Skill-mismatch and the consequences for the labor market: Evidence from Kosovo†

Wolfgang Frimmel[a,b], Fitor Murati[c], and Burim Prenaj[c]

[a]JKU, Johannes Kepler University Linz
[b]Christian Doppler Laboratory Aging, Health and the Labor Market, Linz
[c]University of Prishtina

July 20, 2020

Abstract

Skill-mismatch is a persistent problem in countries with weakly developed labor markets. Evidence on the consequences of skill-mismatch in these countries is scarce and mostly correlational. Using the Kosovo Labor Force Survey data from 2012 to 2017, we empirically analyze the consequences of over-education on individual's employment, hourly wage, working hours, low job satisfaction and the probability of being on short working contracts. We apply a combination of exact and nearest-neighbor propensity score matching to tackle selection into over-education, and find that over-educated workers move on a lower wage trajectory compared to adequately matched workers and incur a wage penalty of 11 percent. This penalty is larger in urban areas and more than twice as large for females. Over-education is also associated with precarious labor contract situations and leads to lower job satisfaction. Our results underline that mismatch is a serious problem in the Kosovo labor markets, which should be tackled by policy-makers to a much greater extent.

JEL Classification: J21, J24, I25
Keywords: Skill-mismatch, over-education, labor market outcomes, propensity-score matching

†Corresponding author: Wolfgang Frimmel, Johannes Kepler University of Linz, Department of Economics, Altenbergerstraße 69, 4040 Linz, Austria; email: wolfgang.frimmel@jku.at. We would like to thank participants of the LSEE EBRD Conference 2019 in Belgrade for comments and Ardiana Gashi and Rudi Kapeller for their valuable assistance. The usual disclaimer applies. Financial support from the HERAS-Project K-02-2018 funded by the Austrian Development Agency (ADA) with funds of Austrian Development Cooperation is gratefully acknowledged.

1 Introduction

A large volume of empirical literature studies the determinants and economic consequences of the mismatch between formal education and labor market requirements. Specifically over-education, a situation in which the formal education of an employee exceeds the requirements of the occupation, is of particular interest to policy-makers and researchers.[1] In this context, the labor market in Kosovo represents an intriguing opportunity for research. The labor market performance lags behind that of comparable neighboring countries and is associated with a pronounced weak link between schooling and labor markets. Figure 1 shows the change in the share of high-educated workers (grey bar) and the change in the share of high-skilled jobs (dark red bar) for selected countries over the period of 2012-2017 based on data from the European Labor Force Surveys and the Kosovan Labor Force Survey. A universal trend across Europe is a significant increase in the share of high-educated people in the labor market. This share increased by more than 6 percentage points for Kosovo, only Sweden and France exhibit comparably large increases. However, while exclusively all EU countries were able to increase the share of high-skilled labor in the economy over the same period, even though to a smaller degree relative to increases in high-educated workers, the situation for Balkan States Serbia, North Macedonia and Kosovo is quite alarming where the share of high-skilled labor even decreased. The exceptional gap for Kosovo with a decline in high-skilled labor of 4 percentage points suggests significant difficulties with skill-mismatch and over-education in the labor market and is considered among the main factors preventing the inactive population from obtaining adequate jobs (Cojocaru, 2017).

We are interested in the consequences of this exceptional development of over-education in Kosovo, which have not been studied so far. A substantial level of over-education in labor markets can have negative effects in several aspects. Firms may suffer in efficiency if mismatched workers are less productive or less motivated; individuals who are mismatched may earn less than their matched peers; and governments may waste scarce financial resources by excessively educating young people without policies for creating adequate jobs. In this paper we focus on the impact of over-education on individuals' labor market outcomes. We define over-education as the difference between workers' completed level of schooling and the level of schooling required for the job and follow the International Standard Classification of Occupations (ISCO-08) (ILO, 2012) and UNESCO's International Standard Classification of Education (ISCED-97) definitions in order to match skill levels with levels of education required to perform a job competently. Using data from

[1] (McGuinness, 2006) and (McGuinness et al., 2017) deliver an extensive review of literature on all forms of mismatch. In terms of over-education researchers mainly studied effects on wages, job satisfaction, contracts as well as productivity (Cutillo and Pietro (2006); Lindley and McIntosh (2010); McGuinness and Sloane (2011);Verhaest and Omey (2009) Sloane (2014); Frenette (2004); Bauer (2002)).

the Kosovo Labor Market Survey, we empirically assess the effects of skill-mismatch on workers' wages, type of contracts and job satisfaction. The identification of a causal effect in this context is challenging. Workers' selection into firms is usually endogenous and workers do not randomly work in a mismatched job. We address this challenge by using a comprehensive data-set offering a wide set of individual characteristics. We aim to eliminate selection bias at least to a large degree by implementing a combination of exact and propensity score matching methods to compare mismatched and matched workers which are identical at least in observable characteristics.

We find that over-educated employees in Kosovo suffer a wage penalty of 11 percent in comparison to their matched peers. This finding for Kosovo is in line with other studies reporting a negative effect of over-education on wages (Comyn et al. (2019);McGuinness and Pouliakas (2017); Cutillo and Pietro (2006); Diem (2015); Levels et al. (2014); McGuinness and Bennett (2007); Robst (2008); Sanchez-Sanchez and McGuinness (2013); Sloane (2014); Verhaest and Omey (2011); Pietro and Urwin (2006)). The wage penalty of over-education is more pronounced for females[2] and in rural areas. Finally, in line with Sanchez-Sanchez and McGuinness (2013), McGuinness and Sloane (2011) or Verhaest and Omey (2009), over-educated worker's job satisfaction significantly suffers as a result of their mismatch.

This paper contributes to the literature in several ways. First, since evidence on the effects of over-education is largely available for high-income countries due to data availability, we provide the first thorough evidence on skill-mismatch in Kosovo and supplement the scarce and inconsistent literature on countries with weak educational and labor market institutions.[3] Second, we extend the existing literature on developing countries[4] by analyzing a broader range of outcome variables such as employment, wages, type of contracts and job satisfaction. Third, we contribute to the literature of over-education by addressing the issue of endogeneity. Several methods have been used to address this issue up to now. Korpi and Tåhlin (2009) used instrumental variable methods to estimate the effects of over-education on wages, while others such as Lindley and McIntosh (2008), Dolton and Silles (2008) and Tsai (2010) applied fixed effect techniques. Likewise, McGuinness (2008))and Karymshakov and Sulaimanova (2019) used propensity score matching to tackle the endogeneity problem. In line with these attempts, we use a matching procedure

[2]compare e.g. Mavromaras et al. (2012); Budria and Moro-Egido (2009); McGuinness (2008); Robst (2008); McGuinness and Bennett (2007)

[3]Handel et al. (2016), for example, finds that over-education is a serious problem across 12 low and middle income countries, while Sparreboom and Staneva (2014) find that under-education is a more serious problem in these countries.

[4]Karymshakov and Sulaimanova (2019); Le Quang and Tran-Nam (2019); Dibeh et al. (2019); Pholphirul (2017); Reis (2017). See also Comyn et al. (2019) for an extensive review of educational mismatch in developing countries

where we combine exact matching and the nearest-neighbor propensity score matching. This study adds to a scarce number of papers addressing endogeneity in the context of under-developed labor markets, where studies are mostly correlational and do not correct for potentially biased estimates.

The paper is organized as follows. Section 2 provides a summary of the labor market in Kosovo. Section 3 presents estimation and matching strategy along with descriptive statistics. Section 4 summarizes the estimation results including a heterogeneity and sensitivity analysis. Finally, section 5 concludes.

2 The labor market in Kosovo

Labor market — Kosovo's GDP growth has been relatively higher compared with its neighbors during the last decade, nevertheless this growth has not translated into similar positive prospects on the labor market. While labor market challenges are a common characteristic in all the countries of Western Balkan, Kosovo lacks behind in almost all labor market indicators. During the period 2012-2017, Labor Force Survey data report an average labor force participation of below 40 percent and employment rate of around 27 percent. Figure 2 compares employment and unemployment rates for males (sub-figures (a) and (b)), females (sub-figures (c) and (d)) and young workers below 24 years of age (sub-figures (e) and (f)) for several countries of the Western Balkan region. In all of these dimensions, Kosovo lacks significantly behind its neighboring countries, where employment rates for males exceed 50 percent and the share of employed women is more than three times as large as compared to Kosovo. We witness similar comparative trends in the unemployment rate. While Kosovo's average trend hovers above 30 percent, almost all other countries in the region have managed to lower unemployment rates up to below 15 percent unemployment rate in 2017.

