Mismatched mismatch measures. Does the definition of over- and under-qualification matter?*

María Ramos

Version: February 2014

Abstract

This paper examines the impact of the definition of qualification mismatch (QM) chosen, not only on the point estimates, but also on the results of the factors explaining the likelihood of being over- or under-qualified. For this aim, three databases are used: the Labour Force Survey (LFS), the Adult Education Survey (AES), and the Programme for the International Assessment of Adult Competences (PIAAC). Moreover, up to five definitions are computed simultaneously: job analysis (JA), mean (Me) and modal (Mo) procedures, as well as two self-assessment measure, whether direct (DSA) and indirect (ISA).

The analyses demonstrates that the definition applied is not neutral. Comparatively, modal procedure and job analysis definition yield high estimates of mismatch, while with the mean procedure, though, the incidence of over-qualification is much lower. In addition, also are relevant the disaggregation of occupational codes, the conversion of levels of qualification into years of schooling and the phrasing in self-assessment questions. With regards to the factors affecting QM, gender, work experience and specially skills are sensitive to the definition applied, although more robust results are found with respect to national origin or size firm.

Keywords: Qualification mismatch, over-qualification, under-qualification.

JEL classification: J24.

*This research was supported by the Spanish Ministry of Science and Innovation (project CSO2010-21004, 'Occupational stratification and educational performance in Spain: job-skills match, immigration, and retirement').

1 Introduction

While there is an extensive literature measuring qualification mismatch (QM) in different ways, less attention has been paid to the discrepancy between definitions.[1] There are some notable exceptions of studies evaluating the measures of mismatch, such as Groot and Maassen Van Den Brink (2000), Hartog (2000) or Rubb (2003), and also deserve to be mention Van der Meer (2006), who evaluate the same definition with two different databases, as well as other works applying more than one definition simultaneously. With respect to the ORU earnings equation, authors like Cohn and Khan (1995), Kiker et al. (1997), Hartog (2000), Verhaest and Omey (2006) or Chiswick and Miller (2010) have studied if results are sensitive to the definition used.

The motivation for this paper is the observation of a lack of reflection on the possible impact of the definition chosen in the study of the determinants of mismatch. The aim of this work is then twofold. First, it seeks to ascertain the impact of the definition of qualification mismatch on point estimates, and second, to assess the impact on the estimations of determinants.

The rest of the paper is organized as follows. Section 2 covers a literature review, revisiting the multiple ways in which qualification mismatch has been measured. Data and methods applied in this work, including the way to compute the definitions of QM, are presented in section 3. Results of the impact of definitions on point estimates and factors determining QM are presented in subsection 4.1 and subsection 4.2, respectively. Finally, section 5 concludes.

2 Literature review

Somewhat, the study of education mismatch resembles a 'polyphonic melody' because more than one simultaneous independent 'melodies' have been played. Several definitions of over- and under-qualification have been in fact simultaneously used in literature, but none of them expresses the idea of mismatch unequivocally. There has not been, in this sense, an unique melody or monophony, and every author has played a different instrument. Rather than a advantage, the proliferation of distinct and not comparable definitions would be challenging from a methodological point of view: the reliability of the measuring instrument is called into question, because, as will be shown in section 4, the levels of mismatch widely differ depending on the measured used.

Following this musical metaphor, below some 'standards' are presented, that is, those definitions of qualification mismatch repeatedly applied, such as the job anal-

[1] Broadly speaking, *qualification mismatch* may be defined as the discrepancy between workers' qualifications and those required by their jobs. It generally occurs when workers possess higher or lower qualifications than those required by their job. More specifically, two kinds of qualification mismatch may be identified: over- and under-qualification. When workers possess *higher* qualifications than those required by their job, they are typed as over-qualified. Thus, over-qualification can be seen as an under-utilization of qualification at work. On the other hand, when they possess *lower* qualifications than those required, workers are described as under-qualified.

ysis definition, the realized matches (both the mean and modal procedure) and the self-assessment ones, whether directly or indirectly computed (subsection 2.1). Some 'variations on a theme' come after, i.e. those standards-based, but modified definitions, such as horizontal mismatch, severe mismatch or what might be called 'tailored suits' (subsection 2.2). Finally, a coda with measures of skill mismatches is presented (subsection 2.3). Taken together, this review illustrates how behind this apparently 'polyphonic chaos' hinted in the literature, an internal order exists: it will be shown that every definition expresses a part of the history of the gap between qualification and jobs.

2.1 *Standards*

Several measures have been proposed to quantify the extension of the qualification (vertical) mismatch, i.e., the discrepancy between worker's qualifications and those required at job. The more frequently used measures, that might be viewed almost as *standards*, are briefly summarized below and then some of the main advantages and disadvantages of each of them are pointed out.[2]

Job Analysis (JA)

With the so-called 'Job Analysis' measure (JA), the required qualification is determined using a systematic classification elaborated *ex ante* by experts, where tasks and knowledge for each occupation are specified in it.[3] Therefore, by comparing worker qualification (q_w) and qualification requirements in his or her occupation (q_r), it is possible establish if the worker is over-qualified ($q_w > q_r$) or under-qualified ($q_w < q_r$).[4]

One problem with this method is that classifications like DOT are difficult to develop and are not always available. Furthermore, when they exist, they quickly become obsolete because they are sensitive to technological change. The assumption of homogeneity is, notwithstanding, the main drawback of this measure: it is assumed that the required qualification is the same for all the jobs within an occupation.

Self-assessment (SA)

Available in some surveys, the 'self-assessment' (SA) definition is another way to measure qualification mismatch. Workers' assessment of their jobs is here key to establishing

[2]For a review of this measures, see Hartog (2000), Rubb (2003) or Leuven and Oosterbeek (2011). Their advantages and disadvantages have been discussed by Groot and Maassen Van Den Brink (2000), Verhaest and Omey (2006) or Quintini (2011)

[3]The *Dictionary of Occupational Titles* (DOT), supplied by the US Department of Labor, and more recently the Occupational Information Network (O*NET) are probably the best examples in America.

[4]Some authors (Battu et al. (2000), Groot and Maassen Van Den Brink (2000), among others) call this measure 'objective'. However, given that analysts may introduce a degree of subjectivity in determining qualification requirements, perhaps 'objective' is not the most appropriate term. Strictly speaking, neither does 'job analysis' seems to be the best description, because occupations –and not jobs– are what are actually analysed.

an eventual mismatch. This assessment may be made directly or indirectly. In the first case (*direct self-assessment*), workers are asked whether they consider themselves to be over- or under-qualified. In the second case (*indirect self-assessment*), however, they are asked which qualification is needed to get or to perform their jobs (q_r). Therefore, by comparing these responses with their level of qualifications (q_w), workers can be classified as over-qualified ($q_w > q_r$) or under-qualified ($q_w < q_r$).

The main advantage of the 'self-assessment' measure is that it takes the heterogeneity of jobs into account: nobody knows a job better than the own worker who performs it. However, precisely because of this subjectivity, it has been found that some workers tend to overstate the requirements of their jobs (Hartog, 2000), and their assessment is strongly determined by their job satisfaction (Verhaest and Omey, 2006). Moreover, these kinds of self-reported questions about mismatch are not always available. Even when they are, the considerable variation in phrasing across studies makes comparisons difficult. As some authors have already pointed out, some variants ask for recruitment standards (i.e. to get the job), while others emphasizes on the requirements to perform in the job (i.e. to do the job). Table 3 in Appendix illustrates with examples this wide variation in phrasing.

Realized matches (RM)

Finally, mismatch is also measured using statistical procedures, such as mean or mode. These are the so-called 'realized matches' measures (RM). In these cases, workers are considered over- or under-qualified in relative terms: their qualifications are compared with the average or modal level of qualification within occupations. The *mean* procedure was widely used after the influential paper by Verdugo and Verdugo (1989).[5] The modal procedure, proposed by Kiker et al. (1997), is quite similar, but in this case, however, the most common value within a given occupation is considered as the required qualification ($q_r = q_{Mo}$).

In general the main advantage of the 'realized matches' measures are their availability: they can always be calculated. However, one problem is that they also assume the homogeneity of qualification requirements for all jobs with the same occupational code. In addition, as mismatch is defined in relative terms, counterintuitive results can be obtained. For instance, if the average or modal level of qualification increases within a given occupation (lets say, waiters) as a consequence of educational upgrading, the required qualification is supposed to be higher, and the worker is thus less likely to be considered as over-qualified (even for graduates working as a waiters, as in this example!).

