# Occupational choice of young graduates, do generic skills matter?

Stijn Rocher*

University of Antwerp†

First Draft

January 31, 2010

**Abstract**

There are several reasons why college students select occupations. This paper examine the extent to which graduates from different disciplines choose different occupations. In order to address this question, it develops an empirical method to identify the relatedness of occupations based on their generic task content. We combine individual education and employment data of UK graduates with ratings on 42 task content areas from the UK Skill Survey. Based on this data, we show that UK graduates direct their search on the job market towards occupations with similar task packages. Furthermore, the wage implications are discussed of entering a very distinct occupation relative to the modal student. As such, our measure can be interpreted within a mismatch context.

*JEL Classification:* I21, J24, J31

*Key words:* Occupational choice, task content, mismatch

## 1 Introduction

The demand for skills in Western economies is bigger than ever. This increase in demand is supported by the remarkable development of educational attainment levels. To illustrate, tertiary attainment for the whole population in OECD countries is at 27.5%, while 38.7% of the students currently graduate in higher education. (OECD, 2009). However, scholars have questioned whether education is providing the right skills to succeed in the job market. An expanding literature on mismatch addresses this question. This strand of research considers workers employed in occupations below their educational qualification as "overeducated". Empirical analysis reveals that the incidence of overeducation is substantial and amounts to 26% of the working population. (Groot & Maassen van den Brink, 2000). Furthermore, it appears to be a stylized fact that the overeducated suffer from a significant wage penalty compared to the earnings of properly matched workers with the same level of education. However, the overeducated earn more than their well-matched co-workers. (Sloane, 2003). Quantity of schooling is only one way to consider the match between graduates and their job. Researchers recently investigated other dimensions of mismatch concerning field of study and competence mismatch. While the different

*e-mail: Stijn.Rocher@ua.ac.be

†Prinsstraat 13, 2000 Antwerp, Belgium.

measures only appear to be weakly related, mismatch based on these indicators is also associated with lower wages. (Allen & De Weert, 2007; García-Aracil & Van der Velden, 2008; Robst, 2007). Moreover, measuring skill mismatch is a rather difficult task and indicators of mismatch are often based on self evaluation. That is, the worker is asked to what extent his educational attainment is related to the job and he can answer on a scale going from complete mismatch to a perfect match. Interpretation of results based on self-reported mismatch is not straightforward. An endogeneity issue could arise if workers rationalize dissatisfaction about their wage or work as skill mismatch. Other underlying reasons as ability or preference of non-pecuniary benefits above wage, could also be responsible for a relation between mismatch and wage. (Nordin et al, 2008; Sloane, 2009). Alternatively, experts classify the match of a worker and his job based on detailed educational and occupational data. Nevertheless, the validity of such a classification is not always clear-cut. In this paper, we address this measurement issue by developing a task-based mismatch indicator.

According to the assignment theory[1], underutilization of acquired human capital may offer an explanation for the reduced wage of mismatched workers relative to their peers. Education provides the skills and knowledge necessary to succeed in the labor market. Yet, returns to educational investment depend on the job in which the worker is eventually employed. While acquired skills are productive and highly valued in some jobs, they may be redundant in others. Therefore, it would seem plausible that graduates direct their search on the labor market towards jobs where their skills are optimally rewarded. Notwithstanding a large literature on the determinants of occupational choice, the extent to which young graduates direct their search on the labor market is largely unknown. Hence, that is a second issue this paper seeks to address. Students entering higher education choose a field of specialization such as Physics, Business or Liberal Arts. Although higher education provides mainly generic skills and knowledge (like problem-solving, numeracy or literacy), each major puts more emphasis on the development of some skills compared to others. For example in the major Mathematics numeracy is accentuated while the major Humanities emphasizes literacy. By definition, generic or transferable skills are skills valued in every job at least to a certain extent. Recent research, however, suggests that significant differences exist in the valuation of these generic skills between (and within) occupations. Autor and Handel (2009) develop a conceptual framework rooted in the task-based approach. They argue that on the one hand jobs differ in their generic task content and that on the other hand workers differ in their skills required to perform these tasks. Since one job can be matched with only one worker, the unbundling of the task content of a job is impossible. Consequently, they conclude that the value of tasks may vary among jobs. Hence, the economy faces an allocation problem. Workers differing in their capability of performing generic tasks have to be associated with jobs rewarding these tasks differently. Income-maximizing young graduates will solve this assignment problem through self-selection into dissimilar jobs. Assuming that task returns reflect marginal task productivity, the graduate will take into account the match between acquired generic skills and job requirements resulting in a non-random allocation towards jobs.

This paper contributes to the literature on two domains. First, it will explore to what extent young graduates direct their search on the labour market. Rooted in the task-approach, an occupation is characterized by its generic task content making use of a unique dataset. Similar to Schönberg and Gathmann (2007), I construct a distance measure based on the generic task vectors to indicate the

[1] See Sattinger, 1993 for an excellent overview.

similarity of occupations. If differences in generic skills acquisitions between young graduates matter for occupational choice, we expect to find that graduates with identical majors self-select into more similar occupations. In particular, we study the relationship between field of study and the task content of jobs. Second, we make a contribution to the literature on skill mismatch. Our distance measure indicates the relatedness of occupations. Hence, we can use this measure as an accurate and objective indicator of mismatch. That is, a graduate entering a very distinct occupation relative to "properly" matched individuals is deemed to be mismatched. We finally investigate if our measure can confirm the obtained results in other studies.

The paper is organized as follows. The following section discusses the data used. Furthermore, I will explain the applied methodology. The obtained results will be presented in the third section and a final section will conclude.

## 2 Data and Methodology

In order to conduct the analysis, two datasets were merged. The first one is the UK Skill Survey taken in 2006. This dataset aims to investigate the employed workforce in the United Kingdom.[2] It provides a resource for analyzing skill and job requirements in the British economy and consists of a representative sample of the employed population aged 20 to 60. In total, 7,762 working individuals across the UK were sampled. The questionnaire has a similar motivation as the US Dictionary of Occupational Titles (DOT) and the Quality and Careers Surveys in Germany. It contains detailed questions on what kind of tasks are important at the current job of the interviewee. This data is used to construct the distance measure between occupations based on the task vector of occupations. Besides the UK Skill Survey, we also make use of the Reflex dataset. This survey contains 70,000 graduates from 16 countries, among which 1,500 British graduates, asked about their qualifications and employment in 2005, five years after graduation. The survey can be used for our purposes because it contains in-depth information on the chosen discipline and the relation between the graduate's job and his education.