Figure 2 also reveals severe gender disparities in the labor market. While male employment in Kosovo slightly increased over time, female employment, mostly of highly educated women, remained steadily just above 10 percent during 2012-2017 period. The unemployment rate is somewhat more balanced between genders and ranges between 30 and 40 percent for females and 25-30 percent for males.

Young people (15-24 years) are a particular disadvantaged group in the Kosovo labor market. They account for 30 percent of all unemployed individuals during 2012-2017. As sub-figures (e) and (f) of Figure 2 show, youth employment rate is dramatically low (below 25 percent) in all Western Balkan countries, but Kosovo exhibits far the worst performance by barely exceeding 10 percent with a youth unemployment rate of above 50

percent during 2012-2017.

Education — In the 2000s, Kosovo experienced a substantial investment in higher education, increasing the number of public universities from a single one in 2000 to 6 and more than 20 private universities until 2012. This led to a huge expansion in the educational attainment, which exceeded by far the absorption capacity of Kosovo's labor market. As Figure 1 points out, the share of highly educated workers has increased at a relative fast pace (6pp), but the availability of high-skill jobs in fact decreased. Still, educational attainment and employment prospects are highly associated. More than half of all highly educated individuals are employed, while 23 percent of people with secondary level of education and only approximately 3 percent of people with primary level of education are working.[5] Accordingly, around 95 percent of inactive individuals have secondary education or less. This suggests that highly educated individuals are accepting jobs that do not match their level of education (Murati and Prenaj, 2020).

3 Research design

3.1 Data

The empirical analysis is based on individual data from six annual survey waves (2012 to 2017) of the Kosovo Labor Force Survey (LFS). Since 2012 the Kosovo LFS is following the Eurostat standards and is conducted by the Kosovo Agency of Statistics. It is the first large-scale nationally representative labor market survey administrated in the country and contains the richest and most reliable data on the labor market in Kosovo. It collects data for approximately 1 percent of Kosovo's population from randomly selected households from all regions of Kosovo[6] and provides internationally comparable information on basic demographic characteristics, employment status, job characteristics and income. Most importantly, the LFS records occupations on the 3-digit ISCO level and educational attainment on the ISCED-97 level, which enables us to measure the extent of skill-mismatch and over-education at individual level.

We restrict our sample to men and women who were interviewed more than once and are employed at the time of the first interview (time t). All covariates, i.e. set of household, individual and job characteristics, and the skill-mismatch indicator are based on information in time t. All outcome variables, i.e. employment status, hourly wage, working hours, type contract and an indicator for low job satisfaction, are measured at the last

[5] It is important to note that jobs held by high-educated individuals are mostly in the public sector, which employs around 75 percent of all workers with tertiary education and offers an approximately 50 percent higher hourly wage than the private sector.

[6] Each annually selected household is followed for one year and usually interviewed once per quarter.

available interview (time $t+1$), hence usually three quarters later. We exclude all individuals with missing information on occupation, e.g workers who are temporarily absent from their job, on employment status at time t or education. The final sample consists of 5,607 individuals.

3.2 Skill-mismatch indicator

Most existing literature use statistical and subjective methods to generate an indicator for skill-mismatch (McGuinness et al., 2017). The subjective approach is based on the respondents' self-reported indicator about their skill usage in their job, which is then interpreted as adequately matched or mismatched depending on the answer, but the LFS does not record any self-reported information on the usage of skills by employees. The statistical approach uses the mean or the mode years of education or mean/mode level of education in each occupational group to derive a measure of skill-mismatch. However, the resulting mismatch indicator is endogenous and could wrongfully assign individuals into mismatch status. By construction, occupations with a larger share of high-educated workers – and therefore with a higher observed mean of years/level of education – are labeled as high-skilled occupation irrespective of the true skill requirement of the occupation. If a low-skill occupation is characterized by a larger share of mismatched individuals, then this method assigns a too high skill requirement to that specific occupation and might significantly underestimate the true level of mismatched workers, i.e. truly mismatched individuals (highly educated) would be considered adequately matched, while truly adequately matched individuals would be considered as under-educated. Since the labor market in Kosovo is characterized with a substantial degree of mismatch, such a statistical approach would introduce an unacceptably large bias to our mismatch indicator.

Instead, we propose a more normative approach to measure skill-mismatch, which we believe is less prone to such a bias mentioned above. The LFS records occupations based on the 3-digit ISCO level and education based on ISCED-97 level. We closely follow the official mapping of occupations to skill levels of the International Labor Organization and use the International Standard Classification of Occupations (ISCO-08), which defines a total of four skill levels, and assign each major occupation to one of these skill levels (see Table 1). Accordingly, each level of education based on the ISCED-97 that is required for competent performance of each major occupational group is mapped to ISCO skill levels. As summarized in Table 2, occupations assigned to the third and fourth skill level, are mapped to individuals with some form of tertiary level of education (5a, 5b and 6 in the ISCED-97 framework), occupations with second skill level are assigned to some form of secondary education (4, 3 and 2 6 in the ISCED-97 framework), and finally, jobs belonging

to first skill level, are mapped to primary education (1 in the ISCED-97 framework). So we end up with a categorical variable assigning either high, medium or low educational attainment to each occupation.

We define a worker to be mismatched if the observed level of education is not at the same level as the required educational attainment of the worker's occupation. Specifically, a worker is denoted as over-educated if observed level of education is above the skill-level of the observed occupation, (e.g. a highly educated individual working in a a medium skill job) and as under-educated if level of education is below the skill level of the occupation (e.g.a medium educated individual holding a high skill job). This leaves us with $1,716$ over-educated and $3,891$ adequately matched workers in our sample.

Table 3 compares our skill-mapping measure of over-education with two common statistical approaches using the mode and mean years of education within each occupation. First, there is little variation across occupations with respect to the mode and mean years of education. So the degree of information derived from the statistical measures is rather limited. For most occupations, the mode is 12 years of education independent of the underlying skill requirements of the occupation. Equivalently the same is also true for mean years of education of approximately 12 years across occupations. This is a clear hint for a severe degree of mismatch in a labor-market. Second, although in some occupations all three measures of over-education yield very similar shares, the degree of over-education is – as expected – substantially under-estimated by both statistical measures as they underlay unrealistically high thresholds for over-education[7].

Our normative approach limits the bias introduced by the statistical approach and yields an unbiased measure of mismatch if formal education requirements are a valid measurement of skill level of an occupation, i.e. the mapping from educational attainment to skills and occupations is an adequate description of real occupational skill requirements. Even if this is arguably plausible for a lot of occupations, informal training or sufficient work experience might substitute for formal education in some occupations. In that case our mapping systematically assigns a higher educational attainment to each occupation than required, leading to a downward bias in the measurement of over-education. So our measure of over-education represents a lower bound estimate of the true extent of over-education but is still more precise than conventional statistical measures. Additionally, since the importance of informal training is at least partly industry-specific (e.g. manufacturing vs. professional services), adding industry fixed-effects to our empirical model will adress this potential drawback.

[7] Also note that by construction there should not be over-education at the highest skill-level.

3.3 Descriptive statistics

In Table 4 we provide descriptive statistics for individual characteristics of matched and mismatched (over-educated) workers. There are substantial differences between these two groups of workers with respect to several individual and household characteristics. Over-educated workers are on average younger, have fewer years of education, lower tenure, are more likely to be male (10 pp difference) and employed in the private sector. They are also less likely to have an employed partner and have on average a higher number of children. So the higher prevalence of breadwinners in mismatched jobs might be an important indicator why workers are willing to accept a job for which they tend to be over-educated.