Finally, from a methodological viewpoint, both procedures (mean and mode) have significant drawbacks. On the one hand, because there is no general agreement about the way to convert qualification levels into years of schooling, and it has been found that the incidence of mismatch also depends on the schooling equivalences used to construct

[5] It is worth emphasizing, however, that Clogg and Shockey (1984) or Shockey (1989) used it before.

the schooling variables (Rubb, 2003).[6] On the other hand, in the case of the mean procedure, the arbitrariness of choosing one standard deviation as a threshold has been criticized.

2.2 *Variations on a theme*

The next three subsections discuss some alterations to the standard measures. They are the so-called horizontal mismatch, definitions focused on show the intensity of the mismatch and those that might be called 'tailored suits', only formulated for a particular study.

Horizontal mismatch: discrepancy between fields of study.

If an engineer works as an economist, with the standard measures of qualification mismatch he/she would be probably classified as 'properly qualified'. The quantity of schooling required and possessed is the same, but a mismatch based on the type of schooling would be identified, since he/she is working in a job unrelated to his/her degree field. This type of imbalance has-been called 'horizontal mismatch', and has been calculated mostly for graduates' samples.

All standards described above may be used, as variations, to explore horizontal mismatches. For instance, some authors have computed them from direct questions of self-assessment (direct-SA). To wit, with questions where workers are asked directly if they consider that another field of education would be most appropriate for the job (a case in point is Allen and Van der Velden ()2001), using *CHEERS* data); or to what extent their job are related to their degree field (like Robst (2007), with the *US National Survey of College Graduates*). Wolbers (2003), meanwhile, using the *LFS ad hoc module on school-to-work transitions* (2000), apply an objective measure, closer to the standard job analysis (JA) definition. Information on the fields coursed is available, and additionally he makes a list with occupations that match every field of education. Then, are in a non-matching job those whose field of education does not match with that considered as proper by the analyst. Finally, a sort of variation of the definition based on realized matches (RM-mode), specifically the modal procedure, have also been used to study horizontal mismatches. Nieto et al. (2013), for instance, use the *Adult Education Survey* to explore differences in the vertical and horizontal mismatch of immigrants and natives in Europe. By doing this, they assume that workers have horizontal mismatch when their field of education diverges from the mode of the native workers' field within each occupation.

Intensity of qualification mismatch: severe vs. weak over-qualification.

Any graduate in medicine working as street sweeper is classified as over-qualified with any of the standard measures. It would also be tagged as over-qualified as well if he/she

[6]This issue is explored in more deep in subsection 4.1.

was working as a nurse. However, it is easy to understand that the degree of mismatch is not the same in both cases: it is not the same to be a doctor working as a nurse that working as a street sweeper.

This is the reason why several authors have calculated variations to the standards considering also the intensity of the mismatch: whether measured by education levels or years of schooling. For example, considering levels of education, Quintini (2011) label as severe qualification mismatch" the situation in which the qualification level of workers is more than one step away (higher for over-qualified, lower for under-qualified) from the required qualification in their job on the 5-point ISCED scale. Meanwhile, García-Montalvo et al. (2003) had introduced some nuances, because they distinguished between 'weak over-qualification' (when they have up to two levels higher), 'strong over-qualification' (between two or four levels higher) or 'very strong over-qualification' (up to four levels higher). A measure derived from this, with the qualification variable expressed in years of schooling (instead of levels) is calculated by García-Montalvo et al. (2006, 2009). García-Serrano and Malo-Ocaña (1996, 1997 and 2002) also compute measure of mismatch taking into account the intensity, both measured by education levels or years of schooling.

Tailored suits.

Other measures of mismatch different from those typically used –but that arise from them- can be calculated. They are what we might call 'tailored suits' that, by definition, are hard to replicate with different datasets, but are still noteworthy. A case in point is the work by Battu et al. (2000). They use an external source -LFS, more specifically- to identify occupations with a low average education level, namely, those in which more than 50 per cent of the workers aged 20-38 are not-graduates. Then, individuals for the main survey -where the target population are graduates- are scored for over-qualification if they are in an occupation with a low average education level. Ortiz and Kucel (2008), Barone and Ortiz (2011) and Ortiz (2010), meanwhile, apply a variation of the realized match (RM) measure with another threshold. Instead of using the standard deviation criteria, what they do is to compute the 80th percentile of education within an occupation according to ISCO-88 one-digit level classification, but excluding the three first great groups (managers, professionals and technicians).[7] Then, all workers beyond that threshold are over-qualified.

2.3 *Coda:* skills beyond formal qualifications

'Education mismatch', 'qualification mismatch', 'schooling mismatch' or 'skill mismatch' are concepts that sometimes have been mistakenly used as interchangeable terms. In like manner, but referring to workers, expressions like 'over-educated', 'over-qualified', 'over-schooled", 'over-skilling' or 'skill surplus' have sometimes been indistinctly applied. But they are different concepts. The vagueness and ambiguity with

[7]The sixth group, agriculture, is also omitted in Ortiz (2010).

which these terms are been used, though, obscures the understanding of the phenomenon.[8]

Both terms –qualifications and skills–, are closely related. Indeed, qualifications have been widely used in the literature as a proxy for skills. Nevertheless, there are good reasons to not consider them as equivalent terms. First of all, because even within the same qualification level workers are very heterogeneous in skills (Chevalier, 2003). Thereupon, there may be over-qualified workers being under-skilled. Secondly, because qualifications only reflect skills learnt in formal and certified education: not all skills are acquired through education, nor education only serves to provide skills. And third, because some of the skills reflected in qualifications may be deteriorate over time (obsolescence), while others may be acquired through on-the-job learning or labour market experience (update). Certification is, therefore, the hallmark to differentiate 'skills' from 'qualifications': those knowledge or abilities used at work that can be certified are considered 'qualifications', while the others –whether innate, acquired through a formal process or by training– are 'skills'.

Skill mismatch, then, can be defined as the discrepancy between those skills possessed by workers and the skills required by their jobs. This mismatch is more difficult to measure, because the use of skills at work is not always available in surveys. Moreover, it is conceptually hard to imagine that may be an 'excess of skills', since the more skilled a worker is, the better he/she perform his/her jobs. Self-assessment measures of skill mismatch have been used in several studies.[9] Direct indicators of skills mismatches have been calculated, however, only a short time ago. Are noteworthy in this sense the *Adult Literacy and Life Skills Survey* (ALL), first conducted in 2003 and used, among others, by Dejardins and Rubenson (2011); the *International Adult Literacy Survey* (IALS), analysed by Quintini (2011); or, more recently, the *Programme for the International Assessment of Adult Competences* (PIAAC), which have been already explored by Pellizzari and Fichen (2013).

3 Data and methods

In the first place, to ascertain the impact of the definitions of qualification mismatch on point estimates three different but complementary databases are used. First, the *Labour Force Survey* (LFS) that allows analyse a period of ten years (all quarters between 2002 and 2012 are considered) and includes occupations of ISCO disaggregated up to a three-digits level. Second, the *Adult Education Survey* (AES) (waves 2007 and 2011), that includes rich information about fields of study and non formal education activities. And third, the *Programme for the International Assessment of Adult Competences* (PIAAC),

[8]For a further discussion and enlightening clarifications of these conceptual issues, see Halaby (1994), Green et al. (1999), Allen and Van der Velden (2001), Dejardins and Rubenson (2011) or Quintini (2011).

[9]See, among others, Allen and Van der Velden (2001) using *CHEERS* (1999), Green and McIntosh (2007) with the *Skills survey* (UK, 2001), Barone and Ortiz (2011) using *REFLEX* data (2005) or Quintini (2011) with the 2005 wave of the *ESWC*.

conducted in 2012, that besides includes scores on standardized tests (used as proxy for skills) and self-assessment measures or qualification mismatch. The descriptive analysis have been focused just on one country (Spain) in order to offer a better understanding of the phenomenon by isolating the possible influence of country particularities. The categories of these variables have been harmonized across the different databases used.

In the second place, in order to assess the impact of the definitions on the estimations of determinants of mismatch, several discrete choice models were run using PIAAC data, and five definitions of qualification mismatch were computed for each model [10] Estimations were focused on European countries to avoid an excessive heterogeneity of the education system and labour market institutions.[11] In the first model, only employees over 25 years old where considered, and great groups 0 and 1 in ISCO-08 (armed forces and managers, respectively) were excluded from the analysis. The second model was restricted to foreign workers, and the third one, to graduate workers. Models 1 and 2 are multinomial regression logit where the probability of being over- or under- qualified is explored. Model 3, in turn, is a binomial logit regression, where only the probability of being over-qualified is explored. The reference category of the dependent variable in the three models is, of course, being adequately qualified for the job. Covariates included in the estimations are listed in Table 4 in Appendix. Given the specific survey design of PIAAC (jackknife replicate procedure for standard error computations), estimates have been computed using replicate sampling weights, and taking into account plausible values of scores provided.

