Measuring Labour Mismatch in Europe

António Morgado¹ · Tiago Neves Sequeira² · Marcelo Santos² · Alexandra Ferreira-Lopes³,⁴ · Ana Balcão Reis⁵

Accepted: 7 September 2015 / Published online: 14 September 2015 © Springer Science+Business Media Dordrecht 2015

Abstract We calculate aggregate and comparable measures of mismatch in the labour market for 30 European countries. These indicators measure vertical mismatch (related to the level of education, e.g. overeducation, and undereducation) and horizontal mismatch (related to the field of education) and are comparable across countries and through time. In European countries, between 15 % to nearly 35 % of workers have a job for which they have more (or less) qualifications than the usual level. Approximately 20 % to nearly 50 % work in a job for which they do not have the usual field qualification. There is a great variability on mismatch across European labour markets. Undereducation affects more workers than overeducation in most European countries. Low correlations between mismatch and unemployment indicate that mismatch should be regarded as an additional informative variable, thus useful to characterize labour markets.

Keywords Education · Overeducation · Undereducation · Mismatch · Labour market

JEL Classification J24 · O50

— Tiago Neves Sequeira
sequeira@ubi.pt

¹ ISCAL - IPL, Politecnic Institute of Lisbon, Lisbon, Portugal
² Departamento de Gestão e Economia and CEFAGE-UBI, Universidade da Beira Interior, Estrada do Sineiro, 6200-209 Covilhã, Portugal
³ BRU-IUL (Business Research Unit), ISCTE Business School Economics Department, Instituto Universitário de Lisboa, ISCTE-IUL, Lisbon, Portugal
⁴ CEFAGE-UBI, Covilhã, Portugal
⁵ Nova School of Business and Economics, Universidade Nova de Lisboa, Lisbon, Portugal
1 Introduction

Mismatch in labour markets is usually associated with the presence of overeducation and undereducation, meaning the extent to which individuals possess a level of education and/or training above or below the one required for their job, respectively. Some researchers also look at field or horizontal mismatch. This has been essentially used to estimate private returns for both types of mismatch. In fact, the existence of overeducation in labour markets may question the benefits of public funding of higher education, even if overeducation may have positive private returns or premia. Until now, mismatch measures have been obtained mainly through individual questionnaires, using subjective measures and for just a few countries.1 Alternative methods2 such as job analysis or realized matches have also been used, but in a minority of studies. While the first method (job analysis) supposes a detailed description of the required skills for each profession, which is available only for a few countries in the world, the second (realized matches) is based on measuring deviations from a central measure of the distribution. Until now, the use of the realized matches method has faced the scarcity of large databases with sufficient information to perform the analysis. Consequently, aggregated labor mismatch comparisons between different countries and across several years are absent from the literature. We fill this gap using the realized matches approach to present measures of mismatch across European countries during a period of nearly 20 years. We obtain new evidence of mismatch using aggregate measures that are comparable across countries and over time. As becomes evident from the survey provided by Leuven and Oosterbeek (2011), there have been no earlier attempts to obtain aggregate and comparable measures of mismatch. Instead, the literature has focused on estimating the returns of mismatch, that is, the conditional change on wages due to mismatch.

We calculate values of several different types of mismatch for 30 countries for the period between 1993 and 2011 using a dataset from the Labour Force Survey. Our main objective in this paper is to provide (equilibrium) measures of mismatch in European countries and thus the realized matches method fits our purposes perfectly well. Our contribution to the literature is twofold: (1) we provide new aggregate evidence on vertical mismatch (total, under- and overeducation, strong under- and overeducation) as for horizontal, or field, mismatch that is comparable by countries and years, presenting a new panel dataset with 30 countries during nearly 20 years; (2) we provide evidence according to which both vertical mismatch and horizontal mismatch may be useful and complementary measures to characterise labour markets and do not overlap with unemployment measures.

This work has the following structure. In Sect. 2, we revise the existing literature on labour mismatch. In Sect. 3 we describe the methodology and data used. Section 4 describes our new measures of mismatch, detailing the evidence they highlight for European countries. Section 5 concludes.

---

1 We have found articles that individually evaluate the effects of mismatch measures in USA, Canada, UK, Netherlands, France, Germany, Spain, Portugal and Hong Kong. Those measures are not comparable due to different sources, year coverage, and methods. Sloane (2002) presents a detailed review of some of these papers.