The bottom panel of Table 4 compares our main outcome variables for matched and over-educated workers and reveals significant differences in means. Over-educated workers have a significant lower hourly wage (€0.579 less), work on average 2 hours more per week, and are 14 percentage points more likely to be employed in short-term contracts. So over-educated workers are at an inferior position in all of these dimensions. While these comparisons provide interesting descriptive insights, we cannot interpret these differences causally due to nonrandom selection of workers into matched and mismatched groups. Also unobservable individual characteristics such as ability or motivation might be crucial in worker selection into mismatch.[8]

3.4 Estimation strategy

To assess the effect of being over-educated, we examine several labor market outcomes, for which we estimate the following equation:

$$Y_{i,t+1} = \beta_0 + \beta_1 M_{i,t} + \Gamma X_{i,t} + \tau_t + \epsilon_{it}$$

The set of outcome variables $Y_{i,t+1}$ consists of employment status, working hours, hourly wage, type of labor contract and and a binary indicator for low job satisfaction[9] in period $t+1$[10]. The treatment is captured by the binary indicator M_{it} , which is equal to one if individual i is over-educated in period t and zero if adequately matched. We include a comprehensive set of covariates capturing individual characteristics in period t, such as an indicator for gender, age, years of education, monthly income, previous employment status, binary indicators for private-sector employees and part-time workers, marital status, partner's employment status, firm tenure, number of children below 6, between 6 and

[8]see (Leuven and Oosterbeek, 2011) for a more detailed discussion.

[9]Low job satisfaction refers to the question in the LFS whether individuals would like to change their current job.

[10]The time span mostly covers three quarters after the first interview.

10 and between 10 and 18, as well as number of household members above age 60. We also include field of education and region fixed-effects as well as wave-fixed effects τ_t to capture time trends and wave-specific effects.

Despite this set of covariates, we cannot fully rule out endogenous sorting of workers into mismatched jobs, leading to a correlation between the treatment status and confounding factors included in ϵ_{it}. Thus, we should be careful with interpreting the coefficient of interest β_1 in a causal way. To improve the identification of the treatment effect, we control for selection based on observables in a more comprehensive way by implementing a matching procedure.

3.5 Matching strategy

To control for sorting into mismatch and allow for a pairwise comparison of mismatched and adequately matched workers, we need to estimate the average treatment effect on the treated (ATT), which is the expected effect for an individual, who is randomly drawn from the population of mismatched employees only. Such an ATT cannot be identified from the data without further assumptions, since the counterfactual labor market outcome for mismatched workers is unobserved. Given the assumption that all factors X jointly influencing the labor market outcomes and the choice of accepting a mismatched job are observable, then – conditional on these factors X – the labor market outcomes and the mismatched status are independent (Lechner et al. (2011), Lechner (2001), Lechner (2002), Imbens (2000)). This conditional independence assumption can be summarized by

$$E\left(Y_{it}^0 \,|M=1, X=x\right) = E\left(Y_{it}^0 \middle|\, M=0, X=x\right)$$

where M is an indicator equal to one if mismatched, and zero otherwise. The plausibility of the conditional independence assumption requires an understanding of the underlying selection process into mismatched jobs and identification of variables describing this selection process. Table 5 summarizes results from a linear probability model for selection into mismatched for different individual, household and firm characteristics. We find that the driving force behind mismatch are significant regional differences. Compared to the capital city of Prishtina, the probability of being in a mismatched job in Mitrovica, Prizren or Ferizaj is 17 to 20 percentage points lower. This points towards a pronounced divergence between labor and skill supply and skill-demand of firms across regions. In contrast, individual or household characteristics only play a minor role in the selection into mismatch[11]. Only younger males appear to be slightly more likely to accept a mismatched job. Nevertheless we use all available information on workers and households

[11]This is also in line with Böheim et al. (2008) who argue that over-education is primarily determined by the institutional setting.

as well as workplace in the data-set as conditioning variables to justify the conditional independence assumption. It is also important to note that all conditioning variables are measured at time t, while all outcome variables are observed in period $t + 1$.

3.6 Matching estimation

An estimator of the ATT requires for both adequately matched and over-educated workers a comparison observation from the other worker type with the same characteristics regarding all factors that jointly influence selection into mismatch and outcomes. To increase the precision for comparisons of both types, and taking the relatively small number of observations into account, we use a combination of exact matching and nearest-neighbor propensity score matching.[12] For a review of advantages and disadvantages of matching estimators, see e.g. Imbens (2004).

We select gender, region, marital status, previous employment, an indicator whether there is an employed partner in the household, a narrow classification of field of education, e.g. social sciences, life sciences, agriculture, health care..., and survey wave indicators as characteristics E_{it} to be exactly identical for both types of workers. In particular, we regard the endogenous choice for the field of education as a useful proxy of the unobserved innate abilities, i.e. technical, craft, language or social abilities. The propensity scores are based on conditioning variables X_{it} which are individual characteristics such as a second-order polynomial of age, years of education, monthly income, tenure and indicators for part-time work,private sector workers and living in a minority city, household characteristics (number of children below 6, between 6 and 10 and between 10 and 18, as well as number of household members above age 60) and firm characteristics of current employment such as firmsize and industry classifications (84 categories). The propensity scores for over-educated workers ($M = 1$) and adequately matched workers ($M = 0$)

$$P^M [M_{it}] = \Gamma X_{it} + \epsilon_{it}$$

are then estimated with a probit model.

We can only estimate the effects of over-education for workers with an overlap in the conditioning variables between both comparison states. We define this common support in terms of the propensity score, and delete observations with scores below the maximum of the minima and above the minimum of the maxima over both treatment states. We

[12] We also show results for nearest neighbor matching based on the actual variables instead of the propensity, however, due to relatively small sample sizes, propensity score matching considerably reduces the dimensionality of the estimation problem (Rosenbaum and Rubin, 1983). We also use radius propensity score matching as a further check. Results remain quantitatively unchanged.

thus delete 36 workers from the sample, so the losses are small.

In the last step, we estimate the counterfactual expectation of the outcome variable for any given pair of over-educated and adequately matched worker. We choose one over-educated worker and find an observation in the subsample of adequately matched workers who is (i) identical in terms of the set of covariates E_{it} and (ii) as close as possible in terms of the propensity scores P^1 for over-educated and P^0 for matched workers, where closeness is based on the Mahalanobis distance. The final treatment effect is the difference in the mean of the outcome variable between the group of over-educated workers from the group of adequately matched workers. The standard errors are obtained by bootstrapping.

Table 6 compares individual characteristics X_{it} of over-educated and adequately matched workers after matching. Ideally, there should be no significant differences in averages between the two groups. For most individual characteristics the matching procedure indeed is able to fully eliminate the bias between the two groups. Only for age and monthly income, a significant difference in means remains, although matching reduced the bias in monthly income by more than 50 percent. By construction there are no differences in E_{it} due to exact matching.

Overall, to interpret the matching estimate in a causal way, we need to rely on the assumption of selection on observables. Potentially unobserved characteristics affecting the mismatch status and labor market outcomes, such as ability, job search effort, motivation or selective outmigration, may result in an upward bias of our estimates, if we assume that e.g. more motivated workers are also willing to accept a (temporary) job with lower skill requirements. However, the results of our selection equation suggest that primarily local labor market conditions are the most important determinant of the mismatch status and individual characteristics affect selection only to a negligible degree. This gives us confidence that our matching procedure considers the most important sources of selection. Given a potential upward bias of an uncorrected estimate, our matching estimates should not only reduce the bias substantially, but ideally represent a lower bound estimate of the true unbiased effect.

4 Estimation results

This section summarizes our main results. Section 4.1 presents results for the effect of over-education on our set of labor market outcomes, Section 4.2 results from a heterogeneity analysis, and Section 4.3 presents results from an analysis on under-education.

4.1 Main results

Table 7 summarizes our estimates for the effect of over-education on labor market outcomes. Column (I) shows the coefficient of being over-educated from an OLS regression with all covariates used in the matching procedure. We do not find a correlation between over-education and employment status in period $t+1$, however, we see a significant negative correlation for the hourly wage, and a significant positive correlation for working hours, the probability of a short-term contract. However, OLS estimates conditional on covariates are methodologically not sufficient to warrant a valid comparison between matched and mismatched workers, so these estimates are likely to be biased.