3.1 Computing definitions of qualification mismatch

In order to allow the comparisons, the five standard definitions of qualification mismatch have been created within each database, although the self-assessment measures are only available in PIAAC. The so-called "variations on a theme", as well as the "tailored suits" are not computed because of their specificity and difficulty to compared across studies.

Job Analysis (JA)

To compute the job-analysis definition of qualification mismatch, listings considering the last two international classifications of occupations (ISCO-88 and ISCO-08) have been *ad hoc* developed for the three data sets.[12] Occupations in the major groups 0 and 1 (Armed forces and Managers, respectively), however, have been excluded because it

[10]Technical details regarding the construction of qualification mismatch definitions are explained below (Subsection 3.1).

[11]Belgium, Czech Republic, Denmark, France, Ireland, Italy, Netherlands, Norway, Poland, Slovak Republic, Spain, Sweden and United Kingdom are the countries included in the estimations. Austria, Estonia, Finland and Germany were excluded because some key variables were are not available.

[12]The ISCO-88 was used to encode occupations in the first wave of the Adult Education Survey (2007), while the ISCO-08 was already apply to the second wave of AES (2011), as well as to the PIAAC survey (2012). The Labour Force Survey in Spain, meanwhile, is encoded with a National Occupation Classification (CNO-2011), coming from ISCO-88.

would be meaningless to speak of 'qualification requirements' in these jobs, given that channels to access are others than formal education. For the remaining occupations, the procedure was as follows. First, it was established the theoretically required formal level of qualification for each occupation.[13] The approach has been quite restrictive. This means that although in practice there is a considerable amount of graduates performing (say) clerk jobs, it is strongly argued that a lower level of qualifications is enough. Regarding vocational education, the distinction between intermediate and upper level was based on the public titles offered, which are directly related to the classification of occupations. Then, once the required qualification have been established, the levels considered higher than that required (surplus) and lower (deficit) were in turn specified. Finally, from these lists (not shown for reasons of space, but available upon request), the translation to the data is quite straightforward: those whose qualification is considered higher for their occupation are over-qualified, and under-qualified in the opposite case.

For the preparation of these lists, academic papers and technical reports have been considered. The skills-based approach of ISCO-08, explained in detail in ILO (2012), it is noteworthy.[14] However, given that formal education is just *one* of the dimensions taken into account by the ILO to arrange occupations into groups, it is only partially helpful in order to elaborate a list matching occupations and required levels of education, which is precisely what is needed to establish the qualification mismatch through a job analyst (JA) definition. In addition, there are situations in which formal education requirements differ between national education systems. Two occupations in particular deserve to be named: nurses and pre-primary school teachers, given that in some countries a university degree is required, while in others this is not the case. Meanwhile, SPEE (2008) includes a list with occupations linked to three levels of formal education, which a panel of experts considered as the most appropriated. The problem is that in some occupations more than one educational level was simultaneously indicated.

Realized Matches (RM)

Realized matches definitions are quite simple to compute. With regards to the mean procedure, workers have been considered over-qualified if their schooling level is more than one standard deviation above the mean qualification for their occupations ($q_{\bar{x}}$). That is to say, they are over-qualified if $q_w > (q_{\bar{x}} + \sigma_{\bar{x}})$. Conversely, they are considered

[13]Six levels of qualification have been taken into account: '**NF**' (No formal qualification, primary or lower secondary education), '**SEC**' (Upper secondary general education), '**VET-Int**' (Vocational Training – Intermediate level), '**VET-Adv**' (Vocational Training – Advanced level), '**TERC-s**' (University education: short-cycle) and '**TERC-l**' (University education: Bachelor, Master or Doctorate).

[14]Specifically, four skill levels are defined, ranging from the more elemental level (coded as '1') to the fourth level, which involves the most complex tasks. Interestingly, one of the dimensions considered is formal education as defined in ISCED-97. This means that equivalences between major groups of occupations and these skills levels are provided. The main drawback if one wants to use it as a reference in the construction of a JA measure, is that it suffers from a lack of unambiguous equivalences. This is especially problematic for the intermediate skill level 2, associated to five ISCO-08 major groups (from 4 to 8).

under-qualified if their schooling level is less than one standard deviation below the mean qualification within their occupations, i.e. $q_w < (q_{\bar{x}}$ - $\sigma_{\bar{x}})$. With regards to the modal procedure, is the most common value within a given occupation the required qualification ($q_r = q_{Mo}$). Thus, workers are classified as over-qualified ($q_w > q_{Mo}$) or under-qualified ($q_w < q_{Mo}$) in turn. By default, mean and modal procedure were computed with a 2-digits level of disaggregation in ISCO. In the case of Austria, Estonia and Finland, only a 1-digit level was available in PIAAC. In addition, 1 and 3-digits level have been also computed with the Spanish Labour Force Survey data for each cycle.

Self-Assessment (SA)

Of the data sets considered, PIAAC is the only one in which the workers are asked about of a possible qualification mismatch, both directly and indirectly. In the first case (dir-SA), the classification is immediate: they are considered over-qualified if they answer 'a lower level would be sufficient', or under-qualified if they recognize that 'a higher level would be needed'.[15] The second case (dir-SA) is not so straightforward, because one needs to compare the answers to the indirect question of mismatch to their qualifications.[16] For this, first, both questions have been expressed using the same classification, in this case with ISCED expressed in 6 groups, as in the job-analysis classification. Then, workers were classified as over-qualified if their qualifications levels are higher than what they consider the most appropriate level to get their job (that is, $q_w > q_r$) or under-qualified in the opposite case ($q_w < q_r$). Below (Table 1) this process is illustrated.

4 Results. Does it really matter the way in which QM is measured?

This section is organized around the two main aims mentioned in section 1: (i) to ascertain the impact of the definition of qualification mismatch on point estimates (subsection 4.1), and (ii) to assess the impact on the estimations of determinants (subsection 4.2)

4.1 Point estimates and effects of the definition chosen

The incidence of over- and under-qualification depends largely on the definition of mismatch used. This can be demonstrated not only with a painstaking review of studies

[15]The question is: *Thinking about whether this qualification is necessary for doing your job satisfactorily, which of the following statements would be most true? 'This level is necessary', 'A lower level would be sufficient' or 'A higher level would be needed'.*

[16]The indirect question included in PIAAC is: *'Still talking about your current job: If applying today, what would be the usual qualifications, if any, that someone would need to get this type of job?".* And the possible answers are one of the ISCED-97 levels disaggregated to 14 categories.

Table 1: **Comparison of qualification levels of workers (q_w) and those considered as usual to get their jobs (q_r).**

		Qualification levels of workers (q_w)					
		NF	SEC	VET-Int	VET-Adv	TERC-s	TERC-l
Required qualification (q_r)	NF						
	SEC				Over-qualified		
	VET-Int						
	VET-Adv						
	TERC-s						
	TERC-l			Under-qualified			

]***Qualification levels:*** **NF** = No formal qualification, primary or lower secondary education. **SEC** = Upper secondary general education. **VET-Int** = Vocational Training – Intermediate level. **VET-Adv** = Vocational Training – Advanced level. **TERC-s** = University education (short-cycle). **TERC-l** = University education (long-cycle): Bachelor, Master or Doctorate.

(see Figures 1 and 2, as well as Table 2 in Appendix), but also with a replication of several definitions with different databases (see Figure 3).[17]

Figures 1 and 2 show, for different years, 118 estimates based on 35 studies, grouped according the definition of mismatch used.[18] As can be seen, figures diverge substantially: over-qualification estimates range from an almost negligible 3% to an exorbitant 61%, and from 3% to 50% for under-qualification. Therefore, is dangerous —and almost impossible– to draw definite conclusions from them, since different time periods are covered, with diverse data and definitions. The purpose is, rather, to show the wide variety of results and to alert the need for care in the analyses.

