–25 (as opposed to 15–29 in other countries). Although the SWT data are by now a few years old, the general lack of publicly available data on earnings or wages in Kosovo means that the availability of SWT data still provides a rare opportunity to look at gender wage differentials in Kosovo.

The results of the analysis reveal that women tend to have lower education attainment not only in the adult population, as noted in a recent (World Bank 2012) report, but also among those in the youngest age group. In general, young women tend to be disproportionately represented among those with low educational attainment. However, this educational disadvantage disappears when one considers only the employed men and women. This provides compelling evidence of the existence of a notable segmentation between employed and non-employed young women in Kosovo. The former are generally much better educated, at the same level or more than their contemporary men, and have higher than average productivity characteristics. The latter, instead, have a much lower educational level and productivity characteristics than their male counterparts.

The consequences of this segmentation on labor market participation and wages in early career are striking. Among employed men and women in their late teens (age 15–19), there is no evidence of a gender wage gap in favor of men. This indicates that in the period prior

to the childbearing years, women tend to fare as well as their male counterparts. The results of Juhn et al. (1993) decomposition analysis reveal that differences in productivity characteristics, rather than in prices or in residual inequality, drive the gender gap. The greater mean educational attainment of employed women tends to compensate for the higher average number of working hours of their male counterparts.

The last part of the analysis attempts to assess the impact of observed and unobserved differences between employed and non-employed women on the returns to education and the gender wage gap. Interestingly, when we control for selection bias, the gender wage gap for the SWT sample as a whole and for the component of the young adults is still not statistically significant, despite the lower productivity characteristics of the non-employed women compared to employed women. This may be due to the relatively small sample size. At the same time, the returns to women's education rise when selection bias is taken into account, confirming the lower productivity-related characteristics of non-employed women compared to employed women.

The rest of the paper is outlined as follows. Section 2 reviews the existing literature on gender disparities in Kosovo's labor market. Section 3 presents the methodology adopted in the econometric analysis. Section 4 discusses the main advantages and shortcomings of the Kosovo SWT survey database. Section 5 reports the main empirical findings. The final section provides some concluding remarks.

2. Literature review: gender issues in Kosovo's labor market

2.1 Educational attainment

Women's educational attainment compares unfavorably with men in Kosovo, according to a recent report (World Bank 2012). The report finds that there are distinct gender differences in literacy and educational attainment in favor of men. In particular, 2.2% of adult men are illiterate compared to 7.2% of women; 37.3% of men have only basic education compared to 61.7% of women; and 11.9% of men possess a university degree compared to only 5.5% among women. More generally, the educational level of Kosovo's population is relatively low by international standards. Half of the population has only 9 years of schooling, equivalent to basic education. Although educational attainment has been improving slightly in recent years, the share of university graduates remains higher among the adult population compared to the youth population, because the educational system was interrupted by the war and because of the high emigration rate of some of the most educated members of the younger generation.

Figure 1 reproduced from (Pastore 2011) provides a graphical representation of gender gaps in educational attainment among young people across eight countries for which SWT data are available, namely Azerbaijan, China, Egypt, Iran, Kosovo, Mongolia, Nepal and Syria. The panels are focused on the following levels of education: panel (a) on primary education or below; panel (b) on low secondary education; panel (c) on high secondary education; panel (d) on tertiary and higher education. The female share is measured along the vertical axis while the male share is measured along the vertical axis. The diagonal line represents the case of exact gender equality. Points above the diagonal line represent countries that offer an advantage to women and those below the diagonal line are those that offer an advantage to men. An advantage at lower education levels of course implies a disadvantage at higher educational attainment levels. China, Syria and Mongolia can be found below the diagonal line in panel (a), which

Figure 1 Educational levels by gender. Source: own elaboration on SWT surveys.
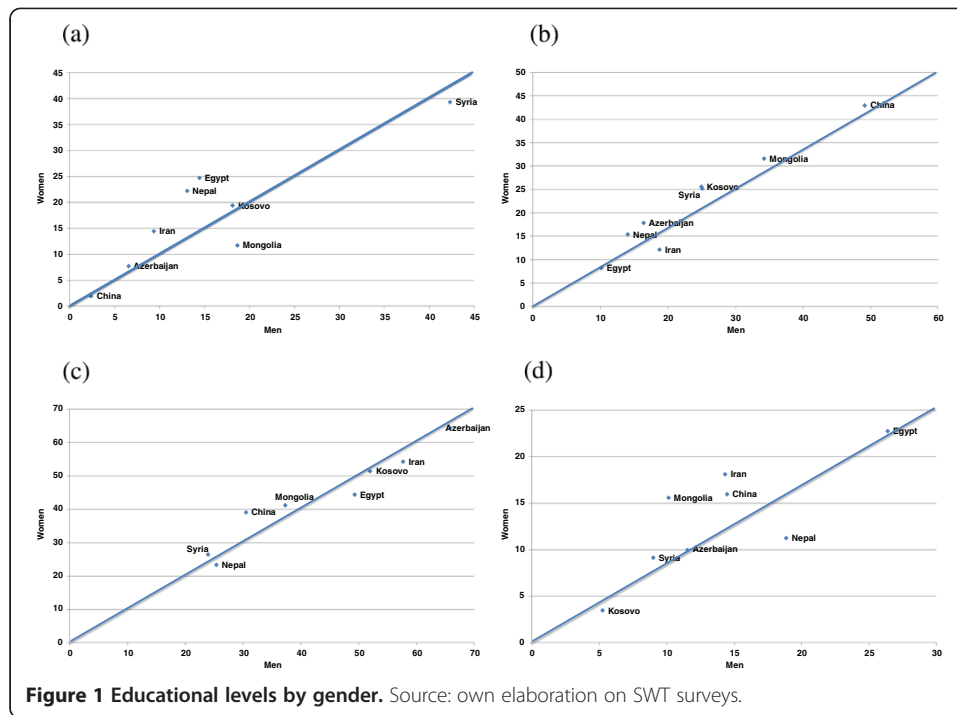
suggests that the shares of women holding primary education or below is lower than aver-age in these countries; in turn, women enjoy some advantage at higher educational levels. In fact, women' advantage in these countries is substantial, starting from low secondary education, but especially in terms of high secondary and tertiary education or above. In the case of Kosovo, the educational disadvantage of women is quickly disappearing among the youngest generation, as the country can be found very close to the diagonal line. The figures also confirm the low average educational level of the youth population. Kosovo children tend to be found on the upper right side of the panels (a), (b) and (c), and at the bottom left side of panel (d).

2.2 Employment

In addition to gender disparities in educational attainment, the report suggests that women in Kosovo are less likely to be employed[4]. More generally, labor market out-comes in Kosovo compare unfavorably with those in Europe and women are particu-larly at a disadvantage. The report indicates that 34.1% of men are employed compared to only 11.5% of women. In addition, women represent less than a tenth of all entrepre-neurs and are under-represented among top management (0.3%). Meanwhile, Kosovo's unemployment rate at 45.4% is at the highest end of the distribution of unemployment rates in the region. Women, young workers, and the least educated have the highest probability of being unemployed. Unemployment rates for women gradually fell in the years leading to the global financial crisis but mostly due to decreasing activity rates ra-ther than due to rising job opportunities. Most of the unemployed have never worked previously. Furthermore, about half have been looking for work for four years or more. Unemployment is very high among young people aged 15–24 which as a group repre-sents close to a fifth of Kosovo's population.

Several studies cited by the report indicate that workers in the private sector typically do not receive a basic benefits package (e.g., maternity leave, paid annual leave, paid sick leave and unpaid leave). This effectively weakens the productivity of all workers and, in particular, women, because they are more likely to benefit from leave options. According to the report, although a new Labor Law was enacted in 2010 that mandates the provision of basic benefits, some worry that the new law is not sufficiently enforced.

2.3 Earnings

In this difficult environment, very little is known about gender wage disparities or even about earnings in general. We know of only a handful of studies that examine wages systematically in Kosovo[6]. These include a labor market study conducted by the World Bank in 2003 based on 2002 Labor Force Survey data; a global study of gender wage disparities (Oostendorp 2009) in 63 countries, including Kosovo, based on 2001 LSMS data; a study of earnings among emigrants, based on a survey conducted by the Riinvest Institute along Kosovo's borders (Havolli 2011); and a very recent paper on returns to education based on the 2002 Riinvest Household and Labor Force Survey (Hoti 2011).

The results of the few studies of Kosovo that exist suggest the following: The overall gender wage differential is either small or insignificant in Kosovo. In the private sector, it is equivalent to 8 percent and in the public sector it is insignificant (World Bank 2003). Although the sample of emigrants in the Riinvest Institute survey is likely very different from the general population of domestic workers, the results of the analysis of emigrant earnings are nonetheless potentially instructive, and they suggest that there are no statistically significant differences in earnings by gender, though there are gender differences in the likelihood of sending home remittances (Havolli 2011). The global study of gender wage gaps also suggests that Kosovo is at the lower end of the distribution (Oostendorp 2009).

Although the general lack of a gender wage gap may be interpreted by some as evidence of gender equality, one possibility – given the differences in educational attainment and employment – is that the empirical study of earnings largely ignores issues related to sample selection biases. In other words, the women who are typically included in empirical analyses of earnings may be systematically different from unemployed women who are, by construction, left out of the sample. The recent study of returns to education cited earlier (Hoti 2011) in fact, provides evidence that workers, particularly women, are not randomly selected into employment. The study used a Heckman procedure to account for this selection bias, but the study does not at all explore gender wage differentials.

Taken all together, this suggests that a new systematic study of gender wage differentials would be a valuable contribution to the literature. Such a study would have to carefully account for likely selection biases. In addition, a study of the first few years spent by young workers in the labor market—as they transition from schools to the labor market—may be warranted given the importance of the relationship between education and employment. A small but growing international literature in fact suggests that wage developments in this "early career" period are critical drivers of the overall gender wage gap. For example, one very recent cross-sectional study of male and

female professionals in the U.S. financial sector suggests that although wages are essentially equal at the beginning of their careers, the women's earnings soon lag behind those of their male counterparts[7]. It finds that the divergence in wages is due in part to differences between men and women in the type of training received prior to the completion of an MBA (and rising returns to this training), differences in career interruptions (and large penalties associated with such interruptions), and differences in the number of hours worked, typically around the childbearing years. However, another recent strand of the literature suggests that there is a large, measurable gender wage gap, already present at entry into the labor market—largely driven by occupational segregation by gender—which then persists throughout people's careers[8]. In any case, a study of this early career period is potentially illuminating.

3. Methodology[9]

To shed light on the stylized facts discussed above, this paper will first estimate extended and augmented wage equations to assess the impact of different determinants of wages on the gender wage gap and quantify returns to education by gender. Second, we implement several decomposition techniques to assess the role of – both quantity and price – determinants of wages at the mean and at different quantiles of the distribution. Third, as noted in the previous sections the participation rates of both men and women are very low and in order to assess the possible impact of the ensuing sample selection bias on their relative wages, we estimate (Heckman's 1979) corrected earnings equations and then assess the importance of selection bias on wages by way of the Neumann and Oaxaca (2004) decomposition.