Column (II) of Table 7 shows the results of the matching estimates described in Section 3.6. As compared to the standard OLS estimators, the matching estimates yield qualitatively identical results. There are, however, quantitative differences, indicating that the selection into mismatch matters. As expected, it turns out that the OLS estimates suffer from an upward bias. While we do not find an effect on the employment status in $t+1$, the hourly wage in the next period - conditional on today's income - is reduced by €0.242. Given the average hourly wage of €2.18 in our sample, this is on average a reduction of about 11 percent. Working hours significantly increase by 0.95 hours or 2.2 percent. Most strikingly, the probability of a short-term contract in the next period increases by 7.9 percentage points. There is also statistically weaker evidence that mismatched workers have a lower job satisfaction (plus 1.1 percentage points). To summarize, over-educated workers significantly suffer from an hourly wage penalty at longer weekly working hours and they are substantially more likely to be trapped in precarious employment relationships. Accordingly this further translates into lower job satisfaction and an increasing desire to change the job in the future.

These findings are robust to different matching methods, i.e. propensity score radius matching or nearest neighbor matching without changes to our conclusions (see Table 9). We also used an alternative definition of mismatch based on the original mapping of skill-levels of the International Labor Organization with 4 educational categories mapped to occupations. Our findings remain unchanged (see Table 10).

4.2 Heterogeneity analysis

In this section we provide a heterogeneity analysis on important subgroups of workers to learn and identify particularly affected groups. Table 8 and Figures 3 to 6 summarize the matching estimates for various sample splits. For each labor market outcome[13] we study separate effects for males and females, workers in rural and urban areas, and workers with

[13]Due to small sample sizes we cannot do a sample-split for job satisfaction.

and without a STEM background[14]. Figure 3 summarizes the results for employment status. For all subgroups we do not find any statistically significant effect of over-education. However, for the hourly wage (Figure 4), working hours (Figure 5) and the probability of a short-term contract (Figure 6), we find a very clear pattern.

The wage penalty of over-education is particularly pronounced for women and workers in urban areas. Women's wage penalty of €0.53 is more than twice as large as the wage penalty of men (€0.19). Hence, over-education is significantly contributing and deteriorating the gender-wage gap. Similarly, the wage penalty in urban regions is twice the penalty on the countryside. This is of course closely related to the availability of jobs and overall labor force participation, which is significantly higher in cities. Also, workers with a non-science background suffer more, while there is no such wage penalty for workers with science or technology background. The gender difference in wages does not translate to working hours or to jobs with short-term contracts. Longer working hours and a higher probability of a short-term contract due to over-education is only present among male workers. Over-educated workers in rural areas are also more likely to work under a short-term contract.

4.3 Under-education

So far, we presented evidence for the consequences of over-education on the labor market. From a policy and worker perspective, over-education is the most relevant case of mismatch. This is of course only one type of mismatch. Another form of mismatch is under-education, where employed workers would not meet the educational requirements to competently fulfill occupational tasks. Under-education is predominantly a problem for firms if they are forced to hire less competent workers and potentially incur productivity losses. A significant share of under-education in a labor market is also a clear signal to policy makers to increase investment in higher educational attainment.

Identifying the extent of under-education is not easy, because on-the-job training, job experience or self-learning may substitute well for a lack of formal education. Thus it is also unclear to (i) define an under-educated worker and (ii) to form expectations about the consequences on labor market outcomes. In this section we aim to shed some light on this issue. Similarly to the construction of our over-education measure, we rely our measure for under-education entirely on formal education requirements. Hence we ignore the upgrading of skills due to experience and training.[15]. This implies that our results might be influence by such skill-upgrading. Methodologically, we follow the exact same

[14]STEM is an acronym for studies in Science, Technology, Engineering and Mathematics.

[15]Our dataset does not provide sufficient information to generate a more precise measure of under-education

estimation strategy as for the case of over-education but replace over-educated workers by under-educated workers in our sample.

Table 11 summarizes our results on the effects of under-education. In contrast to the case of over-education, we find that under-educated workers do not suffer from a wage penalty nor are they more likely stuck in precarious job contract conditions. This result can simply be explained by unobserved worker skills substituting for lower formal education. We find weak evidence for slightly lower weekly working hours (-0.87 hours) and a 1.2 percentage point lower job satisfaction. Both coefficients are, however, only marginally significant. In sum, we find no evidence that formal under-education has any consequences for workers' labor market outcomes.

5 Conclusions

Skill-mismatch is a persistent problem in countries with weakly developed labor markets. Evidence on the consequences of skill-mismatch in these countries is scarce and mostly correlational. Using data from the Kosovo Labor Force survey, we examine the impact of skill-mismatch on workers' labor market outcomes in Kosovo, a country with a significant share of over-education. We apply a combination of exact and nearest-neighbor propensity score matching to reduce and control for selection into over-education, and find that over-educated workers move on a lower wage trajectory compared to adequately matched workers and incur a wage penalty of 11 percent. This penalty is larger in urban areas and more than twice as large for females. Hence, over-education is a significant contributor to the gender-wage gap in Kosovo. We do not find such a penalty for workers with an educational background in STEM-fields. Further results indicate that over-education is also highly associated with precarious labor contract situations and leads to lower job satisfaction. Overall, we show that skill-mismatch and over-education is a serious problem in the Kosovo labor market with lower wage trajectories for affected workers.

Our study carries important implications for public policies in Kosovo: Investments into higher education is generally recommendable but may backfire if it excessively leads to a mass of educated workers without any adequate job prospects. This may lead to significant welfare losses by first leading to lower wage trajectories for over-educated workers, and second by incentivizing highly-educated young people to migrate into countries with better job opportunities. Such a brain drain does not only question high investments in high education but may particularly have unprecedented consequences for the future development of such countries if highly educated workers are absent when urgently needed in the labor market. Experiences of several Eastern European countries show that this is a serious threat to a faster economic convergence.

Policy-makers should therefore focus on supporting the creation of high-skilled labor and put more effort in redirecting young people towards educational fields with promising job prospects such as STEM-fields. A successful reduction of mismatch on the labor market would not only lead to welfare gains by pushing workers to higher wage trajectories with more stable long-term employment prospects, by improving gender equality and female labor force participation with reducing the gender-wage gap, and finally by reducing incentives for outmigration of high-educated workers.

References

Bauer, Thomas (2002), 'Educational mismatch and wages: a panel analysis', *Economics of Education Review* **21**(3), 221–229.

Böheim, René, Iga Magda and Martina Zweimüller (2008), Skill mismatch in Europe, *in* 'Study and Conference on European Labour Market Analysis Using Firm-Level and Linked Employer-Employee Data'.

Budria, Santiago and Ana I. Moro-Egido (2009), 'The Overeducation Phenomenon in Europe', *Revista Internacional de Sociologia* **67**(2), 329–345.

Cojocaru, Alexandru (2017), 'Kosovo Jobs Diagnostic', *World Bank Washington DC* .

Comyn, Paul, Olga Strietska-Ilina, Adele Bergin, Seamus Delaney, Judith M.and Handel. Michael J.and McGuinness, Olga Kupets, Konstantinos Pouliakas and Paul Redmond (2019), 'Skills and jobs mismatches in low-and middle-income countries', *International Labour Organization* .

Cutillo, Andrea and Giorgio Di Pietro (2006), 'The effects of overeducation on wages in Italy: a bivariate selectivity approach', *International Journal of Manpower* **27**(2), 143–168.

Dibeh, Ghassan, Ali Fakih and Walid Marrouch (2019), 'Employment and skill mismatch among youth in Lebanon', *International Journal of Manpower* **40**(8), 1438–1457.

Diem, Andrea (2015), 'Overeducation among Graduates from Universities of Applied Sciences: Determinants and Consequences', *Journal of Economic and Financial Studies (JEFS)* **3**(2), 63–77.

Dolton, Peter J. and Mary A. Silles (2008), 'The effects of over-education on earnings in the graduate labour market', *Economics of Education Review* **27**(2), 125–139.

Frenette, Marc (2004), 'The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification', *Economics of Education Review* **23**(1), 29–45.