The UK Skill Survey contains 42 items that describe the generic task content of jobs. Individuals report on a 5 point scale how essential a certain task is for their job. Although the tasks are generic in the sense that they are valued in every job to a certain degree, each task might be done to a greater or lesser extent among jobs. Therefore the task intensity is job specific. These data based on self-reported response has its drawbacks, but is nevertheless informative. We use explanatory factor analysis to reduce the amount of intercorrelated variables and identify twelve underlying factors. These factors will serve as measures for the intensity of generic tasks executed by the worker on his job and cover the following task fields: Computer, Literacy, Managing, Numeracy, Nurturing, Physical, Problemsolving, Reviewing, Routine, Self-planning, Selling and Teamwork. Table 1 gives an overview of the task item categorization and reports the task units on which each factor loaded strongly. Croanbach's alpha is reported for every generic task component and easily satisfies the acceptable level of 0.7 in almost all cases. (Nunnally & Bernstein, 1994).

[2] For an in-depth analysis see Felstead et al. (2007).

Table 1: Categorization of Task Items

| | |
|---|---|
| **Computer** ($\alpha = 0.7629$) | **Physical** ($\alpha = 0.8158$) |
| Computer use | physical strength |
| Complexity of computer use | physical stamina |
| Internet use | finger and hand dexterity |
| **Literacy** ($\alpha = 0.8787$) | knowledge of use or operation of tools |
| Reading written information | **Problemsolving** ($\alpha = 0.8636$) |
| Reading short documents | spotting problems or faults |
| Reading long documents | working out cause of problems & faults |
| Write forms, notices or signs | thinking of solutions to problems |
| Write short documents | analyzing complex problems in depth |
| Write long documents | **Reviewing** ($\alpha = 0.7763$) |
| **Managing** ($\alpha = 0.8028$) | noticing when there is a mistake |
| teaching people | checking things to ensure no errors |
| persuading or influencing others | paying close attention to detail |
| making speeches or presentations | **Routine** |
| planning the activities of others | short repetitive tasks |
| **Numeracy** ($\alpha = 0.8125$) | **Self-planning** ($\alpha = 0.8336$) |
| basic arithmetic | planning own activities |
| arithmetic involving fractions | organizing own time |
| advanced mathematics and statistics | thinking ahead |
| **Nurturing** ($\alpha = 0.7505$) | **Selling** ($\alpha = 0.5745$) |
| counselling, advising or caring for clients | Knowledge of products or services |
| dealing with people | Selling a product or service |
| handling feelings of others | Specialist knowledge or understanding |
| managing own feelings | **Teamwork** ($\alpha = 0.8505$) |
| | working with a team |
| | cooperating with colleagues |
| | listening carefully to colleagues |

We characterize how similar occupations are in terms of these generic task factors. Within our framework, two occupations are entitled to be similar if they put similar weights on their generic task content. Given that different fields of study put emphasis on the development of other generic skills, graduates will acquire different competencies. (see e.g. Heijke et al, 2003; García-Aracil and Van der Velden, 2008). Likewise, occupations differ in their generic task requirements and the relative valuation of these tasks. (Autor and Handel, 2009). Consequently, the specific job where an individual is employed determines this worker's productivity. To the extent that earnings also rely on the match between the worker's skills and the job requirements, income-maximizing graduates direct their job search. We assume that students graduated in the same major have acquired more similar generic skills and therefore self-select into more similar occupations.

A first thing to do is defining an occupation. The international standard classification of occupations (ISCO) groups jobs in occupations and assigns codes.[3] For our analysis, we keep all 4-digit ISCO codes

[3] The International Standard Classification of Occupations (ISCO) is organised by the ILO and is occasionally updated.

with more than 10 observations. Next, all other observations are grouped in the 3-digit ISCO coding. If individuals report doing a 3-digit ISCO code occupation with less then 10 observations, they are removed from the sample. Based on the task evaluation of 7,651 observations, we are able to identify the task content of 97 occupations.[4]

In order to quantify the similarity of occupations in their generic task content, we construct the Mahalanobis distance (MD) between the 12-dimensional task vectors of each occupation which is given by following formula:

$$MD_x^y = \sqrt{(\mathbf{x}-\mathbf{y})^t \, \Sigma^{-1} \, (\mathbf{x}-\mathbf{y})} \tag{1}$$

In expression (1) $\mathbf{x}$ and $\mathbf{y}$ indicate the task vectors of two random occupation $x$ and $y$ taken from a task vector distribution with covariance matrix $\Sigma$. The MD is in fact the weighted Euclidean distance where the weights are determined by the covariance matrix and for that reason superior in data analysis. The computation of the inverted covariance matrix may cause problems in the case of multicollinearity. Furthermore, the MD is rather sensitive to measurement error. Therefore, we opt to use the 12-dimensional task vector, instead of all 42 task units. We obtain the MD between the 97 occupations described by the task data. The distances varies from 0 to less than 1. The value equals zero for occupations that involve an identical task package and approaches one if the task content is far from similar. The average distance is 0.48 with a standard deviation of 0.10.

A next step consists of merging our distance measure with the Reflex dataset. This dataset contains 1,500 British HE graduates. From all these observations, we managed to match 1,145 individuals with 67 unique occupations for which task content and distance measures are available. This selected sample is used for further analysis. Table 2 comprises some selected examples that illustrates differences in the task measures among commonly selected occupations by HE graduates. Since we obtained our task indices from factor analysis, they have mean zero and standard deviation one. The occupations presented are architects, engineers and related professionals (isco 214), health professionals (isco 222), finance and sales associates (isco 341) and secondary education teaching professionals (isco 2320).

For our analysis, the version of 1988 is used. For more info: www.ilo.org

[4] As a robustness check, we also group observations only per 3-digit code and raise the minimum number of observations towards 25. Although this procedure reduces the number of occupations for which we can identify the task content, the final results are not much altered.