2 The alternative methods are discussed below.
2 Literature Review

Mismatch has been widely referred to in the literature, specifically related to unemployment duration and wage inequality, for which Ordine and Rose (2011) and Budría and Moro-Egido (2008) are good examples. Undereducation and overeducation are specific types of mismatch. We may think that mismatch is a transitory short-term phenomenon, associated with rigidities in flows in and out of employment. Actually, overeducation is referred to by Ordine and Rose (2011) as a voluntary short-term strategy to enter the labour market, but as they point out, it may also be a consequence of low individual ability for that level of education or it may just be a matter of choice related to compensating advantages in the choice occupation/payoff. Also, overeducation can be the outcome of long periods of unemployment, giving rise to an exacerbation of the availability to enter the labour market losing the wage premia associated with higher education. Tracing the reasons for mismatches between labour supply and demand is very important, as they can potentially be very costly to the economy, by restricting productivity growth (see Leuven and Oosterbeek 2011). Hence, measuring mismatch adequately is central to our subsequent analysis. We do not use the mismatch measures to estimate returns to mismatched workers, as earlier literature did. Instead, we aggregate individual mismatch to obtain national mismatch measures.

Following Hartog (2000), we identify and discuss three methods of measuring over and undereducation, which we describe briefly:

1. **Job Analysis**, in which professional job analysts specify the required level and type of education for the job titles in an occupational classification. The Dictionary of Occupational Titles is the result of such an analysis. Detecting mismatch using this method consists basically in comparing a current worker’s situation with the standard match described in, e.g. the Dictionary of Occupational Titles.

2. **Worker Self-Assessment**, in which the worker specifies the education she deems necessary for the job. This can be done directly through the explicit specification of the type of schooling required, or indirectly, through comparison to the workers’ actual education, assessing whether a higher, or lower, or different education is needed.

3. **Realized Matches**, in which the required education for a job or occupation is derived from what workers in that job or occupation usually have attained. Mismatch in this methodology is, therefore, assessed by comparing a worker’s education to the mean or mode of education of workers in the same job or occupation, resulting in a measure similar to the variance of the distribution of education levels for the workers in that job or occupation.

Even though objective, **Job Analysis** is dated, and has a large measurement error in relation to the assignment of job level codes to any survey responses on the type of work. Also, this assignment can be country-specific, amely sensitive to differences in countries’ different educational systems, which may invalidate international comparisons. Furthermore, any error or ad-hoc choice as far as the matching education/occupation/job is concerned, has a contamination effect throughout the whole analysis, as that matching depends on the personal judgement of job analysts. For Portugal, an occupational classification exists in **Quadros de Pessoal**, and a matching of level and type of education for each type of occupation has been done in a very aggregate form (Coelho et al. 1982). This can be seen

---

3 A matched employer-employee dataset with information for all firms with at least one wage earner.
Worker Self-Assessment is an alternative to Job Analysis but is subjective. Even though it deals with the survey respondent’s job specifically and does not seek to match the responses to any pre-existing standard, it rests on workers’ own judgement. The worker may overstate the requirements of education and/or training to the job due to a status effect, or the responses may reproduce the hiring standards, not the match between education/training and job/occupation. Actually, as school levels increase with time, employers tend to adjust hiring standards, even though the nature and skills of a job may not have changed accordingly. Worker Self-Assessment is widely used due to its relative simplicity.

Budría and Moro-Egido (2008) use the worker responses from the survey of European Community Household Panel (ECHP) to cue the following questions: (a) Do you feel that you have skills or qualifications to do a more demanding job than the one you have now? and (b) Have you had formal training or education that has given you skills needed for your present type of work? The answers are then classified according to the following rule: if ‘yes’ to both, the worker is overqualified, if ‘no’ to both, the worker is incorrectly qualified, if ‘yes’ to question (a) and ‘no’ to question (b), the worker is strongly mismatched, and finally, if ‘no’ to question (a) and ‘yes’ to question (b), the worker is correctly qualified.

An example of the use of Realized Matches methodology can be found in Kiker and Santos (1991) for Portugal, where they classify a worker as overeducated if he has the number of years of schooling greater than one standard deviation above the mean for his (essentially) three-digit occupational code, whereas a worker is classified as undereducated if his years of schooling are more than one standard deviation below the mean for his occupational code. Also, Mendes de Oliveira et al. (2000) use this method, but instead of the mean, they use the mode as the reference. This methodology to measure mismatch has been criticized for using the endogenous (equilibrium) assignment of jobs, which depends on hiring standards and labour market conditions, to track an error in that same assignment, that is, mismatch, without using any exogenous information to separate correct matches from mismatches. This means that it does not distinguish if the source of mismatch is the labour demand or the labour supply. It uses observations on the realized equilibrium between demand and supply. Thus, it measures verified mismatch.

There are fewer examples in the literature that measure field, or horizontal, mismatch. Robst (2007) measures horizontal mismatch in the US using the self-assessment approach. Wolbers (2003) studies Field mismatch for 13 European countries using a Job Analysis approach.