Handel, Michael J., Alexandria Valerio and Maria Laura Sanchez Puerta (2016), *Accounting for Mismatch in Low- and Middle-Income Countries: Measurement, Magnitudes, and Explanations*, The World Bank.

Imbens, Guido W (2000), 'The role of the propensity score in estimating dose-response functions', *Biometrika* **87**(3), 706–710.

Imbens, Guido W (2004), 'Nonparametric estimation of average treatment effects under exogeneity: A review', *Review of Economics and statistics* **86**(1), 4–29.

Karymshakov, Kamalbek and Burulcha Sulaimanova (2019), 'The school-to-work transition, overeducation and wages of youth in Kyrgyzstan', *International Journal of Manpower* **40**(8), 1458–1481.

Korpi, Tomas and Michael Tåhlin (2009), 'Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000', *Labour Economics* **16**(2), 183–193.

Le Quang, Huy and Binh Tran-Nam (2019), 'Qualification mismatch in the labor market and the impact on earnings: evidence from Vietnam', *Journal of Economics and Development* **21**(2), 223–233.

Lechner, Michael (2001), Identification and estimation of causal effects of multiple treatments under the conditional independence assumption, *in* 'Econometric evaluation of labour market policies', Springer, pp. 43–58.

Lechner, Michael (2002), 'Program heterogeneity and propensity score matching: An application to the evaluation of active labor market policies', *Review of Economics and Statistics* **84**(2), 205–220.

Lechner, Michael, Ruth Miquel and Conny Wunsch (2011), 'Long-run effects of public sector sponsored training in West Germany', *Journal of the European Economic Association* **9**(4), 742–784.

Leuven, Edwin and Hessel Oosterbeek (2011), Overeducation and mismatch in the labor market, *in* 'Handbook of the Economics of Education', Vol. 4, Elsevier, pp. 283–326.

Levels, Mark, Rolf van der Velden and Jim Allen (2014), 'Educational mismatches and skills: new empirical tests of old hypotheses', *Oxford Economic Papers* **66**(4), 959–982.

Lindley, Joanne and Steven McIntosh (2008), A Panel Data Analysis of the Incidence and Impact of Over-education, Working Papers 2008009, The University of Sheffield, Department of Economics.

Lindley, Joanne and Steven McIntosh (2010), 'Is the Over-Education Wage Penalty Permanent?', *Working Paper. Department of Economics, University of Sheffield ISSN 1749-8368* .

Mavromaras, Kostas, Peter Sloane and Zhang Wei (2012), 'The role of education pathways in the relationship between job mismatch, wages and job satisfaction: a panel estimation approach', *Education Economics* **20**(3), 303–321.

McGuinness, Seamus (2006), 'Overeducation in the Labour Market', *Journal of Economic Surveys* **20**(3), 387–418.

McGuinness, Seamus (2008), 'How biased are the estimated wage impacts of overeducation? A propensity score matching approach', *Applied Economics Letters* **15**(2), 145–149.

McGuinness, Seamus and Jessica Bennett (2007), 'Overeducation in the graduate labour market: A quantile regression approach', *Economics of Education Review* **26**(5), 521–531.

McGuinness, Seamus and Konstantinos Pouliakas (2017), Deconstructing Theories of Overeducation in Europe: A Wage Decomposition Approach, *in* 'Skill Mismatch in Labor Markets', Emerald Publishing Limited, pp. 81–127.

McGuinness, Seamus, Konstantinos Pouliakas and Paul Redmond (2017), How Useful Is the Concept of Skills Mismatch?, IZA Discussion Papers 10786, Institute of Labor Economics.

McGuinness, Seamus and Peter Sloane (2011), 'Labour market mismatch among UK graduates: An analysis using REFLEX data', *Economics of Education Review* **30**(1), 130–145.

Murati, Fitor and Burim Prenaj (2020), 'Overview Of Trends In The Employment Structure In Kosovo', *Balkan Economic Review* **2**, 46–62.

Pholphirul, Piriya (2017), 'Educational mismatches and labor market outcomes: Evidence from both vertical and horizontal mismatches in Thailand', *Education+ Training* **59**(5), 534–546.

Pietro, Giorgio Di and Peter Urwin (2006), 'Education and skills mismatch in the Italian graduate labour market', *Applied Economics* **38**(1), 79–93.

Reis, Mauricio Cortez (2017), 'Educational mismatch and labor earnings in Brazil', *International Journal of Manpower* **38**(2), 180–197.

Robst, John (2008), 'Overeduction and College Major: Expanding the Definition of Mismatch between Schooling and Jobs', *The Manchester School* **76**(4), 349–368.

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

Sanchez-Sanchez, Nuria and Seamus McGuinness (2013), 'Decomposing the impacts of overeducation and overskilling on earnings and job satisfaction: an analysis using REFLEX data', *Education Economics* **23**(4), 419–432.

Sloane, Peter (2014), 'Overeducation, skill mismatches, and labor market outcomes for college graduates', *IZA World of Labor* .

Sparreboom, Theo and Anita Staneva (2014), 'Is education the solution to decent work for youth in developing economies', *Work4Youth Publication Series* **23**, 23.

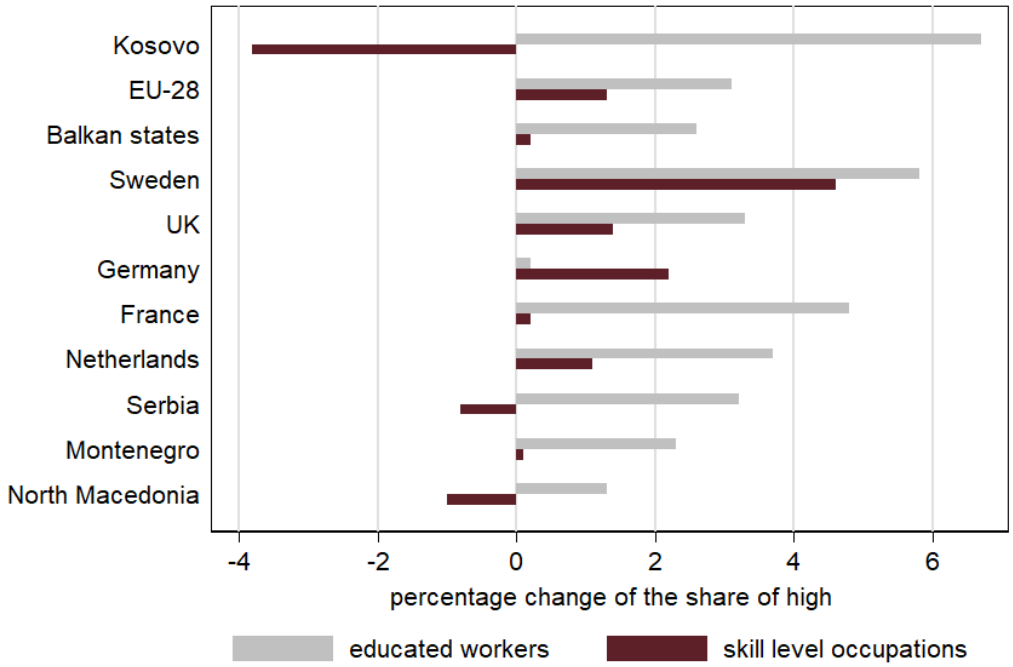
Tsai, Yuping (2010), 'Returns to overeducation: A longitudinal analysis of the US labor market', *Economics of Education Review* **29**(4), 606–617.

Verhaest, Dieter and Eddy Omey (2009), 'Objective over-education and worker well-being: A shadow price approach', *Journal of Economic Psychology* **30**(3), 469–481.

Verhaest, Dieter and Eddy Omey (2011), 'Overeducation, Undereducation and Earnings: Further Evidence on the Importance of Ability and Measurement Error Bias', *Journal of Labor Research* **33**(1), 76–90.