3.1 Baseline approach: Mincerian wage equations

The Mincerian approach to estimating the returns to human capital and the gender wage gap is used as baseline framework. As is well known, the *basic* Mincerian earnings equation amended to include also control for gender differences is as follows:

$$lnw_i = \alpha + \gamma W_i + rS_i + \delta EXP_i + u_i \tag{1}$$

where w_i are monthly labor earnings for an individual i, W_i represents the gender dummy variable, S_i represents a measure of the years of schooling, EXP_i is a measure of work experience and u_i is a disturbance term representing other variables that are not explicitly measured and assumed to be independent of the other regressors. Work experience is typically included also as a quadratic term to account for the concavity of the earnings profile, which would be consistent with decreasing returns to human capital. However, because this sample covers younger workers alone, the quadratic term is omitted here. The parameters of interest are clearly γ and r, namely the coefficient of the gender dummy and the private financial return to schooling or, equivalently, the proportionate effect on wages of incremental increases in S.

The above earnings function is in fact a log-linear transformation of an exponential function and can be estimated by OLS, with the coefficients requiring a semi-elasticity interpretation. That is, as is well known, they measure the percentage change in the dependent variable for any unit change in any independent variable, holding other things constant[10].

The paper also provides estimates of the *augmented* version of the Mincerian specification of the earnings equation:

$$lnw_i = \alpha + \gamma W_i + rS_i + \delta EXP_i + \sum_{j=1}^{n} \beta X_{ij} + \mu_i \tag{2}$$

where $\sum_{j=1}^{n} \beta X_{ij}$ is a set of other variables assumed to affect earnings. Among such variables, we include mainly pre-work individual characteristics, such as civil status, having children, ethnic differences, type of job search adopted, time spent waiting before finding a job, aspirations in life. In the final specification, we also include such job characteristics as whether the young person has a work contract, the type of contract (whether written or oral), the duration of the contract, the ownership, size and industry of the firm where the worker is employed, and the type of on-the-job training received if any.

3.2 The Juhn, Murphy and Pierce decomposition

This paper uses the Juhn et al. (1993) methodology to disentangle the components of the wage gap[11]. It differs from the classical Oaxaca and Ransom decomposition because it decomposes the wage gap not only at the mean, but along the entire wage distribution, thereby accounting for role of the residual unexplained component, measured by the residual wage distribution. In this study, these properties of the Juhn et al. (1993) decomposition may be useful, considering the absence of a gender wage gap at the mean of the distribution and toned to ascertain whether this holds along different quantiles of the distribution.

Without discussing all the analytical details, it is sufficient to mention that in order to capture gender differences at different quintiles of the wage distribution, two different equations will be estimated as a first step, namely the male and the female wage equations, both of the type described in equation [2], but without the gender dummy. Intuitively the decomposition allows disentangling the following components:

$$\Delta lnw_m = lnw_m - lnw_w = \hat{\beta}_m \left(\bar{X}_m^q - \bar{X}_f^q \right) + \bar{X}_m^q \left(\hat{\beta}_m - \hat{\beta}_f \right) + \left(\hat{\mu}_m - \hat{\mu}_f \right) \tag{3}$$

where the hat denotes estimated coefficient, the horizontal bar signifies the mean characteristics and q indicates the corresponding quintile of the wage distribution. The row wage gap (Δlnw_m) is decomposed into the following: a) an endowment effect, measured by differences in observed characteristics weighted by the estimated coefficients of the male earnings equation; b) a remuneration effect, measured by the estimated differences in prices of the observed characteristics; c) an unobservable effect, measured by the difference in the estimated residuals. These are usually called, respectively, the quantity effect (Q), the price effect (P) and the residual wage distribution, which is unobserved. The male distribution is used as a reference category. Findings are presented at the mean and at different percentiles of the wage distribution, namely 25, 50 and 75.

3.3 The Machado and Mata decomposition

A shortcoming of the Juhn et al. (1993) decomposition is that it does not allow an assessment of the size of the gender wage gap along the entire wage distribution. As such, we also implement the (Machado and Mata 2005) decomposition analysis. It allows the definition of different counterfactual distributions on the basis of conditional wage distributions at different percentiles.

We attempted to implement the van Albrecht et al. (2009) version of the (Machado and Mata 2005) decomposition. This version controls for sample selection bias along the entire wage distribution. We expect that this might be the case considering how gender differences vary along the distribution of educational attainment. Nonetheless, likely due to the small number of observations, we failed to obtain any useful results and therefore we turn now to the more traditional analysis of sample selection bias using the Heckit methodology, which looks only at the mean value of the wage distribution.

3.4 Sample selection correction

The last part of the analysis aims to correct for sample selection bias the baseline estimates drawn from the Mincerian earnings equation. Toward this end, the standard (Heckman 1979) methodology is implemented. In brief, the basic intuition of this model is the following: the baseline earnings equation is missing some variables that may explain women's (labor force) participation. If one finds factors that help explain female participation but are not statistically related to their wages, it makes it possible to estimate a specific equation, the so-called participation or selection equation. This equation has, as a dependent variable, a dummy variable taking a value of one if an individual is employed and zero otherwise and, thus, yields a measure of the probability of participating in the labor market (the inverse Mills ratio). Adding this variable to the main equation—in this case, the earnings equation—would provide a test of the existence of sample selection bias and correct the coefficients in case the added variable proves to be statistically significant. Analytically, model [2] consists now of two equations, whereas the main equation is the usual earnings equation, which can be estimated by OLS, but there is a selection equation of the following type that can be estimated by Probit:

$$LS = 1 \, if \, \sum_{j=1}^{m} \gamma_j Z_j + v \geq 0 \tag{4}$$

where LS represents labor supply. Sample selection arises when there is correlation between u and v. In this case, and assuming that u is independent of $z\gamma$, namely the covariates of the selection equation, the expected value of lnw_i, will be:

$$E(\ln w_i | \Sigma Z_j, v) = \alpha + rS_i + \delta EXP_i + \Sigma_{j=1}^{n} \beta X_{ij} + \rho \lambda \left(\Sigma_{j=1}^{m} \gamma_j Z_j \right) \tag{5}$$

where ρ is the correlation between the error terms of the main and of the participation equation and λ is the inverse Mills ratio evaluated at the mean of the covariates $(z\gamma)$. This equation shows that when there is sample selection bias, one should include the term $\lambda(z\gamma)$ as an additional regressor in order to obtain unbiased estimates of the parameters of interest. If $\rho=0$, $\lambda(z\gamma)$ does not appear in the main

equation, and OLS will provide consistent estimates of the parameters of interest. However, if ρ is statistically different from 0, ignoring the λ(zγ) term is equivalent to omitting a statistically significant variable, and the coefficients will be biased. The inverse Mills ratio can be estimated by Probit in the first step and in the second step the correlation with the γ variable can be estimated by OLS, a two-step method that is sometimes called the Heckit. In addition to this two-step procedure, the two equations above can be simultaneously estimated by maximum likelihood, as done in this study.

Following Wooldridge (2003, p. 588) and Cameron and Trivedi (2009), the independent variables in the main equation, x, should be a strict subset of the independent variables in the selection equation, z. In other words, the selection equation should include all (and only) the variables contained in the main equation – because the explanatory variables in the main equation usually affect also the probability of selection through wages, plus instrumental variables that affect participation but not earnings. Otherwise there could be a high correlation between λ and x, leading to multicollinearity in the main equation.

Among the main weaknesses of the Heckit procedure is the lack of suitable instruments for labor force participation. Though this shortcoming cannot be overcome, we experiment with different solutions and instrumental variables and compare and contrast findings based on different sets of estimates in order to gain more information on the underlying mechanisms at work.

As a first step, we provide maximum likelihood estimates with no exclusion restrictions. As explained in Cameron and Trivedi (2009, p. 543), among others, in this case, we exploit the exogenous variation due to the high non-linearity of the selection equation as a determinant of labor force participation. This is not without cost. The maximum likelihood estimates are based on a bivariate normality assumption. This is the reason why a two-step procedure, based on a univariate normality assumption is considered to be more reliable. Nonetheless, the inverse Mills ratio might be highly collinear with the independent variables of the main equation if the set of variables in the selection equation are exactly the same. This suggests that it is appropriate to use instruments to identify the selection equation.

As a second step, following existing approaches in the literature, we use as instruments the number of sons/daughters. We also test other instrumental variables, such as the number of brothers and the number of sisters. Generally speaking, these variables should be taken with caution, especially in the case of a sample of young people. In the case of the number of children, for example, women may start to earn less over time because they pay a price for child-bearing (they earn less because they take more absences, have to take maternity leave, for example) as documented in the literature. If this is the case, then the number of children might be correlated with wages. In fact, for this young age group, the number of children is more likely than usual to be correlated with wages: women that have children at a very young age are probably a very different group from those who do not, in ways that probably reflect expected payoffs from the labor market. In addition, in the case under scrutiny, the sample includes only few women with children[12]. Third, this variable might be irrelevant in the case of men.

Third, we use responses to a survey question on the main aspirations in life of young people as possible determinants of their labor supply. More specifically, we use the answer: "making a contribution to social life" and "gaining all types of experiences in life". On one hand, they can easily be capturing issues of motivation, attitudes, non-cognitive skills (like team work, persistence, etc.) that would be correlated with wages; on the other hand, we think that the link is likely not very strong at the age of individuals in the sample and might instead reflect unobserved characteristics that affect the decision to work/not work, rather than the type of job and therefore the wage/job package per se.

All this considered, we treat the sample selection correction estimates based on the use of instruments only as a robustness check of the maximum likelihood estimates.

4. Data

As previously discussed, this paper uses Kosovo's SWT data. The data is at an individual level [13], with a sample size of 1,336 young people aged 15 through 25 years. The survey thus includes "teenager workers", aged 15–19 years, and "young adults" (20–25 years)[14]. The sample size is the smallest of all the ILO SWT surveys (SWTS), but still is quite numerous for this specific segment of the labor force, generally under-represented in labor force surveys of the type.

By construction, the survey is statistically representative of the age group considered and by gender in the provinces considered. Not all provinces are included in the sample, however, and, unfortunately, no information is available on the location of respondents.

For all the above reasons, considering also the dramatic lack of information on the countries considered, the SWTS represents a valuable source of information on every aspect of the demand and supply of youth labor in Kosovo. A key advantage of the SWTS is the availability of information on earnings, which are missing in other sample surveys recently run in the country, the only exception being the Riinvest Household and Labor Force Survey, which was conducted in December 2002 (used in Hoti, 2011).

The natural logarithm of monthly wages is the dependent variable of our Mincerian earnings equations [15]. The weekly hours worked are used as a regressor, rather than as the denominator of monthly wages, in order to control for gender differences in the number of hours worked. In addition, hourly wage rates are not widely used in Kosovo.

Table S1 and S2 in the Additional file 1 provides descriptive statistics for the sample of men and women in wage employment and non-employment, respectively. The tables provide an overview of the variables available in the data.

Moreover, this paper uses a definition of actual, rather than potential work experience, which is in fact more properly a definition of job tenure in the available data. In particular, we use the following question: "When were you recruited in your current job?" Considering the young age of individuals in the sample and the relatively small share of temporary work arrangements, this is a quite accurate proxy also of work experience.

The data allows the measurement of "overeducation" and "overskilling." The former is based on question g30 ("Do you feel the education/training you received in the past was useful in getting your present job?") and the latter is based on question g34

("Do you feel the past education/training you received has been useful in your current activity?"). Those answering "not useful" to the former question are considered overeducated and those answering "not really relevant" to the latter question are considered overskilled.