Specifically, Job Analysis detects that, of all workers, 33.1% are overeducated, 37.5% are undereducated and 29.4% are correct matches. The methodology of Realized Matches methodology, using the mean as the reference, detects 85.6% of correct matches (9.4% overeducation and 5% of undereducation), but using the mode as the reference, correct matches are no more than 57.5% (25.5% overeducation and 17% undereducation). These results are in line with those we obtain for Portugal (in 1995, near 12% overeducation and 4% undereducation, compares with the mean results of 9.4 and 5%, respectively), although we analyse a more recent period. This comparative study was performed using Quadros de Pessoal for the period 1985–1991. Comparison with Worker Self-Assessment method is not feasible.
3 Methodology and Data

Using the Labour Force Survey (LFS) we propose to measure mismatch using the *Realized Matches* approach, following the works of Kiker et al. (1997), Mendes de Oliveira et al. (2000) and Kiker and Santos (1991). We choose this methodology as we wish to guarantee international comparison of results as well as time-coherence.

The Labour Force Survey (LFS) database contains worker specific information in relation to (reference to Eurostat variable code):

- highest level of education or training successfully completed (HATLEVEL)
- field of highest level of education or training successfully complete (HATFIELD)
- occupation (ISCO4D)
- economic activity of the local unit (NACE3D, NA113D).

There is information for the following 30 countries: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovak Republic, Slovenia, Spain, Sweden, Switzerland and United Kingdom. Most countries have information available for the HATLEVEL variable from 1993 to 2011 (19 years) and for the HATFIELD variable from 2003 to 2011 (9 years). In the “Appendix” we include information about the number of observations for both variables and each country. Definitions for both variables are also in the “Appendix”.

We start by calculating the average highest level of education, measured in years of education, attained by workers in a given occupation in a certain economic activity in each country. Because in different waves of the LFS, the included ISCED levels in the database are different, we had to consider a correspondence scale between ISCED levels and years of education. In the database the HATLEVEL variable (after conversion into years of education) oscillates between 0 and 19 years, with an average of around 11 years of education.

If a worker lies beyond one standard-deviation from the mean, we consider that there is vertical mismatch and signal it with 1 or 2 depending on being above or below the high or low limits of the interval. In those cases we say that the individual is overeducated or undereducated, respectively. We also signal the cases that lie beyond two standard-deviations from the mean, and signal those with 2 or 2, respectively depending on whether the observation is over or under the limit of the interval. In these cases, we say that the individual is strongly overeducated and strongly undereducated, respectively. Our measure of vertical mismatch is the proportion of cases signaled 2, 1, 1 and 2 in relation to the total number of observations for a given country in a given year. They represent respectively undereducation and overeducation. That measure will have a value between 0 and 1, such that 0 means no mismatch and 1 means that all workers are mismatched. Overeducation and undereducation are additive and total the percentage of vertical mismatch.

More precisely, let $e_i$ represent the highest level of education for worker $i$ ($\bar{e}$ is the average and $s_e$ is the standard-error of $e_i$). Let $o_i$ be the occupation of that worker and $a_i$ the
economic activity according to NACE classification. We define the indicator function, specific for country $c$ and year $y$ as:

$$AE_{icy}(e_i|a_i, a_i, c_i, y_i) = \begin{cases} -2 & \text{if } e_i - \bar{e} < -2s_e \\ -1 & \text{if } -2s_e < e_i - \bar{e} < -s_e \\ 0 & \text{if } -s_e < e_i - \bar{e} < s_e \\ 1 & \text{if } s_e < e_i - \bar{e} < 2s_e \\ 2 & \text{if } e_i - \bar{e} > 2s_e \end{cases}$$

Thus, our measure of mismatch for a given country $c$ and year $y$ is

$$ME_{c,y} = \frac{\sum_{i: AE_{icy} \neq 0} |AE_{icy}|}{n_{cy}}$$

An overeducation measure can thus be obtained using just $AE = 1$ or $AE = 2$ and an undereducation measure using $AE = -1$ or $AE = -2$. A strongly overeducated measure is obtained using just $AE = 2$ and a strongly undereducated measure is obtained using just $AE = -2$. This methodology allows us to identify the proportion of workers that have less education or more education than what is typically required for the job in a given occupation, economic activity, and country (which is the interpretation of the value for the mean level of education). Hence, the sum of both proportions yields the level of mismatch, which is the value of indicator $ME_{c,y}$.

We also apply this methodology to measure the mismatch in the field of education, i.e., if a worker has not completed the usual field of education to perform a given job. We establish variable $f_i$, representing the field of education for worker $i$ ($\bar{f}$ is the average and $sf$ is the standard-error of $f_i$). Therefore, the objective is to understand whether a worker possesses adequate training for her job or if she is performing a task completely unrelated to the training she received, which will be a case of horizontal mismatch. Our measure of horizontal mismatch is thus:

$$HE_{icy}(f_i|a_i, a_i, c_i, y_i) = \begin{cases} 0 & \text{if } sf < f_i - \bar{f} < sf \\ \end{cases}$$

$$MF_{c,y} = \frac{\sum_{i: HE_{icy} \neq 0} |HE_{icy}|}{n_{cy}}$$

This measure is also country and year-specific. These measures are comparable across country and over time.