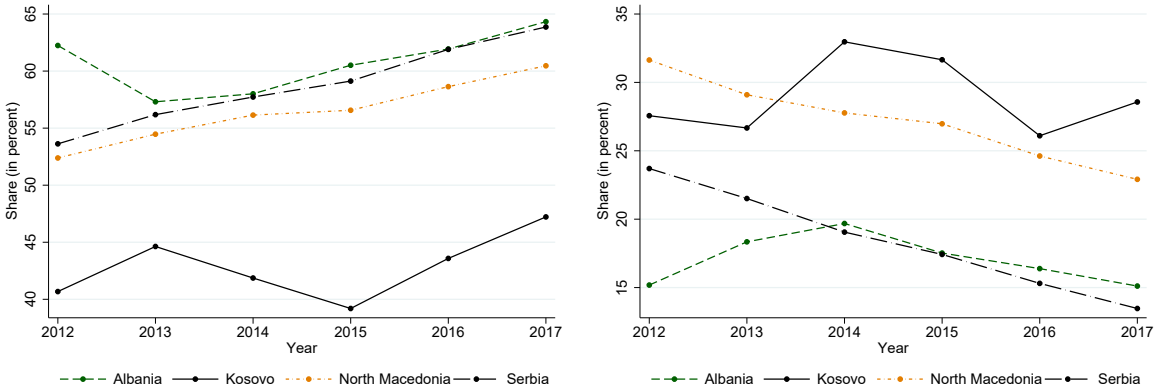
6 Tables and figures (to be placed in the article)

Figure 1: Percentage change in shares of high-skill and high-education labor over 2012-2017



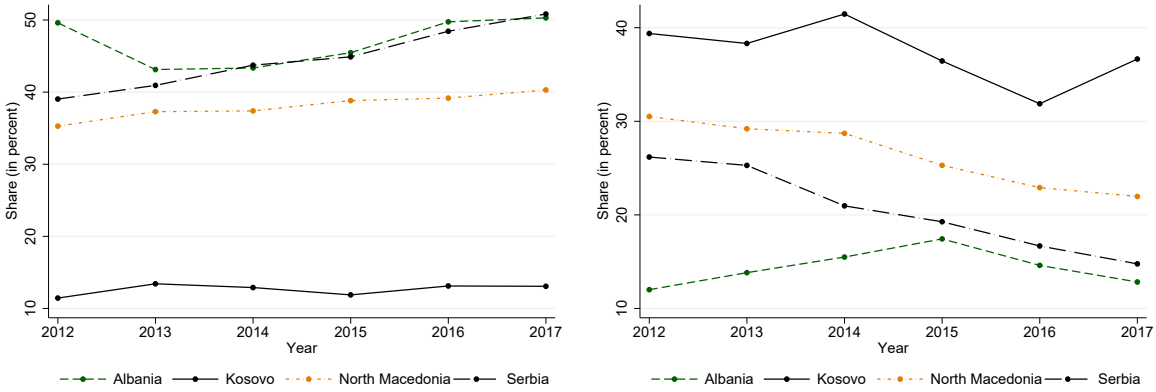
Notes: Own calculations based on data from the European Labor Force Survey and the Kosovo Labor Force Survey

Figure 2: Share of employed and unemployed in Western Balkan countries



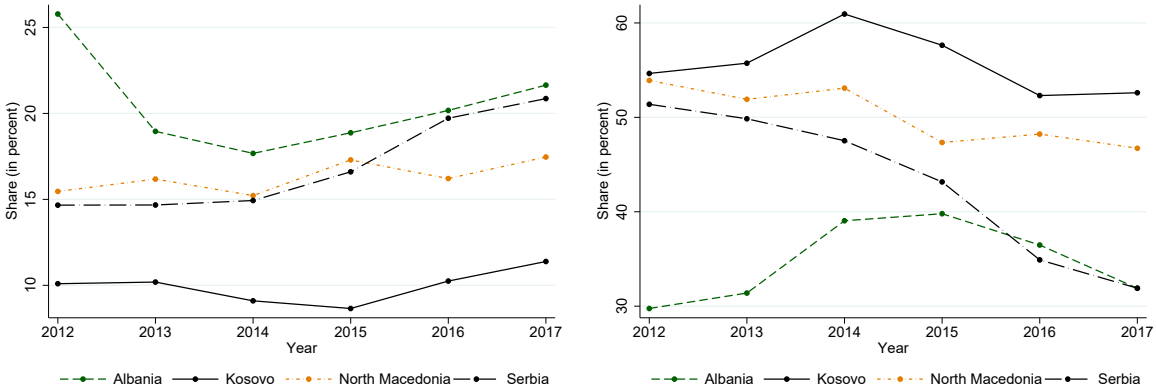
(a) Share of employed men

(b) Share of unemployed men



(c) Share of employed women

(d) Share of unemployed women

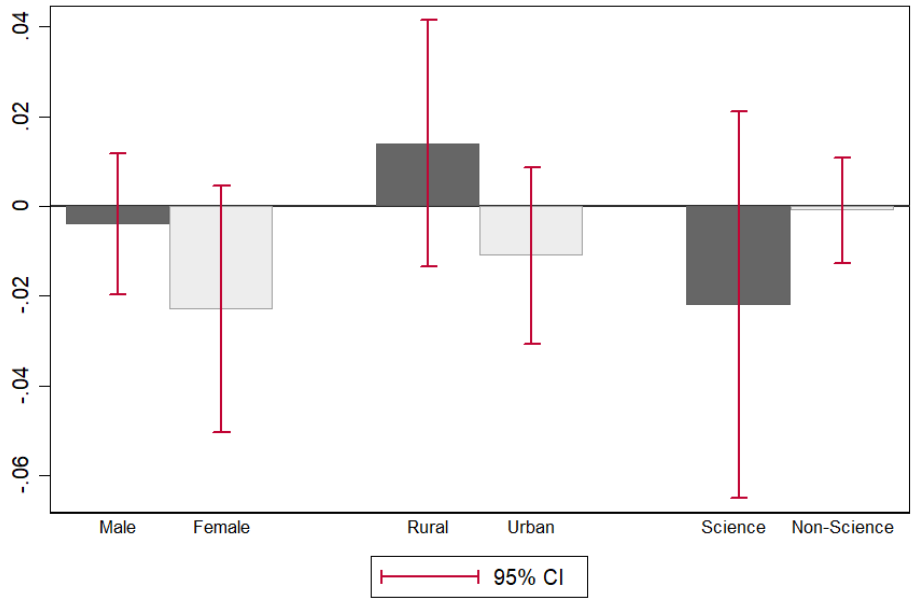


(e) Share of employed young

(f) Share of unemployed young

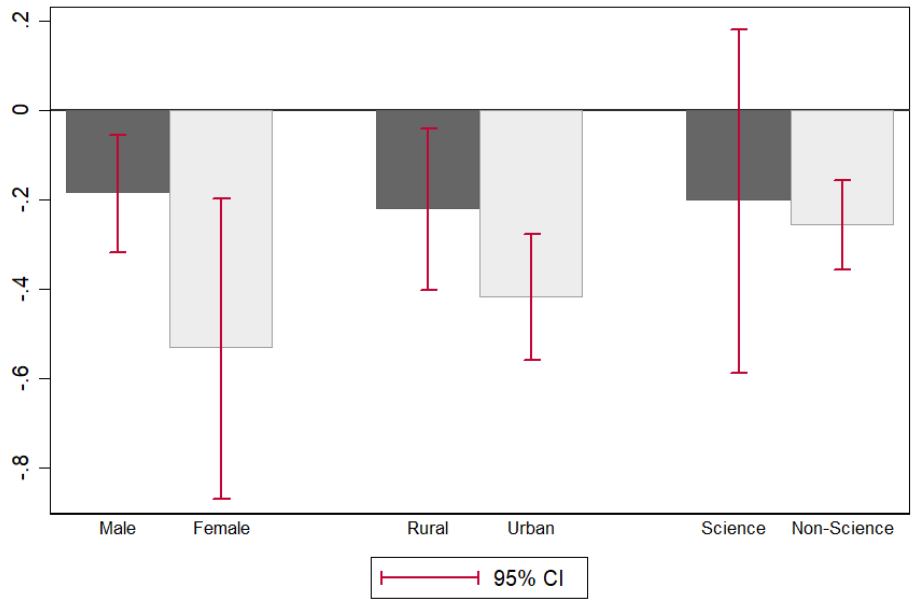
Notes: Own calculations based on Labor Force Survey data.