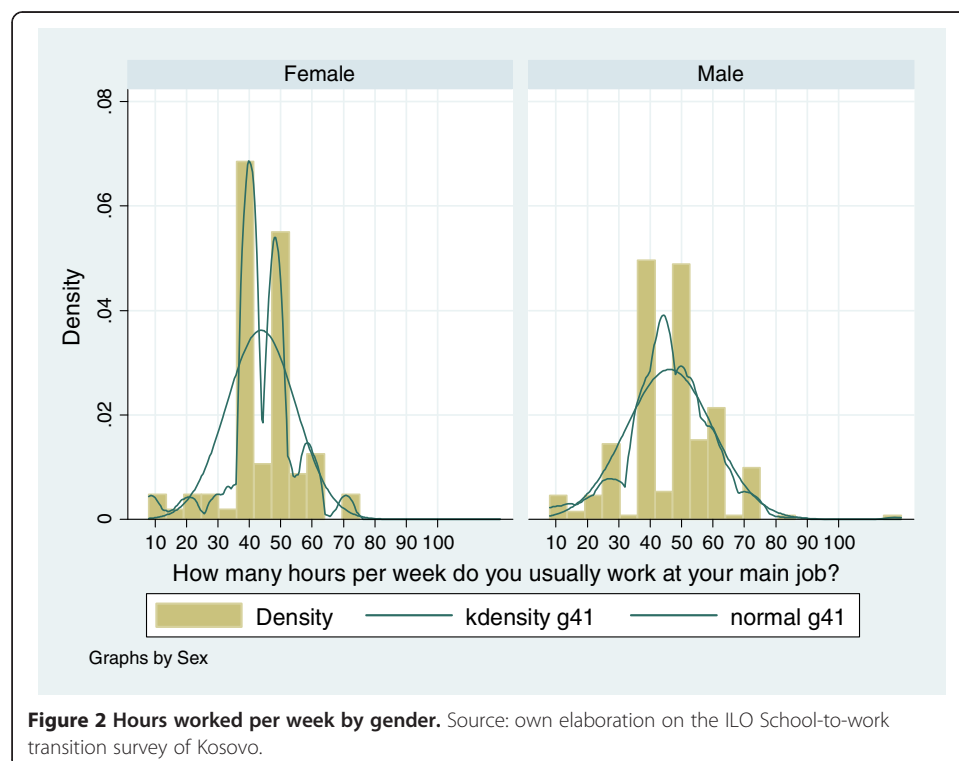The other variables used in the estimates are self-explanatory.

5. Main findings

5.1 Descriptive analysis

Table S1 in the Additional file 1 provides descriptive statistics for dependent and independent variables by gender relative to the sample of wage employees. With a share as high as 34.5%, men's employment rate exceeds that of women by about 6 percentage points.

On average, wages are slightly higher for women, while the number of hours worked is slightly higher for men [16]. Already at this early stage in their career, women tend to work about two and a half hours fewer than men: 43.79 weekly hours against 46.18 hours, with an average of 45.12 hours. This is apparent from inspection of Figure 2, which shows the distribution of weekly hours worked by men and women. The Figure indicates that women tend to work mainly around 40 and 50 weekly hours, whereas men tend to be more frequent in correspondence to a higher number of working hours.

Going back to Table S1 in the Additional file 1 it is noticeable that, while receiving higher wages, employed women also have a much higher educational attainment level than men. This confirms the impression that there is a strong difference between the

Figure 2 Hours worked per week by gender. Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

average educational attainment of the sample population (Figure 1) and that of the employed women (Table S1).

The share of women holding a university degree is three times greater than that of men, although both are very low not only by EU standards, but also by the standards of other developing countries. Only 3% of employed men and about 9% of employed women have university degrees, In addition, the share of employed women holding a high school diploma (61.9%) is higher than that of men (58.2%). Of course, it follows that men are disproportionately represented among those with the lowest levels of educational attainment, particularly among the illiterate and those with primary education only. The difference in the shares of individuals with elementary education only is about 7%.

This pattern is interesting considering that, as shown in Figure 1 above, for the general youth population, women tend to have the same educational levels as men or only slightly lower. In turn, this suggests that there might be sample selection mechanisms at work—in other words, women possessing higher educational attainment have a much greater chance of finding employment. This might be due either to their higher productivity, as postulated in the human capital model, or to some unobserved characteristics, such as their greater motivation and commitment to work. Testing these alternative hypotheses will be one of the objectives of the next section.

Men are also found to be more frequently "overeducated" and, to a greater extent, "overskilled." In other words, they report themselves to be holding jobs for which their degree was neither necessary *to get* the job, nor useful *to do* the job. Almost one of every two young employed men declares himself as overskilled; the comparable figure for women is only 33%. These are both phenomena that denote a lower human capital level. As recent theories of overeducation argue (see for a recent overview, Leuven and Oosterbeek 2011), individuals accepting jobs that are below their own educational level likely signals an awareness that they possess a lower level of human capital or skills than what their educational attainment otherwise signifies. Holding constant the level of education attainment, the lower human capital might depend either on the lower quality of their education or on the lack of the work-related component of human capital, such as work experience. Clearly, in the case of young people, previous work experience is very important to access jobs.

Despite this, in the sample under consideration, men report slightly longer job tenures. On average, they have been working in the same job for 1.8 years, whereas women have been working in the same job for 1.6 years. In general, the average job tenure is quite low because of the nature of the sample. The range of variation of job tenure is between 0.01 and about 8 years in the case of men and between 0 and 7.3 years in the case of women.

Young women tend to be more frequently engaged in a relationship with a partner or to be married compared to men of the same age, which suggests that they likely choose older partners. Over 25% of women are reportedly in a stable relationship against 11.5% for men. Most women who are not single tend to be engaged rather than married. The gender difference in the share of workers engaged is about 10% (13.7% of women and 3.6% of men). The corresponding difference in terms of the share of workers who are married is only 4% (11.3% of women and 7.3% of men). Not surprisingly, at that age, the share of individuals who are divorced or separated is very small.

Consequently, employed young women are also more likely to have children than their male counterparts: 8.3% against 4.8%. The information on when respondents had their first child is not available, but the survey does contain a question regarding the age at marriage. (This information is used in the next section).

As reported in Table S1, Albanians are about 90.4% of the population, the second largest ethnic groups being Serbians and Turkish with shares of 4% and 3% respectively. There are more Roma men than women in the sample population, likely reflecting the difficulty of finding Roma women available for the interview. Bosnians are less than one per cent of the sample.

Men and women use slightly different search methods to find jobs. Considering the entire sample of young men and women, the vast majority of young people (53.5%) rely on their network of family and friends. Direct contacts with prospective employers (19.8%) and answering advertisements (15.0%) are the second largest methods of job search. This is a sign of the lack of (public) intermediaries in the labor market able to bring vacancy news to unemployed young people. Men tend to use more frequently than women their network of friends and relatives. Similarly, they also tend to find a job through direct contacts with employers. Women found their job more frequently through advertisements and via information gathered at educational institutions.

Women are more likely to have contract jobs than men and they tend to have written rather than oral contracts compared to men. Only 19.6% of women accept to work without contract, against a share of 21.8% in the case of men. About 65% of women work under a written contract against 43.6% of men. One possibility might be that women tend to be more "risk averse" in terms of labor market experience and contract jobs may be thought to provide better maternity leave benefits and other forms of assistance compared to non-contract jobs or jobs based on oral contracts only. In addition, women tend to choose jobs in the state sector, where legal arrangements are naturally more binding. In addition, about 25% of employed women in the sample have a job in the state sector; as a result, there are of course fewer women in the private sector. Women are also likely to be found in larger firms: 32.1% of women work in firms with more than 20 employees against a share three times smaller of men. If women are indeed more risk averse, one possibility is that smaller firms are perceived to provide lower job security.

Not surprisingly, women tend to prefer longer term work arrangements than men, if they have a temporary contract. In fact, while the share of women with a permanent contract is relatively low, that of women holding a temporary contract of between 1 and 3 years of length is greater.

Women also find jobs within a shorter time period than men. About 18.5% of women found their job in less than a month, against a share of 10.3% for men.

There are notable gender differences in industry of employment. Female-dominated industries include trade fair, finance, public administration, social services; male-dominated industries include manufacturing, trade, hotels and restaurants. Other industries have more or less an equal share of young people by gender. The most striking difference with respect to other transition countries in the region is perhaps the tendency of men in Kosovo to be more likely employed in the trade, hotels and restaurants sectors, those that have experienced high income growth over the transition period.

Finally, as much as 34.3% of employed women tend to participate in training programs compared to only 27.6% of men. Women are especially more likely to engage in non-apprenticeship training.
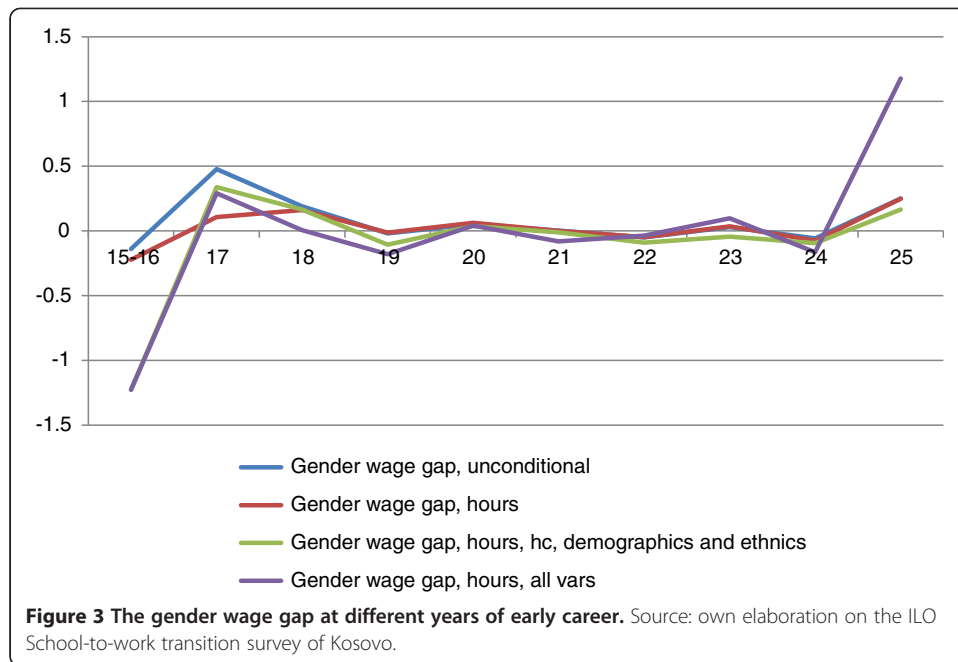
5.2 The gender wage gap

The unconditional gender wage gap for the entire sample of employed workers – including wage employees and self-employed – is −0.0842 and is statistically significant with a significance level of 10%.

When the sample includes only individuals in wage employment, the (unconditional) coefficient of the gender gap turns positive and equal to 0.025 and is not statistically significant. This gender gap is lower than expected, but one should consider that the sample includes only young people, surveyed at a time when they had only recently started working. This is also a period when, as noted in the previous section, only less than 5% of men and 8.3% of women have children, and therefore the female disadvantage at the labor market is less pronounced.

Figure 3 shows the evolution of the gender wage gap by years of age. The unconditional gender gap is compared with different conditional estimates, although there are little differences among them. The figure highlights the fact that among the youngest age segments, the gender gap tends to be positive, and hence in favor of women, whereas at later ages it approaches zero.