---

6 The use of one and two standard deviation from the mean is based on the 95% confidence intervals (2 standard-deviations from the mean)—see e.g. Kiker et al. (1997).

7 An alternative approach here would be to consider any deviation from the mode. However, given that the definitions of the source variable HATFIELD have a certain notion of ‘proximity’ between the needed skills to attain different fields of study [e.g. Humanities, languages and arts (200) is closer to Foreign languages (222) than to Computer science (481)] we choose the approach that measures distance from the average, which we think better captures this notion of ‘proximity’ between fields of study.
4 Mismatch in Europe

In this section we will concentrate on describing the mismatch measures in Europe. First, we analyse vertical mismatch (under and overeducation) and then we analyse horizontal mismatch.

4.1 Under and Overeducation (Vertical Mismatch) in Europe

There is a great diversity in the level and evolution of vertical mismatch across countries and also on the prevalence of under and overeducation. This may reinforce the evidence for specific labour markets in each of the European countries and thus the possibility that those labour markets are affected by idiosyncratic shocks. The following figures show the evolution of vertical mismatch (the sum of under- and overeducation) by country throughout the years (Fig. 1). The first insight from the figures is the great diversity in the incidence of vertical mismatch across countries. Switzerland and UK tend to have the highest level within the Northern and Central Europe countries, Italy the highest amongst the Southern countries and Estonia and Romania the highest amongst the Eastern European countries. Denmark, Portugal, and Slovak Republic tend to have the lowest level of vertical mismatch in their respective groups. From the first inspection of the figures it is not possible to anticipate any relationship with income levels, or even growth or unemployment. There seems to be a convergence of vertical mismatch levels amongst Northern and Central European countries, with a slight decrease in mismatch after 1997 in some countries, a pattern that appears to occur in the Southern and Eastern European countries only after 2001.

When under- and overeducation are dissected by country (see Fig. 2), there is again evidence according to which vertical mismatch is a country-level phenomenon. Undereducation is a more prevalent phenomenon than overeducation; considering all the country and year pairs for which we have data; in 76.5 % of the cases undereducation is higher than overeducation. For almost half of the countries undereducation is stronger than overeducation for the whole period. In a smaller set of countries there are mixed results and only for Portugal is there more incidence of overeducation than of undereducation.8 Also here, it is not possible to anticipate a clear relationship with income, growth, or unemployment although we can note that southern European countries tend to be in the group of countries in which overeducation tends to dominate and the richest and northern European countries tend to be in the group of countries in which undereducation dominates.

---

8 Undereducation is always higher than overeducation in Austria, Bulgaria, Denmark, Estonia, Finland, France, Hungary, Latvia, Netherlands, Norway, Poland, Sweeden, and Switzerland. Undereducation is higher than overeducation in a majority of the years in the sample in Belgium, Czech Republic, Ireland, Lithuania, Luxembourg, Slovenia and United Kingdom. In Iceland, Cyprus, Germany, Malta, Slovak Republic, and Spain there are mixed results (undereducation and overeducation levels are very close and there are switches of the one that prevails along the time series). In Italy, Greece and Romania overeducation is higher than undereducation in a majority of the years in the sample and Portugal is the only country that presents higher levels of over- than of undereducation in the whole country time-series.
Fig. 1 Vertical mismatch for a set of regions in Europe
Fig. 2 Undereducation and overeducation across time for a set of 30 countries
4.2 Horizontal Mismatch in Europe

Regarding horizontal (or field) mismatch, there is also great diversity across countries. Norway, Poland, Spain, and Estonia present the highest levels of horizontal mismatch with nearly 40% of workers not matching the average field for each occupation and activity. The lowest levels of horizontal mismatch are seen in Germany, Denmark, Portugal, Cyprus, Greece, and Slovak Republic. There is an evident time pattern for a drop in horizontal mismatch after 2008, which is more evident for Northern, Central, and Southern European countries than for Eastern European Countries (see Fig. 3).

4.3 Mismatch by Sector and Occupation

In the whole database vertical mismatch is higher in “household production” and “agriculture, forestry and fishery”. There are a few sectors with more overeducated workers than undereducated, which are “agriculture, forestry and fishing” (A), “information and communication” (J), “financial and insurance” (K) “real estate” (L), “administrative and support services” (N), “public administration and defense” (O), and “extraterritorial organizations” (U). When comparing strong undereducation and strong overeducation, strong undereducation is higher for all economic sectors of activity. Interestingly, horizontal mismatch tends to be quite correlated with vertical mismatch across sectors of economic activity, with quite similar levels. Exceptions are for “agriculture, forestry and fishery” where horizontal mismatch is quite smaller and for “Transportation and Storage”, “Accommodation and Food Service”, “Professional, Scientific and Technical” and “Arts, Entertainment and Recreation”, for which horizontal mismatch is much higher than vertical mismatch.