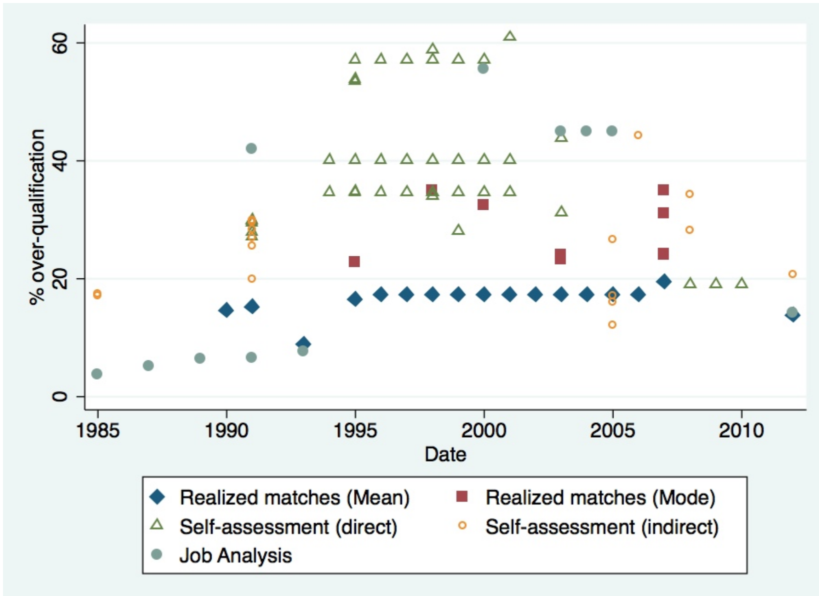
(In)consistencies and correlations between definitions.

Correlations between the qualification mismatch definitions shown in Figure 3 are rather weak. These correlations are larger between JA and Mo or between statistical measures (Mean and Mode). Still, the fact is that they are never higher than 0.60. This suggests

[17]For reviews for other countries, see Hartog (2000), Groot and Maassen Van Den Brink (2000), or, more recently, Quintini (2011).

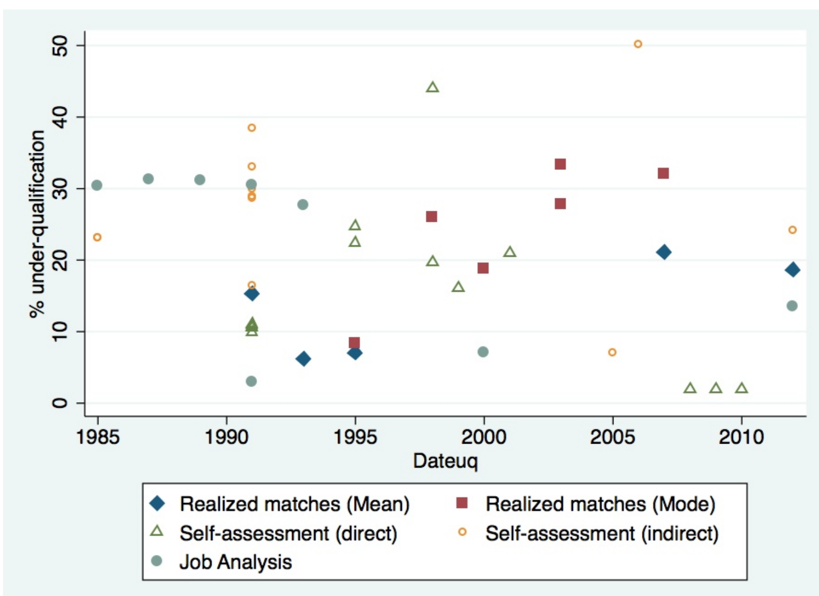
[18]For further information, consult Table 2 in the Appendix. It provides, for each study, the source, the year of data collection, the definition of mismatch used and the incidence of qualification mismatches (both over- and under-qualification). When available, the number of digits in the occupational classification used is also reported. Generally, the target population is active workers (all ages and educational level). When a different subsample is used in a study, it is advised: for example, surveys focused exclusively on young people, on graduates or including also inactive workers who had a job in the past. Additionally, other specific estimates are also specified when available. This occurs, for example, if the incidence of over- and under-qualification was calculated separately for men and women, natives and immigrants, or at different moments in the working life.

Figure 1: Incidence of over-qualification depending on the definition of mismatch. Review of studies (only Spain).



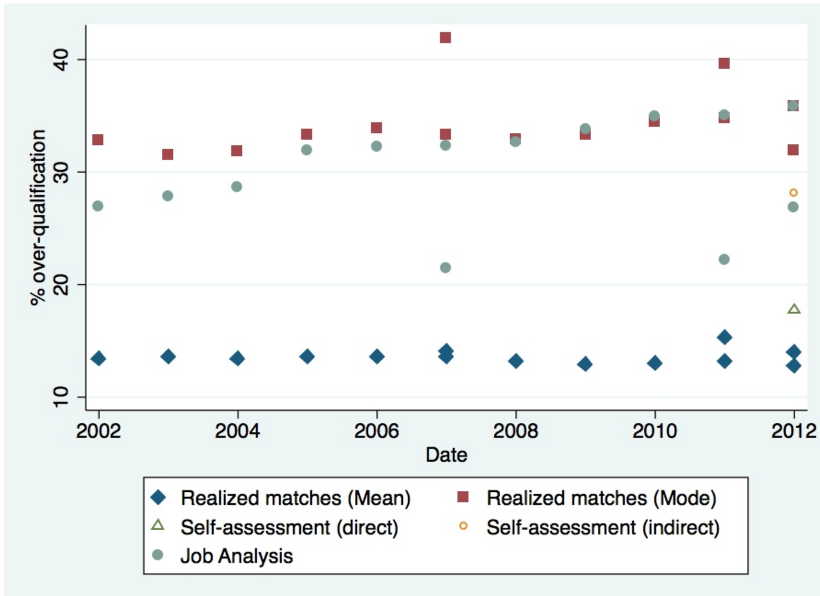
Source: Own elaboration (from Table 2 in Appendix).

Figure 2: Incidence of under-qualification depending on the definition of mismatch. Review of studies (only Spain).



Source: Own elaboration (from Table 2 in Appendix).

Figure 3: **Incidence of over-qualification depending on the definition of mismatch and the database used. (Spain, all workers).**



Source: Own elaboration (Labour Force Survey 2002-2012; Adult Education Survey, 2007 and 2011; Programme for the International Assessment of Adult Competencies, 2012).

that they are not measuring the same and, importantly, it reinforces the idea that the choice of the definition matters. In fact, as other authors already confirmed, there are appreciable inconsistencies between measurements. This means that not all workers are classified under the same category of (mis)match when more than one measure is used: one person would be classified as 'over-qualified' according with a certain definition, but the same person can be labelled as 'adequately qualified' with another. Less common is, however, to find cases tagged as 'over-qualified' are 'under-qualified' simultaneously. Furthermore, the percentage of consistency between measures is certainly low. At best, only 72% of workers are classified under the same category of mismatch using two different classifications[19].

General patterns of incidence of OQ and UQ among definitions.

Certain regularities can be identified concerning the relative levels of mismatch obtained depending on the definition used. On the one hand, there is a fairly widespread agreement that self-assessment measures (whether direct-SA or indirect-SA) yield the highest estimates of over-qualification. The reason is easy to understand: workers tend to overestimate their own knowledge, abilities and skills. In the review of studies (see Figures 1 and2) it is found that the incidence of OQ is far higher with self-assessment

[19]Detailed figures not shown for reasons of space, but available upon request

(whether direct or indirect measured).[20] However, in the case of UQ the pattern is rather the opposite, since it is not so high the percentage of people who claim to have lower qualifications than required, especially with the direct-SA definition. The percentages of OQ and UQ are also high when calculated with the self-assessment (indirect) definition included in the PIAAC data (28 and 34%, respectively), as can be seen in Figure 3. Interestingly, with the direct self-assessment definition the figure of OQ is not as high in relative terms: nearly 18% say openly that 'a lower level [of a formal qualification] would be sufficient', but the figure is even higher with Job Analysis or Modal definition of mismatch (27 and 32% respectively). By contrast, only 9.5% that 'a higher level would be needed'. Taken together, these findings suggest the existence of a psychological trend of workers to emphasize the excess of qualifications with the self-assessment definitions, but the lesser extent to recognize their lack.

On the other hand, it is expected that estimates based on a statistical measure (realized matches) yield relatively lower levels of mismatch. Within them, it is assumed that is more likely to obtain higher levels of mismatch with the measure based on the mode than with that based on the mean. This regularity has been observed, among others, by Battu et al. (2000), Leuven and Oosterbeek (2011), Nieto and Ramos (2013) or Iriondo and Pérez-Amaral (2003). The explanation follows directly by their expressions: it is more likely to be labelled adequately qualified with the mean measure (those with $\pm \sigma q_{\bar{x}}$) than with the mode (only those with $q_{M}o$). Consequently, the probability of being considered over- or under-qualified is higher with the modal procedure. This regularity has been empirically observed when both procedures are used simultaneously: systematically the mismatch is higher when calculated with the mode definition compared to the mean one (See Figure 3).