Table 2: Occupations and their task content

| **Occupation** | **architects & engineers** | **health prof. (except nursing)** | **finance & sales ass.** | **sec. educ. teachers** |
|---|---|---|---|---|
| ISCO - code | 214 | 222 | 341 | 2320 |
| | | | | |
| Computer tasks | 0.444 | -0.482 | 0.656 | 0.074 |
| Literacy Tasks | 0.069 | -0.007 | -0.267 | 0.967 |
| Managing Tasks | -0.107 | 0.281 | -0.182 | 1.952 |
| Numeracy Tasks | 1.232 | -0.135 | 0.869 | -0.197 |
| Nurturing Tasks | -1.029 | 1.187 | -0.334 | 1.421 |
| Physical Tasks | 0.079 | 0.637 | -0.782 | 0.112 |
| Problemsolving Tasks | 0.678 | 0.964 | -0.590 | -0.493 |
| Reviewing Tasks | 0.147 | 1.052 | -0.139 | -1.175 |
| Routine Tasks | -0.817 | 1.094 | -0.125 | -1.142 |
| Self-Planning Tasks | -0.237 | -0.669 | 0.229 | 1.270 |
| Selling Tasks | 0.177 | 1.179 | 1.132 | -0.821 |
| Teamwork | 0.035 | 1.380 | -0.589 | 1.272 |
| | | | | |
| share of workers | 1.49 | 1.08 | 3.05 | 1.15 |
| Amount of workers | 114 | 83 | 233 | 88 |

*note:* Based on author's calculations

It is clear from table 2, that the occupations selected by graduates vary in their task content. According to our task data, the main tasks of architects and engineers comprise non-routine numeracy and problem-solving tasks often supported by computers. Nurturing tasks are rather exceptional. Also finance and sales associates spend a great deal of their time on numeracy using computers. However, for this occupational group, selling is much more important than problem-solving. They also perform relatively little physical tasks. Furthermore, our task measures indicate that health professionals work a lot in teams and mainly perform nurturing tasks. In addition, they spend much time on reviewing and routine tasks. Also very plausible is the task description of secondary education teachers. They manage, nurture and often work in groups. Furthermore, planning and organizing their own time is reported to be relatively important. To better grasp the distance measure, we report table 3. This table gives details on our examples (column 1) and their distance to closely linked (column 2) and distant occupational titles (column 3). The generic tasks performed by architects and engineers relates to the tasks of physics and engineering science technicians (isco 311). This is indicated by a rather low MD of 0.180. The distance measured between architects and engineers and library and filling clerks is rather high and takes a value of 0.580. Also our other examples of occupational titles have close links with some occupations and are very distinct from others. The MD ranges from 0.200 to 0.600.

Table 3: Close and distinct occupations

| occ. 1 | occupation 2 | dist. | occupation 3 | dist. |
|---|---|---|---|---|
| 214 | physical and engineering science technicians (311) | 0.180 | library and filing clerks (4141) | 0.580 |
| 222 | health associate professionals (322) | 0.237 | government tax and excise officials (3442) | 0.612 |
| 341 | other specialist managers (123) | 0.192 | products machine operators (826) | 0.617 |
| 2320 | college, university and HE teaching professionals (2310) | 0.204 | writers and creative or performing artists (245) | 0.609 |

*note:* based on author's calculations

# 3 Results

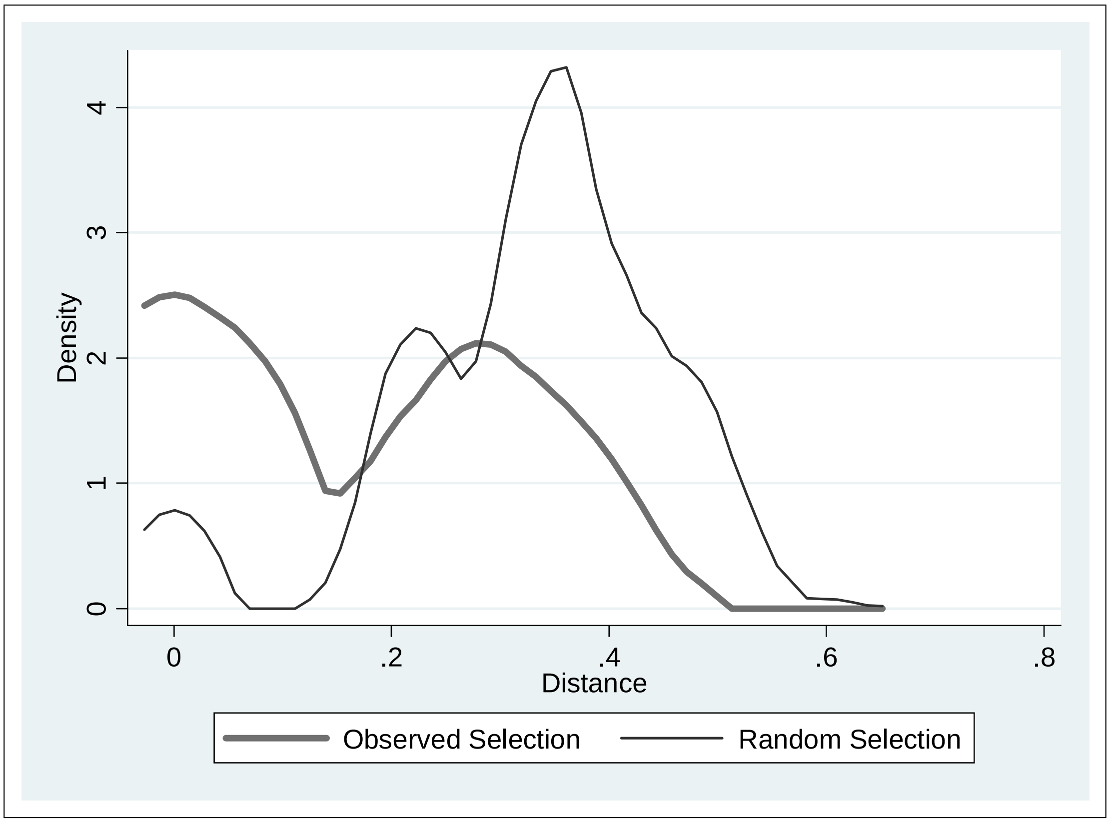
## 3.1 Occupational choice and the task distance between occupations

Now, we will use the dataset constructed in the previous section to demonstrate that the generic task content of jobs influences the occupational choice of HE graduates. As illustrated in table 2, occupations vary in their generic task content. Together with the finding that different fields of study in higher education put different emphasis on the development of generic skills, we claimed that graduates will self-select into occupations optimizing the match between their skills and the task content of the job. The assignment theory supports our statement and argues that the productivity of a job not only depends on the skills of the workers, but also on the skill matching of the worker to his job. On the contrary, we could assume that the productivity only depends on the workers' human capital ignoring the different task weights of occupations. If this would be the case, then the job allocation of graduates would not be influenced by the distance in generic task content of jobs. Every job would equally value the generic skills of the labour market entrant. In this case, the allocation of workers to jobs is random and only determined by the relative size of an occupation. Therefore, to test our hypothesis, we compare two distributions of distances, the observed and the random distribution. Under our hypothesis that generic skills matter for occupational selection, we expect to find that the observed distances are distributed closer to the null than under random allocation.