When analysed by professions or occupations, vertical mismatch is higher for “agricultural, forestry and fishery” workers, “elementary occupations”, “services and sales”, “technicians”, and “clericals”. On the contrary, mismatch is lower amongst “managers”
Fig. 3 Horizontal mismatch for a set of regions in Europe
and “professionals”. Overeducation is prevalent among “technicians”, “clerical”, “agricultural, forestry and fishery” workers and undereducated workers prevail in “managers”, “professionals”, “service and sales”, “craft and related trades”, “plant and machine operators” and “elementary occupations”. Horizontal mismatch tends to be higher for “elementary occupations” and “service and sales” and lower for “crafts and related trades”. Data of mismatch by sector and occupation are plotted in Fig. 4.

Interestingly, correlations between vertical mismatch by sector and occupations between countries are low and differ considerably between countries. Table 1 shows the correlations between each country vertical mismatch level and the European level given by the whole dataset obtained for the year 2011. The table reveals a great heterogeneity between countries meaning that the distribution of vertical mismatch proportions between sectors of activity and occupations is quite diverse across European countries. It is not even possible to identify groups of countries in which both correlations (sectoral and occupational) are high or low, nor by income level, geographic proximity or any other criteria. This evidence reinforces the great heterogeneity in country specific labor markets in Europe.

4.4 Mismatch per Country: Summary Statistics

In this section, we analyse some of the properties of the newly created mismatch series. Table 2 presents some statistics. On average, for European countries 25.9 % workers are mismatched, i.e., have quite different levels of education than the average for the profession, sector, and country. Of those, more are undereducated (near 15 %) and the rest (near 11 %) are overeducated. Strong over and undereducation is low when compared to the total, meaning that, on average amongst the European countries, strong undereducation is 3.6 % while strong overeducation is 1.6 %. Additionally, horizontal mismatch is around 30 %. The averages for the variables are very close to the median, meaning that the data are quite evenly distributed around the mean. Those variables are relatively persistent with autocorrelation coefficients around 0.4. Most countries present high autocorrelation coefficients. The Pesaran (2007) test for unit roots indicates that the null of no-unit root is rejected, with one exception: the test without trend for the strong overeducated measure. Thus, there is evidence against stationarity in mismatch variables.

In what follows we will relate our different measures of mismatch amongst themselves and with unemployment (from the Eurostat). Correlations between the incidence of mismatch according to the different measures are shown in Table 3. Undereducation and strong overeducation are negatively correlated and the same is found for strong undereducation and overeducation. Undereducation is positively correlated with horizontal mismatch while strong overeducation is negatively correlated with horizontal mismatch.

Correlations with unemployment are also interesting. First, these correlations are low (below 0.2). This makes our mismatch measures interesting complements of unemployment to characterize the labour market, giving additional information to politicians and analysts concerned with labour market features and distortions. Second, undereducation is negatively correlated with unemployment, while strong overeducation is positively correlated with unemployment (the correlation of overeducation with unemployment is not significantly different from zero). This suggests that undereducation and (strong) overeducation occur at different moments of the business cycle and/or have different meanings in

---

9 We do not perform this analysis for horizontal mismatch due to the small number of time-series observations per country.
terms of situation in the labor market. Undereducation seems to occur when unemployment is low, suggesting that it is linked to periods or countries with tight labour markets. On the other hand, overeducation, or at least strong overeducation seems to be linked to periods of high unemployment, suggesting that it is in these periods (and/or countries) that workers are willing to accept jobs for which they are overqualified.
We provide new evidence on aggregate mismatch levels (total, under- and overeducation) comparable by countries and years, presenting a new panel dataset with 30 countries for approximately 20 years, between 1993 and 2011. We also provide a measure of horizontal mismatch for the same 30 countries for a period of 9 years. Additionally, we study the relationship between different measures of mismatch and between mismatch and unemployment. The weak correlation we obtain between mismatch and unemployment implies that mismatch measures may be regarded as supplying additional information about labour markets when compared to the simple analysis of the unemployment rate and thus do not overlap information on this widely used statistic for the labour market.

In European countries, between 15 % to nearly 35 % of workers have a job for which they have more (or less) qualifications than the usual level, depending on the country. Approximately 20 % to nearly 50 % work in a job for which they do not have the usual qualifications.