Finally, the job analyst measure (JA) depends by definition on the analyst's strictness to establish the required education for the job. With an equal target population, over-qualification will be higher if the requirements for the job established by the analyst are lower. In fact, some authors believe that JA tends to give lower estimates, while some others consider that the opposite is the case In the review of studies conducted in Spain (see Table 2 in Appendix), some authors set higher formal qualification requirements and therefore obtain a relatively low percentage of OQ and high of UQ (a case in point is García-Montalvo (1995), while others get the opposite, i.e., higher percentages of OQ and relatively low of UQ, because lower formal requirements are established (Lassibille et al. 2001, Rahona, 2007 or Ortiz and Kucel, 2008). The results of the analysis carried out with the LFS, AES and PIAAC data (see Figure 3) is closer to this second pattern, because in the construction of the JA definition only the strictly necessary formal requirements for the job had been considered (see section 3.1).

[20] OQ figures are especially high in studies using the *ECHP*, even to exceed 50% (Alba-Ramírez and Blázquez, 2003; Aguilar and García, 2008).

Some other nuances.

Non neutral choices when computing qualification mismatch measures may affect the point estimates figures. Most times decisions such as the disaggregation of occupational codes, the conversion of levels of education into years of schooling or the phrasing in self-assessment questions are driven by data availability. Notwithstanding, when more than one alternative exist, is desirable the to reflect on the impact of each decision.

- The role of the disaggregation of occupational codes.

The incidence of mismatch may differ depending on the disaggregation of the classification of occupations used.[21] Higher measurement errors are expected -and therefore less precision in estimating- if qualification requirements of jobs are establish using a one-digit level classification of occupations. The reason is the heterogeneity of occupations and jobs included under a major heading. To opt for the three-digits level disaggregation neither seems to be the best alternative. As Bauer (2002) already pointed out, mistaken estimates might be obtained if there are only few cases in each occupation.[22] Indeed, when the Mean and Modal procedures were calculated with three different levels of disaggregation in the ISCO using the LFS data, it is found that both definitions one and three digits show erratic trends and are more dependent of changes in the composition of the sample (Figures not shown for reasons of space, but available upon request). A compromise solution seems to be, then, to rely on a two-digits level classification. Indeed, Quintini (2011) found little difference in the incidence of qualification mismatch whether one-digit or two-digit level are used.

- The role of the conversion of levels of education into years of schooling.

An additional related and challenging issue is the existence of workers who studied under different education systems. This implies variations in the years of schooling compulsory to achieve a given level of education.[23] The approach taken here was to group under one unique label all equivalent qualification levels, although they correspond to different educational systems. For PIAAC data an equivalence is already provided for each

[21]This issue has been addressed by Bauer (2002), Verhaest and Omey (2010) or Quintini (2011). Table 2 in Appendix shows the digit level of disaggregation employed. It can be observed that both occupations disaggregated on a 1-digit level (Lassibille et al., 2001; Rahona, 2007; Murillo et al., 2012) as on a 2-digit level (García-Montalvo, 1995; Badillo-Amador et al., 2005; Quintini, 2011; Ghignoni and Verashchagina, 2013 or Ramos and Sanromá, 2013) has been used.

[22]This can be especially problematic when the sample size is small.

[23]In Spain, for instance, the latest adaptation of the old bachelor degrees into the new ones, as a consequence of the Bologna process, will also have to be taken into account as the number of new graduates increases in the labour force. Previous 3-year ('Diplomaturas') and 5-year degrees ('Licenciaturas') will be nominally equivalent to new degrees ('Grados'). Not to mention the peculiarities of titles such as Engineering, Medicine or Architecture, which are slightly different from the traditional 'Licenciaturas', given that 6 years was the minimum formal requirement to pursuit it.

country. For data from the LFS and AES, though, the conversion of levels of education into equivalent years of schooling is an own adaptation.

- The role of the variation in phrasing in self-assessment questions.

Measures of mismatch based on self-assessment answers are, by nature, subjective and hard to extrapolate. Not only for the manifest subjectivity of workers' answers, but specially because of the variation in phrasing.[24] It is easy to understand, therefore, that the same worker would be differently labelled in terms of mismatch depending of the survey question. It is expected, in this sense, the incidence of over-qualification to be higher with questions about the qualification requirements to do the job than with those where the requirements to get the job are asked. The explanation is quite simple: educational upgrading (but also the growing credentialism) has shift up qualification requirements to be hired (to get a job), while in practice a lower level of qualification would be enough to properly do the job. Unfortunately, with the available data cannot be compared the responses of workers to various subjective questions of mismatch. All in all, it is understandable that the figure of mismatch obtained with the PIAAC data is less than that of other databases such as ECHP, where in addition to asking for qualifications, workers were asked about 'skills or qualifications' to do a more demanding job.

4.2 Determinants of qualification mismatch and effects of the definition chosen

Gender

In a seminal paper, Frank (1978) pointed out the existence of a differential over-qualification by gender. He argued that when men are the main income earners in a household, women's labour market prospects are more restricted. Accordingly, it is expected to find a higher risk of mismatch for women. After this contribution, several studies have confirmed that the incidence of OQ is indeed higher among female workers than among males, while the opposite holds for UQ.[25] It seems reasonable to also take into account family composition (mostly marital status and/or the presence of children), the size of the local labour market, as well as the possible selection bias in labour market participation of females, that varies with age. In this sense, Quintini (2011) includes in her model interactions of gender and marital status and finds that these variables do not play a role for OQ, while apparently having young children increases the probability of UQ. García-Serrano and Malo (1997), meanwhile, found that interactions of gender with age or with marital status also play a significant role on mismatch, and Ramos and Sanromá (2013), who explore the relationship between

[24]Table 3 in the Appendix illustrates with examples this wide variation.

[25]For a review of studies, see Groot and Maassen Van Den Brink 2000.

gender and the size of the labour market, find that the spatial dimension is indeed more relevant than the differential OQ for women.

With respect to the definition used, appreciable differences have been found in the probability of mismatch by gender. For instance, Verhaest and Omey (2010) observe that women are more likely to be over-qualified if measured by job analysis definitions compared with the self-assessment measures. Battu et al. (2000), meanwhile, concluded that self-assessment definitions tend to bring lightly higher over-qualification figures among females, whilst the opposite is found with the mean procedure. And García-Montalvo (2009), using two self-assessment measures of mismatch, get inconsistent results on the probability of over-qualification for men and women. In a bivariate analysis with Spanish data, OQ is higher among men with the Modal procedure in all the databases or, when available, also with the self-assessment definitions. On the contrary, with the JA definition, qualification mismatch (both OQ and UQ) is higher among women. Altogether, these considerable qualitative discrepancies suggest that the choice of the mismatch definition is not neutral in terms of gender.

Estimates from pooled Model 1 show that the effect of gender are not robust. Coefficients are in fact not always statistically significant, and predictions by Frank (1978) are only confirmed with the modal procedure. Results are substantively equal whether with or without controls of family composition (such as living with spouse or partner or presence of children) or when an interaction term between gender and age is considered.[26].

Age and work experience

The theory of career mobility, first proposed by Sicherman and Galor (1990) states that workers may temporarily work in jobs for which they are over-qualified in order to acquire skills and knowledge that can be used in higher-level positions or jobs in their future careers. If this theoretical hypothesis holds true, it is then expected to find a negative impact of experience and age, suggesting that the probability of OQ decrease as the workers gains experience. According to previous evidence, the probability of being over-qualified is in fact higher for young workers and those with less experience in the firm, suggesting that mismatches are reduced over time.[27]. Meanwhile, García-Montalvo (2009) also find differences in the coefficients using two diverse self-assessment definitions.

In bivariate terms, qualitatively the effect of age and work experience are consistent:

[26] In other specifications without control for cohabitation status or presence of children in the household (not shown, but available upon request), nor gender is found to play a robust role in mismatch.

[27] For a review, see Groot and Maassen Van Den Brink (2000), Rubb (2003) or Leuven and Oosterbeek (2011).