Table 1 reports the gender wage gap by age group and with varying levels of control for explanatory variables. The upper panel of the Table uses monthly wages and the lower panel uses hourly wages. The Table shows that, whatever the measure adopted, although differences are not statistically significant, women fare slightly better than their male counterparts when both are under the age of 20, whereas women fare slightly worse than their male counterparts, when they are in their twenties. The gender premium in hourly wages goes up from about 8% (or zero, because it is significant) to 52% among teenagers, when including all the available control variables. However, this last finding may be affected by the small number of observations.

The differences in gender wage gaps between teenagers and young adults point to possible differences, at least for some young people, in family life and decisions to marry and have children. As already noted above, the data set allows knowing the age at first marriage, but not that of having children. In fact, as noted previously only about 8.3% of employed women and 10.4% of non-employed women have children. Assuming that women and men tend to have children soon after marriage, the variable age at marriage might shed useful information on the small number of married women. Figure 4 shows the age distribution by gender of the first marriage. In panel (b) and (c), the Figure also distinguishes employed and non-employed women. The Figure shows that women marry much earlier than men, especially the non-employed ones, which suggests that for at least some of the young female population, marrying (and having children) early in life might not be conducive to productive employment. The average age at marriage is 19.1 for men and 19.7 for women. This rises to 20.5 for the employed men and 20.1 for the employed women: employed women marry later than their male counterparts. In contrast, age at marriage is much lower for non-employed women (18.7) than for non-employed men (19.2). Taken

Figure 3 The gender wage gap at different years of early career. Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

all together, this suggests a strong correlation between the decision to enter the labor market and the decision to start a family. Women who work tend to postpone marriage for work-related reasons, more so than men.

Considering the age at marriage of all the workers in Kosovo, it is worth noting that the decision to marry is taken not much later than the age at marriage of this sample of young men and women. According to the Statistical Office of Kosovo (SOK), the average age at marriage in Kosovo is 27 for women and 30 for men. The highest number of marriages is

Table 1 The gender wage gap by age group

	All	Teenagers	Young adults
	Monthly wages		
Monthly wages, unconditional	0.025	0.0624	0.0064
Monthly wages, lhours	0.0323	0.0657	0.0106
Monthly wages, lhours, education	−0.0068	0.0502	−0.0187
Monthly wages, lhours, education, human capital	−0.0113	0.0205	−0.0261
Monthly wages, lhours, education, human capital, demographics	−0.0211	−0.031	−0.0215
Monthly wages, lhours, all controls[1]	−0.0066	0.4002	−0.043
Number of observations	315	58	257
	Hourly wages		
Hourly wages, unconditional	0.0624	0.0769	0.055
Hourly wages, lhours, education	0.0226	0.1006	0.0125
Hourly wages, lhours, education, human capital	0.014	0.1294	−0.0066
Hourly wages, lhours, education, human capital, demographics	−0.0153	0.1143	−0.0373
Hourly wages, lhours, all controls[1]	−0.0025	0.5201**	−0.0706
Number of observations	315	58	257

Note: * p<.1; ** p<.05; *** p<.01.
[1] The controls included in the last row of each group of estimates are the same as those included in the augmented wage equations in Table 4.
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

Figure 4 Age at marriage by labor market status and gender. Source: own elaboration on the ILO School-to-work transition survey of Kosovo.
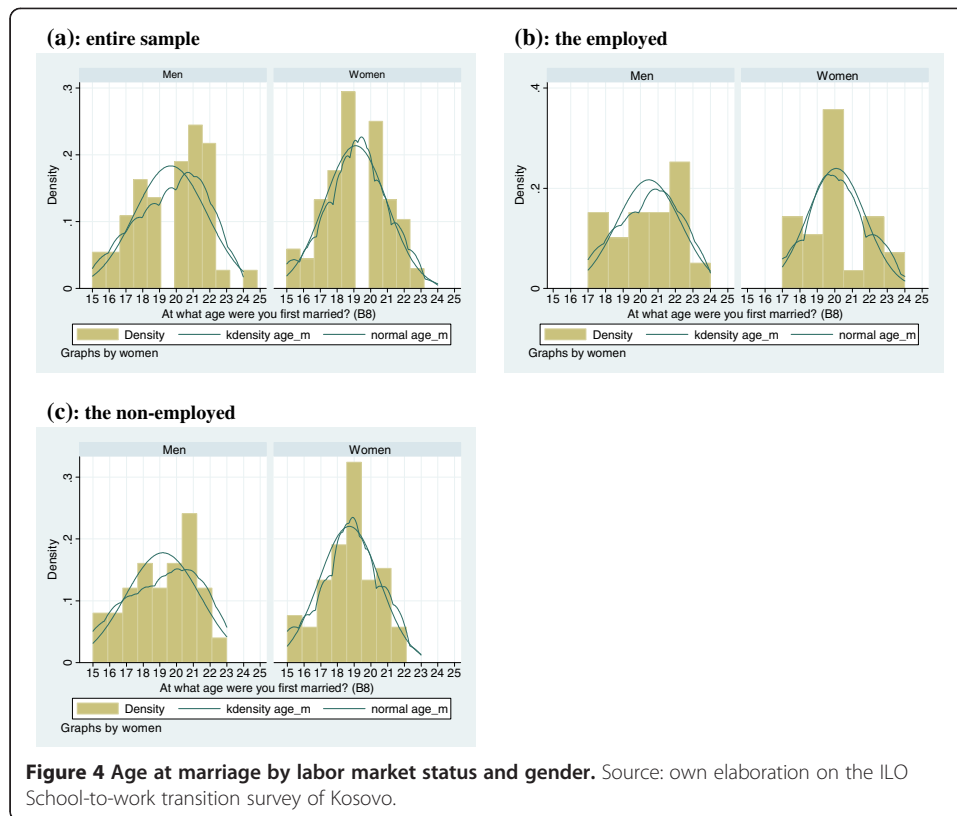
for women in the age group 20–29 and for men in the age group 25–34. According to statistics for 2005, the average age for marriage recently increased by 11 months for men and by 14 months for women. Nonetheless, in Kosovo the practice of marrying prior to reaching the majority age of 18 is still relatively common, especially in rural areas.

5.3 Mincerian earnings regressions: returns to education

Table 2 presents the first set of estimates of the returns to years of education by age group and gender. Each row represents a different set of estimates, with an increasing number of controls. The first row refers to the basic earning equations, the second row to the extended version and the following rows to different versions of the augmented earnings equations. The Table reports only the returns to years of schooling. It shows that the returns to education are quite low in Kosovo, at about 5% per year of schooling. Similar to other countries, the annual rate of return is higher for women than for men in the entire sample and in each of the age groups considered. However, the returns to education might be far from linear and returns might differ by type of degree.

Table 3 reports the results of the augmented Mincerian earnings equations[17]. The coefficient of determination is higher for women this time. The coefficient of the gender dummy does not seem to be affected by inclusion of all observed characteristics of men and women, remaining statistically insignificant. But this does not necessarily imply the absence of a gender gap; as noted previously, employed women have a higher educational level than men and, therefore, are expected to have higher wages.

Table 2 Returns to years of education by age group

	All	Women	Men	Teenagers	Teenagers (women)	Teenagers (men)	Young adults	Young adults (women)	Young adults (men)
Monthly wages, lhours, years of education	0.0489***	0.0655***	0.0364***	0.0137	0.0471	−0.0038	0.0516***	0.0656***	0.0407***
Monthly wages, lhours, years of education, human capital	0.0458***	0.0575***	0.0394***	0.0043	0.0301	−0.0076	0.0498***	0.0616***	0.0451***
Monthly wages, lhours, years of education, human capital, demographics	0.0460***	0.0591***	0.0386***	0.0171	0.0647*	−0.0101	0.0487***	0.0610***	0.0433***
Monthly wages, lhours, years of education, all controls	0.0358***	0.0275	0.0353**	−0.0208	−0.2983	0.1363	0.0309***	0.0271	0.0271

Note: * p<.1; ** p<.05; *** p<.01.
The controls included in the last row of each group of estimates are the same as those included in the augmented wage equations in Table 4.
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

Table 3 Augmented Mincerian earnings equations (aged 15–25 years)

	All	Men	Women	Teenagers	Young adults
Women	−0.0006	(omitted)	(omitted)	0.2898	−0.0355
Natural log of weekly hours	0.2146***	0.3200***	0.0945	0.2893	0.2270**
Education (baseline: elementary or below)					
Vocational	0.1095	0.0982	−0.0884	0.9722	0.0866
Highsec	0.2001***	0.1824	0.0942	0.493	0.1716**
University or above	0.3666***	0.6241**	0.0881	(omitted)	0.2955**
Overeducation	−0.085	−0.0973	−0.096	−0.1951	−0.0914
Overskilling	0.0758	0.1812*	−0.0361	0.3605	0.0829
Job tenure	−0.0157	0.0078	−0.0406*	−0.0025	−0.0082
Civil status (baseline: single with no children)					
Women with children	−0.0461	(omitted)	−0.0535	(omitted)	−0.0219
Men with children	−0.038	−0.1014	(omitted)	(omitted)	−0.0504
Women engaged	0.0941	(omitted)	0.0685	2.7038**	0.0953
Men engaged	0.0986	0.0087	(omitted)	(omitted)	0.1212
Women married, separated, divorced	0.1093	(omitted)	0.1379	(omitted)	0.0835
Men married, separated, divorced	0.0154	0.0121	(omitted)	−0.6939	−0.0104
Ethnic group (baseline: albanian)					
Serbian	−0.5123***	−0.2827*	−0.6319***	−0.822	−0.4077***
Roma	0.0902	0.0644	0.5021*	−0.3615	0.2258*
Turkish	−0.0681	−0.0923	0.0079	0.3328	−0.0285
Bosnian	0.0758	(omitted)	0.1312	−0.8624	(omitted)
Registered at the Public employment Services when unemployed	0.0133	0.1056	−0.0228	−0.4964	−0.0373
Type of job search used to find the current job (baseline: through own' network of relatives and friends)					
Educational institutions	−0.1028	−0.0712	0.0115	1.2673	−0.0231
Direct contact with employer	−0.0693	−0.0472	−0.0164	−0.4706	−0.0748
Search through the PES	−0.0325	(omitted)	−0.0766	1.063	0.0085
Labor contractor	−0.1084	−0.3537	0.0735	0.3732	0.0638
Other search method	−0.1611	0.0146	−1.1659***	0.6252	(omitted)
Job advertisement	−0.0146	0.0997	−0.0437	−1.1517	0.0089
Type of working contract (baseline: no contract)					
Written contract	0.0891	0.0177	−0.0684	0.1381	0.0049
Oral contract	0.0486	−0.0035	−0.0823	−0.4332	−0.0044
Length of contract (baseline: no contract)					
Unlimited	0.0844	0.1804	0.1689	0.0317	0.1341*
1 year or less	0.0823	0.1651	0.2345	0.3315	0.1019
From 1 to 3 years	0.0919	0.1472	0.2238	0.1663	0.0948
Firm's ownership (baseline: private company)					
Family business	0.008	0.1596	−0.0463	0.2726	−0.0621
Government / public sector	−0.2194***	0.0837	−0.3152***	−2.175	−0.1652*
Multinational corporation	0.5323**	(omitted)	0.4287*	−1.7853	0.6153**
Joint venture	0.0327	−0.0135	0.1385	1.6849	0.302