### Table 1 Correlations between each country vertical mismatch and European mismatch by sector and occupation

| Country        | Sectoral mismatch | Occupational mismatch |
|----------------|-------------------|-----------------------|
| Austria        | 0.20              | 0.25                  |
| Belgium        | 0.23              | 0.60                  |
| Bulgaria       | 0.21              | 0.12                  |
| Cyprus         | 0.41              | 0.77                  |
| Czech Republic | −0.02             | 0.00                  |
| Germany        | 0.22              | 0.37                  |
| Denmark        | 0.26              | 0.70                  |
| Estonia        | −0.11             | 0.70                  |
| Finland        | 0.44              | 0.25                  |
| France         | 0.66              | 0.75                  |
| Greece         | 0.38              | 0.21                  |
| Hungary        | 0.38              | 0.68                  |
| Iceland        | 0.33              | 0.73                  |
| Ireland        | 0.57              | 0.48                  |
| Italy          | 0.11              | 0.31                  |
| Latvia         | 0.28              | 0.71                  |
| Lithuania      | 0.47              | 0.56                  |
| Luxembourg     | 0.60              | 0.43                  |
| Malta          | −0.11             | 0.20                  |
| Netherlands    | 0.51              | 0.42                  |
| Norway         | 0.44              | 0.65                  |
| Poland         | 0.57              | 0.66                  |
| Portugal       | −0.09             | 0.16                  |
| Romania        | 0.44              | 0.65                  |
| Slovenia       | 0.41              | 0.31                  |
| Slovak Republic| −0.14             | 0.02                  |
| Spain          | 0.61              | 0.72                  |
| Sweden         | 0.47              | 0.82                  |
| Switzerland    | 0.14              | 0.52                  |
| United Kingdom | 0.83              | 0.82                  |

Correlations are country data correlations with EU data in 2011. Example: the vertical mismatch by sectors in Austria has a correlation of 20 % with the vertical mismatch by sectors in Europe.
field qualification. There is a great difference in mismatch amongst European labour markets. This may reinforce the evidence for specific labour markets in each European country and thus the possibility that those labour markets are affected by idiosyncratic shocks. Undereducation affects more workers than overeducation in most European countries. Among sectors of activity, “agriculture, forestry and fishery” and “household production” are the ones in which there is the highest proportion of mismatch. Among occupations, it is also among “agricultural, forestry and fishery” workers that both vertical and horizontal mismatch is higher. However, this also differs considerably among European countries.

The presentation of new and comparable mismatch measures for European countries opens a wide avenue of empirical research on the features of European labour markets and also on the determinants of labour mismatch. The data presented here highlight the low integration of labour markets in Europe and a significant number of workers who are mismatched regarding level and field of education.

Table 2 Descriptive statistics and unit root tests

| Variable     | Vert. mismatch | Undered. mismatch | Overed. mismatch | Hor. mismatch | St. under | St. over |
|--------------|----------------|-------------------|------------------|---------------|-----------|----------|
| Average      | 0.259          | 0.146             | 0.113            | 0.300         | 0.036     | 0.016    |
| Median       | 0.261          | 0.150             | 0.112            | 0.287         | 0.036     | 0.015    |
| SD           | 0.053          | 0.042             | 0.036            | 0.061         | 0.017     | 0.013    |
| Maximum      | 0.388          | 0.281             | 0.207            | 0.461         | 0.093     | 0.062    |
| Minimum      | 0.135          | 0.026             | 0.038            | 0.190         | 0.001     | 0.000    |
| Autocorrelation | 0.33        | 0.44              | 0.43             | 0.45          | 0.36      | 0.44     |
| Panel UR (2 lags) | 0.042      | 0.285             | 0.678            | –             | 1.929     | 1.811**  |
| (p value)    | (0.517)        | (0.612)           | (0.751)          | (0.973)       | (0.035)   |          |
| Panel UR (2 lags) w. trend | 1.678    | 2.635             | 1.294            | –             | 2.385     | 1.781    |
| (p value)    | (0.953)        | (0.996)           | (0.902)          | (0.991)       | (0.963)   |          |

* Means significance at 10 % level, ** at 5 % level and *** at 1 % level. Panel Autocorrelation is the mean of the AC(1) coefficient for the 30 countries in the database. Panel UR tests are the Pesaran (2007) panel unit root test in which the null is that the series are I(1). Variables: Vert. Mismatch = vertical mismatch; Undered = undereducated; Overed = overeducated; Hor. Mismatch = horizontal mismatch; St. Under = strongly undereducated and St. Over = strongly overeducated

Table 3 Correlations between mismatch measures

| Variable     | V. mismatch | Undered. | Overed. | H. mismatch | St. under | St. over | Unemp. |
|--------------|-------------|----------|---------|-------------|-----------|----------|--------|
| V. mismatch  | 1           | –        | –       | –           | –         | –        | –      |
| Undered.     | 0.722***    | 1        | –       | –           | –         | –        | –      |
| Overed.      | 0.6129***   | 0.618*** | –0.098**| 1           | –         | –        | –      |
| H. mismatch  | 0.229***    | 0.361*** | –0.032  | 1           | –         | –        | –      |
| St. undered. | –0.342***   | 0.072    | –0.563***| –0.099      | 1         | –        | –      |
| St. overed.  | –0.373***   | –0.612***| 0.147***| –0.191***   | –0.339*** | 1        | –      |
| Unemp.       | –0.139**    | –0.176***| 0.011   | –0.003      | 0.099     | 0.199*** | 1      |