Figure 3: Heterogeneous treatment effects: Employed



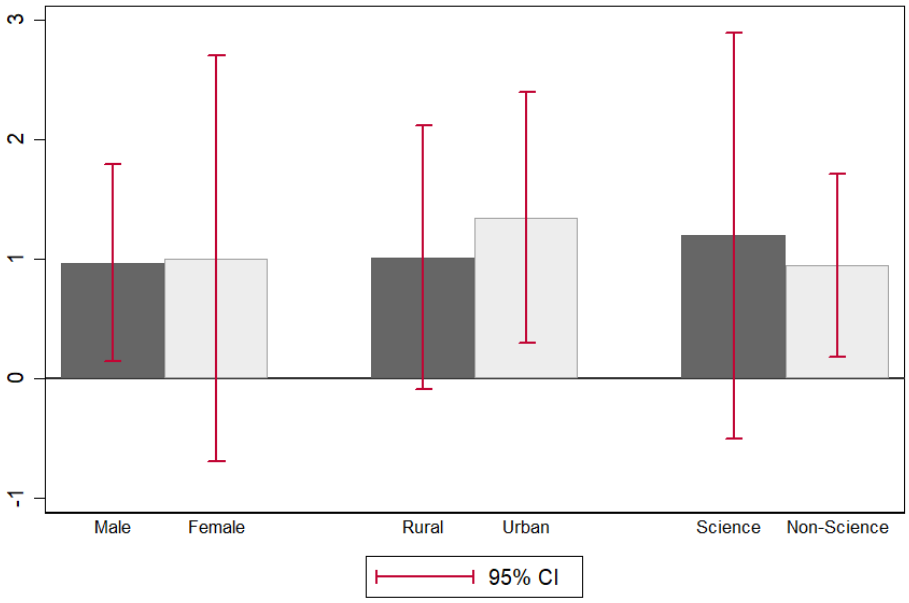
Notes: Coefficients are based on the propensity-score nearest neighbor matching described in Section 3.6 and separately estimated for different subsamples. Standard errors are obtained by bootstrapping.

Figure 4: Heterogeneous treatment effects: Hourly wage



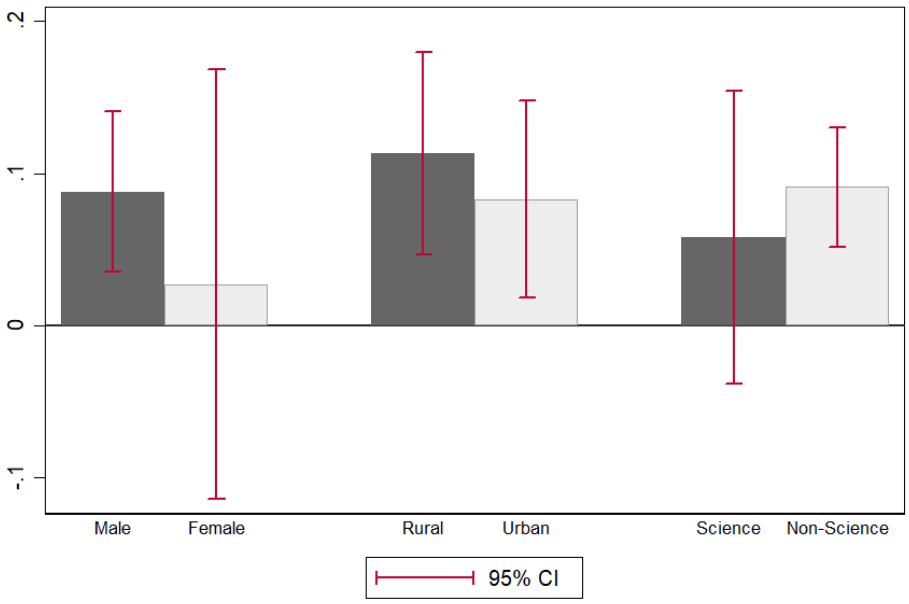
Notes: Coefficients are based on the propensity-score nearest neighbor matching described in Section 3.6 and separately estimated for different subsamples. Standard errors are obtained by bootstrapping.

Figure 5: Heterogeneous treatment effects: Working hours



Notes: Coefficients are based on the propensity-score nearest neighbor matching described in Section 3.6 and separately estimated for different subsamples. Standard errors are obtained by bootstrapping.

Figure 6: Heterogeneous treatment effects: Short-term contract



Notes: Coefficients are based on the propensity-score nearest neighbor matching described in Section 3.6 and separately estimated for different subsamples. Standard errors are obtained by bootstrapping.

Table 1: Mapping of $ISCO - 08$ to occupations

Major occupation groups	(a) $ISCO$ **skill level**	(b) **Skill level**
1. Managers	3 and 4	3
2. Professionals	4	3
3. Technicians and Associate Professionals	3	3
4. Clerical Support Workers	2	2
5. Service and Sales Workers	2	2
6. Skilled Agricultural, Forestry and Fishery Workers	2	2
7. Craft and Related Trade Workers	2	2
8. Plant and Machine Operators, and Assemblers	2	2
9. Elementary Occupations	1	1

Notes: We reduced the official ILO-mapping of skill-levels to occupations (column (a)) to three different skill-levels (column (b)).

Table 2: Mapping of ISCO-08 to ISCDED-97

(a) **Skill level**	(b) *ISCO*	(c) *ISCED*	**Education groups**
3	4	6	Second stage of tertiary education (advanced research qualification)
		5b	First stage of tertiary education, 1st degree (medium duration)
	3	5a	First stage of tertiary education (short or medium duration)
2	2	4	Post-secondary, non-tertiary education
		3	Upper secondary level of education
1		2	Lower secondary level of education
	1	1	Primary level of education

Notes: We reduced the official ILO-mapping of skill-levels (column (b)) to educational attainment (column (c)) to three different skill-levels (column (a))

Table 3: Comparison of over-education shares (ISCO skill mapping method vs statistical method)

	ISCO-skill mapping		**Years of education**			
ISCO (two digits)	Skill-level	Share	Mode	Share	Mean	Share
Armed forces	3	0.000	12	0.225	12.39	0.197
Legislators and senior officials	3	0.000	16	0.030	15.31	0.030
Corporate managers	3	0.000	12	0.342	13.27	0.319
Managers of small enterprises	3	0.000	12	0.286	12.64	0.286
Physical, mathematical and engineering science professionals	3	0.000	12	0.452	13.62	0.387
Life science and health professionals	3	0.000	12	0.450	14.14	0.146
Teaching professionals	3	0.000	16	0.032	15.43	0.095
Other professionals	3	0.000	16	0.010	14.30	0.080
Physical and engineering science associate professionals	3	0.000	12	0.362	13.36	0.319
Life science and health associate professionals	3	0.000	12	0.333	13.07	0.272
Other associate professionals	3	0.000	16	0.000	14.11	0.066
Office clerks	2	0.486	12	0.432	13.69	0.351
Customer services clerks	2	0.318	12	0.318	13.05	0.264
Personal and protective services workers	2	0.076	12	0.076	11.69	0.076
Models, salespersons and demonstrators	2	0.092	12	0.092	11.66	0.092
Skilled agricultural and fishery workers	2	0.053	12	0.018	10.56	0.053
Extraction and building trade workers	2	0.020	12	0.011	10.58	0.020
Metal, machinery and related trades workers	2	0.046	12	0.046	11.38	0.046
Precision, handcraft, craft printing and related trades workers	2	0.138	12	0.138	11.84	0.138
Other craft and related trades workers	2	0.013	8	0.373	9.46	0.340
Stationary plant and related operators	2	0.182	12	0.127	11.88	0.127
Machine operators and assemblers	2	0.129	12	0.109	11.64	0.129
Drivers and mobile plant operators	2	0.026	12	0.026	11.26	0.026
Sales and service elementary occupations	1	0.748	12	0.043	11.10	0.050
Agricultural, fishery and related laborers	1	0.527	12	0.034	10.31	0.034
Laborers in mining, construction, manufacturing and transport	1	0.793	12	0.049	11.20	0.059
Unclassified elementary occupations	1	0.815	12	0.112	11.75	0.112