We obtain the first distribution in a few steps. The first step is to identify a proper match between a graduate from a certain discipline and an occupational field. We use the mode to assign such a "core occupation" to each field of study. Occupations that attract the largest share of graduates from a certain field is regarded as the core occupation. Based on this statistical method to define (mis)match (see also Hartog, 2000), arts graduates are assigned to the occupation "artistic, entertainment and sports associate professionals" (isco 347), law graduates to "legal professionals" (isco 242) and computing graduates to "computer associate professionals" (isco 3120). A complete list can be found in table A1 in the appendix. While identifying a proper match is not an exact science, the result seems plausible. Then, we plot the distribution of the distance between the core occupation and the observed occupational choice of our sample of HE graduates. We refer to it as the observed distribution. For the construction of the random distribution, the workers are randomly allocated to an occupation. The probability that a worker enters a certain occupation is only determined by the relative size of this occupation. Furthermore, we weight the distances between the core and randomly assigned occupation with the relative size of the major.

Figure 1 shows the observed and random distribution of graduates majored in "Engineering and engineering trade". From this picture, we clearly see that the distribution of the distances under observed occupational selection has more mass closer to the origin than under random assignment. Hence, we may conclude that "Engineering and engineering trade" graduates direct their search and choose more similar occupations than can be expected from random assignment.

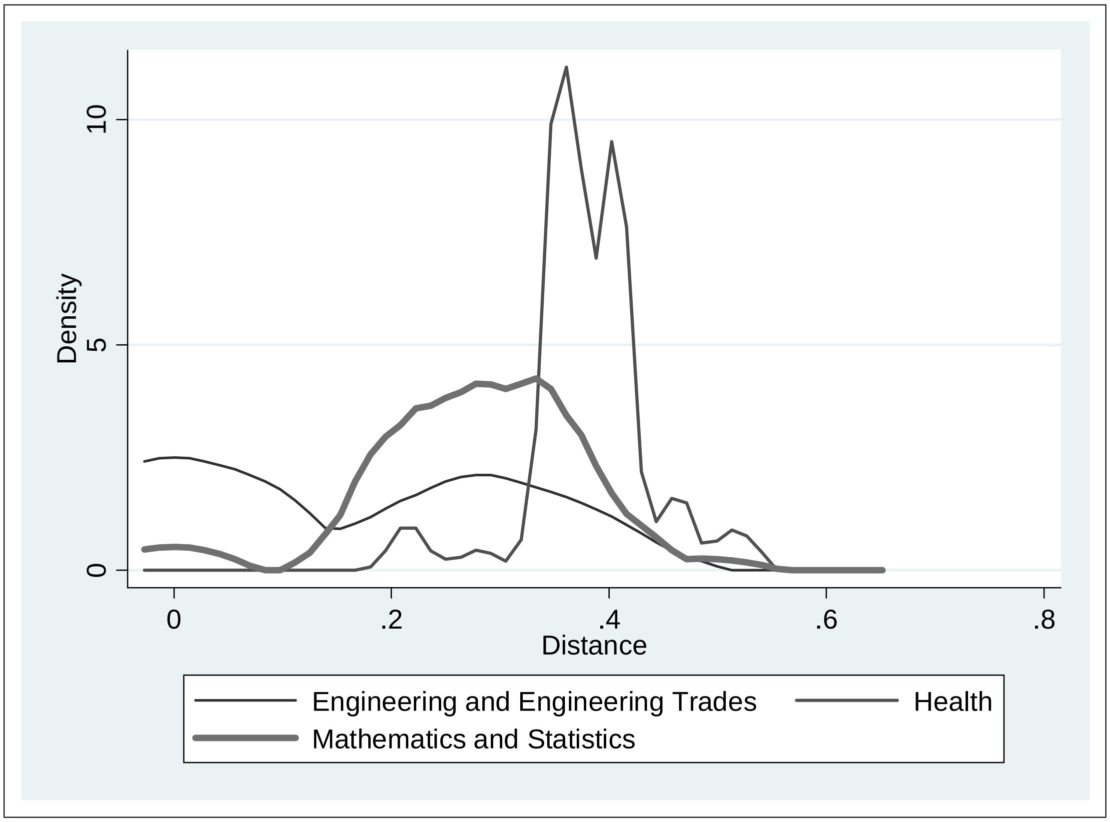
Figure 1: Engineering and engineering trade: observed and random selection




*note:* based on author's calculations

Our framework also predicts that graduates from fields of study that put similar weights on the development of generic skills, will self-select into occupations with a similar task content. A related major to "Engineering and engineering trade" is "Mathematics and Statistics", while "Health" is a very distinct field of study. The distances between the core occupation of "Engineering and engineering trade" graduates, being "architects, engineers and related professionals", and the observed occupational selection of these three groups of graduates differ. The distribution of these distances is given in figure 2 per field of study. As expected, "Mathematics and Statistics" graduates choose more similar occupations than graduates in "Health".