* Means significance at 10 % level, ** at 5 % level and *** at 1 % level. Variables: V. Mismatch = vertical mismatch; Undered = undereducated; Overed = overeducated; H. Mismatch = horizontal mismatch; St. Under = strongly undereducated; St. Over = strongly overeducated and Unemp = unemployment
Acknowledgments  We kindly acknowledge financial support by Fundação para a Ciência e Tecnologia, under project Education Mismatches and Productivity Differences (PTDC/EGE-ECO/112499/2009). The raw micro-data used in this paper are from the Labour Force Survey (LFS) and were supplied by the Eurostat, under contract LFS/2012/22, which we acknowledge. The responsibility for the conclusions in this paper is the authors’ and not of Eurostat, the European Commission, or any of the national authorities whose data have been used. An earlier version of this work has circulated under the same title as a Working-Paper (CEFAGE-UE Working-Paper 2014/13).

Appendix

Definitions of the Source Variables

| HATLEVEL | Definition                                                                 | Years of schooling |
|----------|---------------------------------------------------------------------------|--------------------|
| 0        | No formal education or below ISCED 1                                      | 0                  |
| 10       | ISCED 0–1 (pre-primary education)                                         | 1                  |
| 11       | ISCED 1 (primary education or first stage of basic education)             | 5                  |
| 21       | ISCED 2 (lower secondary education or first stage of basic education)     | 8                  |
| 22       | ISCED 3 (upper secondary education; access to labour market)-shorter than  | 10                 |
|          | 2 years                                                                   |                    |
| 31       | ISCED 3c (2 years and more)                                               | 13                 |
| 32       | ISCED 3a,b (upper secondary education providing access to level 5)        | 13                 |
| 30       | ISCED 3 (without distinction a,b, or c possible, 2 years + )              | 13                 |
| 33       | ISCED 3c (3 years or longer) or ISCED 4c                                  | 13                 |
| 34       | ISCED 3b or ISCED 4b                                                      | 13                 |
| 35       | ISCED 3a or ISCED 4a                                                     | 13                 |
| 36       | ISCED 3 or 4 (without distinction a, b or c possible)                     | 13                 |
| 41       | ISCED 4a, b (post secondary, non-tertiary education giving access to level 5) | 14                 |
| 42       | ISCED 4c                                                                 | 14                 |
| 43       | ISCED 4 (without distinction a, b or c possible)                          | 14                 |
| 51       | ISCED 5b (first stage of tertiary education; provides access to an occupation) | 17                 |
| 52       | ISCED 5b (first stage of tertiary education theoretically based; provides access to research programmes) | 17                 |
| 60       | ISCED 6 (second stage of tertiary education, leading to advanced research qualification) | 19                 |

| HATFIELD | Definition                                                                 |
|----------|---------------------------------------------------------------------------|
| 0        | General programmes                                                        |
| 100      | Teacher training and education science                                     |
| 200      | Humanities, languages and arts                                            |
| 222      | Foreign languages                                                         |
| 300      | Social sciences, businesses and law                                       |
| 400      | Science, mathematics and computing                                        |
| 420      | Life sciences (including biology and environmental science)               |
| HATFIELD | Definition |
|----------|------------|
| 440      | Physical science (including physics, chemistry and earth science) |
| 460      | Mathematics and statistics |
| 481      | Computer science |
| 482      | Computer use |
| 500      | Engineering, manufacturing and construction |
| 600      | Agriculture and veterinary |
| 700      | Health and welfare |
| 800      | Services |

| Economic activity | Definition |
|-------------------|------------|
| A                 | Agriculture, forestry and fishing |
| B                 | Mining and quarrying |
| C                 | Manufacturing |
| D                 | Electricity, gas and steam |
| E                 | Water and waste |
| F                 | Construction |
| G                 | Wholesale and retail trade, repair of vehicles |
| H                 | Transportation and storage |
| I                 | Accommodation and food service |
| J                 | Information and communication |
| K                 | Financial and insurance |
| L                 | Real estate |
| M                 | Professional, scientific and technical |
| N                 | Administrative and support service |
| O                 | Public administration and defense |
| P                 | Education |
| Q                 | Human health and social work |
| R                 | Arts, entertainment and recreation |
| S                 | Other service activities |
| T                 | Households production |
| U                 | Extraterritorial activities |
| Occupation   | Definition                        |
|-------------|-----------------------------------|
| 100         | Managers                          |
| 200         | Professionals                     |
| 300         | Technicians                       |
| 400         | Clerical                          |
| 500         | Service and sales workers         |
| 600         | Agricultural, forestry and fishery workers |
| 700         | Craft and related trades          |
| 800         | Plant and machine operators       |
| 900         | Elementary occupations            |