Using different definitions of mismatch, the trade-off between schooling and experience is partly confirmed: more years of experience increase the likelihood of being under-qualified. However, Kiker et al. (1997) find that this is not the case when the mean procedure is used: the coefficient on the probability of being over- and under-qualified turns then, as opposed to what is usually found in the literature.

across all the databases and definitions computed, the incidence of OQ decreases with age, and UQ is generally greater as the age workers increases. Quantitative differences in the point estimates values between definitions noticeably remain, though.[28] With respect to the multivariate analysis, estimates from pooled Model 1 mostly confirm what theory of career mobility predicts regarding work experience (although not with all definitions). The effect of this variable on the probability of OQ is actually negative in four of the definitions, and positive with three of them in the case of the likelihood of UQ. This findings them seem to give support to those who argue that OQ is corrected through working lives.

Country of birth and time of residence

Foreign workers face worse prospects in the labour market compared to natives. Qualification mismatch is not an exception: most studies have found, indeed, that non-native workers are more likely to be over-qualified while, conversely, UQ levels are higher among natives. This stylized fact has been explained as a sign of discrimination or rather as a consequence of a lack of transferable human capital. Time of residence is thus an important covariate to take into account. In this sense, following the seminal work of Chiswick (1978), it is expected to find a positive assimilation of foreign workers in the labour market. The probability of OQ would then decline with length of residence in the host country. The geographic origin might also play a role on qualification mismatch given the heterogeneity of educational systems around the world. However, going beyond formal qualifications, Dumont and Monso (2007) suggest that the effect associated with the immigrant variable diminishes once literacy level (measured though scores in standardized tests) is controlled for.[29]

In a bivariate analysis with Spanish data, qualitatively most definitions and databases analysed agree with respect to the trends in mismatch depending on the country of birth: higher levels of OQ are found among workers born in other countries compared to natives. And, conversely, UQ levels are higher among natives. However, some noteworthy differences are detected according to the definition applied. For example, since 2010, the pattern of UQ is not so obvious both using the LFS and the AES (2011) and PIAAC (2012): with some definitions the incidence of UQ is higher among natives, while with others the opposite is the case. Moreover, although systematically is higher the relative differences in mismatch among native workers and those born abroad with

[28]Tenure in the firm have been considered as another covariate elsewhere, suggesting that those workers who remain for longer in the firm improve their chances of getting a better match within it (see, among others, Duncan and Hoffman, 1981 or Wolbers, 2003). However, in the model this potential variable has not been included, given the possible problem of reverse causality: those workers in a non-matching job have a strong incentive to change to another job that fits better their qualifications, making almost impossible to infer causality.

[29]With regards to horizontal mismatch, however, Nieto et al. (2013) found not significant differences between immigrants and natives once individual characteristics are controlled for. This suggests that are precisely the diverse characteristics of natives and immigrants what explain their different risk of mismatch.

the average definition with the Adult Education Survey (in 2007 as well as in 2011), the JA is the definition which reflects a greater relative differences in OQ between the two groups of workers.

The negative outcomes of non-natives in terms of qualification mismatch are robust: even controlling for other covariates, all models and definitions show that the likelihood of being over-qualified is higher for immigrants. In Model 2 (not shown), only taking into account non-native workers, partial support to the positive assimilation hypothesis is found: the time of residence in the country reduces the risk of over-qualification, although only with the mean procedure and the indirect self-assessment definition. The impact of the region of origin, though, only is significant with some definitions of mismatch.

Fields of study

Studies that focus only on graduate workers find that the risk of mismatch varies across fields of study. In general, the probability of OQ seems to be higher for graduates in humanities or social sciences, while graduates in more technical or scientific fields face a lower risk of mismatch. These differences can be explained through at least three different but complementary mechanisms. First, by distinguishing between general and specific skills, it might be argued that the risk of OQ will be higher for those graduates from degrees that emphasize general skills (such as humanities), while the opposite holds for degrees providing very specific skills (such as engineerings). Second, the relative supply and demand of graduates across fields may explain their different risk of mismatch: an excess of supply (and thus a lack of demand) may imply a downgrading for those "surplus" graduates. And third, it is important to stress that the choice of degree is not exogenous. Rather, individuals self-select themselves into fields of study. Controlling for skills (measured though scores in standardized tests) are thus essential.[30]

Participation in non-formal education

A trade-off between formal and informal education has been identified in literature: a lack of formal qualification may be compensated by skills obtained through on-the-job training or experience. Less clear is, however, the effect of participation in non-formal education activities (on-the-job training) on the probability of mismatch. Lassibille et al. (2001) find that participation in non-formal education programs has no significant impact on the job match. Nevertheless, by interacting formal and non-formal education they observe that the likelihood of being OQ decreases for those who posses a low level of education but are enrolled in a non-formal education program. When the relation between these two variables is considered in the other direction (the effect of mismatch on participation in non-formal education activities) results are not more conclusive: Beneito et al. (2000) consider that the impact of OQ on the probability of participation

[30] For a discussion of these competing explanations, see Robst (2007) , Ortiz and Kucel (2008) , Barone and Ortiz (2011) , Quintini (2011) or Ghignoni and Verashchagina (2013).

in specific training is negative and positive in the case of UQ; Wolbers (2003) detect that mismatched workers (measured as horizontal mismatch) participate less in additional training, while Nieto and Ramos (2013) found that over-qualified workers participate more in non-formal education activities.

Estimates from Model 1 show that participation in non-formal education activities have no significant effects on the probability of OQ, although it increases the likelihood of being UQ in two of the specifications shown (those using the job analysis or the indirect-self assessment definition). This suggests that formal on non-formal education are partially substitutes for under-qualified workers.

Job characteristics

Beyond personal characteristics of workers, job characteristics are essential to explain the likelihood of mismatch. Besides, it is reasonable to think that the distinctive features of the labour market in each country (regarding for instance to the firm size structure or the stability in the employment) may partly account the different incidence of qualification mismatch across countries. The size of the firm is of course crucial: the risk of qualification mismatch (either OQ, UQ or horizontal mismatch) is higher in small firms compared to the medium or big ones (Wolbers, 2003; Quintini, 2011; Ghignoni and Verashchagina (2013).

With respect to the type of contract, several studies have observed that temporal workers face higher likelihood of mismatch compared to those with a more permanent contract. Quintini (2011), however, found no significant differences in QM between permanent workers and workers on fixed-term or temporary contracts, and Ortiz (2010) punctuate that the effect of job security on over-qualification is not uniform across countries.

As can be seen in Model 1, the negative effect of the size firm on the likelihood of being over-qualified is only statistically significant with three definitions of qualification mismatch. Paradoxically, when controls of type of contract and hours worked per week are included, it is found that with some of the definitions having an indefinite contract or a full-time job increase, rather than reduce, the likelihood of being over-qualified. It cannot be argued, then, that OQ is mostly associated to temporary or part-time jobs.

Social background

Although they have not been very studied because of a lack of data, variables of family background, such as parent's occupation or their educational attainment have an impact on the probability of qualification mismatch. (Barone and Ortiz, 2011) found that this was robust across the different definitions of OQ they use. However, as can be seen in Model 1, a higher educational attainment of the parents (at least one parent attained tertiary education) only has a significant effect reducing the likelihood of over-qualification with the job analysis and the indirect-self assessment definitions.

Skills

When ability or skills are measured, several studies find that, controlling for other variables, scores in standarized tests to are negative correlated to over-qualification. This means that those workers less skilled (i.e., with lower scores in standarized tests) are more likely to be over-qualified (Green et al., 1999; Chevalier, 2003; Green and McIntosh, 2007; Quintini, 2011). Interestlingly, estimates from Model 1 show that this is only the case of the direct self-assessment definition, suggesting that those workers who consider themselves as over-qualified are actually less skilled. When other definitions of mismatch are used, though, over-qualification and higher skills are positive associated.

5 Conclusion

Given the divergences obtained by different measures, two general conclusions are drawn. First, and obvious, the need for cautious in interpreting empirical results concerning the extent of mismatch. And second, the desirability of measure qualification mismatch in more than one way if reliable results about over- and/or under-qualification want to be obtained. No general agreement exists regarding to which one is the best measure of mismatch. To a large degree, it may be argued, in fact, that there are not any 'better' than another: each of them contains a different perspective and probably the key is to justify the choice, that should be not only dictated by data availability. All in all, there are good theoretical and empirical reasons to prefer the job analysis definition or the self-assessment ones. The statistical measures -specially the definition based on the mean procedure-, however, seem to be the more arbitrary and less recommendable, although can be advisable when the evolution of the mismatch is analysed over time with the same database. The figures of qualification mismatch are also affected by other apparently less important aspects such as the level of disaggregation of occupational codes, the conversion of levels of qualification into years of schooling or the phrasing in self-assessment questions.