Table 4: Characteristics of matched and mismatched individuals

	Matched		**Mismatched**		**Difference**
	Mean	S.D	Mean	S.D	
Individual characteristics in period t					
Age	48.79	(9.011)	47.72	(9.262)	*
Male	0.713	(0.453)	0.819	(0.385)	***
Married	0.962	(0.190)	0.935	(0.246)	*
Years of education	13.58	(2.432)	12.13	(2.112)	***
Tenure	15.10	(10.57)	11.58	(9.452)	***
Part time job	0.010	(0.100)	0.019	(0.135)	
Private sector employee	0.340	(0.474)	0.572	(0.495)	***
Employed one year ago	0.963	(0.190)	0.933	(0.251)	**
Employed Partner	0.329	(0.470)	0.201	(0.402)	***
Minority city	0.057	(0.232)	0.046	(0.210)	
Household characteristics in period t					
Number of children $10-18$ years	1.037	(1.142)	1.215	(1.250)	**
Number of children $6-10$ years	0.354	(0.723)	0.361	(0.743)	
Number of children < 6 years	0.179	(0.505)	0.181	(0.486)	
Number of household members > 60 years	0.376	(0.610)	0.308	(0.549)	*
Firm characteristics in period t					
Firm size	2.656	(1.105)	2.500	(1.190)	*
Labor market outcomes in period $t+1$					
Hourly wage	2.389	(1.263)	1.810	(0.891)	***
Monthly income	379.2	(156.6)	315.5	(124.5)	***
Short contract	0.621	(0.485)	0.762	(0.426)	***
Low job satisfaction	0.023	(0.151)	0.036	(0.187)	
Working hours	43.14	(10.58)	45.23	(7.997)	***
Number of observations	3,891		1,716		

Notes:* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Selection into mismatch

	P(Mismatched)	
Individual characteristics		
Age	0.001	(0.006)
Squared age	-0.000	(0.000)
Male	0.045***	(0.013)
Years of education	0.008**	(0.003)
Tenure	-0.003***	(0.001)
Monthly income	-0.001***	(0.000)
Employed one year ago	0.026	(0.028)
Private sector employee	0.024	(0.021)
Married	0.010	(0.026)
Employed Partner	-0.021*	(0.012)
Part time job	0.007	(0.042)
Minority city	0.054**	(0.025)
Household characteristics		
Number of children age $10-18$ years	-0.004	(0.006)
Number of children age $6-10$ years	0.004	(0.012)
Number of children <6 years	-0.021	(0.015)
Number of household members >60 years	-0.016*	(0.009)
Firm characteristics		
Firm size	0.001	(0.006)
Region effects		
Prishtina (baseline)		
Peja	-0.177***	(0.044)
Mitrovice	0.011	(0.049)
Prizren	-0.171***	(0.035)
Ferizaj	-0.209***	(0.044)
Gjilan	-0.091*	(0.049)
Gjakove	0.008	(0.049)
Fixed effects		
Field of study FE	Yes	
Industry FE	Yes	
Wave FE	Yes	
Number of observations	5607	
R-squared	0.20	
Mean of dep. var.	0.22	

Notes: Estimation based on a linear probability model with a binary outcome variable being 1 if mismatched, and zero otherwise. Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Balancing test: Individual characteristics after matching

	Matched	**Mismatched**	**Difference after matching**	
	Mean	Mean	t-stat	
Age	49.603	47.426	-2.73	***
Years of education	13.072	12.92	-0.96	
Tenure	13.439	12.241	-1.35	
Part time job	0.008	0.008	0.00	
Private sector employee	0.506	0.569	1.38	
Minority city	0.025	0.051	1.44	
Monthly income	366.03	321.2	-4.53	***
Number of children $10 - 18$ years	0.987	1.198	1.97	*
Number of household members > 60 years	0.232	0.274	0.93	
Firm size	2.713	2.595	-1.08	
Number of observations	3,891	1,716		

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: The effect of being mismatched on labor market outcomes

	(I) OLS		(II) PS-NN	
A. Employed in *t+1*				
Over-educated	0.000	(0.006)	-0.007	(0.008)
B. Hourly wage in *t+1*				
Over-educated	-0.069**	(0.029)	-0.242***	(0.074)
C. Working hours in *t+1*				
Over-educated	1.013***	(0.241)	0.951***	(0.391)
D. Short term contract in *t+1*				
Over-educated	0.040**	(0.016)	0.079***	(0.022)
E. Low job satisfaction in *t+1*				
Over-educated	0.002	(0.005)	0.011*	(0.006)
Observations	5,607		5,582	

Notes: Standard errors obtained by bootstrapping in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Heterogeneous effects of mismatch

	(1) Male	(2) Female	(3) Rural	(4) Urban	(5) Science	(6) Non-Science
A. Employed in *t+1*						
Over-educated	-0.004	-0.023	0.014	-0.011	-0.022	-0.001
	(0.008)	(0.014)	(0.014)	(0.010)	(0.022)	(0.006)
B. Hourly wage						
Over-educated	-0.186***	-0.533***	-0.222**	-0.418***	-0.203	-0.257***
	(0.067)	(0.172)	(0.092)	(0.072)	(0.196)	(0.051)
C. Working hours						
Over-educated	0.968***	1.005	1.016*	1.349**	1.199	0.950**
	(0.420)	(0.866)	(0.564)	(0.533)	(0.866)	(0.389)
D. Short term contract						
Over-educated	0.088***	0.027	0.113***	0.083**	0.058	0.091***
	(0.027)	(0.072)	(0.034)	(0.033)	(0.049)	(0.020)
E. Low job satisfaction in *t+1*						
Over-educated	0.011*	0.023	0.011	0.006	0.024	0.010*
	(0.006)	(0.015)	(0.008)	(0.006)	(0.020)	(0.005)
Observations	4153	1353	2792	2742	1297	4221

Notes: Estimates based on propensity score nearest neighbor matching.Standard errors obtained by bootstrapping in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Robustness: alternative matching methods

	(I) PS-Radius matching		(II) NN	
A. Employed in *t+1*				
Over-educated	-0.011	(0.012)	0.008	(0.021)
B. Hourly wage in *t+1*				
Over-educated	-0.521***	(0.094)	-0.245**	(0.099)
C. Working hours in *t+1*				
Over-educated	1.995***	(0.670)	0.297	(0.973)
D. Short term contract in *t+1*				
Over-educated	0.153***	(0.044)	0.108**	(0.046)
E. Low job satisfaction in*t+1*				
Over-educated	0.027*	(0.016)	0.004	(0.015)
Observations	5,582		5,582	

Notes: Standard errors obtained by bootstrapping in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Robustness: alternative mismatch definition

	(I) OLS		(II) PS-NN	
A. Employed in *t+1*				
Over-educated	-0.001	(0.005)	-0.004	(0.005)
B. Hourly wage in *t+1*				
Over-educated	-0.064**	(0.027)	-0.176***	(0.061)
C. Working hours in *t+1*				
Over-educated	0.902***	(0.218)	1.046***	(0.293)
D. Short term contract in *t+1*				
Over-educated	0.038**	(0.015)	0.072***	(0.019)
E. Low job satisfaction in*t+1*				
Over-educated	-0.003	(0.004)	0.004	(0.006)
Observations	5,607		5,582	

Notes: Standard errors obtained by bootstrapping in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: The effect of being mismatched on labor market outcomes: Under-education

	(I) OLS		(II) PS-NN	
A. Employed in *t+1*				
Under-educated	-0.003	(0.007)	0.005	(0.009)
B. Hourly wage in *t+1*				
Under-educated	-0.016	(0.025)	0.031	(0.065)
C. Working hours in *t+1*				
Under-educated	-0.894***	(0.280)	-0.867*	(0.507)
D. Short term contract in *t+1*				
Under-educated	0.002	(0.021)	-0.027	(0.033)
E. Low job satisfaction in *t+1*				
Under-educated	0.006	(0.006)	0.012*	(0.007)
Observations	5,607		5,463	

Notes: Standard errors obtained by bootstrapping in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$