Number of Observations by Country for HATLEVEL Variable

| Country          | Obs.  | Country  | Obs.  | Country  | Obs.  | Country    | Obs.  |
|------------------|-------|----------|-------|----------|-------|------------|-------|
| Austria          | 1,288,556 | France  | 3,087,955 | Lithuania  | 341,138 | Slovak Rep.  | 672,149 |
| Belgium          | 942,070 | Germany  | 1,130,341 | Luxembourg  | 296,124 | Slovenia  | 464,774 |
| Bulgaria         | 531,795 | Greece  | 2,232,709 | Malta  | 52,721 | Spain  | 2,433,548 |
| Cyprus           | 234,948 | Hungary  | 1,830,379 | Netherlands  | 1,293,398 | Sweden  | 1,647,500 |
| Czech Rep.       | 1,342,063 | Iceland  | 94,197 | Norway  | 351,623 | Switzerland  | 491,162 |
| Denmark          | 574,749 | Ireland  | 1,562,564 | Poland  | 1,643,658 | United Kingdom  | 1,435,754 |
| Estonia          | 147,989 | Italy  | 3,858,707 | Portugal  | 1,143,686 |
| Finland          | 488,471 | Latvia  | 216,173 | Romania  | 1,437,500 |

Number of Observations by Country for HATFIEL Variable

| Country          | Obs.  | Country  | Obs.  | Country  | Obs.  | Country    | Obs.  |
|------------------|-------|----------|-------|----------|-------|------------|-------|
| Austria          | 711,155 | France  | 1,251,307 | Lithuania  | 237,564 | Slovak Rep.  | 460,423 |
| Belgium          | 322,322 | Germany  | 649,987 | Luxembourg  | 145,196 | Slovenia  | 281,122 |
| Bulgaria         | 281,195 | Greece  | 821,746 | Malta  | 15,973 | Spain  | 306,184 |
| Cyprus           | 141,340 | Hungary  | 1,043,699 | Netherlands  | 614,926 | Sweden  | 1,139,786 |
| Czech Rep.       | 749,145 | Iceland  | 43,033 | Norway  | 173,987 | Switzerland  | 269,252 |
| Denmark          | 327,599 | Ireland  | 431,997 | Poland  | 1,002,464 | United Kingdom  | 477,807 |
| Estonia          | 84,242 | Italy  | 1,606,183 | Portugal  | 221,632 |
| Finland          | 233,986 | Latvia  | 117,825 | Romania  | 836,328 |
# References

Biagetti, M., & Scicchitano, S. (2011). Education and wage inequality in Europe. *Economics Bulletin, 31*(3), 2620–2628.

Budría, S., & Moro-Egido, A. (2008). Education, educational mismatch, and wage inequality: Evidence for Spain. *Economics of Education Review, 27*, 332–341.

Coelho, H. M., Soares, L. M., & Feliz, M. I. B. (1982). *Os Níveis de Qualificação na Contratação Coletiva, Sua Aplicação a Algumas Empresas Públicas do Sector dos Transportes e Comunicações*. Ministério do Trabalho.

den Haan, W., Ramey, G., & Watson, J. (2000). Job destruction and propagation of shocks. *American Economic Review, 90*(3), 482–498.

EUROSTAT. (2012). *Labour Force Survey Microdata, supplied by the EUROSTAT under contract LFS/2012/22*.

Glocker, D., & Steiner, V. (2011). *Returns to education across Europe*. CEPR discussion paper series no. 8568.

Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review, 19*, 131–147.

Kiker, B. F., & Santos, M. C. (1991). Human capital and earnings in Portugal. *Economics of Education Review, 10*(3), 187–203.

Kiker, B. F., Santos, M. C., & Mendes de Oliveira, M. (1997). Overeducation and undereducation: Evidence for Portugal. *Economics of Education Review, 16*(2), 111–125.

Leuven, E., & Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. *Handbook of the Economics of Education, 4*, 283–326.

Mauro, L., & Carmeci, G. (2003). Long run growth and investment in education: Does unemployment matter? *Journal of Macroeconomics, 25*(1), 123–137.

Mendes de Oliveira, M., Santos, M. C., & Kiker, B. F. (2000). The role of human capital and technological change in overeducation. *Economics of Education Review, 19*, 199–206.

Ordine, P., & Rose, G. (2011). Educational mismatch and wait unemployment. Alma Laurea working papers 19.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265–312.

Robst, J. (2007). Education and job match: The relatedness of college major and work. *Economics of Education Review, 26*(4), 397–407.

Sloane, P. (2002). Much ado about nothing? What does the over-education literature really tell us? Keynote address, *International Conference on Over-education in Europe*.

Wolbers, M. (2003). Job mismatches and their labour market effects among school leavers in Europe. *European Sociological Review, 19*(3), 249–266.