On the other hand, the analysis of the determinants of qualification mismatch allows confirm some of the hypotheses suggested in literature. It is worth noting, though, non negligible differences in the effects depending on the definition of qualification mismatch used. The possible presence of differential over-qualification by gender is only partially confirmed with the modal procedure. The negative outcomes of non-natives in terms of qualification mismatch, though, are robust, although only partial support to the positive assimilation hypothesis of non-natives is found with some definitions of mismatch. Regarding age and work experience, although not with all definitions, predictions of theory of career mobility are mostly confirmed. OQ seems to be in fact a rather temporary phenomenon, because those workers with more work experience face less likelihood of being over-qualified while the opposite holds for UQ. The existence of certain substitution between formal on non-formal education is also partially confirmed, given that the likelihood of being under-qualified increases, at least with two of the definitions of qualification mismatch, for workers who participate in non-formal education activities. With regards to job characteristics, a higher firm size reduces, as expected, the likelihood of being over-qualified, while having an indefinite contract or a full-time job paradoxically increase, rather than reduce, that likelihood. All in all, these results are not completely robust with the all five definitions. Finally, the inclusion of direct measures of skills, shows an interesting effect of the definitions. Specifically, it was found that skills only have a negative effect on the likelihood of over-qualification with the direct self-assessment definition. This suggests that those workers who consider themselves as over-qualified through direct questions are actually less skilled. On the contrary, with other definitions over-qualification and higher skills are positive associated, giving more support to theories that stress on the existence of signalling effects of qualifications in the labour market.

6 References

M. I. Aguilar and D. García. Desajuste educativo y salarios en España: nueva evidencia con datos de panel. Estadística española, 50(168): 393–426, 2008.

A. Alba-Ramírez. Mismatch in the Spanish labor market: Overeducation? Journal of Human Resources, 259–278, 1993.

A. Alba-Ramírez and M. Blázquez. Types of job match, overeducation and labour mobility in Spain. In Overeducation in Europe. Current issues in theory and policy, pages 65–92. Edward Elgar: Cheltenham, 2003.

J. Allen and R. Van der Velden. Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. Oxford Economic Papers, 53(3): 434–452, 2001.

L. Badillo-Amador, A. García-Sánchez, and L. E. Vila. Mismatches in the Spanish labor market: Education vs. competence match. International Advances in Economic Research, 11(1):93–109, 2005.

C. Barone and L. Ortiz. Overeducation among European University Graduates: a comparative analysis of its incidence and the importance of higher education differentiation. Higher Education, 61(3): 325–337, 2011.

H. Battu, C. R. Belfield, and P. J. Sloane. How well can we measure graduate over-education and its effects? National Institute Economic Review, 171(1): 82–93, 2000.

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

P. Beneito, J. Ferri, M.-L. Moltó, and E. Uriel. Over/undereducation and specific training in Spain: complementary or substitute components of human capital? In Education and Training in a Knowledge-Based Economy, pages 191–214. Macmillan Press, 2006.

S. Budría and A. I. Moro-Egido. Education, educational mismatch, and wage inequality: Evidence for Spain. Economics of Education Review, 27(3):332–341, 2008.

A. Chevalier. Measuring over-education. Economica, 70(279): 509–531, 2003.

B. R. Chiswick. The effect of americanization on the earnings of foreign-born men. The Journal of Political Economy, pages 897–921, 1978.

B. R. Chiswick and P. W. Miller. Does the choice of reference levels of education matter in the oru earnings equation? Economics of Education Review, 29(6): 1076–1085, 2010.

C. C. Clogg and J. W. Shockey. Mismatch between occupation and schooling: A prevalence measure, recent trends and demographic analysis. Demography, 21(2): 235–257, 1984.

E. Cohn and S. P. Khan. The wage effects of overschooling revisited. Labour Economics, 2(1): 67–76, 1995.

R. Dejardins and K. Rubenson. An analysis of skill mismatch using direct measures of skills. Education Working Papers, No. 63, OECD Publishing, 2011.

G. J. Duncan and S. D. Hoffman. The incidence and wage effects of overeducation. Economics of Education Review, 1(1): 75–86, 1982.

J.-C. Dumont and O. Monso. Matching educational background and employment: A challenge for immigrants in host countries. In International Migration Outlook. OECD.

C. Fernández and C. Ortega. Labor market assimilation of immigrants in spain: employment at the expense of bad job-matches? Spanish Economic Review, 10(2): 83–107, 2008.

M. Fernández-Enguita. Educación, formación y empleo. EUDEMA, 1992.

R. H. Frank. Why women earn less: the theory and estimation of differential overqualification. The American Economic Review, 360–373, 1978.

I. García-Espejo and M. Ibáñez. Educational–skill matches and labour achievements among graduates in Spain. European Sociological Review, 22(2):141–156, 2006.

J. García-Montalvo. Empleo y sobrecualificación: el caso español. FEDEA Working Paper, 95-20, 1995.

J. García-Montalvo. La dinámica de la sobrecualificación en España. In Análisis de la sobrecualificación y la flexibilidad laboral. Bancaja, IVIE, 2009.

J. García-Montalvo, J. M. Peiró, and A. Soro Bonmatí. Capital humano. observatorio de la inserción laboral de los jóvenes: 1996-2002. 2003.

J. García-Montalvo, A. S. Bonmatí, and J. M. P. Silla. Los jóvenes y el mercado del trabajo en la España urbana: resultados del observatorio de inserción laboral 2005. 2006.

C. García-Serrano and M. Á. Malo-Ocaña. Educational mismatch and internal labour markets: is there any relationship? Working papers of the ESRC Research Centre on Micro-social Change. Paper 96-16. Colchester: University of Essex, 1996.

C. García-Serrano and M. Á. Malo-Ocaña. ¿Es diferente el desajuste educativo de las mujeres? Información Comercial Española, ICE: Revista de economía, (760): 117–128, 1997.

C. García-Serrano and M. Á. Malo-Ocaña. Over-education, under- education and expected promotions. Revista Universitaria de Ciencias del Trabajo, (5): 145—-160, 2002.

E. Ghignoni and A. Verashchagina. Educational qualifications mismatch in Europe. is it demand or supply driven. Journal of Comparative Economics, 2013.

F. Green and S. McIntosh. Is there a genuine under-utilization of skills amongst the over-qualified? Applied Economics, 39(4): 427–439, 2007.

F. Green, S. McIntosh, and A. Vignoles. Overeducation and skills. Clarifying the concepts. Centre for Economic Performance, LSE. Discussion paper, 435, 1999.

W. Groot and H. Maassen Van Den Brink. Overeducation in the labor market: a meta-analysis. Economics of Education Review, 19(2): 149–158, 2000.

C. N. Halaby. Overeducation and skill mismatch. Sociology of Education, pages 47–59, 1994.

J. Hartog. Over-education and earnings: where are we, where should we go? Economics of Education Review, 19(2): 131–147, 2000.

ILO. International Standard Classification of Occupations (ISCO-08). structure, group definitions and correspondence tables. International Labour Office, Geneva, 2012.

I. Iriondo and T. Pérez-Amaral. The effect of educational mismatch on wages using European panel data. Working Paper No. 700. Queen Mary University of London, 2013.

B. F. Kiker, M. C. Santos, and M. M. De Oliveira. Overeducation and undereducation: evidence for Portugal. Economics of Education Review, 16(2): 111–125, 1997.

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

S. Nieto, A. Matano, and R. Ramos. Skill mismatches in the EU: Immigrants vs. natives. IZA Discussion Paper, 2013.

S. Nieto and R. Ramos. Non-formal education, overeducation and wages. Revista de Economia Aplicada, XXI(61): 5–28, 2013.

M. Madrigal-Bajo. Una revisión de los métodos de medición del desajuste educativo: ventajas e inconvenientes. In Hacienda pública y convergencia europea: X Encuentro de Economía Pública, Santa Cruz de Tenerife 2003, page 75. Universidad de La Laguna, 2003.

J. S. Martínez. Sobrecualificación de los titulados universitarios y movilidad social. In PIAAC. Programa Internacional para la Evaluación de las Compentencias de la población adulta. 2013. Volumen II. Análisis secundario. 2013.

I.P. Murillo, M. Rahona-López, and M.-d.-M. Salinas-Jiménez. Effects of educational mismatch on private returns to education: An analysis of the Spanish case (1995–2006). Journal of Policy Modeling, (34): 646–659, 2012.