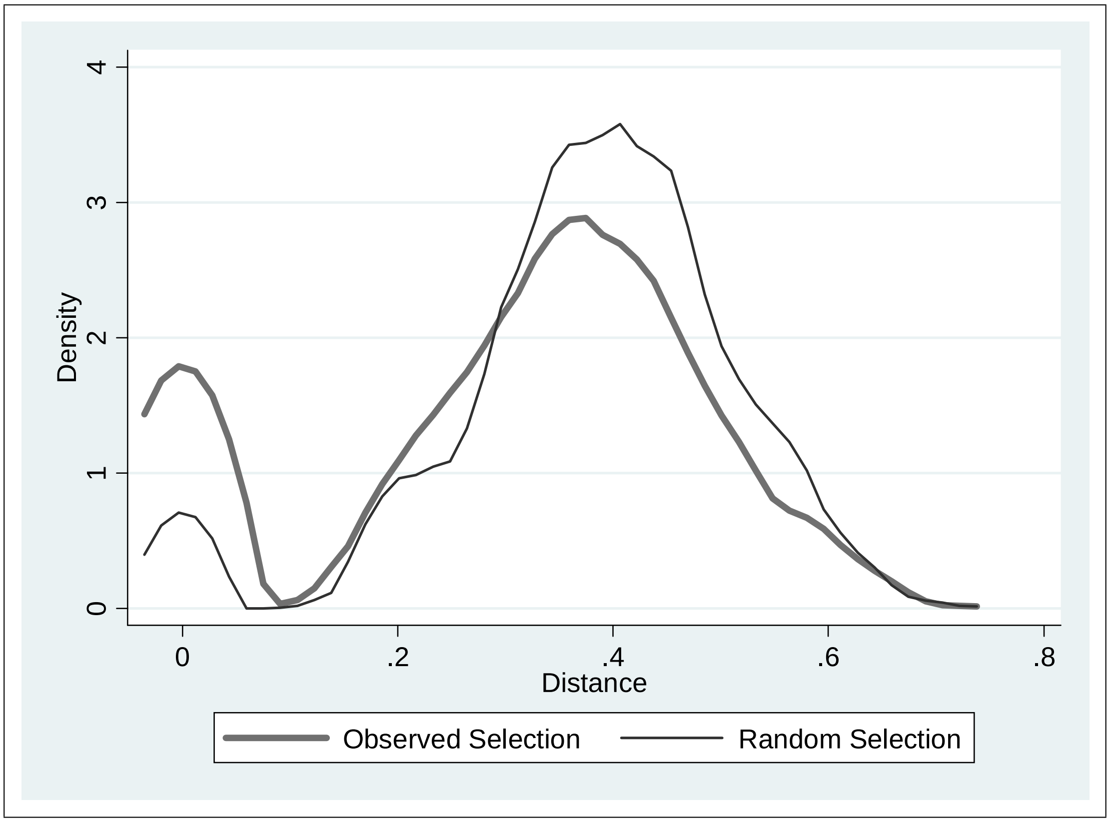
Figure 2: Related and distinct occupational selection




*note:* based on author's calculations

If we look at the observed and random distribution of distances for our whole sample, the picture is less clear in figure 3. However, the peak in the distribution around zero is much larger under observed selection than under random assignment. In table 4, we report some descriptive statistics. These figures indeed indicate that the observed occupational selection of young graduates is directed towards more similar occupations. Although the standard deviation is larger indicating a larger spread, the 25th percentile, the mean, the 90th percentile and the maximum are all lower for the density distribution of the observed distances. Since both distributions are bimodal, these summary statistics should be interpreted with care. Therefore, a Kolmogorov-Smirnov equality-of-distributions test (p-value =0.000) is performed. Also this test rejects the null that both distributions are equal and favours that the distribution of the distance measure under observed selection is more right-skewed than this under random selection.

Figure 3: Distance under observed and random selection




*note:* based on author's calculations

Table 4: Selected moments of the distance density function under observed and random selection

| | **Observed Mobility** | **Random Mobility** |
|---|---|---|
| Minimum | 0 | 0 |
| 25th percentile | 0.208 | 0.319 |
| Mean | 0.306 | 0.377 |
| 90th percentile | 0.493 | 0.538 |
| Maximum | 0.702 | 0.854 |
| Standard deviation | .177 | .134 |
| | | |
| Kolmogorov-Smirnov Test | p-value = 0.000 | |

*note:* based on author's calculations

### 3.2 Mismatch among UK graduates

Since the early nineties, the large incidence of educational mismatch among young graduates has raised a lot of concern. Mismatched graduates are not making use of all their acquired skills in their jobs. This alleged waste of public money has provoked controversy. Different indicators measuring the degree of overeducation show substantial negative wage effects of this type of mismatch. However, differences in measurement result in a range of estimated magnitudes of the wage penalty. (Sloane, 2003). Empirics rely on three methods to define overeducation. First, a subjective measure is based on the selfassessment of the worker who is asked to evaluate his own qualification with the job requirement. Second, labor market experts identify the sufficient educational degree required to perform a certain occupation. A third method consists of statistical measurement. Workers who possess more (typically one standard deviation) years of schooling than the mean or median worker in a certain occupation are classified as overeducated. Besides different methods to measure overeducation, researchers recently argued that also other dimension of mismatch may exist, like competence mismatch or mismatch concerning field of study. (Allen and De Weert, 2007; García-Aracil and Van der Velden, 2008; Robst, 2007). While also these dimensions of mismatch are penalized in terms of reduced wages, these results should be treated with a degree of caution as they make use of subjective indicators. In a meta-analysis, Groot and Maassen van den Brink (2000) find that measures based on self-reports result in higher incidence of mismatch than more objective measures. Our distance measure can be interpreted as an improved statistical measure of mismatch that takes into account the generic task similarity of jobs. Here, we will discuss how our distance measure relates to the other mismatch measures and to what extent we can replicate the finding that the skill match between a worker and his job matters for earnings.

Table 5: Our mismatch measure compared to other measures

| **Variables** | **(1)** | **(2)** | **(3)** | **(4)** |
|---|---|---|---|---|
| *horizontal mismatch* | | | | |
| exclusively own field | ref. | | | ref. |
| own or relative field | 0.062*** | | | 0.060*** |
| | (0.016) | | | (0.016) |
| very different field | 0.180*** | | | 0.171*** |
| | (0.018) | | | (0.019) |
| no particular field | 0.182*** | | | 0.165*** |
| | (0.016) | | | (0.019) |
| *vertical mismatch* | | | | |
| higher level | | ref. | | ref. |
| same level | | -0.012 | | -0.036 |
| | | (0.02) | | (0.020) |
| lower tertiary level | | -0.092* | | -0.109* |
| | | (0.046) | | (0.043) |
| below tertiary level | | 0.097*** | | -0.006 |
| | | (0.022) | | (0.024) |
| *skill mismatch* | | | | |
| intensive skill use | | | ref. | ref. |
| standard skill use | | | 0.061*** | 0.009 |
| | | | (0.014) | (0.014) |
| low skill use | | | 0.152*** | 0.033 |
| | | | (0.016) | (0.021) |
| constant | 0.217*** | 0.297*** | 0.276*** | 0.248*** |
| | (0.013) | (0.019) | (0.008) | (0.021) |
| adj. R-squared | 0.128 | 0.042 | 0.037 | 0.135 |
| correlation: | 0.358 | 0.205 | 0.192 | 0.367 |
| N | 988 | 988 | 988 | 988 |

*note:* stars indicate significance levels: * p<0.05, ** p<0.01, *** p<0.001; robust standard errors are given between parentheses

Since our statistical measure of mismatch is new, we report its relation with other measures of mismatch in table 5. The correlation with subjective measures of mismatch concerning years of schooling, field of study and competence mismatch are modest and range from 0.192 to 0.358. This is not surprising since researcher also already demonstrated weak correlations between different dimensions of mismatch.[5] More crucial is the fact that our measure behaves well and is positively related with higher degree of mismatch following the other mismatch indicators. Furthermore, it is an important finding that our measure is related the strongest with mismatch concerning educational field and the job. Hence, if a graduate reports a mismatch between his field of study and his job, it is likely that he is doing a job requiring very different tasks compared to jobs of workers with similar majors. This supports our hypothesis that the different task content among occupations attracts graduates from distinct fields since these fields emphasized the development of some, but not all, skills.

Next, we examine in table 6 which personal characteristics determine the degree of mismatch according to our measure. We estimate four specifications. In the first specification we only control for gender, age, marital status and ethnic background. The findings are that women are mismatched significantly (at the 0.1% level) more than men. Furthermore, Asians are matched better than the European reference group. Other ethnic backgrounds are not mismatched significantly different than the European reference group. Age or marital status do not seem to matter. Subsequently, we also control for current degree, experience (in months), past unemployment spells, public sector employment and whether he participated in an internship or not. The likelihood that a graduate accept a distinct job relative to the job of his peers is higher if the graduate has suffered from an unemployment spell. Internship participation makes graduates choose less for jobs that do not relate to their skills. Because vocationally oriented study programmes often provide internships, this could explain why internship participation enhances "good" matches among graduates. The last two specifications show that the degree of mismatch greatly vary across fields of study. Relative to the omitted category "humanities", none of the other fields is significantly worse matched on average except for the major "services". Graduates from "engineering, manufacturing and construction" choose the most similar jobs. The related major "architecture and building" follow this example to a certain extent. Furthermore, majors "arts", "business and administration", "law", "science, mathematics and computing", "life sciences", "agriculture, forestry and fishery" and "health" choose occupations with a more related task content compared to "humanities". Remarkably, these coefficients remain robust after taking into account control variables like "internship participation". Therefore, we believe that not only the degree of vocational orientation of the different fields of study matters for occupational choice. In contrast, it strengthens our hypothesis that the different fields emphasize the development of very different generic skills directing occupational choice.

[5] E.g. Allen and de Weert (2007) show educational mismatches in no way imply mismatches between available and required knowledge and skills. Moreover, graduates indicating an educational match do report skill mismatches.

Table 6: Spread according to different fields

| **Variables** | **(1)** | **(2)** | **(3)** | **(4)** |
|---|---|---|---|---|
| bachelor | | ref. | | ref. |
| master | | -0.006 | | -0.003 |
| | | (0.018) | | (0.015) |
| phd | | 0.052 | | 0.062 |
| | | (0.042) | | (0.048) |
| internship | | -0.087*** | | -0.061*** |
| | | (0.015) | | (0.014) |
| experience (months) | | -0.001 | | -0.000 |
| | | (0.001) | | (0.000) |
| unemployed | | 0.040** | | 0.028* |
| | | (0.013) | | (0.012) |
| public | | 0.018 | | -0.026* |
| | | (0.012) | | (0.012) |
| humanities | | | ref. | ref. |
| education and training | | | 0.014 | 0.051 |
| | | | (0.070) | (0.071) |
| arts | | | -0.092*** | -0.085*** |
| | | | (0.024) | (0.024) |
| social and behavioral science | | | 0.019 | 0.022 |
| | | | (0.020) | (0.020) |
| business and administration | | | -0.142*** | -0.123*** |
| | | | (0.019) | (0.020) |
| law | | | -0.146*** | -0.141*** |
| | | | (0.037) | (0.037) |
| science, mathematics and computing | | | -0.183* | -0.190* |
| | | | (0.071) | (0.077) |
| life science | | | -0.062** | -0.060** |
| | | | (0.021) | (0.022) |
| physical science | | | -0.050 | -0.045 |
| | | | (0.030) | (0.029) |
| mathematics and statistics | | | -0.102*** | -0.101*** |
| | | | (0.028) | (0.029) |
| computing | | | 0.021 | 0.043 |
| | | | (0.046) | (0.046) |
| eng., man. and constr. | | | -0.375*** | -0.340*** |
| | | | (0.015) | (0.019) |
| architecture and building | | | -0.284*** | -0.262*** |
| | | | (0.024) | (0.026) |
| agriculture, forestry and fishery | | | -0.123** | -0.071 |
| | | | (0.038) | (0.037) |
| health | | | -0.164*** | -0.117*** |
| | | | (0.024) | (0.025) |
| services | | | 0.033 | 0.074* |
| | | | (0.036) | (0.033) |
| environmental protection | | | -0.004 | 0.016 |
| | | | (0.077) | (0.067) |
| constant | 0.242*** | 0.331*** | 0.380*** | 0.368*** |
| | (0.025) | (0.031) | (0.014) | (0.039) |
| controls | yes | no | no | yes |
| adj. R-squared | 0.045 | 0.049 | 0.250 | 0.280 |
| N | 988 | 988 | 988 | 988 |

*note:* stars indicate significance levels: * $p<0.05$, ** $p<0.01$, *** $p<0.001$; robust standard errors are given between parentheses; controls are gender, age, marital status and ethnic background.

Finally, table 7 reports results from a wage regression where the dependent variable is the log of gross hourly wage. We estimate three specifications. The first specification contains only standard controls, specification (2) includes our distance measure and the last specification also contains various subjective measures of mismatch. These relate to the degree of field mismatch, skill mismatch and educational mismatch. Turning to the first specification, all coefficients have expected signs. We shortly summarize the main findings. Graduates in "business and administration", "mathematics and statistics" and "health" earn a significant wage premium compared to "humanities" graduates. At the other hand, "arts" graduates have significant lower wages than the reference category. Women were found to earn nine percent less than men. Older workers and those living with a partner have higher expected earnings. Ethnic background, current obtained degree, public sector employment or experience are not found to matter significantly for wages. Other variables of interest included a post-graduation unemployment spell which depress earnings by 15 per cent. Furthermore, we proxy ability by the relative mark compared to other students. This turns out to be significant. Graduates indicating they had average grades earn 19 per cent more than below average graduates. Those with relatively high grades are paid even 26 per cent more. Finally, we also included a dummy variable indicating the status of the higher education institute. Graduating from a prestigious university tend to significantly increase earnings with almost 17 per cent.