Table 3 Augmented Mincerian earnings equations (aged 15–25 years) *(Continued)*

Nonprofit organization	0.1493	0.2569	0.0211	−0.4166	0.3106**
Other	0.2184	0.2457	0.1746	(omitted)	0.1321
Size of firm (baseline: less than 5 workers)					
From 5 to 9 workers	0.2094***	0.2988***	0.1662*	0.3255	0.2596***
From 10 to 19 workers	0.1159	0.1436	0.1810*	0.0128	0.1595**
More than 20 workers	0.1865***	0.1645	0.2921***	1.1227*	0.1620**
Does not know the size of the firm	0.1495	−0.0285	0.2955**	0.492	0.1449
Sector of activity (baseline: trade fairs)					
Finance	0.4182***	0.7041***	0.2201	(omitted)	0.3760**
Public administration	−0.0105	−0.0981	−0.0104	(omitted)	−0.0326
Manufacturing	0.0883	0.1935	−0.0011	0.4889	0.0755
Trade, hotels and restaurants	0.0519	0.0738	0.1108	0.4532	−0.0235
Real estate	0.1611	0.1448	(omitted)	(omitted)	0.044
Transportation and telecommunications	0.2294**	0.3692**	0.1032	0.0259	0.1809*
Constructions	−0.0231	−0.1436	0.0943	3.0073**	−0.0376
Social services	−0.0798	0.0009	−0.1504*	0.6861**	−0.1259*
Electricity	0.0647	0.1288	(omitted)	(omitted)	0.0269
Moonlighting	0.1984*	0.2965	0.1202	(omitted)	0.145
Job satisfaction (baseline: mainly satisfied)					
Partly satisfied	−0.1663***	−0.1414	−0.1767***	−0.3472	−0.1414***
Not satisfied	−0.2897***	−0.3642*	−0.2343*	−0.7205*	−0.2048**
Waiting time before finding the current job (baseline: less than a week)					
1 through 4 weeks	−0.0222	−0.0799	0.0337	−0.8459**	−0.0239
From 1 to 2 months	−0.07	−0.0547	−0.0201	0.5689*	−0.1912**
From 3 to 6 months	0.0019	−0.0321	−0.0262	0.7232	−0.0556
From 6 to 12 months	−0.1017	−0.1858	−0.0388	0.4599	−0.1953**
More than 13 months	−0.0844	−0.1092	−0.0792	0.4616	−0.1682*
On-the-job training (baseline: no training)					
Apprenticeship type	0.0453	0.0541	0.1271	0.8838	0.1072
Non apprenticeship type	−0.0501	−0.1088	0.0538	−1.0425	−0.0248
Non job related training	−0.025	−0.1755	−0.0167	0.6032	−0.0606
Business development	0.0375	−0.0901	0.052	(omitted)	0.058
Accounting	0.173	(omitted)	0.326	(omitted)	0.1082
Union membership	0.0459	0.0909	0.1196	2.4004	0.0485
Constant	4.1311***	3.5603***	4.8066***	2.7633*	4.2217***
Nobs	315	154	161	58	257
R²	0.507	0.519	0.676	0.942	0.549

Note: * $p<.1$; ** $p<.05$; *** $p<.01$.
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

Interestingly, the hours worked weekly affect wages in a statistically significant way only in the case of men, not women. This is likely driven in part by the higher share of state sector jobs among women compared to men and the tendency of state sector jobs to pay a fixed salary independent of the hours worked.

For the sample of women, the educational variables are not statistically significant. This suggests that education provides access to jobs with better characteristics and, once we control for such characteristics, the relative return to education shrinks. More generally, the new controls for individual and job characteristics previously excluded are able to account for at least part of the impact of education on earnings. Again, for reasons that have already been discussed previously, job tenure does not seem to matter in terms of wages.

Overeducation generates a wage penalty of between 8% and 10%, similar to those found in other EU countries using the REFLEX data (McGuinness and Sloane 2010) though in the case of Kosovo it is not statistically significant.

Meanwhile, job tenure does not seem to affect wages in a statistically significant manner. One possibility is that average job tenure may be too short to have any impact on wages. Recall from the previous section that it equals on average 1.7 years. Interestingly, using the same type of data for Mongolia, Pastore (2011, p. 248) finds that declared work experience bears a wage premium only when it is long 5 years or more, which is rarely the case in the sample under consideration.

Civil status and children do not yield statistically significant wage effects, neither for women nor for men, although coefficients have the expected signs.

Wages are, ceteris paribus, much lower for Serbians, by about a half, a substantial amount considering the relatively low earnings of young people in Kosovo. This wage penalty is highly significant from a statistical point of view for men, at 65%, but less so for women. Other ethnic groups do not seem to suffer statistically significant wage differences, except for Roma adult women who declare on average higher than average wages.

The search method used to find a job does not seem to affect wages either. The baseline is having found their own job through the network of friends and relatives. Differences are small also for those who hold a contract versus their counterparts with no contract, although contracts of unlimited length and written contracts seem to correlate with higher than average wages.

Working in the governmental sector implies a wage penalty of about 20% for the entire sample. The effect is mainly due to women who bear a wage penalty of over 30%. This might depend on the type of job women tend to choose in the state sector, jobs that trade off lower pay for greater job security and legal commitment to maternity leave arrangements and maternity benefits in general. Working in a multinational corporation increases wages by about 50%. It is essentially women who work in multinational corporations when they are above 20 years old. Bigger firms pay higher wages, in general, although there are some fluctuations across firm groups.

The finance and transportation and telecommunication industries pay higher-than-average wages. One possibility is the presence of more foreign firms in these industries and foreign firms are documented to pay higher wages. In both cases the wage premium is essentially a male effect; the impact on female wages is not statistically significant.

Being dissatisfied with one's own job is correlated with a strong wage penalty, though some endogeneity might exist here. On-the-job training does not seem to affect wages in any measurable way.

5.4 Decomposition analysis

Table 4 reports the results of the Juhn et al. (1993) decomposition of the gender wage gap at the mean and at different percentiles of the wage distribution, using

Table 4 Juhn et al. (1993) decomposition

	T	Q	P	U
Mean	−0.0250	−0.0342	−0.0040	−0.0051
25th percentile	.0408	.0530	−.0063	−.0059
50th percentile	.0124	−.0541	.0876	−.0211
75th percentile	−.0488	−.1184	.0378	.0319

Note: T = total gender wage gap; Q = impact on the wage differential of individual characteristics; P = impact on the wage differential of prices to individual characteristics; U = impact on the wage differential of the residual wage distribution. The male distribution is used as a reference category.
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

the estimated male distribution as the one providing the baseline values of the estimated coefficients. Only wage employees are included in the estimates. Interestingly, the average effect, which is close to zero, is, in fact, the algebraic sum of gaps of different signs along the wage distribution. Although small, the gender gap is positive (in favor of women) at low wage levels, when the educational level of women is lower than that of men and becomes negative (in favor of men) and wider at higher wage levels, when the educational level of women is higher than that of men.

When looking at the mean gender wage gap, all of the effect seems to come from differences in characteristics, while differences in prices and the impact of the residual wage distribution seem to be irrelevant. The impact of the quantity effect is higher at all percentiles, but especially at the highest percentile considered, the 75th, where it is greater than 10% in the direction of a reduction of the gap. This effect is quite high considering the low incomes of young people in Kosovo. Assuming a monotonic relationship between education and earnings it is no surprise that at low wage levels, differences in individual characteristics are in favor of men: in fact, as documented above, men tend to have higher educational levels at low income (and therefore educational) levels and women at high income (and therefore educational) levels.
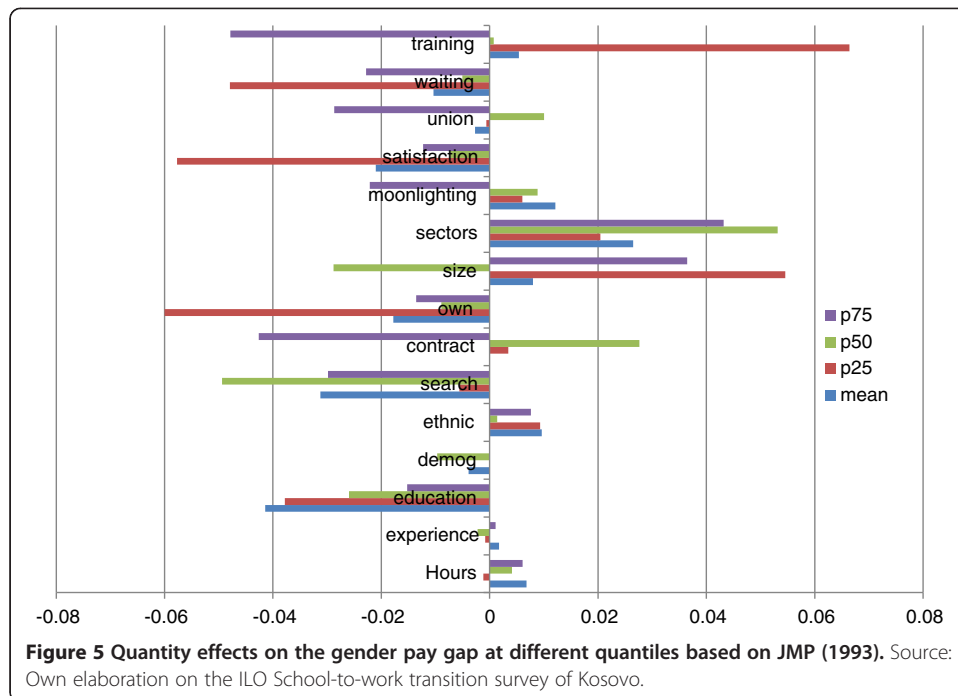
One can disentangle these quantity effects, as illustrated by Table 5 and Figure 5. The Figure represents the factors that tend to either increase or reduce the gap at the mean and at different percentiles of the wage distribution. Factors associated with a right-flagging bar tend to widen the gender gap, while factors associated with a left-flagging bar tend to narrow the gender gap. Among the factors that tend to increase the gap, there are training experiences (at high levels of the wage distribution), moonlighting, the sector of activity, the firm's size, the ethnic group and the weekly hours worked, although the impact is not very strong, but in few cases. In other words, if these were the only characteristics able to affect wages, the gender gap should increase overall by about 7% in favor of men at the mean of the wage distribution. In fact, men tend to work more hours, are more likely to moonlighting, include a smaller number of Serbian workers, and work in sectors that tend to provide wage premiums, such as private sector jobs in the finance' and the transportation and communication industry.

On the other hand, young women tend to be associated with factors and characteristics that likely decrease the gender wage gap. Among the factors that tend to reduce the gender gap include the educational qualifications of employed women, the greater degree of job satisfaction, the type of job search adopted to find a job, and the waiting time to find a job. Overall, these factors offset the gender wage gap by about 12.9 per cent.