L. Ortiz. Not the right job, but a secure one. Over-education and temporary employment in France, Italy and Spain. Work, Employment & Society, 24(1): 47–64, 2010.

L. Ortiz and A. Kucel. Do fields of study matter for over-education? the cases of Spain and Germany. International Journal of Comparative Sociology, 49(4-5): 305–327, 2008.

M. Pellizzari and A. Fichen. A new measure of skills mismatch: Theory and evidence from the Survey of Adult Skills (PIAAC). OECD Social, Employment and Migration Working Papers, No. 153, OECD Publishing, 2013.

G. Quintini. Right for the job: Over-qualified or Under-skilled? OECD Publishing, 2011.

M. Rahona. Desajuste educativo en el primer empleo significativo. In Investigaciones de Economía de la Educación, 375–386. 2007.

M. Rahona-López and C. Pérez-Esparrells. Educational attainment and educational mismatch in the first employment in Spain. ISRN Education, 2013.

R. Ramos and E. Sanromá. Overeducation and local labour markets in Spain. Tijdschrift voor Economische en Sociale Geografie, 104(3).

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

S. Rubb. Overeducation in the labor market: A comment and re-analysis of a meta-analysis. Economics of Education Review, 22(6): 621–629, 2003.

N. Sánchez and S. McGuinness. Decomposing the impacts of overeducation and overskilling on earnings and job satisfaction: An analysis using reflex data. ESRI working paper, No. 393, 2011.

J. W. Shockey. Overeducation and earnings: A structural approach to differential attainment in the us labor force (1970-1982). American Sociological Review, pages 856–864, 1989.

N. Sicherman and O. Galor. A theory of career mobility. Journal of Political Economy, pages 169–192, 1990.

SPEE. Prospectiva del empleo cualificado en España. Servicio Público de Empleo Estatal. Observatorio Ocupacional. 2008.

P. H. Van der Meer. The validity of two education requirement measures. Economics of Education Review, 25(2): 211–219, 2006.

R. Van der Velden and M. Van Smoorenburg. The measurement of overeducation and undereducation: self-report vs. job-analyst method. Research Centre for Education and the Labour Market, Faculty of Economics and Business Administration, Maastricht University, 1997.

R. R. Verdugo and N. T. Verdugo. The impact of surplus schooling on earnings: Some additional findings. Journal of Human Resources, XXIV(4).

D. Verhaest and E. Omey. Discriminating between alternative measures of overeducation. Applied Economics, 38(18):2113–2120, 2006.

D. Verhaest and E. Omey. The determinants of overeducation: different measures, different outcomes? International Journal of Manpower, 31(6): 608–625, 2010.

D. Verhaest and E. Omey. The impact of overeducation and its measurement. Social Indicators Research, 77(3): 419–448, 2006.

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

Table 2: INCIDENCE OF OVER- AND UNDER-QUALIFICATION IN SPAIN DEPENDING ON THE DEFINITION OF MISMATCH. REVIEW OF STUDIES.

Study	Data	Date	Measure	OQ (%)	UQ (%)	Digit level
Fernández-Enguita (1992)	LWCS	1985	SA (ind)	17.4	23.1	n.a.
Alba-Ramírez (1993)	LWCS	1985	SA (ind)	17.1	23.1	n.a.
García-Montalvo (1995)	LFS	1985	JA	3.7	30.4	2
		1987	JA	5.1	31.2	2
		1989	JA	6.3	31.1	2
		1991	JA	6.6	30.5	2
		1993	JA	7.7	27.6	2
		1993	RM (mean)	8.9	6.2	2
García-Serrano and Malo (1996)	SCCB	1991	SA (ind)	28.4	30	n.a.
			SA (dir)	29.4	11	n.a.
	SCCB***	1991	SA (ind)	29.6	28.7	n.a.
			SA (dir)	29.9	10.6	n.a.
García-Serrano and Malo (1996)	SCCB***	1991	SA (ind)	26.9	33	n.a.
			SA (dir)	27.1	9.9	n.a.
García-Serrano and Malo (1997)	SCCB***	1991	SA (ind)	29.9	28.9	n.a.
			SA (dir)	29.7	10.5	n.a.
Beneito et al. (2000)	SCCB	1991	SA (ind)	25.6	16.5	n.a.
			SA (dir)	27.9	10.9	n.a.
			RM (mean)	15.2	15.3	
Lassibille et al. (2001)	SDS	1991	JA	42	3	1
García-Serrano and Malo (2002)	SCCB	1991	SA (ind)	20	38.5	n.a.
Alba-Ramírez and Blanco (2003)	ECHP	1995	SA (dir)	53.8	n.a.	n.a.
Madrigal-Bajo (2003)	ECHP	1995	SA (dir)	34.7	22.3	n.a.
			RM (mean)	16.5	7	
			RM (mode)	22.8	8.4	
Badillo-Amador et al. (2005)	ECHP	1998	SA (dir)	34	44	n.a.
			RM (mode)	35	26	2
Fernández and Ortega (2006)	LFS	1996-2006	RM (mean)	17.3a	n.a.	
				39.0b		
García-Espejo and Ibáñez (2006)	UOS*	2003	SA (dir)	43.8c	n.a.	n.a.
				31.2d		
Rahona (2007)	LFS**	2000	JA	55.5c	7.1c	1
Aguilar and García (2008)	ECHP	1995	SA (dir)	53.4	24.6	n.a.
		1998	SA (dir)	58.8	19.7	n.a.
		2001	SA (dir)	60.9	20.9	n.a.
Budría and Moro-Egido (2008)	ECHP	1994-2001	SA (dir)	34.6e	n.a.	n.a.
				40.1f		
Galasi (2008)	ESS*	2006	SA (ind)	50.2	44.3	n.a.
Ortiz and Kucel (2008)	LFS	2003-05	JA	45	n.a.	
García-Montalvo et al. (2009)	IVIE**	2008	SA (dir)	34.3c	n.a.	n.a.
				28.2g		

Incidence of over- and under-qualification in Spain depending on the definition of mismatch. Review of studies. (Continuation)

Study	Data	Date	Measure	OQ (%)	UQ (%)	Digit level
Barone and Ortiz (2011)	REFLEX*	2005	SA (ind)	17.1	n.a.	n.a.
			SA (ind)	12.2h	n.a.	n.a.
Quintini (2011)	ESWC	2005	RM (mode)	32	31	2
Sánchez and McGuinness (2011)	REFLEX*	2005	SA (ind)	16	7.1	n.a.
Murillo et al. (2012)	SES	1995	RM (mode)	35.3	20.8	1
		2002	RM (mode)	31.9	25.6	1
		2006	RM (mode)	37.2	23	1
Ghignoni and Verashchagina (2013)	LFS*	2003	RM (mode)	23.2e	33.3e	2
				24.0f	27.8f	
Nieto et al. (2013)	AES	2007	RM (mode)	24.0a	n.a.	
				31.0i		
				35.0j		
Nieto and Ramos (2013)	AES	2007	RM (mean)	19.5	21.1	
		2007	RM (mode)	24.2	32	
Rahona and Pérez-Esparrells (2013)	LFS**	2000	RM (mode)	32.4c	18.8c	2
Ramos and Sanromá (2013)	BFS	1990/91	RM (mean)	14.6	n.a.	2
Martínez (2013)	PIAAC	2012	JA	14.1	13.5	2
			SA (ind)	20.7	24.2	n.a.
			RM (mean)	13.8	18.6	2
Van der Velden and Van Smoorenburg (2013)	REFLEX*	2005	SA (ind)	45.0k	n.a.	n.a.
				26.7l		

Data: AES = Adult Education Survey. BFS = Budget Family Survey. CHEERS = Careers after Higher Education: a European Research Study. ECHP = European Community Household Panel. ESS = European Social Survey. ESWC = European Survey of Working Conditions. EU-SILC = European Union Statistics on Income and Living Conditions. IVIE = Observatorio de Inserción Laboral de los Jóvenes (IVIE-Bancaja). LFS = Labour Force Survey. LWCS = Living and Working Conditions Survey. PIAAC = Programme for the International Assessment of Adult Competencies. REFLEX = The Flexible Professional in the Knowledge Society. SCCB = Structure, Consciousness and Class Biography Survey. SDS = Sociodemographic Survey. SES = Structure of Earnings Survey. UOS = University of Oviedo Survey.

Target population: * Only graduates. ** Only young people. *** Including also unemployed people who worked the year before.

Measures: JA = Job analysis. SA (dir) = Self-assessment (direct). SA (ind) = Self-assessment (indirect). RM (mean) = