Institut für Ost- und Südosteuropaforschung
Institute for East and Southeast European Studies

Arbeitsbereich Ökonomie

IOS Working Papers

No. 347 July 2015

Education mismatch, human capital and labour status of young people across European Union countries

Fabrizio Pompei* and Ekaterina Selezneva†

*University of Perugia, Perugia, Italy, email: fabrizio.pompei@unipg.it
†Institute for East and Southeast European Studies, Regensburg, Germany, email: selezneva@ios-regensburg.de

Institut für Ost- und Südosteuropaforschung
Institute for East and Southeast European Studies

Landshuter Straße 4
93047 Regensburg

Telefon: (0941) 943 54-10
Telefax: (0941) 943 94-27
E-Mail: info@ios-regensburg.de
Internet: www.ios-regensburg.de

Contents

List of Tables

List of Figures

Abstract

This paper analyses the influence of country-level education mismatch on the individual-level relationship between education and the probability of being unemployed or staying in alternative labour statuses, for young people aged 15–34 in 2006, 2008 and 2010, living in 21 EU countries. We assume that young people may fall in five labour market statuses: 1) Employee; 2) Self-employed; 3) Unemployed; 4) In Education; 5) Inactive, and perform a multinomial logit model to study the effects of years of education on relative probability of being in labour statuses 2, 3, 4, or 5, compared to the base category (Employee). Afterwards, we interact the individual-level years of education with a country-level indicator of education mismatch in order to identify the heterogeneous effects of the aggregate mismatch among people with different educational attainments. Results show that more years of education: *i)* reduce the relative probability of being unemployed; *ii)* have a cumulative effect by extending the period of education; *iii)* slightly raise the relative probability to be self-employed. As regards country-level education mismatch, we found that only after 2008 it produces an additional effect on better educated young people by further reducing their relative unemployment risk when it is compared to that of low educated youngsters. This outcome tells us that improving access to university degrees remains the main road to tackle youth unemployment caused by education mismatch, even after the outburst of the current financial and economic crisis.

JEL-Classification: I20; J24; Z13

Keywords: youth unemployment; education mismatch; multinomial logit model

We wish to thank the discussant and other participants in the 13th EACES biannual conference (4–6 September 2014) for their useful comments to a previous version of this paper.

1 Introduction

The labour market mismatch hypothesis is recently regaining ground as a possible explanation, at the macro level, of the severe youth unemployment which is plaguing Europe (OECD, 2014; European Commission, 2013; ILO, 2013; ECB, 2012). Indeed, the dramatic structural changes caused by the Great Recession, have contributed to the creation of the multiple forms of skill mismatch in the European labour markets. The process of labour reallocation – from declining to emerging sectors – is not free from frictions and may be a cause of long-lasting unemployment, especially among the youth. However, it is not clear yet if this aggregate mismatch negatively affects all young people in the same way or, instead, heterogeneous and counterbalancing effects emerge as a result of the different education profiles of individuals. There is, in fact, theoretical and empirical evidence reporting a positive effect of human capital accumulation (i.e., increasing years of education)[1] in reducing unemployment or inactivity risk, albeit this literature is not largely developed (Spence, 1973; Mincer, 1991; Trostel and Walker, 2006). Consequently, if the aggregate skill mismatch means a lack of higher educated young people on the supply side of the labour market, the higher the mismatch, the more favored the scanty better educated youngsters should be. On the other hand, job polarization theories (Acemoglu and Autor, 2012) question a positive monotonic relationship between formal education and the probability of being employed, hence the effects of aggregate education mismatch on the relationship between human capital accumulation and youth labour status remains a testable empirical question.

Starting from the considerations above, this paper aims to bridge the gap between the macro- and micro-level literature on education mismatch, human capital (years of education) and individual labour status of people aged 15–34 in the European Union (EU). More precisely, after investigating whether formal education mattered to avoid youth unemployment /inactivity or to favoring other labour statuses across EU countries over the period 2006–2010, we analyse if the relationship education-labour status is affected by different intensities of the country-level education mismatch and if the current crisis played some role on it, directly or indirectly (i.e., through the aggregate education mismatch). We take into account the multilevel character of the data and the possible cross-level effects (interactions at the country-individual level). In order to do so, we implement the methodology described by Bryan and Jenkins (2013). For dealing with endogeneity between education and labour statuses we apply the two-stage residual inclusion approach (2SRI), suggested by Terza et al. (2010) and Bollen et al. (1995). The policy implications of the study are of interest because understanding whether the crisis changed or not the urgency of improving access to tertiary education in order to reduce structural and long-term unemployment remains a crucial question. The paper is structured as follows. In the next section, a thorough discussion of theoretical and empirical background supporting the hypotheses that drive the empirical analysis is reported. Section 3 presents the econometric

[1] Notwithstanding we are aware they are not exactly the same thing, in this paper we use the terms *education* and *human capital* interchangeably.

strategy, whereas section 4 shows data sources and variables that we use in the estimations. After a brief summary of statistics (section 5), a detailed discussion of the econometric results is given in section 6. The last section is dedicated to the final remarks.

2 Background and Hypotheses

An immense literature addresses a beneficial impact of human capital, and in particular of education, on labour market outcomes as well as a variety of problems caused by a mismatch between acquired and demanded skills at micro and macro-levels.[2]

At the micro-level, the majority of studies combine the analysis of mismatch-caused problems with that of private returns to schooling by challenging the human capital theory of Becker (1964) and Mincer (1974), albeit results on the effects of education on wages and productivity are still far from providing coherent conclusions and a clear-cut picture (Desjardins and Rubenson, 2011; Quintini, 2011; Mac Guinness, 2006). On the contrary, rather few studies have investigated the impact of individual education on the unemployment and inactivity incidence in the working age population;[3] despite this, a clearer beneficial role of education in reducing the unemployment risk is established (Spence, 1973; Nickell, 1979; Mincer, 1991; Blöndal et al., 2002). Highly educated job seekers signalise to employers – through education level achieved – their potentially greater productivity, and thus have more chances to be hired (Spence, 1973). When hired, college graduates demonstrate better ability to acquire firm-specific knowledge during on-the-job training, than less educated workers; for this reason the former experience less job turnover and unemployment (Mincer, 1991).

Trostel and Walker (2006) conclude on a set of both developed and emerging countries that in the first part of the life-cycle individual decision to invest in own's human capital improves both intensive and extensive margins of employment, namely it increases the hours worked and reduces the probability of being unemployed, respectively. In particular, the study: i) clearly shows that the choice to invest in human capital is endogenous to the labour market status (employee, unemployed, etc.) in the first part of the life; ii) demonstrates that the impact of education on inactivity is significant. Trostel and Walker's article fits the strand of the literature on youth unemployment that investigates the chances of being in different labour market statuses, both within and between EU countries. In these studies a beneficial impact of secondary and tertiary educational attainment is established also for a broader set of labour market statuses including self-employment (Millàn et al., 2012), and decisions to continue education (Styczynska, 2013). Alongside individual education, household structure (marital status, presence of children), civic engagement and social activities (association memberships, meeting friends) help to reduce the unemployment and inactivity risk (Millàn et al., 2012; Dietrich, 2012; Pfeiffer and Seiberlich, 2010).

Recent studies also take into account macro-level factors to explain individual labour status, citing among others labour market institutions, unemployment rate and international trade (see De Lange et al., 2014). The higher expected – after completing a university degree – lifetime-

[2] For more details see reviews in Acemoglu and Autor, 2012; Leuven and Oosterbeek, 2011; Desjardins and Rubenson, 2011; Quintini, 2011; Mac Guinness, 2006.

[3] In this case it is more difficult to analyse the effect of individual education mismatch on unemployment due to unavailability of data on vacancies by education composition. However, also studies concerning the simple effect of educational attainment on the probability of being unemployed are much less numerous than those addressing wages and productivity.

earnings and a higher unemployment rate at the regional level, support a decision to continue studies and co-reside with parents over the alternative to work and live alone after having completed high school among Italians aged 18–32 (Giannelli and Monfardini, 2003). In Poland, education also reduces the unemployment risk. However, unlike in Italy, higher regional unemployment rates do not support the further education decision (Pastore, 2012).

Matching problems in the labour markets at the macro level are among the causes of the structural unemployment in both the European Union and the United States since the recent crisis outbreak (Pissarides, 2000, 2013; ECB, 2012; ILO, 2013; European Commission, 2013; OECD, 2014). The crisis has aggravated the education mismatch between labour demand and supply, especially in the EU countries. These imbalances in the educational composition of the labour supply and demand sides are among the causes of appearance of frictions in the process of labour reallocation among sectors. Currently, the labour supply side is characterized by an excess of low educated people and a shortage of those highly educated (European Commission, 2013). One may deduce that people with higher educational attainment should be favored by education mismatch. Marsden et al. (2002) provided indirect evidence on this point by showing the positive effect of tertiary education on the reduction of education mismatch. It means that a higher education mismatch relates to a higher labour demand for educated workers. However, this does not automatically mean that there is a simple monotonic relationship between education and employability, and thus the higher education level lowering the risk of being unemployed. Indeed, a discussion on job polarization is gaining momentum. Concentration of employees in either high paying cognitive occupations or in lower paying manual-service jobs has been observed in both US and EU-countries (Autor et al., 2003; Goos and Manning, 2007; Dustmann et al., 2009; Acemoglu and Autor, 2012). Unlike in 1995–2008, the latest crises has caused extensive changes in the European job structure (Eurofound, 2013). An accelerated task-biased technological change together with a profound institutional transformation of the labour markets have lead to contraction of the mid-pay routine job segment both in manufacturing and in service sectors.

We link together the above-mentioned considerations and investigate whether country-level education mismatch could modify the relationship between educational attainment and the probability of being in different labour status for people aged 15–34 across EU countries over 2006–2010. The aggregated measure of mismatch allows us to study its overall impact not only on employees but on other possible individual labour statuses in which young people could be, especially unemployment and inactivity. In other terms, we want to test the following hypotheses.

H1: the higher the length of individual education, the lower the risk of being unemployed or inactive and the higher the probability of being employee or self-employed. This hypothesis stems directly from predominant evidence on a positive effect of education on employment, even though the job polarization theories are currently challenging this view;

H2: the higher the country-level education mismatch, the stronger the positive effect of education length on reducing unemployment and inactivity risk, at the individual level. As discussed above, the overall effects of education mismatch on unemployment could be different from the specific effects calculated along individual profiles. For example, education mismatch means that highly educated young people are on higher demand than less educated ones, thus the former should be in advantageous position in countries with higher education mismatch;

H3: the higher the country level education mismatch at the moment of the crisis, the stronger the positive effect of education in reducing unemployment risk or inactivity choice at the individual level. The crisis, by increasing education mismatch, should have supported more highly educated youngsters, even though different results could emerge along the different years of education accumulated.

3 Method

Based on the conceptual framework discussed above, we assume that young people may fall in five mutually exclusive unordered labour market statuses: 1) Employee; 2) Self-employed; 3) Unemployed; 4) In Education; 5) Inactive. As indicated in the literature previously discussed, the multinomial logit model (MNL) is the econometric specification that best fits in studying determinants of multiple categorical outcomes:

$$Pr(Y = m|X) = \frac{Exp(X\beta_{m|b})}{\sum_{j=1}^{J} Exp(X\beta_{j|b})}$$

where $J = 1..., 5$, b is the base category (1 Employee), m and j are respectively the specific outcome (labour status) to be examined and the generic outcome, X is the matrix of regressors.

According to Luce (1959) and Cameron and Trivedi (2005, 2009), a restriction is needed to ensure the model identification, namely the sum of probabilities of alternatives has to equal 1. For the MNL the comparison is to a base category which is the alternative normalized to have coefficients equal to zero, $\beta_{b|b} = 0$. In our case, this leads to estimation of four binary logit models for choice between a labour market status m and the base status (1 Employee).

$$\frac{Pr(Y = m|X)}{Pr(Y = b|X)} = \frac{Pr(Y = m|X)}{Pr(Y = 1)} = Exp(X\beta_{m|b}) \tag{1}$$

A positive value of the estimated parameter $\beta_{m|b}$, means that the higher the value of the regressor, the higher the likelihood of being in an alternative labour status m with respect to the probability of being employed. Therefore, the coefficient indicates a change in the relative probability for an outcome and not for the outcome itself.

In order to ensure that the relative probabilities of the alternatives are not correlated among themselves (for example, for statuses Inactivity and Unemployed), we test the validity of the irrelevant alternatives assumption (IIA). This is done with the help of the Small-Hsiao test, which does not reject the hypothesis for any set of outcomes (see table A.1 in the appendix). This means that the alternative-specific errors are uncorrelated and that the odds-ratios for pairs of alternatives are invariant with respect to the expansion (and contraction) of the alternatives set.

Our baseline specification aims to test the validity of hypothesis 1 (H1). Therefore, by taking the log odds version of equation (1), we estimate the following equation (2) on a pooled sample of data taking into account country-level fixed effects and time dummies:

$$\ln \Omega_{m|b} = \alpha_{m|b} + EduYrs\beta_{1,m|b} + P\beta_{2,m|b} + F\beta_{3,m|b} + S\beta_{4,m|b} + \delta_{t,m|b} + \eta_{c,m|b} \tag{2}$$

where m are the four outcomes alternative to the base category b (Employee).[4] *EduYrs* is the main variable of interest at the individual level (years of education); P, F and S are matrices including other personal, family and socio-political characteristics of young people,[5] $\delta_{t,m|b}$ are time dummies $(t = 2006, 2008, 2010)$, $\eta_{c,m|b}$ are country fixed effects and $c = 1..., 21$.

Afterwards, we augment equation (2) with the interaction term of the aggregated country-level education mismatch (EMI) and years of education completed by individuals in order to test the second hypothesis, i.e. the impact of the EMI on the relationship between the individual-level education and relative probabilities of choice of the outcomes:

$$\begin{aligned}\ln \Omega_{m|b} = \alpha_{m|b} + (EMI * EduYrs)\beta_{1,m|b} + EduYrs\beta_{2,m|b} \\ + P\beta_{3,m|b} + F\beta_{4,m|b} + S\beta_{5,m|b} + \delta_{t,m|b} + \eta_{c,m|b}\end{aligned} \tag{3}$$

We include EMI in the interaction term only, but not as alone standing term. In our specification, country dummies are intended to capture all other possible country-specific variables, such as business cycle and institutions.[6] However, as we will see below, the main effect of EMI will be thoroughly studied and it will play a central role in the econometric strategy.

Eventually, as explained in the previous section, we also pay attention to the impact of crisis on the picture described above, by adding the following interactions:

$$\begin{aligned}\ln \Omega_{m|b} = \alpha_{m|b} + (EMI * EduYrs * 2010)\beta_{1,m|b} + (EduYrs * 2010)\beta_{2,m|b} \\ + (EMI * EduYrs)\beta_{3,m|b} + X\beta_{4,m|b} + \delta_{t,m|b} + \eta_{c,m|b}\end{aligned} \tag{4}$$

where now X is a matrix including all personal, family and socio-political characteristics reported in the previous terms P, F and S.

Two main problems undermine these specifications, the endogeneity of education with respect to labour status (Riddell and Song, 2011; Trostel and Walker, 2006) and the multilevel nature of the data that we use for the econometric analysis (Bryan and Jenkins, 2013).

As regards endogeneity, we follow several authors (Ivlevs and King, 2012; Wooldridge, 2010; Terza et al., 2010; Bollen et al., 1995) and prefer a 2-stage residual inclusion regression (2RSI) in place of the conventional 2-stage predictor substitution approach, given that all simulation studies conducted by the authors above confirm the superiority of 2RSI in the non-linear models. The 2RSI method consists in setting up an OLS regression in the first stage in which, similarly to the conventional 2-stage predictor substitution approach, we regress our continuos endogenous variable *years of education* on instrumental variables.

$$EduYrs = \alpha + IV\beta_1 + P\beta_2 + F\beta_3 + S\beta_4 + \delta_t + \eta_c \tag{5}$$

[4] Hence our multinomial model includes 4 regressions to study the probability of being in labour market status $2; 3; 4; 5$ (Self-employed, Unemployed, In education, Inactive) compared to the probability to be in 1 (Employee, the base category)

[5] The components of these matrices are described in detail in Section 4.

[6] Needless to say that including EMI alone, besides country dummies, would have caused multicollinearity problems.

where IV is a matrix containing a set of excluded instruments that we thoroughly discuss in the next section, P, F and S are the same matrices of equation (2), containing all the individual level control variables (included instruments), δ_t and η_c are time and country dummies respectively.

In the second-stage regression, however, the endogenous variables are not replaced. Instead, the first-stage residuals are included as additional regressors in second-stage estimation, besides the actual value of $EduYrs$ (Terza et al., 2010).

$$\begin{aligned}\ln \Omega_{m|b} = \alpha_{m|b} + EduYrs\beta_{1,m|b} + \textbf{1-stage Resid}\beta_{2,m|b} \\ +P\beta_{3,m|b} + F\beta_{4,m|b} + S\beta_{5,m|b} + \delta_{t,m|b} + \eta_{c,m|b}\end{aligned} \tag{6}$$

where **1-stage Resid** are the included residuals stemming from the first stage.

According to Bollen et al. (1995) and Ivlevs and King (2012), we test the relevance of instruments in the first stage by means of an F-test, discuss the endogeneity/exogeneity of $EduYrs$ by simply reporting the Wald test for the coefficients of **1-stage Resid** and take into account the exclusion restrictions by comparing the log-likelihood between the reduced form and the structural equation in the second stage MNL model. As regards the exclusion restrictions, in the reduced form we replace $EduYrs$ with the set of instruments IV, whereas in the structural equation we only include the predicted value for $EduYrs$ and omit instruments. If the instruments only indirectly influence the labour status, through their effects on $EduYrs$, the log-likelihood of the reduced and structural equations should be similar (Bollen et al., 1995). Hence, we conduct this test on the identifying assumptions to prove the exogeneity of instruments. Moreover, we find $EduYrs$ as being endogenous; consequently we include the residuals in specifications (2), (3) and (4).

Bryan and Jenkins (2013) highlighted problems arising with multi-country datasets in which there are observations at the individual level nested within a higher level (countries). On the one hand, this multi-level structure provides useful information about *country effects* as well as *individual effects*, and also about interactions between them (*cross-level effects*); on the other hand, the drawback due to the small number of groups (countries) is not alleviated by the large size of the sample at the individual level (thousands of observations). This means that the desirable properties of regression model parameter estimates for individual-country level interactions, such as consistency and efficiency, are questionable when the number of countries is below 30. For this reason we follow Bryan and Jenkins (2013) in performing a two-step approach that is useful to disclose the statistical significance of the variable of interest at country-level. In other terms, we consider the baseline specification in our analysis with correction for endogeneity (equation (6)) as the first step. The only difference is that we estimate separately three regressions of equation (6) for each year (2006, 2008, 2010). Afterwards, we take the country-intercepts from these three regressions and express them as a linear function of EMI at country level. In the second step estimation we therefore regress coefficients of country intercepts on the country-level variable EMI, using OLS. We repeat this regression for each outcome m (Unemployed, Self-employed, Education, Inactive) stemming from the first step.

$$\hat{\eta}_{c,t} = \alpha + EMI_{c,t}\beta_1 + Lab.MarketLiberal_{c,t}\beta_2 + GDP\_Shock_{c,t}\beta_3 + \delta_t + \varepsilon_{c,t} \qquad (7)$$

where $c = 1, ...21$ countries and $t = 2006, 2008, 2010$ years;[7] $\hat{\eta}_{c,t}$ are the estimated parameters for the country intercept c and year t, describing the relative probability to be in labour status m; EMI is the same proxy for the education mismatch used in equations (4) and (5); $Lab.MarketLiberal_{c,t}$ and $GDP\_Shock_{c,t}$ are two country-level control variables that take into account labour market institutions and business cycles, respectively.

This supplementary approach offers two advantages to our econometric analysis: *(i)* we have a preliminary assessment concerning the reliability of EMI as country-level effect, namely a significant coefficient for EMI means that its main effect on the average relative probability of being in a labour status is binding; *(ii)* it provides useful information on the sign (direction) of the main effect of EMI in order to clarify the interpretation of cross-level effects (interaction terms) in the main specifications (equations (4) and (5)). Concerning this last point, independently from the statistical significance of the EMI coefficients that we obtain in regressions (7), we again follow a Bryan and Jenkin's suggestion and perform a graphical analysis by plotting $\hat{\eta}_{c,t}$ on EMI for each labour status.

[7] This leads having approximately 60 observations or slightly under due to missing data.

4 Data sources and variables

We drew all the individual level variables from the European Social Survey (ESS).[8] The cumulative data files integrate cross-section information collected in 2006, 2008, and 2010, respectively. Unfortunately, the first edition of the 6th round (2012) excludes a large number of countries we wanted to take into account in this study, so we limited our investigation to the period 2006–2010. In any case, also for this period there are data missing for some countries, thus we considered only 21 European Union members and excluded Italy, Austria, Malta, Luxembourg, Latvia, Lithuania, Romania.

The key variable of interest represents a self-reported labour status at the moment of the interview for young individuals aged 15–34. More precisely, the status of *Employee* is our base outcome and includes all young employees (contract with limited and unlimited duration). *Self-employment* is the second status and includes self-employed and persons working for their own family business. Unemployed actively looking for a job is the third status (*Unemployed*), whereas youngsters still in education is the fourth one (*Education*). Eventually, *Inactive* is a residual category that includes unemployed young people who are not actively looking for a job and are not in education, and young people who are inactive for different reasons (permanently sick or disabled, community or military service, housework, looking after children, other).

As regards the key explanatory variable at the individual level, we took the full time completed *years of education* that includes compulsory schooling. In addition, according to the literature mentioned above, we considered as controls a set of variables describing personal characteristics (age and gender), family characteristics (number of family members, presence of children, labour/capital income as main source of the household income), political rights (citizenship), social relationships (frequency of meetings with friends or colleagues, taking part of events with other people, membership in trade unions).

Additionally, we also drew four binary variables from ESS to instrument the *years of education* at the first stage of the 2SRI approach. These variables are the *father's tertiary education level* and three proxy variables for *altruism*, *equalitarianism* and *environmentalism* from the section of ESS database dedicated to the human scale values [9]. We assume that these four binary variables are correlated to the *years of education* while not having an impact on probability of being in any of the five labour statuses considered. The parents' educational attainment is largely used in literature as an instrument for education (Ivlevs and King, 2012; Parker and Van Praag, 2006; Trostel et al., 2002) even though its exogeneity with respect to income or labour status has been questioned (Card, 1999). We only use the father's education due to the excessive number of missing data for the mother's education. As regards human scale values, there is a growing consensus in defining basic values as cognitive representations of desirable

[8] ESS is an academically-driven multi-country survey aiming at developing a series of European socio-economic indicators.