The second specification also includes our distance measure. Following this measure, young graduates performing very different tasks in their job than most graduates from their field are classified as being mismatched. According to the literature on skill mismatch, we find that the mismatched graduates have significantly lower returns to schooling than properly matched individuals with the same major. A one standard deviation more mismatch results in a wage penalty of 3.4% significant at the five per cent level. However, when the other mismatch indicators are added to the model, our distance measure becomes insignificant. It turns out that it is standard or low skill use and lower than tertiary educational requirements that are heavily penalized. Similar estimates for overeducation and overskilling for the UK were found by other studies. (see e.g. Allen and de Weert, 2007; Sloane, 2003). Furthermore, also the subjective measure for field mismatch remain insignificant. Since our measure is the strongest correlated with this mismatch indicator, we exclude the distance measure from the regression. However, this does not alter the finding that field mismatch has no significant impact on earnings. Like Sloane (2003), we therefore conclude that it is not necessary field mismatch, but rather the feeling of being overeducated or overskilled for the job that results in lower wages. Obviously, our mismatch measure, although related as shown in table 5, does not automatically imply overeducation or overskilling. While selecting an unrelated job does not necessarily result in lower wages, overeducation and overskilling apparently do. In addition, adding these vertical mismatch measures has a negative impact on the proxies for ability, i.e. the graduate's perceived grade relative to other students and the prestigious institute dummy. This indicates that low ability is one of the reasons for overqualification.

Table 7: The wage effect of different types of mismatch

| **Variables** | **(1)** | | **(2)** | | **(3)** | |
|---|---|---|---|---|---|---|
| | *b* | *robust se* | *b* | *robust se* | *b* | *robust se* |
| humanities | ref. | | ref. | | ref. | |
| education and training | -0.101 | 0.188 | -0.095 | 0.186 | -0.109 | 0.159 |
| arts | -0.167* | 0.077 | -0.181* | 0.077 | -0.103 | 0.071 |
| social and behavioral science | 0.050 | 0.050 | 0.052 | 0.05 | 0.057 | 0.048 |
| business and administration | 0.124* | 0.053 | 0.103 | 0.054 | 0.105* | 0.05 |
| law | 0.076 | 0.076 | 0.054 | 0.078 | 0.047 | 0.075 |
| science, math. and comp. | 0.072 | 0.243 | 0.034 | 0.232 | 0.042 | 0.206 |
| life science | 0.052 | 0.062 | 0.043 | 0.063 | 0.055 | 0.063 |
| physical science | 0.179 | 0.095 | 0.171 | 0.096 | 0.157 | 0.093 |
| mathematics and statistics | 0.252*** | 0.068 | 0.233*** | 0.07 | 0.233*** | 0.066 |
| computing | -0.100 | 0.103 | -0.098 | 0.101 | -0.065 | 0.1 |
| eng., man. and constr. | 0.181* | 0.084 | 0.121 | 0.091 | 0.128 | 0.089 |
| architecture and building | -0.016 | 0.071 | -0.062 | 0.075 | -0.099 | 0.075 |
| agriculture, forestry and fishery | -0.201 | 0.139 | -0.214 | 0.137 | -0.253 | 0.139 |
| health | 0.164** | 0.054 | 0.144** | 0.055 | 0.102 | 0.055 |
| services | -0.016 | 0.068 | -0.005 | 0.069 | -0.032 | 0.071 |
| environmental protection | -0.109 | 0.210 | -0.106 | 0.215 | -0.032 | 0.177 |
| bachelor | ref. | | ref. | | ref. | |
| master | 0.062 | 0.048 | 0.061 | 0.048 | 0.074 | 0.05 |
| phd | -0.023 | 0.109 | -0.016 | 0.108 | -0.001 | 0.115 |
| internship | 0.090* | 0.036 | 0.080* | 0.036 | 0.076* | 0.035 |
| experience (months) | 0.007 | 0.007 | 0.007 | 0.007 | 0.003 | 0.007 |
| experience squared | -0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| unemployed | -0.154*** | 0.033 | -0.150*** | 0.033 | -0.118*** | 0.031 |
| public | -0.035 | 0.032 | -0.039 | 0.032 | -0.042 | 0.031 |
| female | -0.087** | 0.031 | -0.081** | 0.031 | -0.106*** | 0.029 |
| age | 0.007** | 0.002 | 0.007** | 0.002 | 0.008*** | 0.002 |
| partner | 0.095** | 0.029 | 0.095** | 0.029 | 0.072* | 0.028 |
| European | ref. | | ref. | | ref. | |
| Asian | 0.047 | 0.062 | 0.036 | 0.062 | 0.029 | 0.059 |
| Black | 0.120 | 0.111 | 0.126 | 0.11 | 0.058 | 0.12 |
| Mixed | 0.092 | 0.097 | 0.08 | 0.098 | 0.019 | 0.087 |
| lower than average grades | ref. | | ref. | | ref. | |
| average grades | 0.193* | 0.097 | 0.182 | 0.094 | 0.172 | 0.135 |
| higher than average grades | 0.261* | 0.114 | 0.249* | 0.111 | 0.227 | 0.146 |
| prestigious institute | 0.168*** | 0.046 | 0.174*** | 0.046 | 0.132** | 0.045 |
| field | | | -0.034* | 0.017 | -0.011 | 0.017 |
| exclusively own field | | | | | ref. | ref. |
| own or a related field | | | | | 0.000 | 0.036 |
| a completely different field | | | | | 0.055 | 0.052 |
| no particular field | | | | | 0.026 | 0.05 |
| intensive skill use | | | | | ref. | ref. |
| standard skill use | | | | | -0.169* | 0.073 |
| low skill use | | | | | -0.106** | 0.033 |
| higher educ. level required | | | | | ref. | ref. |
| same educ. level required | | | | | 0.08 | 0.056 |
| lower tertiary level required | | | | | -0.103 | 0.121 |
| below tertiary level required | | | | | -0.243** | 0.075 |
| constant | 1.999*** | 0.198 | 2.026*** | 0.199 | 2.108*** | 0.222 |
| | | | | | | |
| adj. R-squared | 0.137 | | 0.14 | | 0.222 | |
| N | 890 | | 890 | | 890 | |

*note:* stars indicate significance levels: * p<0.05, ** p<0.01, *** p<0.001.

In a merely indicative exercise, we put table 6 and 7 together. Although with some exceptions, graduates from those fields that do relatively very similar jobs (see table 6) seemingly tend to earn more (see table 7). Graduates majored in "business and administration", "mathematics and statistics" and "health" earn more and also have more similar jobs than "humanities" graduates. As an additional test, we replace our distance measure by the mean distance per field of study and then estimate a wage regression reported in appendix table A2. The added variable has a significant negative effect on wages. Hence, this provides evidence that weakly matched majors earn less than more closely matched fields. We found that selecting relatively unrelated fields does not imply skill underutilization and thus lower earnings, but we do find that majors leading their graduates into rather distinct jobs have lower returns even after controlling for skill underutilization. Following our assumption, each discipline puts the emphasis on the development of certain generic skills. Occupations differ in the valuation of these different skills and therefore they attract only graduates from some fields. Moreover, our findings suggest that graduates with a specific skill package seem to optimize income and not the match between their acquired skills and the job requirement. It are workers majored in fields with relatively low returns that tend to seek their luck more often in very distant jobs. Hereby they possibly forego an optimal skill match, but on the other hand are able to get better paid positions.

# 4 Conclusion

In this paper, we make use of a novel approach to relate occupations. This method rooted in the task approach (see e.g. Autor and Handel, 2009) classifies the similarity of occupations based on their generic task content. Exploiting the relatedness of many occupations, we are able to show that UK graduates direct their search on the job market towards more similar occupations. This finding improves our understanding of the occupational choice of young graduates. Given that higher education is organized around disciplines each accentuating particular generic skills like numeracy or literacy, graduates from distinct fields excel in other generic tasks. As a result, they self-select into very different jobs and tend to match their skills and the task content of the job. In contrast to the common believe that generic skills are valued equally in every job, these skills apparently do matter for occupational choice.

Further analysis relates our distance measure with measures of skill mismatch. There exists considerable controversy over the extent to which graduates are mismatched and what dimensions of mismatch are harmful for earnings. At first sight, our results suggests that graduates entering a very distinct occupation relative to properly matched individuals are deemed to suffer from a wage penalty. However, after taking into account other mismatch measures, it seems that it is overeducation and to a lesser extent overskilling that depresses wages. Moreover, we find in a suggestive exercise that workers majored in fields with relatively low returns tend to look for jobs outside their task specialization more often. Although this action results in higher incidence of overeducation or overskilling, graduates may also find better paid positions. This would be consistent with an efficient functioning labour market rather than indicating market failure in assigning heterogeneous workers to heterogeneous jobs.

# Bibliography


1. Allen, J. and E. De Weert (2007). "What do Educational Mismatches tell us about Skill Mismatches? A Cross-country Analysis, European Journal of Education", Vol. 42 (1), pp. 59-73.

2. Autor, D. and M. Handel (2009). "Putting Tasks to the Test: Human Capital, Job Tasks and Wages", NBER working paper, nr. 15116.

3. Felstead, A., D. Gallie, F. Green and Y. Zhou (2007). "Skills At Work, 1986 to 2006". University of Oxford, SKOPE.

4. García-Aracil, A. and R. Van der Velden (2008). "Competencies for young European higher education graduates: labor market mismatches and their payoffs", Higher Education, vol. 55 (2), pp. 219-239.

5. Gathmann, C. and U. Schönberg (2007). "How general is Human Capital? A Task-Based Approach", IZA discussion paper, nr. 3067.

6. Groot, W. and H. Maassen van den Brink (2000). "Overeducation in the labor market: a meta-analysis", Economics of Education Review, vol. 19 (2), pp. 149-158.

7. Hartog, J. (2000). "Overeducation and Earnings: where are we, where should we go?", Economics of Education Review, vol. 19 (2), pp.131-147.

8. Heijke, H., C. Meng and C. Ris. (2003). "Fitting to the job: The role of generic and vocational competences in adjustment and performance", Labour Economics, vol. 10 (2), 215-229.

9. Nordin, M., I. Persson and D. Rooth. (2008). "Education-Occupation Mismatch: Is there an income penalty?", IZA Discussion paper, nr. 3806.

10. Nunnally, J.C., and I.H. Bernstein. (1994). "Psychometric theory", (3rd ed). New York: McGraw-Hill.

11. OECD (2009). "OECD in figures 2009", OECD publishing, Paris.

12. Robst, J. (2007). "Education and job match: the relatedness of college major and work", Economics of Education Review, vol. 26 (4), pp. 397-407.

13. Sattinger, M. (1993). "Assignment models of the distribution of earnings", Journal of Economic Literature, vol. 31 (2), pp.831-880.

14. Sloane, P. (2003). "Much ado about nothing? What does the Overeducation Literature Really Tell Us?" in F. Büchel, A. de Grip and A. Mertens (Eds.), Overeducation in Europe: current issues in theory and policy (pp. 11-48), Edward Elgar Publishing.

15. Sloane, P. and S. McGuinness. (2009). "Labour Market Mismatch Among UK Graduates; An Analysis Using REFLEX Data", IZA Discussion Papers, nr. 4168.

# Appendix

Table A1: Identification of proper match between field of study and occupation

| **Field of study** | **Core Occupation (isco)** | **N** |
|---|---|---|
| education | secondary education (2340) | 14 |
| arts | artistic, entertainment and sports associate prof. (347) | 65 |
| humanities | sec educ teaching professionals (2320) | 186 |
| social and behavioral science | social science and related professionals (244) | 171 |
| business and administration | other department managers (123) | 126 |
| law | legal professionals (242) | 47 |
| science, mathematics and computing | business professionals (2419) | 6 |
| life science | physical and engineering science technicians (311) | 64 |
| physical science | physicists, chemists and related professionals (211) | 41 |
| mathematics and statistics | finance and sales associate professionals (341) | 34 |
| computing | computer associate professionals (3120) | 20 |
| engin., manuf. and constr. | architects, engineers and related professionals (214) | 43 |
| architecture and building | architects, engineers and related professionals (214) | 33 |
| agriculture, forestry and fishery | architects, engineers and related professionals (214) | 23 |
| health | health professionals (except nursing) (222) | 80 |
| personal services | social science and related professionals (244) | 29 |
| environmental protection | managers of small enterprises (131) | 6 |
| | | |
| total nr. of observations | | 988 |

Table A2: Mean distance measure per field of study in wage regression

| **variables** | **(1)** |
|---|---|
| | *b/se* |
| mean distance | -0.046** |
| | (0.016) |
| controls | yes |
| | |
| adj. R-squared | 0.107 |
| N | 769 |

*note:* stars indicate significance levels: * p<0.05, ** p<0.01, *** p<0.001;
dependent variable is the log of hourly wage; controls are degree, internship participation, experience, past unemployment spells, public sector employment, gender, age, partner, ethnic background, relative mark, prestigious institute.