Table 5 Juhn et al. (1993) quantity effects

	Hours	Experience	Education	Demog	Ethnic	Search	Contract	Own	Size	Sectors	Moonlighting	Satisfaction	Union	Waiting	Training
p25	−0.00118	−0.00082	−0.03779	0	0.009334	−0.0056	0.003418	−0.05999	0.054583	0.020446	0.006041	−0.05772	−0.00061	−0.04793	0.06642
p50	0.004138	−0.00225	−0.02595	−0.00966	0.001394	−0.0494	0.02765	−0.0089	−0.02883	0.053167	0.008843	−0.00727	0.010051	−0.00503	0.000748
p75	0.006075	0.00109	−0.01521	0	0.007633	−0.02982	−0.0426	−0.01356	0.036464	0.043174	−0.02212	−0.01228	−0.0287	−0.02281	−0.04787

Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

Figure 5 Quantity effects on the gender pay gap at different quantiles based on JMP (1993). Source: Own elaboration on the ILO School-to-work transition survey of Kosovo.
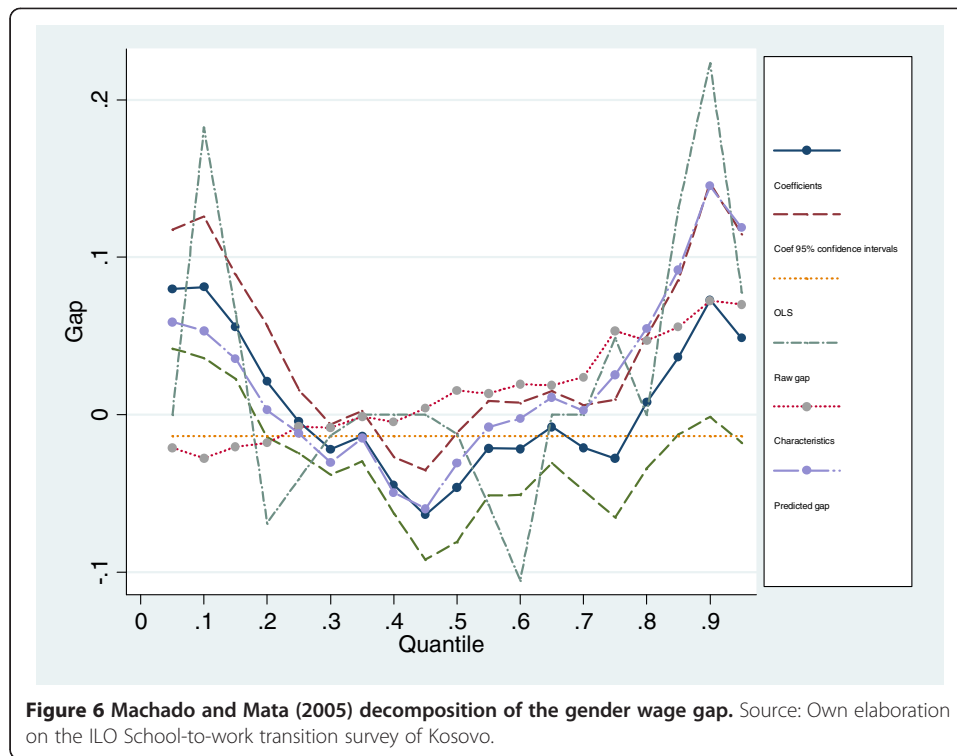
Results of the Juhn, Murphy and Pierce decomposition at different percentiles of the wage distribution suggest that the gender gap, which is not statistically significant at the mean, might be so at different percentiles of the wage distribution. A way to assess this is by estimating counterfactual distributions as based on Machado and Mata (2005). Figure 6 presents the results. It shows that while in fact the OLS measure of the gender gap is close to the horizontal axis, instead, the distribution of coefficients at different percentiles of the distribution tend to be higher than zero at the bottom and upper part of the distribution and to be negative between the 40th and 50th percentiles. In other words, women tend to earn more than men at both the lower and upper tail of the distribution. The higher wages of women in the upper tail of the distribution are due both to the better characteristics of women and partly to their better remuneration in that part of the distribution. In fact, the row gap is strongly in favor of women in that upper part of the wage distribution.

5.5 Sample selection procedure

The last step of the analysis is an attempt to assess whether some of the findings relative to employed young people apply also to the non-employed. In other words, given the differences in characteristics between employed and non-employed young people that we noted above, the returns to education, for example, may be different once we control for the heterogeneity of the non-employed.

Figures 7 and 8 confirm striking differences that exist between the employed and non-employed women and men. Figure 7 shows that the probability of falling into unemployment or inactivity is very high among those women who have low educational attainment, especially primary or below, but also vocational. The

Figure 6 Machado and Mata (2005) decomposition of the gender wage gap. Source: Own elaboration on the ILO School-to-work transition survey of Kosovo.
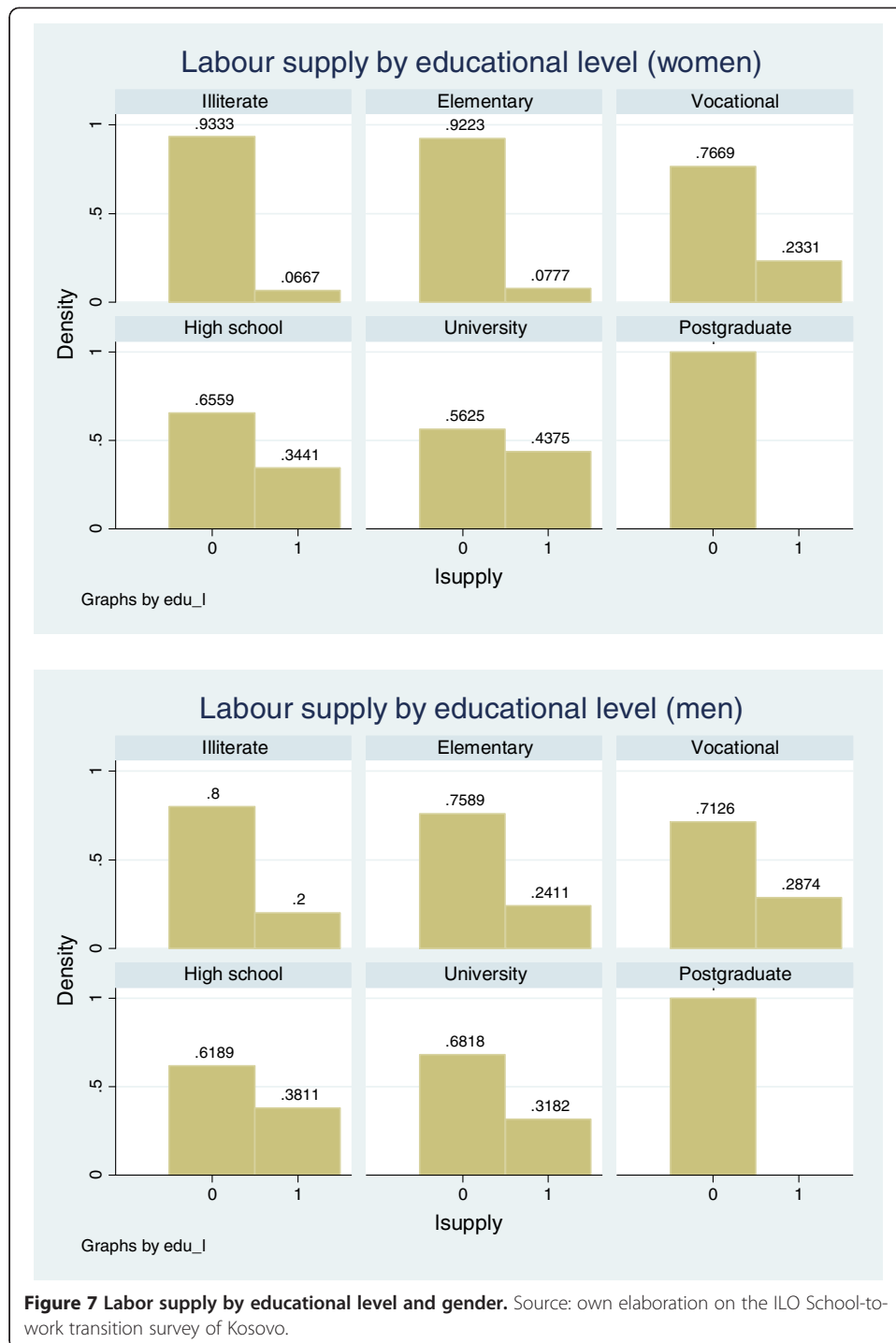
opposite holds true for women holding a university degree. This is due only in part to the fact that some of the women who hold the lowest educational attainment are current students. Not surprisingly, education is also positively associated with employment among men, but the effect is much less striking. Figure 8 shows that female labor supply is much higher among young adults compared to teenagers. The share of employed men is higher for every age group.

As a first step, we study the impact of different factors on labor supply decisions by gender by means of the (Fairlie 2005) decomposition analysis. For the sake of brevity, we do not discuss fully the methodology here[18]. It is sufficient to say that the method aims to assess the contribution of each determinant of labor supply on the gender gap. Since the probability of labor supply is estimated by a nonlinear model (LOGIT), the Blinder and Oaxaca decomposition provides unsatisfactory results. In contrast, the Fairlie decomposition was especially designed for nonlinear estimates. This preliminary step of the analysis will provide information useful also for the implementation of the ensuing sample selection correction estimates of the gender wage gap.
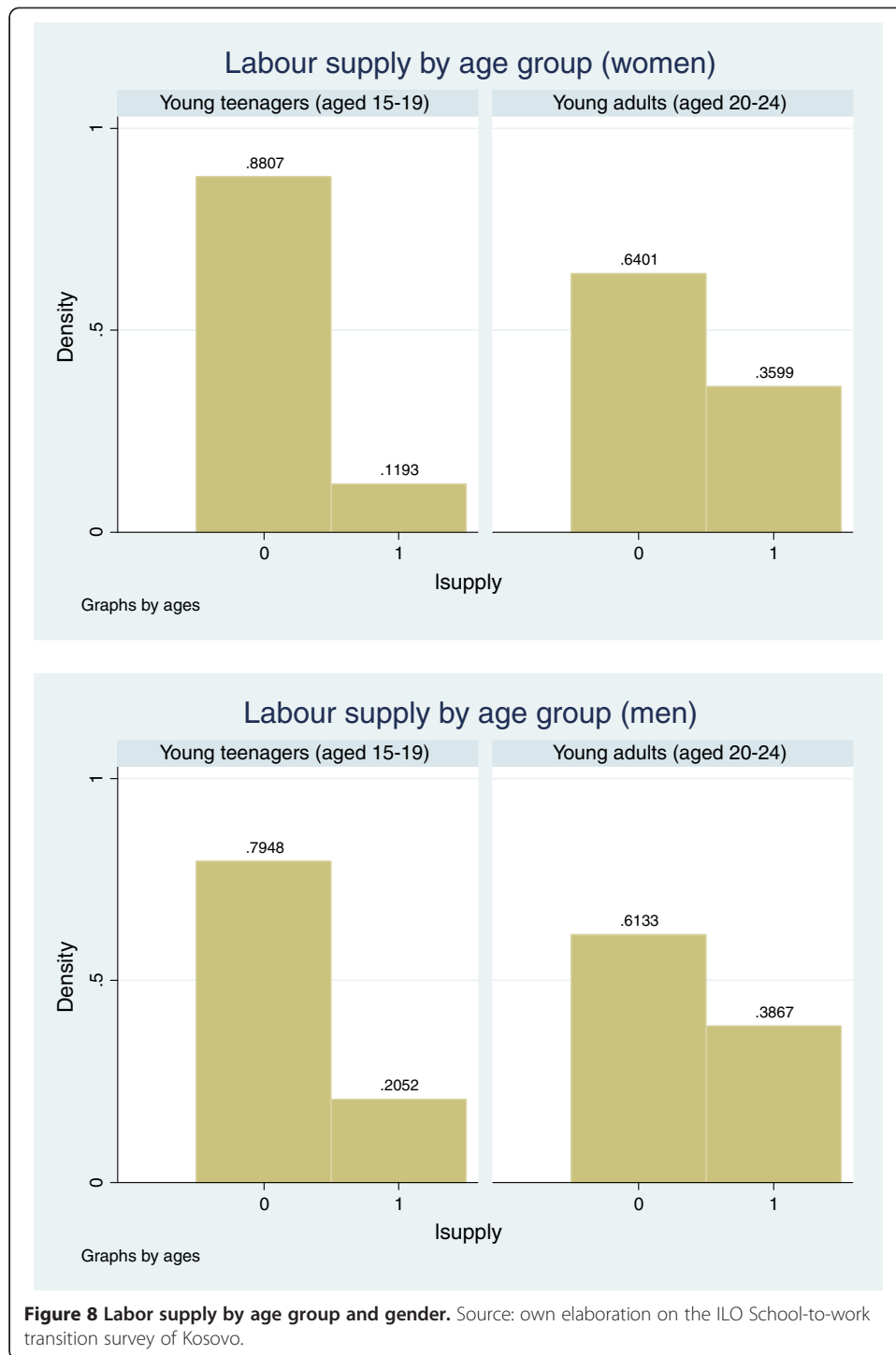
Table 6 reports the results of the (Fairlie 2005) decomposition. The probability of participating in the labor market among men is on average about 5.5% higher than women. The variables included in the model represent educational attainment (vocational education and high secondary education or above, the baseline being represented by education attainment below vocational education), ethnic group (Serbians, Roma, Turkish, Bosniac, the baseline being Albanians) and demographic factors (being engaged to be married, married, separated, divorced or widow, the baseline being the condition of single) seem to explain 100% of the gap. Demographic factors account for a sizeable share (about 85%) of the gender

Figure 7 Labor supply by educational level and gender. Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

gap in labor supply. Age explains about 9% of the gap and the ethnic composition of the sample by gender explains about 6% of the gap[19]. These findings can be interpreted as confirming women's lower likelihood of participating in the labor market, largely for family reasons.

This suggests that returns to education of men and women may be different once we control for sample selection bias. If the non-employed have lower

Figure 8 Labor supply by age group and gender. Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

productivity characteristics than the employed men and women, then, controlling for selection bias would yield higher returns to education, compared to the results of the standard OLS earnings equations. Similarly, the gender wage gap estimated by OLS might not be statistically significant because of the presence of sample selection bias.

Table 6 Fairlie (2005) decomposition of factors affecting the gender gap in female labor supply

Sample used for coefficients	Women/men pooled
	(1)
Number if observations	1336
Women	679
Men	657
Probability of labor supply of men	0.3255
Probability of labor supply of women	0.2709
Difference between men and women	0.0546
Explained component of the gap	0.0546
Variables:	
Age	.0047***
s.e.	(.0017)
% contribution to the gender gap	8.6%
Education	.0004
s.e.	(.770)
% contribution to the gender gap	0.8%
Ethnic composition	.0033***
s.e.	(.0013)
% contribution to the gender gap	6.0%
Demographic factors	.0461**
s.e.	(.0232)
% contribution to the gender gap	84.4%

Note: The educational variables include vocational education and high secondary education or above. The ethnic variables include dummies for Serbians, Roma, Turkish, Bosniac. The demographic variables include the number of children and civil status (engaged to be married, married, separated, divorced or widow).
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

As explained in the methodology section, the Heckman procedure to correct for sample selection bias is implemented using maximum likelihood estimates. The first exercise consists of using the simplest possible specification, one where we control only for age, education and ethnicity. The reason is that we should have the same set of variables in the main and in the selection equation and no job characteristics would be defined for the non-employed. In addition, we do not use instruments, but exploit the high non-linearity of the selection equation to predict labor supply decisions. Table 7 reports the findings.

Note that for the teenagers in the sample, the small number of observations available does not allow running separate estimates for men and women.

We find evidence of sample selection bias in almost all cases, but the gender wage gap seems to be unaffected by this in all the pooled estimates. This suggests that whatever the direction of the bias is–negative in some cases and positive in others– it is not changing in a statistically significant way the relative wage of men and women.

The returns to education are now higher for women than for men, especially the returns to tertiary education. This provides evidence of the lower educational attainment (as well as lower motivation and skill level) among non-employed

Table 7 Maximum likelihood estimate of the selectivity corrected earnings equation (Heckman procedure) – no instruments

Variable	All	Women	Men	All teenagers	All adults	Adult women	Adult men
Main equation. Dependent variable is hourly wages							
Women	0.0367			−0.026	0.0494		
Age	0.0860***	0.1308***	0.0082	−0.0208	0.0923***	0.1554***	0.0403
Education (baseline: Elementary or below)							
Vocational	0.2637**	0.6488***	0.0665	−0.0499	0.4569***	0.8431***	0.2342
High Secondary	0.4249***	0.8367***	0.1603	0.4166	0.5259***	0.8629***	0.2951*
University or above	0.7268***	1.0824***	0.7276***	(omitted)	0.8160***	1.1172***	0.7224***
Ethnic group (baseline: Albanian)							
Serbian	−0.7956***	−0.6752***	−0.6802***	−0.8126*	−0.7876***	−0.6139***	−1.0195***
Roma	0.4514*	0.9109**	0.3221	0.6510*	0.2229	0.2056	0.1375
Turkish	0.115	0.2555	−0.1675**	0.7622*	−0.0419	−0.0248	−0.051
Bosnian	−0.3634	−0.0566	(omitted)	0.1085	(omitted)	(omitted)	(omitted)
Constant	−2.9336***	−4.4493***	−0.4023	−1.128	−3.2014***	−5.0084***	−1.6971**
Selection equation. Dependent variable is the decision to participate to the labor market							
Women	0.0202			−0.2024	0.0799		
Age	0.1548***	0.1622***	0.1455***	0.1014	0.1767***	0.1977***	0.1418***
Education (baseline: Elementary or below)							
Vocational	0.3418**	0.6958***	0.1138	0.0413	0.5319***	0.8765***	0.2753
High Secondary	0.4410***	0.7975***	0.2141	0.7221***	0.4373***	0.7298***	0.2386
University or above	0.3191	0.8646***	−0.2526	(omitted)	0.3486	0.8419***	−0.2213
Ethnic group (baseline: Albanian)							
Serbian	−0.5306***	−0.258	−0.8551***	0.1696	−0.6355***	−0.3339	−1.0323***
Roma	0.1029	0.1203	−0.0207	0.4683	−0.111	−0.5477	−0.1254
Turkish	0.6036**	0.4961	0.647	1.5941***	0.2254	0.1108	0.2246
Bosnian	−0.2245	0.0231	−5.8114***	0.7713	−5.8998***	−5.5195***	−5.9794***
Constant	−4.2419***	−4.7359***	−3.8125***	−3.2221**	−4.7821***	−5.4903***	−3.7677***
athrho							
Constant	1.4061***	2.0961***	0.1174	1.4905**	1.6171***	2.0918***	1.3142***
lnsigma							
Constant	−0.3900***	−0.2229**	−0.8711***	−0.1494	−0.4095***	−0.2941**	−0.5743***
N	1250	640	610	454	796	399	397

Note: * p<.1; ** p<.05; *** p<.01.
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

women. Once this is taken into account, it is apparent that the return to education of women is much higher and might explain the recent tendency of young women to reach the same levels of education as men. Omitted estimates suggest that also the returns to a year of education increase in the sample selection corrected estimates relative to women, up to 8% per year in the estimates including all the women and up to about 9.8% per year for the young adult women. In

the case of men, the sample selection corrected estimates the returns to a year of education is not statistically significant for men and equal to 5.4% for the young adult' segment.

Finally, we estimate the Heckit using the instruments mentioned in the methodological section, namely controls for household characteristics and declared goals in life. We check in omitted estimates that goals in life do not affect wages. We test a number of instruments and estimate both the maximum likelihood and the two-step versions, but, to conserve space, we report only the former in Table 8[20]. We find evidence of sample selection bias in the case of both men and women; the sign of the correlation is negative in the case of women and positive in the case of men. The instruments used are mostly statistically significant. The main goals in life considered tend to negatively affect employment chances, especially in the case of men: if they think that contributing to social life and/or gaining all sorts of experiences is the main goal in life, then they tend to have lower employment chances. Not surprisingly, the number of kids affects negatively the employment chances of women, not men. Having more brothers is also associated with lower employment chances for young women, but not for young men.

Nonetheless, the gender wage gap is still not statistically significant, though the returns to education are higher for women than for men, in contrast to the OLS estimates.

The results of the two-step version of the sample selection procedure, which are not reported here, are partly similar and partly different. In contrast to the results of the maximum likelihood estimate, the two step procedures do *not* detect the existence of sample selection bias, likely because of the high standard errors typically found in this type of procedure. Nonetheless, the results relative to the gender gap and the returns to education by gender are similar to the maximum likelihood estimates: the former is not statistically significant, while the latter are higher for women.

6. Conclusion

This paper uses an unusual database, the SWT survey conducted by the ILO in Kosovo in 2004, to shed light on gender wage differentials. Although by now the database is several years old, the data collected by the ILO on the experiences of young workers at the beginning of their career, as they left school and entered the labor market are rare. They are also extremely useful for policy considering that to date little is known about the gender wage gap in Kosovo. In addition, in Kosovo and in many other countries more generally, little is known about the evolution of gender differences in wages (as well as in labor supply) over the life cycle and the SWT survey provides opportunity to analyze the earnings differentials between men and women, particularly in their early labor market experience.

The results of the analysis suggest that young women tend to have lower educational attainment than young men. Women are also less likely to be employed and those that do find employment tend to be the most educated. As a result, the disparities in educational attainment tend to be masked when analyzing only employed men and women in their first years in the labor market.

The consequences of this segmentation on wages and gender wage gaps in early career are striking. On average, there is little or no gender wage gap. The results of the Juhn et al. (1993) decomposition analysis reveal that gender wage differences are almost

Table 8 Maximum likelihood estimate of the selectivity corrected earnings equation (Heckman procedure)

Variable	All	Women	Men	Teenagers (All)	Young adults (All)	Young adults (Women)	Young adults (Men)
Main equation							
Dependent variable: hourly wages							
Women	−0.053			−0.1243	−0.0513		
Age	−0.0057	−0.015	0.0664**	−0.1396*	0.0029	−0.0195	0.0881**
Education (baseline: Elementary or below)							
Vocational	0.0867	0.143	0.1987	−0.0521	0.1880*	0.188	0.3729**
High Secondary	0.2307**	0.3872***	0.3032**	0.0418	0.3581***	0.4915***	0.4088**
University or above	0.5488***	0.6705***	0.389	(omitted)	0.6575***	0.7490***	0.4096
Ethnic group (baseline: Albanian)							
Serbian	−0.4983***	−0.4674***	−0.8247***	−0.9117***	−0.4147***	−0.3753**	−0.9040***
Roma	0.3592*	0.7149***	0.1064	0.0865	0.3325	0.5200***	−0.0107
Turkish	−0.2666**	−0.4273***	0.0983	−0.3897	−0.1936	−0.4317***	0.1148
Bosnian	0.484	−0.2791**	0.9941**	0.0582	0.8914*	(omitted)	1.0969**
Constant	0.1199	0.1974	−2.2302***	2.7556	−0.1647	0.2672	−2.9048***
Selection equation							
Women	−0.1407*			−0.3786**	−0.0678		
Age	0.1711***	0.1992***	0.1457***	0.1032	0.1911***	0.2419***	0.1701***
Education (baseline: Elementary or below)							
Vocational	0.3595***	0.8007***	0.1193	0.1777	0.4952***	0.9949***	0.2173
High Secondary	0.4145***	0.7777***	0.2468	0.7556***	0.3700**	0.7132***	0.2335
University or above	0.1833	0.6754**	−0.2097	(omitted)	0.1522	0.6413**	−0.1527
Ethnic group (baseline: Albanian)							
Serbian	−0.4498**	−0.1505	−0.6671***	0.2207	−0.5591***	−0.2078	−0.7729***
Roma	−0.0315	0.2148	−0.182	0.4004	−0.2892	−0.545	−0.3702
Turkish	0.6105*	0.8234	0.4349	1.7287***	0.2551	0.3639	0.2535
Bosnian	0.3313	0.0022	0.6972	1.0258	−0.1028	−5.4657***	0.8664
Main goal in life is:							
Contributing to social life	−0.4297**	−0.3254	−0.4601**	−5.4936***	−0.2353	−0.1576	−0.3939*
Gaining all types of experiences	−0.6477**	−0.7631*	−0.3404	−0.5639	−0.6681**	−0.6457	−0.4337*
Number of sons and daughters	−0.2574**	−0.4424***	−0.1231	−5.2412***	−0.2568**	−0.4902***	−0.0945
Number of brothers	−0.0578	−0.1120*		−0.066	−0.0597	−0.1511**	0.0497
Constant	−4.2022***	−5.2032***	−3.5979***	−2.9289*	−4.6713***	−6.0678***	−4.2588***
Athrho							
Constant	−0.1775	−0.2811**	1.2599***	−0.1402	−0.2468	−0.4143***	1.7034***
Lnsigma							
Constant	−0.6436***	−0.6573***	−0.3105***	−0.5576***	−0.6784***	−0.6512***	−0.2735***

Table 8 Maximum likelihood estimate of the selectivity corrected earnings equation (Heckman procedure) *(Continued)*

Statistics							
N	1336	657	679	472	864	414	450
Rho	−0.18	−0.27	0.85	−0.14	−0.24	−0.39	0.94
Sigma	0.53	0.52	0.73	0.57	0.51	0.52	0.76
Lambda	−0.92	−0.14	0.62	−0.08	−0.12	−0.2	0.72
Likelihood ratio test of independent equations	1.14	4.84	24.3	0.06	2.51	10.34	75.6
Probability > χ^2	0.29	0.03	0.00	0.81	0.11	0.00	0.00

Note: * p<.1; ** p<.05; *** p<.01.
Source: own elaboration on the ILO School-to-work transition survey of Kosovo.

entirely driven by differences in characteristics rather than either the returns to those characteristics or the residual. The greater mean educational attainment of employed women, among other characteristics, tends to fully offset the gender wage gap. In particular, the greater mean educational attainment of employed women tends to compensate for the higher average number of working hours of their male counterparts.

The Machado and Mata (2005) decomposition of the gap along the entire wage distribution shows that in fact women located in the upper part of the wage distribution should receive a higher wage than their male colleagues, considering their higher educational level and, therefore, productivity characteristics.

When the analysis accounts for selection bias and the heterogeneity of non-employed women, the gender wage gap turns negative (though insignificant) while the returns to education become higher for women than for men, confirming the lower productivity-related characteristics of non-employed versus employed women.

Despite the general absence of a gender wage gap among Kosovo's employed men and women in early career, some evidence suggests that a gender wage gap does begin to emerge among the young adults (age 20–25) compared to teenage workers (age 15 to 19), consistent with a small but growing literature on gender wage gaps during the first few years in the labor market. However, the relatively small sample constrains a fuller analysis of this phenomenon.

The policy implications of the paper are clear. The general absence of a gender wage gap among the young sample of under-25 provides indirect confirmation of the importance of providing childrearing facilities and arrangements to prevent the emergence of the gap at a later age, documented in previous studies (World Bank 2003, 2010, 2012). More directly, the Fairlie analysis shows the primary of demographic factors among the drivers of gender differences in labor supply. The small number of women who marry and give birth at the age covered by the SWT survey may help explain why the gender gap in labor supply is relatively small, while suggesting that it might dramatically increase if more women had children at a young age.

Endnotes

[1] See, in particular, a recent comprehensive report (World Bank 2012).
[2] See Manning and Swaffield (2008).

[3] For further information on the SWT survey, see the following website: http://www.ilo.org/employment/areas/WCMS_159352/lang−en/index.htm.

[4] Women seem to receive higher returns to education than men in terms of employment probability. In other words, men with high educational attainment tend to have higher employment rates than men with low educational attainment, but this education-employment relationship is much stronger for women.

[5] World Bank 2010.

[6] In fact, there is no discussion of wages or earnings in the World Bank (2012) report.

[7] See Bertrand et al. (2010).

[8] See for example (Kunze 2005) based on data from West Germany. Manning and Swaffield (2008) use UK data to look at gender wage disparities in "early career", seeking to explain how the gender gap rises from zero to 25 log points, after the first 10 years in the labor market. They find that a number of theories can explain the gap, but up to a third is still left unexplained.

[9] The discussion in this section follows closely the lead author's previous exposition (Pastore 2009, 2010, 2011, and Pastore and Verashchagina 2011).

[10] When the regressor is a continuous variable, such as years of work experience, the elasticity at the mean of the covariates, namely the percentage change in the regressand for a percentage change in the regressor, can be computed multiplying the coefficient by the mean of the regressor: $\beta \bar{X}$. In the case of independent dummy variables, such as for example the levels of educational attainment, the semi-elasticity interpretation is flawed and, following Halvorsen and Palmquist (1980), it should be computed as: $(e^{\beta} - 1) * 100$. This formula measures the percentage change in the median wage which is less affected by outliers. Nonetheless, many authors interpret the estimated coefficients of dummy variables directly as semi-elasticity as well. This is acceptable when the estimated coefficient is sufficiently close to zero.

[11] See also Blau and Kahn (1996) for an application of this methodology and the STATA help file for an introductory presentation of the methodology (command jmpierce).

[12] We thank Indhira Santos for pointing out the caveats related to the use of the instruments actually available in the survey at our disposal.

[13] In fact, the survey includes not only a questionnaire administered to young people, but also one administered to a sample of firms. The questionnaire of the Employer and Managers module survey contains detailed information on the ownership and industrial sector of the firm, union membership, number and skill requirements of employees, vacancies, recruitment methods, criteria of selection of the personnel, training practices in the firm and so on. Unfortunately, the two questionnaires cannot be linked with each other because there is no connection between the individuals and the firms.

[14] Together with that of Syria, the SWT survey of Kosovo does not include the oldest age segment, aged 25–29 years, which is instead available in the other countries.

[15] The question asked (G42) is: "On average what is your total income from work per month?" It is not clear from the wording of the questionnaire whether

this is the gross wage or the wage net of taxes. In the other surveys, it was the net monthly wage. In addition, it is not clear whether the declared earnings are coming from the main job or from any working activity. Based on personal communications, the ILO staff in charge of the survey does not have any additional information on the exact definition of wages used in the survey.

[16] The wage variable has only few missing observations.

[17] No estimates are provided by gender for different age groups, due to the small number of observations available in the sample for the level of detail that is requested in the augmented version of the estimates.

[18] For a presentation in Stata of this type of decomposition, see Jann (2006).

[19] For shortness sake, the results of the logit equation of the probability of labor supply are omitted and available on request from the authors. The same applies also to the labor supply equation estimate by gender, which is discussed next.

[20] Other results, including the two-step version and the maximum likelihood with different instruments, are available on request from the authors.

Additional file

Additional file 1: Table S1. Descriptive statistics: the employed. **Table S2.** Descriptive statistics: the non-employed.

Author details
¹Seconda Università di Napoli, Italy. ²World Bank, Washington, DC, USA.

References

Bertrand M, Goldin C, Katz LF (2010) Dynamics of the gender gap for young professionals in the corporate and financial sectors. American Economic Journal: Applied Economics 2(28):228–255

Blau FD, Kahn LM (1996) International differences in male wage inequality: institutions versus market forces. J Polit Econ 104(4):791–837

Cameron AC, Trivedi PK (2009) Microeconometrics Using Stata. Stata Press, Texas

Fairlie RW (2005) An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. J Econ Soc Meas 30:305–316

Halvorsen R, Palmquist R (1980) The interpretation of dummy variables in semilogarithimic equations. Am Econ Rev 70(3):474–475

Havolli S (2011) Determinants of Migrants' Earnings and Remittances: Evidence from Kosovo. Oesterreichische Nationalbank Focus on European Economic Integration (Austria) Issue 1:90–98

Heckman JJ (1979) Sample selection bias as a specification error. Econometrica 47(1):153–162

Hoti A (2011) Returns to Education in Kosovo: Estimates of Wage and Employment Premia. South East European Journal of Economics and Business 6(1):71–84

Jann B (2006) fairlie: Stata module to generate nonlinear decomposition of binary outcome differentials. Available from http://ideas.repec.org/c/boc/bocode/s456727.html

Juhn C, Murphy KM, Pierce B (1993) Wage inequality and the rise in returns to skill. J Polit Econ 101(3):410–442

Kunze A (2005) "The evolution of the gender wage gap. Labor Economics 12(1):73–97

Leuven E, Oosterbeek H (2011) "Overeducation and Mismatch in the Labour Market," forthcoming in the *Handbook of the Economics of Education* (also available as IZA discussion paper, n. 5523)

Machado J, Mata J (2005) Counterfactual decompositions of changes in wage distributions using quantile regression. J Appl Econ 20(4):445–465

Manning A, Swaffield J (2008) The Gender Gap in early-career wage growth. Econ J 118(530):983–1024

McGuinness SPJ (2010) "Esiste overeducation? Un'analisi comparata", in AlmaLaurea, XII Rapporto sulla condizione occupazionale dei laureati. Investimenti in capitale umano nel future di Italia ed Europa, Il Mulino, Bologna

Neuman S, Oaxaca R (2004) Wage decomposition with selectivity corrected wage equations: a methodological note. J Econ Inequal 2:3–10

Oostendorp RH (2009) Globalization and the Gender Wage Gap. The World Bank Economic Review 23(1):141–161

Pastore F (2009) School-to-Work transitions in Mongolia. Eur J Comp Econ 6(2):245–264

Pastore F (2010) The Gender Gap in Early Career in Mongolia. Int J Manpow 31(2):188–207, also available as IZA Discussion Paper No. 4480, November

Pastore F (2011) Returns to Education of Young People in Mongolia. Post-Communist Economies 22(2):247–265

Pastore F, Verashchagina A (2011) When Does Transition Increase the Gender Wage Gap? An Application to Belarus. Econ Transit 19(2):333–369, also available as: IZA Discussion Paper No. 2796, May 2007

UNFPA (2007) Gender-Based Violence in Kosovo. A Case Study, Prishtina, Kosovo

van Albrecht J, Vuuren A, Vroman S (2009) Counterfactual distributions with sample selection adjustments: econometric theory and an application to the Netherlands Labour Economics. Labour Econ 16(4):383–396

Wooldridge J (2003) Introductory Econometrics: A Modern Approach. Thomson Learning, Mason. Ohio

World Bank (2003) "Kosovo Labor Market Study: Policy Challenges of Formal and Informal Employment," Report No. 25990. The World Bank, Washington

World Bank (2010) "Kosovo: Unlocking Growth Potential: Strategies, Policies, Actions. A Country Economic Memorandum," Report No. 53185-XK. The World Bank, Washington

World Bank (2012) "Gender issues in Kosovo", mimeo. The World Bank, Washington