[9] These three proxy variables are coded as unity when an individual responds *i) very much like me; ii) like me; iii) somewhat like me* to a relevant question, and zero in case of other answer. We deduced *altruism* from the question *important to help people and care for others' well-being*; *equalitarianism* from *important that people are treated equally and have equal opportunities*; and *environmentalism* from *important to care for nature and environment*.

goals that serve as guiding principles in the life of a person. We defined as *altruism*, *equalitarianism* and *environmentalism* orientations that Piurko et al. (2011) grouped in the broader categories of *benevolence* and *universalism*, namely self-transcendence values located at the opposite of self-enhancement values (power, achievement) that encourage and legitimize the pursuit of self-interest. Therefore, it is plausible to guess that the presence/absence of these values could be correlated with increasing years of education to support mental openness and the desire to understand the world; instead, they should have nothing to do with the behavior in pursuance of self-interest, that could be correlated with the probability of being employed, unemployed, inactive or self-employed.

As far as *EMI* is concerned, we followed the approaches of ILO (2013); European Commission (2013) and ECB (2012), and constructed the country-level educational mismatch as a dissimilarity index. In particular, the index compares the differences in the educational attainment (coded as three levels of education completed) between two groups, of employed and unemployed (or labour force). Indeed, the index is estimated on two proxies of the labour supply, namely on the pools of unemployed (*EMI-un*) and of labour force (*EMI-lf*):

$$EMI - un = \frac{1}{2}\sum_{i=1}^{3}|\frac{E_i}{E} - \frac{U_i}{U}|$$
$$EMI - lf = \sum_{i=1}^{3}\frac{LF_i}{LF}|\frac{E_i}{E} - \frac{LF_i}{LF}|$$

where *i* is an indicator for the level of education coherent with the International Classification of Education 2011 (ISCED, 2012);[10] $\frac{E_i}{E}$ is the proportion of the employed with education level *i*; $\frac{U_i}{U}$ and $\frac{LF_i}{LF}$ the proportion of the unemployed and labour force respectively, with education level *i*.

According to ILO (2013), if the unemployment rate in *EMI-un* is the same among the primary, secondary and tertiary education level graduates, the index equals to zero; no dissimilarity between groups is observed. The index equals unity in the case of complete dissimilarity among groups; that is, for example, when all primary and tertiary education graduates are employed, while those with secondary education are unemployed. The index can also be interpreted as the percentage of unemployed individuals that should be reallocated across skill levels to balance labour supply and demand. *EMI-lf*, instead, does not range from zero to one, albeit also in this case the score of the indicator is low if the skill composition of the employed reflects the labour force's skill composition, while the value is high if the education groups that are highly represented in the labour force are not in terms of employment (European Commission, 2013). Our calculations, presented in the following section, show that the ranking of countries differs for these two definitions of EMI, therefore, both indices are used later in order to test the robustness of our results. Data on shares of employed, unemployed and labour force come from the Eurostat database (country-level labour force survey).[11]

[10] 1) Primary or less and lower secondary education (levels 1–2); 2) Upper secondary and post-secondary non tertiary education (levels 3 and 4); 3) from short-cycle tertiary education on, i.e. bachelor, master (levels 5–8)

[11] It must be noticed that due to data availability, the cohort upon which these EMIs are built (persons 15–39 years old) is slightly different from the one we use at the micro-level (persons 15–34 years old). In addition, we suppose one year delay in the effect of aggregate mismatch on individual outcomes, therefore EMIs referring to 2006, 2008 and 2010 are actually calculated on 2005, 2007 and 2009 respectively.

Lastly, as regards the country-level variables that we inserted as control in equation (7), $GDP-Shock$ was calculated from Eurostat data and it is the difference between the annual variation of GDP (e.g. 2005–2004 for the first year) and the 5-years annual average of GDP (chain-linked volumes, reference year 2005);[12] the proxy for labour market liberalization is a composed indicator that comes from the Fraser Institute database and combines six different components of country-level labour market institutions: *i)* hiring regulation and minimum wage; *ii)* hiring and firing regulations; *iii)* centralized collective bargaining; *iv)* hours regulation; *v)* mandated cost of worker dismissal; *vi)* military conscription.

[12] We successfully confirmed the robustness of the estimation results with other several versions of this index, namely a 3-year annual average and the annual change. These results are available upon request.

5 Descriptive statistics

Table 1 reports descriptive statistics for the whole sample that refers to the 3 rounds under scrutiny (2006, 2008 and 2010). The overall number of observations varies on average from about 9,000 in 2006 to more than 10,000 in 2008 and 2010. The first five rows of the Table describe the five labour statuses of interest, whereas the remaining ones are the explanatory variables at the individual and country level. As expected, over the 5 years that include the outburst of the recent crisis, the employment rate remarkably decreased from 48.6 to 43.53%, whereas the percentage of unemployed on total population (15–34) increased from 5.27 to 7.64%.[13] At the same time, the percentage of young people in education increased from 29.20 to 32.92% and the share of inactivity slightly decreased from 12.08 to 11.42%. Both average age and average years of education remained stable at around 24.7 years old and 13.30 years, respectively. According to ISCED (2012), the latter number corresponds approximately to the end of upper secondary education. Indeed, about half of all the young people in the sample have a secondary level of educational attainment, one third shows only the primary education level, whereas the share of highly educated people varies between 16.26 and 18.20%. Both the nature of household income and the number of family members are important determinants of inactivity and other labour statuses; table 1 shows that the majority of youngsters in the sample live in households in which labour income is the main source of wealth and the average number of family members is slightly above 3. However, the share of young people with children is not negligible, even though it decreased from 26.70 to 24.33%.

Eventually, in line with the previous literature, we observe an overall increase in the two indices of the educational mismatch between 2006 and 2010, for both average and standard deviation. In particular, according to the ILO (2013) interpretation of the *EMI-un*, in 2010 about 20% of youth unemployment needed to be reallocated according to the educational attainment composition of employment, to reduce the mismatch to zero. Figures 1 and 2 chart more detailed levels and variations of *EMI-un* and *EMI-lf* across countries. For the majority of them we observe a similar ranking between the two indexes (see the country localization in the quadrants). Sweden, Finland, France, Belgium and Estonia maintained a very high education mismatch, whereas for Greece, Portugal, Cyprus, Czech Republic, Slovenia and Denmark the two EMIs remained at the bottom of the ranking.[14] As regards the changes over time, the bulk of countries experienced an increase in EMI, especially Spain, Ireland, Portugal and Estonia. According to the theory discussed above, this fact adds to the output drop of the current crisis in explaining the rise in the ratio of unemployed to population for young people (see figure B.1 in the appendix). However, both levels and variations of education mismatch could have had

[13] It is worth noting that this is not an unemployment rate that normally corresponds to the ratio of unemployment to the labour force but an unemployment to population ratio. However, the share of unemployment on population aged 15–34 that we show in Table 1 is not much different from that coming from Eurostat aggregate labour force survey data referring to the same countries: 7.74% in 2006, 6.79% in 2008 and 9.43% in 2010.

[14] The main differences between the two indexes concern the Netherlands, Germany and Poland. Especially the latter two countries show a slight increase in the mismatch according to the unemployment based *EMI-un* (see also ILO, 2013, p.93) and a reduction in the labour force version represented by *EMI-lf*. These differences support the choice to use both of them.

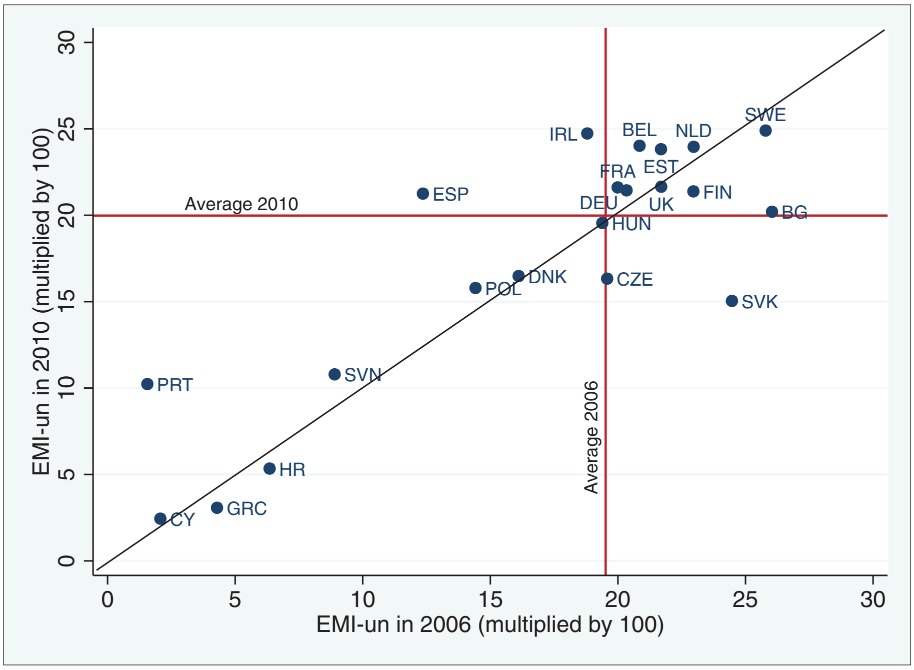
a different impact on individual labour statuses if youngsters decided to invest in education. Indeed, in all countries under scrutiny education mismatch means that besides an excess of low educated people there is an important deficit of highly educated people on the labour supply side, as the graphical representation of *EMI-un* and *EMI-lf* for 2010 clearly discloses (see figures B.2 and B.3 in the appendix). Therefore, the higher the education mismatch, the lower the risk of unemployment or inactivity should be for people with better education.

Table 1: Summary Statistics for Variables used in the econometric analysis

	2006			2008			2010		
	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.
Employees	8933	48.60	49.98	10789	47.12	49.92	10882	43.53	49.58
Self-employed	8933	4.84	21.45	10789	5.33	22.47	10882	4.50	20.72
Unemployed	8933	5.27	22.35	10789	6.07	23.88	10882	7.64	26.57
In Education	8933	29.20	45.47	10789	30.13	45.89	10882	32.92	46.99
Inactive	8933	12.08	32.60	10789	11.35	31.72	10882	11.42	31.80
Age	9050	24.75	5.75	10862	24.75	5.68	10952	24.55	5.73
Gender (male=1)	9041	49.70	50.00	10858	48.77	49.99	10944	48.44	49.98
Years of Education	8949	13.30	3.37	10794	13.32	3.29	10836	13.39	3.31
Primary Ed.	5731	30.91	46.22	7294	31.02	46.26	10899	33.52	47.21
Secondary Ed.	5647	53.20	49.90	7153	51.95	49.97	10899	48.31	49.97
Tertiary Ed.	5647	16.26	36.91	7153	17.56	38.05	10899	18.20	38.59
Citizenship	9044	94.59	22.62	10857	93.94	23.85	10949	93.30	25.01
Disconnected	9036	3.67	18.79	10848	3.72	18.92	10940	4.16	19.98
No Social Activ.	8924	27.46	44.63	10734	27.95	44.88	10876	28.03	44.92
Children(yes=1)	9002	26.70	44.24	10826	24.36	42.93	10946	24.33	42.91
H.Labour Income	8730	88.26	32.19	10625	88.74	31.62	10594	86.98	33.65
H. Capit. Income	8730	2.75	16.36	10625	2.75	16.34	10594	2.89	16.77
Trade Un. Member	8994	13.26	33.91	10815	11.99	32.49	10906	11.71	32.16
Family Members	9038	3.41	1.50	10856	3.37	1.47	10947	3.41	1.44
EMI-un	9050	19.64	4.26	10862	18.64	6.34	10952	19.98	4.64
EMI-lf	9050	1.10	0.38	10862	0.94	0.31	10952	1.22	0.52
GDP_shock (t-1)	9050	42.60	50.50	10862	29.73	92.07	10952	–605.66	309.17
Lab.Market Liberal.	8443	6.03	1.46	9536	5.95	1.33	9789	6.46	1.13

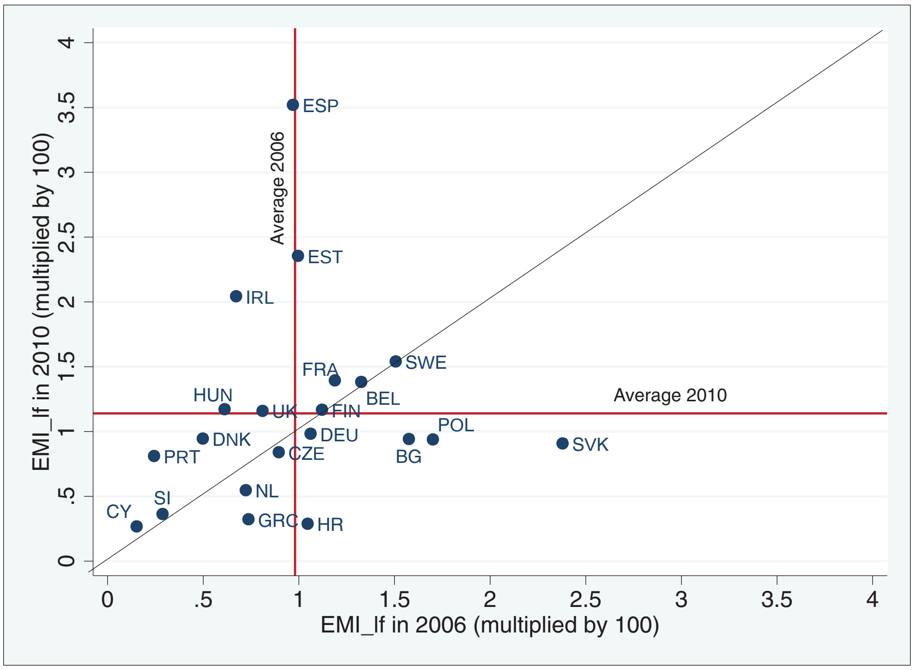
Weighted statistics according to the ESS sample weights. All variables are percentages, with the exception of Age, Family Members, Years of Education and Labour Market Liberal.

Figure 1: Education Mismatch Index (Employment vs Unemployment) between 2006 and 2010



Notes: BG – Bulgaria, CY – Cyprus, CZE – Czech Republic, DEU – Germany, ESP – Spain, EST – Estonia, FIN – Finland, FRA – France, HR – Croatia, HUN – Hungary, GRC – Greece, NLD – The Netherlands, POL – Poland, PRT – Portugal, SVK – Slovak Republic, SVN – Slovenia, SWE – Sweden, UK – United Kingdom.

Figure 2: Education Mismatch Index (Employment vs Labour Force) between 2006 and 2010



Notes: see notes under figure 1.

6 Estimation results

The table 2 shows the coefficients for the MNL model of the baseline specification (equation (2)), estimated on the pooled sample with time and country dummies (27,887 observations). These results partially support the hypothesis 1, namely an extra year of education reduces the probability of being unemployed or inactive, compared to the probability of being employed. Apparently, the number of years of education has no effect on the relative probability of being self-employed, while it has a remarkable impact on the probability of staying in education. The latter is likely to signalise a persistence in the choice of education for the cohort aged 15–34. Almost all other control variable coefficients are significant with the expected sign, according to the literature on youth unemployment (Dietrich, 2012). In particular, until a certain threshold, an additional year of age reduces the relative likelihood of being unemployed, inactive or in education. The risk of being unemployed or inactive is lower for men compared to that of women. When capital income is the main source of the household′s wealth, we observe higher relative probability of being self-employed, in education or inactive, whereas the opposite holds for cases where the labour income is the main source. Eventually, we also find that a relative probability of being unemployed or staying in education is remarkably higher for two years after the beginning of the crisis, in 2010. However, the coefficient of our main interest, related to the years of education, may be subject to a bias due to endogeneity (Trostel and Walker, 2006). As discussed in section 3, we implemented the 2SRI method to tackle this problem and present the respective estimation results in table 3.

First of all, the instruments (father's tertiary education, equalitarianism, altruism and environmentalism) were proved to be relevant and positively correlated to years of education, as the F-statistic value reported at the bottom of table 3 suggests.[15] The test for validity of exclusion restrictions signals that the instruments are also exogenous with respect to the outcomes in the second stage.[16] In addition, the significant coefficients for *1-stage Residuals* indicate that the variable *years of education* is endogenous (Terza et al., 2010; Bollen et al., 1995). As we can see, controlling for endogeneity makes the impact of education in reducing the probability of unemployment stronger. Differently from results of table 2, now a significant and positive effect of education on self-employment emerges and the impact of education on the relative probability of being inactive is slightly positive and significant. This last result apparently contradicts hypothesis 1 and needs some additional discussion. Indeed, by inserting the *1-stage Residuals* it is possible of us to take into account omitted variables such as innate ability, qualitative aspects of education and tasks, as job polarisation theories predict. These aspects might play a role in discouraging youngsters who made the wrong choice in the education field, from entering the labour market. Instead, if higher educated young people decide to seek a job, it may suggest a previous choice of a 'right education', hence the risk of being unemployed is inversely related to the years of education.

[15] See also table A.2 in the appendix. According to Bollen et al. (1995), excluded instruments are relevant when the adjusted R-squared in the first stage OLS regression is above 0.30.

[16] The reduced and structural equations mentioned in section 3 show a very similar Log-Likelyhood and H0 cannot be rejected. It means that instruments influence the probability of being in an alternative labour status only through years of education and do not have any direct effect on these outcomes.

Table 2: Effects of education on labour status of young people aged 15–34. MLN Model: baseline specification; raw coefficients. (Base category: Employee)

	Self-employed	Unemployed	Education	Inactive
Years of Education	–0.002	–0.069***	0.202***	–0.052***
	(0.009)	(0.009)	(0.010)	(0.008)
Age	–0.469***	–0.264***	–1.658***	–0.814***
	(0.075)	(0.060)	(0.048)	(0.047)
Age^2	0.009***	0.003***	0.025***	0.013***
	(0.001)	(0.001)	(0.001)	(0.001)
Gender (male=1)	0.676***	–0.153***	–0.544***	–1.136***
	(0.061)	(0.057)	(0.043)	(0.047)
Citizenship	0.396***	–0.059	0.010	0.002
	(0.137)	(0.108)	(0.109)	(0.092)
Disconnected	–0.234*	–0.304**	–0.702***	–0.060
	(0.138)	(0.127)	(0.141)	(0.089)
No Social Activities	–0.016	0.243***	–0.100**	0.339***
	(0.064)	(0.061)	(0.049)	(0.047)
Children	0.313***	–0.080	–0.647***	1.372***
	(0.076)	(0.077)	(0.078)	(0.059)
H.Labour Income	–0.272	–3.582***	–2.789***	–3.066***
	(0.190)	(0.092)	(0.096)	(0.089)
H.Capital Income	1.228***	–0.384*	1.432***	0.409*
	(0.335)	(0.219)	(0.205)	(0.210)
Trade Un. Member	–0.952***	–0.374***	–0.684***	–0.657***
	(0.092)	(0.086)	(0.073)	(0.068)
Family Members	0.013	0.190***	0.153***	0.213***
	(0.025)	(0.021)	(0.017)	(0.017)
Year-2008	0.062	0.243***	0.025	0.022
	-(0.073)	(0.078)	(0.054)	(0.055)
Year-2010	0.132*	0.659***	0.423***	0.060
	(0.074)	(0.075)	(0.053)	(0.057)
Country dummies	Yes	Yes	Yes	Yes
Obs		27887		
p-value-Overall Model		0.000		
Pseudo-R^2		0.33		

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses.

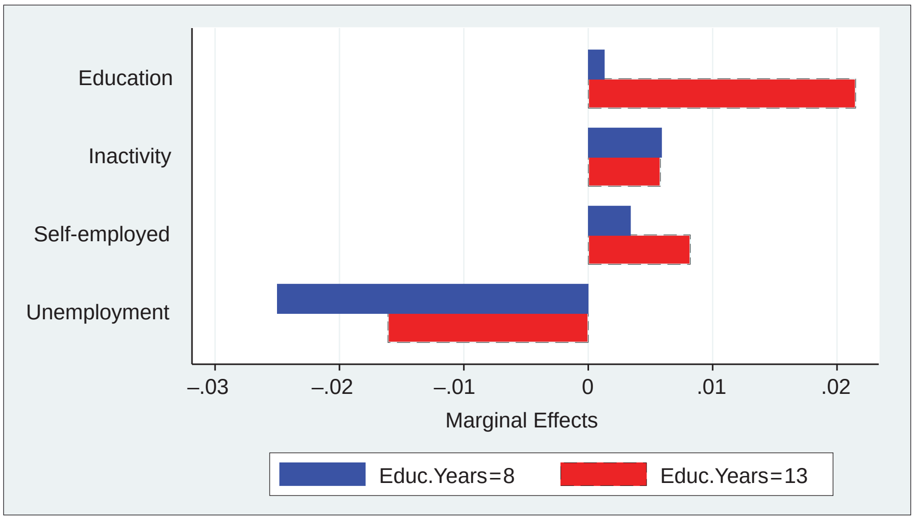
Table 3: Effects of education on labour status of young people aged 15–34. MNL model: endogeneity control with 2-Stages Residual Inclusion method; raw coefficients. (Base category: Employee)

	Self-employed	Unemployed	Education	Inactive
Years of Education	0.219***	–0.149***	0.573***	0.097**
	(0.049)	(0.056)	(0.035)	(0.041)
1-stage Resid.	–0.230***	0.083*	–0.389***	–0.155***
	(0.051)	(0.047)	(0.036)	(0.042)
Age	–0.784***	–0.141	–2.193***	–1.041***
	(0.108)	(0.105)	(0.072)	(0.079)
Age^2	0.014***	0.001	0.033***	0.017***
	(0.002)	(0.002)	(0.001)	(0.001)
Gender (male=1)	0.800***	–0.207***	–0.366***	–1.080***
	(0.065)	(0.065)	(0.047)	(0.053)
Citizenship	0.337**	–0.081	–0.315***	–0.089
	(0.149)	(0.117)	(0.115)	(0.102)
Disconnected	–0.055	–0.328**	–0.428***	0.086
	(0.146)	(0.138)	(0.150)	(0.096)
No Social Activities	0.047	0.220***	0.029	0.382***
	(0.067)	(0.065)	(0.052)	(0.050)
Children (yes=1)	0.653***	–0.196*	–0.135	1.635***
	(0.102)	(0.110)	(0.094)	(0.083)
H.Labour Income	–0.463**	–3.447***	–3.140***	–3.152***
	(0.201)	(0.108)	(0.105)	(0.100)
H.Capital Income	1.016***	–0.145	0.912***	0.288
	(0.348)	(0.242)	(0.221)	(0.227)
Trade Un. Member	–1.049***	–0.346***	–0.868***	–0.705***
	(0.096)	(0.092)	(0.078)	(0.072)
Family Members	0.030	0.189***	0.196***	0.224***
	(0.026)	(0.022)	(0.018)	(0.018)
Year-2008	0.018	0.253***	–0.044	–0.019
	(0.075)	(0.080)	(0.055)	(0.057)
Year-2010	0.068	0.652***	0.298***	–0.010
	(0.078)	(0.080)	(0.056)	(0.061)
Country dummies	Yes	Yes	Yes	Yes
Obs.	26245			
p-value-Overall Model	0.000			
Pseudo-R^2	0.33			
F test for the relevance of instruments in the first stage: F(4)=314.35; p-value=0.000				
Test for validity of exclusion restrictions: *Log-Likelihood Reduced-Form Equation(a)= –23120.932*; *Log-Likelihood Structural Equation(b)= –23127.321*; *Ho: a=b; p-value=0.385*				

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses. Significant t-test for the coefficients of 1-stage Resid. indicates that education is endogenous.

The coefficients reported in table 3 only tell us about relative probabilities, whereas they say nothing in terms of real magnitude and sign of the effects of years of education (Long and Freese, 2006; Cameron and Trivedi, 2009). For this reason, we calculated the marginal effects of the latter[17] and considered individuals in the sample at two different levels of accumulated education: 8 years of education, that approximately corresponds to the end of primary (lower secondary) education and the beginning of upper secondary education; 13 years of education, that is the end of upper secondary and the beginning of tertiary education (ISCED, 2012). As regards the other independent variables, we took age at 26 and the sample mean for all other regressors. At this stage of life, if we take into account *years of education=8* we are studying the effect of an additional year of education for people aged 26 who completed at least primary education; whereas if we consider *years of education=13* we are evaluating people aged 26 with more years of education than those with upper secondary educational attainment. Marginal effects, depicted in the figure 3, suggest that an additional year of education for those with secondary education (*years of education=8*), reduces the probability of being unemployed by –0.025 (2.5 percentage points). This value largely offsets the positive effects on probabilities of being in alternative statuses and indicates that employment (the base category), as complement of unemployment, is the most probable outcome. In the case of *years of education=13*, an extra year of education reduces the probability of being unemployed by –0.016 and increases the probability of being in education by 0.021. However, the size of the effects on inactivity and self-employment is negligible and leads us to conjecture that human capital accumulation is especially important to reduce the risk of unemployment and to reinforce the choice to extend the period of education for higher educated people.

Figure 3: Marginal effects for years of education at age=26 and at the sample mean for other regressors



[17] We used regression results from table 3.

Now, when we have established a beneficial impact of human capital, in the form of formal education, on youth employability, the question remains whether this strategy is successful for young people in countries heavily affected by educational mismatch (hypothesis 2). In addition, we want to investigate if the beginning of the current crisis introduced any changes in this relationship, given that in almost all countries there was a remarkable surge in the education mismatch (hypothesis 3). As Bryan and Jenkins (2013) suggest, we start with presenting the main effects of education mismatch at country level. Tables 4 and 5 report the results of OLS estimates in which we regressed the country intercepts of the MNL model in table 3 on EMI-un and EMI-lf, respectively.[18] According to the macro-level theoretical and empirical evidence we mentioned in section 2, EMIs positively and significantly affect the average probability of being unemployed at the country level. This holds especially for EMI-un, where a one point increase in the mismatch index boosts the average relative probability of being unemployed by 0.015 both with and without macro-level control variables (GDPshock and labour market institutions). This result is less robust for EMI-lf, where the coefficient is not statistically different from zero if control variables are included (table 5). As regard other labour statuses, education mismatch exerts a significant and positive impact only on the average relative probability of staying in education, whereas no significant influence has been found on self-employment and inactivity. Liu (2012) provides support for the positive relationship between country level mismatch and probability of extending education by pointing out that very often people, experiencing job search difficulties due to education mismatch, decide to acquire new skills trough vocational training or more formal education. A graphical analysis in the appendix (see figure B.4 and B.5) basically corroborates the evidence of OLS regressions.

Table 4: Education Mismatch effects at country level. EMI unemployment version (OLS regression)

	Unempl.	Self-empl.	Educ.	Inact.	Unempl.	Self-empl.	Educ.	Inact.
EMI-un	0.015^{**}	-0.001	0.014	0.003	0.015^{***}	0.008	0.021^{*}	0.010
	(0.006)	(0.006)	(0.009)	(0.006)	(0.005)	(0.006)	(0.011)	(0.006)
Year-2008	0.813^{***}	2.510^{***}	-0.440^{**}	3.178^{***}	0.787^{***}	2.454^{***}	-0.437^{*}	3.138^{***}
	(0.104)	(0.086)	(0.205)	(0.114)	(0.107)	(0.089)	(0.225)	(0.115)
Year-2010	0.312^{***}	1.547^{***}	1.082^{***}	2.272^{***}	0.055	1.448^{***}	0.742^{**}	2.204^{***}
	(0.110)	(0.093)	(0.183)	(0.113)	(0.173)	(0.162)	(0.337)	(0.176)
GDPshock					-0.037	-0.011	-0.054	-0.007
					(0.027)	(0.018)	(0.041)	(0.023)
Lab.Mark.Lib.					0.004	-0.054	-0.148^{*}	0.027
					(0.035)	(0.041)	(0.079)	(0.049)
Constant	4.240^{***}	4.527^{***}	19.900^{***}	8.199^{***}	4.195^{***}	4.718^{***}	20.677^{***}	7.882^{***}
	(0.109)	(0.103)	(0.213)	(0.132)	(0.215)	(0.241)	(0.590)	(0.253)
Adj. R^2	0.50	0.92	0.55	0.94	0.54	0.92	0.55	0.94
	60	60	60	60	52	52	52	52

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses.

[18] It must be remarked that we estimated three MNL models, one per year, similar to that of equation (2) and table 3.

Table 5: Education mismatch effects at country level. EMI labour force version (OLS regression)

	Unempl.	Self-empl.	Educ.	Inact.	Unempl.	Self-empl.	Educ.	Inact.
EMI-lf	0.079*	–0.028	0.204***	0.015	0.047	–0.010	0.235***	0.042
	(0.040)	(0.037)	(0.050)	(0.039)	(0.041)	(0.045)	(0.062)	(0.036)
Year-2008	0.843***	2.505***	–0.393*	3.184***	0.816***	2.465***	–0.378*	3.157***
	(0.105)	(0.086)	(0.199)	(0.112)	(0.106)	(0.092)	(0.215)	(0.113)
Year-2010	0.268**	1.565***	0.958***	2.264***	0.051	1.452***	0.706**	2.199***
	(0.125)	(0.099)	(0.174)	(0.118)	(0.185)	(0.171)	(0.327)	(0.182)
GDPshock					–0.034	–0.012	–0.036	–0.005
					(0.029)	(0.020)	(0.044)	(0.025)
Lab.Mark.Lib.					0.031	–0.037	–0.124	0.043
					(0.033)	(0.041)	(0.078)	(0.045)
Constant	4.364***	4.556***	19.804***	8.228***	4.212***	4.765***	20.506***	7.879***
	(0.090)	(0.077)	(0.178)	(0.105)	(0.207)	(0.255)	(0.553)	(0.262)
Adj. R^2	0.48	0.92	0.59	0.94	0.51	0.92	0.60	0.94
	60	60	60	60	52	52	52	52

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses.

Further in this section, we concentrate on the discussion of the combined effects of macro- and micro-level variables for which we obtained significant results. This means that we concentrate on the interpretation of the interactions with EMI-un, whereas the respective results for EMI-lf are presented in the appendix.

Despite that both the country-level effects of education mismatch (table 4) and the individual effects of years of education (table 3) are significant, their cross level effect, that table 6 exhibits, is not. This holds for all alternative labour statuses. Especially for unemployment, the interaction term *EMI-un*Years of Education*, shows the right expected negative sign but it is not significantly different from zero. It means that higher levels of individual human capital help to lower the risk of unemployment, regardless of the level of education mismatch in a country. Therefore hypothesis 2 is not confirmed after the empirical test.

Before considering hypothesis 3, we cannot neglect that the crisis, started in 2008, could have had an influence on the relationships between education and alternative labour statuses (especially unemployment), independently of country-level education mismatch. The results for this test are reported in table 7, where the coefficient of the variable capturing the point under scrutiny (i.e., the interaction term *Years of Education*Year–2010*) is not significant. This result is somehow coherent with recent empirical evidence (OECD, 2014; ILO, 2013) in which better-educated young people continue to be favored with respect to their low-educated peers,regardless of the crisis. Conversely, as stated in hypothesis 3, the crisis could have affected the relationships above trough the education mismatch. Indeed, if we combine these three terms in the interaction *EMI-un*Years of Education*Year–2010*, significant coefficients emerge for unemployment, education and inactivity (see table 8).[19] More precisely, an extra year of education in countries that experienced a remarkable increase in mismatch after the beginning of the crisis has an additional effect (–0.150) in reducing the relative probability of

[19] We also obtain very similar results with EMI-lf, see table A.3 in the appendix.

being unemployed. Therefore, as stated in hypothesis 3, it seems that the crisis, by aggravating education mismatch is, on the one hand, worsening the position of low educated youngsters and, on the other hand, favoring the employability of their better-educated peers. It is also worth noting that, in this case, an extra year of education, combined with mismatch and crisis, significantly decreases the relative probability of being in education (–0.209). This could be another sign for better employability of highly educated youth who, in these conditions, prefer to start seeking a job rather than extend their education period. Moreover, table 8 shows a positive and significant sign of *EMI-un*Years of Education*Year–2010* for inactivity. However, we should be very cautious in interpreting the result for inactivity status because of the lack of significance of the EMIs main effect at the country level on this outcome (see tables 4 and 5).

The coefficients we present in table 8 for *EMI-un*Years of Education*Year–2010* only give us the information concerning relative probabilities, but tell us nothing about the real magnitude and sign of the effect of this combined variable on unemployment status.

Unfortunately, the computation of marginal effects for interaction terms in non linear models is affected by many drawbacks that severely limit their interpretation (Greene, 2010; Ai and Norton, 2003). For this reason, we follow Greene (2010) and use a graphical representation of the sole marginal effects for *Years of Education* on unemployment, based on the model of table 8 and conditional to *EMI-un* and *Year–2010*. Figure 4 depicts this interpretation by recovering the framework already used in figure 3, where we take into account young people aged 26 with *years of education=8* and *years of education=13*. In addition, these marginal effects are calculated at three different values of the distribution of EMI-un, the bottom decile (5.3), the median (16.3) and the top decile (21.4), in both the pre-crisis (*Year–2010=0*) and post-crisis period (*Year–2010=1*). As expected, in every category the marginal effects are negative, even though some differences are worth noting. Indeed, if we focus on the effect of an extra year of education for people aged 26 and provided with 8 years of education, we can see that the negative impact decreases as the EMI-un increases. In 2010, the line connecting marginal effects for this cohort remarkably shifts upward. This indicates that, for people who presumably have a secondary education, the guarantee to escape unemployment attenuates.

On the contrary, an extra-year of education at the end of secondary education (*years of education=13*), for people currently at the age of 26, makes the risk of being unemployed lower when the education mismatch is higher. Moreover, in this case the crisis shifts the line connecting marginal effects downward. The final result is that, in 2010, the risk of being unemployed for higher educated people is lower than that of medium educated people, as outlined by the two solid lines in figure 4. This is an additional reinforcement for hypothesis 3 that is coherent with figures B.2 and B.3. In other terms, if a high education mismatch means lack of young people with tertiary education on the labour supply side, by boosting mismatch the crisis aggravated the employment prospects for low and medium educated youngsters and favored their highly educated peers.

Table 6: Combined effects of country-level EMI-un and individual education on labour status of young people aged 15–34. MNL model: endogeneity control with 2-Stages Residual Inclusion method; raw coefficients. (Base category: Employee)

	Self-employed	Unemployed	Education	Inactive
EMI-un*Education Years	–0.101	–0.037	–0.049	0.100
	(0.068)	(0.074)	(0.055)	(0.069)
Years of Education	0.237***	–0.143**	0.582***	0.077*
	(0.051)	(0.058)	(0.037)	(0.043)
1-stage Resid.	–0.231***	0.083*	–0.389***	–0.153***
	(0.051)	(0.047)	(0.036)	(0.042)
Age	–0.786***	–0.142	–2.194***	–1.036***
	(0.108)	(0.105)	(0.072)	(0.079)
Age^2	0.015***	0.001	0.033***	0.017***
	(0.002)	(0.002)	(0.001)	(0.001)
Gender (male=1)	0.803***	–0.207***	–0.365***	–1.083***
	(0.065)	(0.065)	(0.047)	(0.053)
Citizenship	0.333**	–0.081	–0.315***	–0.086
	(0.149)	(0.117)	(0.116)	(0.102)
Disconnected	–0.052	–0.327**	–0.424***	0.081
	(0.146)	(0.138)	(0.150)	(0.096)
No Social Activities	0.046	0.219***	0.028	0.383***
	(0.067)	(0.065)	(0.052)	(0.050)
Children (yes=1)	0.654***	–0.196*	–0.134	1.631***
	(0.102)	(0.110)	(0.094)	(0.083)
H.Labour Income	–0.461**	–3.446***	–3.141***	–3.151***
	(0.201)	(0.108)	(0.106)	(0.100)
H.Capital Income	1.023***	–0.145	0.911***	0.289
	(0.348)	(0.242)	(0.222)	(0.227)
Trade Un. Member	–1.050***	–0.346***	–0.867***	–0.707***
	(0.096)	(0.092)	(0.078)	(0.072)
Family Members	0.031	0.189***	0.196***	0.223***
	(0.026)	(0.022)	(0.018)	(0.018)
Year-2008	0.031	0.255***	–0.037	–0.040
	(0.075)	(0.080)	(0.056)	(0.058)
Year-2010	0.081	0.660***	0.307***	–0.021
	(0.079)	(0.082)	(0.058)	(0.061)
Country dummies	Yes	Yes	Yes	Yes
Obs	26245			
p-value-Overall Model	0.000			
Pseudo-R^2	0.33			

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses. Significant t-test for the coefficients of the first-stage resid.indicates that education is endogenous.

Table 7: Effects of education in crisis time on labour status of young people aged 15–34. MNL model: endogeneity control with 2-Stages Residual Inclusion method; raw coefficients. (Base category: Employee)

	Self-employed	Unemployed	Education	Inactive
Years of Education*Year-2010	0.015	–0.021	–0.015	–0.014
	(0.018)	(0.017)	(0.014)	(0.015)
Years of Education	0.214***	–0.141**	0.578***	0.101**
	(0.050)	(0.057)	(0.036)	(0.041)
1-stage Resid.	–0.230***	0.083*	–0.389***	–0.155***
	(0.051)	(0.047)	(0.036)	(0.042)
Age	–0.782***	–0.143	–2.195***	–1.042***
	(0.108)	(0.105)	(0.072)	(0.079)
Age^2	0.014***	0.001	0.033***	0.017***
	(0.002)	(0.002)	(0.001)	(0.001)
Gender (male=1)	0.800***	–0.208***	–0.366***	–1.081***
	(0.065)	(0.065)	(0.047)	(0.053)
Citizenship	0.335**	–0.078	–0.314***	–0.087
	(0.148)	(0.117)	(0.115)	(0.102)
Disconnected	–0.055	–0.326**	–0.426***	0.088
	(0.146)	(0.138)	(0.151)	(0.096)
No Social Activities	0.047	0.219***	0.029	0.382***
	(0.067)	(0.065)	(0.052)	(0.050)
Children (yes=1)	0.652***	–0.196*	–0.137	1.635***
	(0.102)	(0.110)	(0.094)	(0.083)
H.Labour Income	–0.462**	–3.447***	–3.139***	–3.152***
	(0.201)	(0.108)	(0.105)	(0.100)
H.Capital Income	1.010***	–0.145	0.912***	0.289
	(0.348)	(0.242)	(0.221)	(0.227)
Trade Un. Member	–1.049***	–0.347***	–0.869***	–0.705***
	(0.096)	(0.092)	(0.078)	(0.072)
Family Members	0.030	0.189***	0.196***	0.224***
	(0.026)	(0.022)	(0.018)	(0.018)
Year-2008	0.018	0.251***	–0.044	–0.019
	(0.075)	(0.080)	(0.055)	(0.057)
Year-2010	–0.136	0.926***	0.500**	0.183
	(0.261)	(0.238)	(0.205)	(0.211)
Country dummies	Yes	Yes	Yes	Yes
Obs	26245			
p-value-Overall Model	0.000			
Pseudo-R^2	0.33			

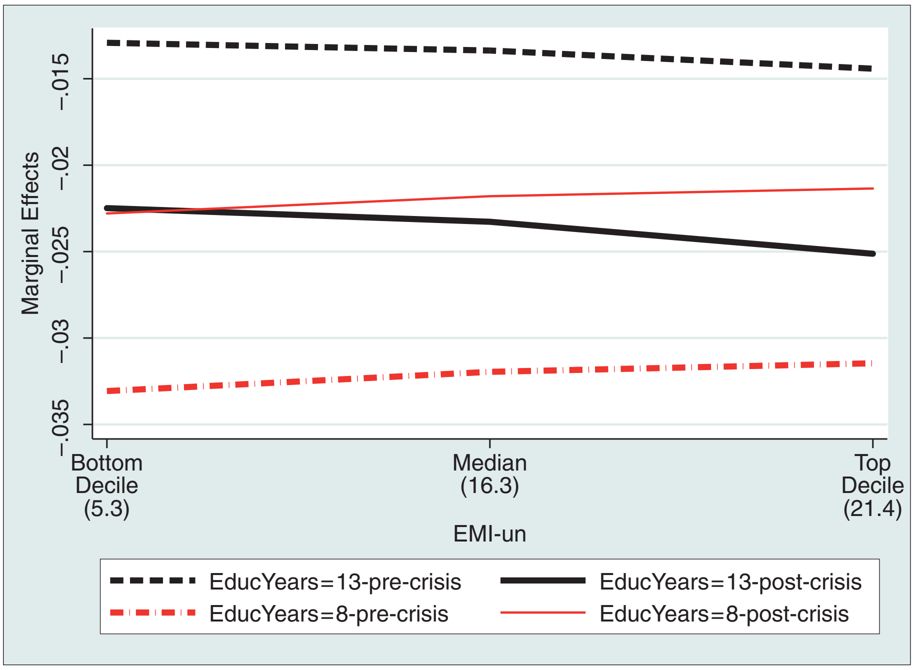
Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses. Significant t-test for the coefficients of the first-stage resid.indicates that education is endogenous.

Table 8: Country-level EMI-un and individual education in crisis time, effects on labour status of young people aged 15–34. MNL model: endogeneity control with 2-Stages Residual Inclusion method; raw coefficients. (Base category: Employee)

	Self-employed	Unemployed	Education	Inactive
EMI-un*Education Years*Year-2010	–0.071	–0.150**	–0.209***	0.115**
	(0.058)	(0.066)	(0.053)	(0.051)
EMI-un*Education Years	–0.077	0.013	–0.003	0.055
	(0.071)	(0.077)	(0.057)	(0.061)
Years of Education*Year-2010	0.029	0.009	0.029	–0.035*
	(0.022)	(0.022)	(0.018)	(0.019)
Years of Education	0.228***	–0.143**	0.580***	0.089**
	(0.051)	(0.058)	(0.037)	(0.043)
1-stage Resid.	–0.232***	0.082*	–0.391***	–0.152***
	(0.051)	(0.046)	(0.036)	(0.042)
Age	–0.785***	–0.144	–2.200***	–1.033***
	(0.108)	(0.105)	(0.072)	(0.079)
Age^2	0.014***	0.001	0.034***	0.017***
	(0.002)	(0.002)	(0.001)	(0.001)
Gender (male=1)	0.804***	–0.207***	–0.365***	–1.085***
	(0.065)	(0.065)	(0.047)	(0.053)
Citizenship	0.328**	–0.081	–0.318***	–0.083
	(0.148)	(0.117)	(0.116)	(0.102)
Disconnected	–0.050	–0.323**	–0.416***	0.080
	(0.146)	(0.138)	(0.150)	(0.096)
No Social Activities	0.046	0.218***	0.028	0.382***
	(0.067)	(0.066)	(0.052)	(0.050)
Children (yes=1)	0.653***	–0.197*	–0.134	1.630***
	(0.102)	(0.110)	(0.094)	(0.083)
H.Labour Income	–0.462**	–3.448***	–3.145***	–3.153***
	(0.201)	(0.108)	(0.106)	(0.100)
H.Capital Income	1.014***	–0.154	0.889***	0.287
	(0.349)	(0.241)	(0.221)	(0.227)
Trade Un. Member	–1.050***	–0.345***	–0.864***	–0.708***
	(0.096)	(0.092)	(0.078)	(0.072)
Family Members	0.032	0.189***	0.197***	0.222***
	(0.026)	(0.022)	(0.018)	(0.018)
Year-2008	0.039	0.274***	–0.020	–0.050
	(0.075)	(0.081)	(0.056)	(0.059)
Year-2010	–0.155	0.900***	0.445**	0.180
	(0.268)	(0.240)	(0.206)	(0.211)
Obs.		26245		
p-value-Overall Model		0.000		
Pseudo-R^2		0.33		

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses. Significant t-test for the coefficients of 1-stage Resid. indicates that education is endogenous.

Figure 4: Marginal effects of years of education with relation to EMI-un and crisis effect (at age=26 and at the sample mean for other regressors)



7 Conclusions

This paper analyses the influence of country level education mismatch on the relationship between education and the probability of being unemployed or staying in alternative labour statuses for young people aged 15–34 in 2006, 2008 and 2010, and living in 21 EU countries. Normally, education mismatch studies concentrate on structural unemployment at the aggregate level or focus on wages and productivity of mismatched workers at the individual level. Our article's contribution is pointing out the effects of country level education mismatch along individual profiles of young people that differ according the years of education they have accumulated. Unlike previous studies, we attempted to take into account how an aggregate phenomenon like education mismatch influences not only young employees or the unemployed, but also the conditions of all the population aged 15–34.

First of all, we provided new evidence that increasing years of education still matters to reduce the probability of being unemployed; it is important to lengthen the period of education, and slightly raises the probability of being self-employed with respect to the probability of being employed. On the other hand, raising the years of education also slightly increases the probability of not entering the labour market (inactivity), even though this effect is not predominant and is completely offset by the former. Therefore, our hypothesis 1, that we raised with the support of previous theoretical and empirical evidence upon the role of education, remains valid, especially for unemployment and self-employment.

Secondly, we investigated whether education mismatch, measured at the country level, weakens or reinforces the results we found above. More precisely, we hypothesized that education mismatch should affect differently youngsters by only favouring better educated people (hypothesis 2). We did not find support for this evidence. This means that independently of the relevance of education mismatch, investing in education has the same effects we found for hypothesis 1 in all countries .

Finally, we considered the effect of the crisis on the relationships studied above. It is worth noting that the crisis *per se* did not significantly influence the impact of increasing years of education on labour status; in particular it did not exert any additional effect on the probability of being unemployed. Instead, the crisis acted through education mismatch as we guessed in hypothesis 3. In other terms, an extra year of education is particularly effective in reducing the probability of being unemployed in countries that experienced a higher mismatch after 2008. This result particularly holds for young people who go beyond the secondary education level, whereas for those that do not, the guarantee of avoiding unemployment attenuates, especially after the beginning of the crisis.

Policy-relevant implications arise from these results and tell us that tertiary formal education still matters to avoid the negative effects of country level education mismatch boosted by the recent crisis. Thus, improving access to university degrees remains the main road to tackle unemployment caused by education mismatch. At the same time, we need further research to prove that increasing education mismatch keeps qualitative changes in the labour demand for skilled workers out of sight. These changes, in turn, might affect the unemployment risk both for young people with intermediate education levels (aimed to routinary tasks) and for graduates who have acquired obsolete knowledge.

References

Acemoglu, D., Autor, D., 2012. What does human capital do? a review of goldin and katz's the race between education and technology. Journal of Economic Literature 50 (2), 426–463.

Ai, C., Norton, E., 2003. Interaction terms in logit and probit models. Economics Letters 80, 123–129. 13

Autor, D., Levy, F., Murnane, J., 2003. The skill content of recent technological change: An empirical exploration. Quarterly Journal of Economics 118 (4), 1279–1333.

Becker, G., 1964. Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education. Columbia University Press, New York.

Blöndal, S., Field, S., Girouard, N., 2002. Investment in human capital through upper-secondary and tertiary education. OECD Economic Studies 2002 (1), 41–89.

Bollen, K., Guilkey, D., Mroz, T., 1995. Binary outcomes and endogenous explanatory variables: Tests and solutions with an application to the demand for contraceptive use in tunisia. Demography 32 (1), 111–131.

Bryan, M., Jenkins, S. P., 2013. Regression analysis of country effects using multilevel data: A cautionary tale. IZA Discussion Papers 7583, Institute for the Study of Labor (IZA).

Cameron, C., Trivedi, P., K., 2009. Microeconometrics using STATA. STATA Press.

Cameron, C., Trivedi, P., 2005. Microeconometrics. Cambridge University Press, Cambridge.

Card, D., 1999. The causal effect of education on earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labour Economics. Elsevier Science, pp. 1801–1862.

Desjardins, R., Rubenson, K., 2011. An analysis of skill mismatch using direct measures of skills.

Dietrich, H., 2012. Youth unemployment in europe. theoretical considerations and empirical findings.

Dustmann, C., Ludsteck, J., Schönberg, U., 2009. Revisiting the german wage structure. The Quarterly Journal of Economics 24, 843–881.

ECB, 2012. Euro area labour markets and the crisis. october 2012.

Eurofound, 2013. Employment polarisation and job quality in the crisis: European jobs monitor 2013.

Giannelli, G., Monfardini, C., 2003. Joint decisions on household membership and human capital accumulation of youths. the role of expected earnings and local markets. Journal of Population Economics 16, 265–285.

Goos, M., Manning, A., 2007. Lousy and lovely jobs: The rising polarization of work in britain. Review of Economics and Statistics 89 (1), 118–133.

Greene, W., 2010. Testing hypotheses about interaction terms in nonlinear models. Economic Letters 107, 291–296.

ILO, 2013. Global employment trends for youth 2013: A generation at risk.

ISCED, 2012. International standard classification of education 2011.

Ivlevs, A., King, R., 2012. Does more schooling make you run for the border? evidence from postindependence kosovo. Journal of Development Studies 48 (8), 1108–1120.

Leuven, E., Oosterbeek, H., 2011. Overeducation and mismatch in the labor market.

Liu, D., 2012. Education as unemployment insurance: A model with endogenous educational requirement for job application and its policy implications.

Long, S., Freese, J., 2006. Regression Models for Categorical Dependent Variables Using STATA. STATA Press.

Luce, R., 1959. Individual Choice Behavior: A Theoretical Analysis. JohnWiley, New York.

Marsden, D., Lucifora, C., Oliver-Alonso, J., Guillotin, Y., 2002. The economic costs of the skills gap in the EU. Istituto per la ricerca sociale, Milano, Italia.

De Lange, M., Gesthuizen, M., Wolbers, M. H., 2014. Youth labour market integration across europe: The impact of cyclical, structural, and institutional characteristics. European Societies 16 (2), 194–212.

European Commission, 2013. Labour market developments in europe 2013.

Mac Guinness, S., 2006. Overeducation in the labour market. Journal of Economic Survey 20 (3), 387–418.

Millàn, M., Congregado, E., Romàn, C., 2012. Determinants of self-employment survival in europe. Small Business Economics 38, 231–258.

Mincer, J., 1974. Schooling, Experience, and Earnings. NBER Press, New York.

Mincer, J., 1991. Education and unemployment.

Nickell, S., 1979. Education and lifetime patterns of unemployment. Journal of Political Economy 87 (5), S117–S131.

OECD, 2014. Oecd employment outlook 2014.

Parker, S., Van Praag, C., 2006. Schooling, capital constraints and entrepreneurial performance: the endogenous triangle. Journal of Business and Economic Statistics 24 (4), 416–431.

Pastore, F., 2012. To study or to work? Education and labour market participation of young people in poland. Eastern European Economics 50 (3), 49–78.

Pfeiffer, F., Seiberlich, R., 2010. A socio-economic analysis of youth disconnectedness.

Pissarides, C., 2000. Equilibrium unemployment theory. MIT Press, Cambridge, Massachusetts.

Pissarides, C., 2013. Unemployment in the great recession. Economica 80, 385–403.

Piurko, Y., Schwartz, S., Davidov, E., 2011. Basic personal values and the meaning of left-right political orientations in 20 countries. Political Psychology 20 (20), 1–25.

Quintini, G., 2011. Over-qualified or under-skilled: A review of existing literature. OECD Social, Employment and Migration Working Papers 121, OECD Publishing.
URL http://dx.doi.org/10.1787/5kg58j9d7b6d-en

Riddell, C., Song, X., 2011. The impact of education on unemployment incidence and re-employment success: evidence from the u.s. labour market. Labour Economics 18, 453–463.

Spence, M., 1973. Job market signalling. Quarterly Journal of Economics 87 (3), 355–374.

Styczynska, I., 2013. Enhancing youth opportunities in employment: Determinants and policy implications. Intereconomics 4, 216–223.

Terza, J., Basu, A., Rathouz, P., 2010. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. Journal of Health Economics 27, 531–543.

Trostel, P., Walker, I., 2006. Education and work. Education Economics 14 (4), 377–399.

Trostel, P., Walker, I., Wooley, P., 2002. Estimates of the economic return to schooling for 28 countries. Labour Economics 9 (1), 1–6.

Wooldridge, J. M., 2010. Econometric analysis of cross section and panel data. MIT press.

Appendix

A Tables

Table A.1: Small-Hsiao tests of IIA assumption

Omitted	lnL(full)	lnL(omit)	chi2	df	p-value	evidence
Self-employed	–9593.874	–9544.271	99.205	102	0.560	for Ho
Unemployed	–9486.677	–9440.177	93.000	102	0.727	for Ho
Education	–7968.380	–7916.704	103.352	102	0.444	for Ho
Inactive	–8068.264	–8020.706	95.116	102	0.672	for Ho

Ho: Odds(Outcome-J vs Outcome-K) are independent of other alternatives.

Table A.2: 2SRI-First Stage estimation for results in table 3: Instrumentalisation of Years of Education (OLS)

	Dependent Variable: Years of Education
Father's Tertiary Education	1.444***
	(0.042)
Equalitarianism	0.082*
	(0.048)
Environmentalism	0.291***
	(0.048)
Altruism	0.152***
	(0.058)
Age	1.502***
	(0.029)
Age^2	–0.025***
	(0.001)
Gender (male=1)	–0.462***
	(0.035)
Citizenship	0.775***
	(0.098)
Disconnected	–0.673***
	(0.093)
No Social Activities	–0.261***
	(0.039)
Children	–1.261***
	(0.053)
H.Labour Income	0.871***
	(0.064)
H.Capital Income	1.226***
	(0.116)
Trade Un. Member	0.396***
	(0.054)
Family Members	–0.087***
	(0.013)
Year-2008	0.182***
	(0.044)
Year-2010	0.331***
	(0.044)
Constant	–9.659***
	(0.388)
Country dummies	Yes
Adj. R^2	0.32
Obs.	26410

Notes: excluded instruments in bold.Robust standard errors in parentheses. ***significant at 1% level; **significant at 5% level; *significant at 10% level.

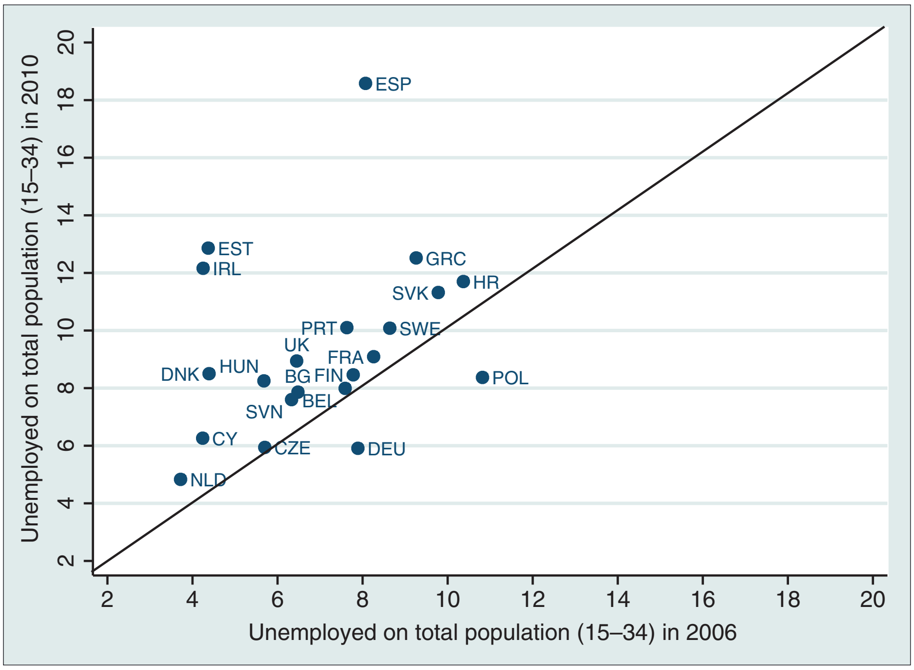
Table A.3: Country-level education mismatch (labour force based) and individual education in crisis time, effects on labour status of young people aged 15–34. MNL model: endogeneity control with 2-Stages Residual Inclusion method; raw coefficients. (Base category: Employee)

	Self-employed	Unemployed	Education	Inactive
EMI-lf*Years of Education*Year-2010	0.105	–1.164**	–0.723*	1.393***
	(0.475)	(0.533)	(0.378)	(0.393)
SMI-lf*Years of Education	–0.198	1.717***	0.870**	–0.097
	(0.479)	(0.516)	(0.362)	(0.385)
Years of Education*Year-2010	0.014	–0.009	–0.006	–0.051***
	(0.021)	(0.021)	(0.016)	(0.017)
Years of Education	0.217***	–0.165***	0.567***	0.101**
	(0.050)	(0.057)	(0.036)	(0.042)
1-stage Resid.	–0.230***	0.081	–0.390***	–0.152***
	(0.051)	(0.057)	(0.036)	(0.042)
Age	–0.784***	–0.148	–2.199***	–1.038***
	(0.108)	(0.105)	(0.072)	(0.079)
Age^2	0.014***	0.001	0.034***	0.017***
	(0.002)	(0.002)	(0.001)	(0.001)
Gender (male=1)	0.801***	–0.208***	–0.366***	–1.086***
	(0.065)	(0.065)	(0.047)	(0.053)
Citizenship	0.333**	–0.078	–0.316***	–0.091
	(0.148)	(0.117)	(0.115)	(0.102)
Disconnected	–0.054	–0.328**	–0.426***	0.074
	(0.146)	(0.138)	(0.151)	(0.097)
No Social Activities	0.047	0.225***	0.031	0.383***
	(0.067)	(0.065)	(0.052)	(0.050)
Children (yes=1)	0.652***	–0.191*	–0.133	1.631***
	(0.102)	(0.110)	(0.094)	(0.083)
H.Labour Income	–0.462**	–3.448***	–3.144***	–3.149***
	(0.201)	(0.108)	(0.106)	(0.100)
H.Capital Income	1.010***	–0.159	0.897***	0.295
	(0.349)	(0.241)	(0.221)	(0.227)
Trade Un. Member	–1.049***	–0.344***	–0.868***	–0.700***
	(0.096)	(0.092)	(0.078)	(0.072)
Family Members	0.030	0.188***	0.196***	0.223***
	(0.026)	(0.022)	(0.018)	(0.018)
Year-2008	0.011	0.302***	–0.018	–0.045
	(0.077)	(0.082)	(0.056)	(0.058)
Year-2010	0.000	–0.141	0.955***	0.531***
	(0.100)	(0.264)	(0.238)	(0.206)
Obs.		26245		
p-value-Overall Model		0.000		
Pseudo-R^2		0.33		

Notes: ***significant at 1% level; **significant at 5% level; *significant at 10% level. Robust standard errors in parentheses. Significant t-test for the coefficients of 1-stage Resid. indicates that education is endogenous.

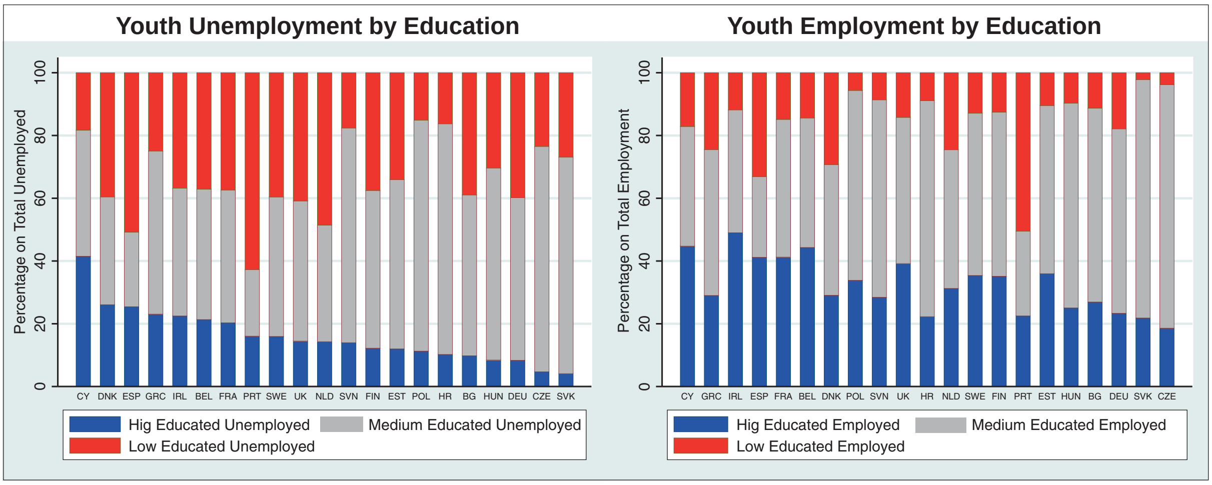
B Figures

Figure B.1: Unemployment-to-population ratio between 2006 and 2010



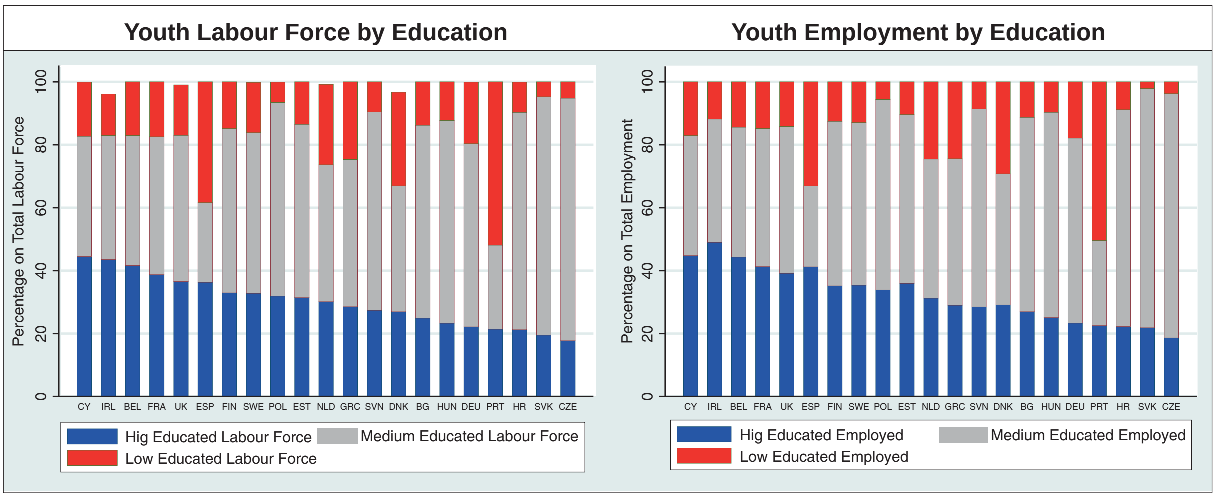
Notes: see notes under figure 1.

Figure B.2: Education Mismatch composition in 2010 (Employment vs Unemployment)



Notes: see notes under figure 1.

Figure B.3: Education Mismatch composition in 2010 (Employment vs Labour Force)



Notes: see notes under figure 1.

Figure B.4: EMI-un and country effects on probability to be in different labour status

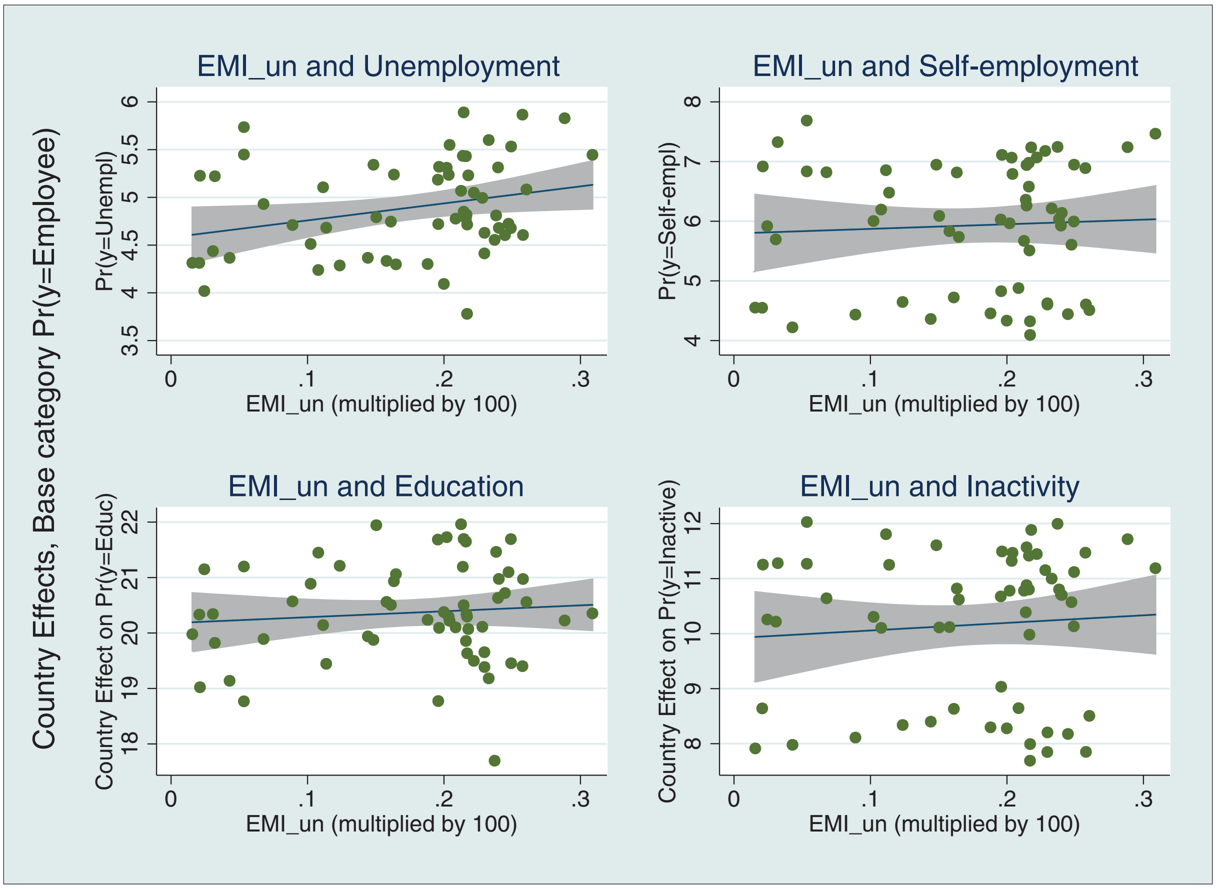


Figure B.5: EMI-lf and country effects on probability to be in different labour status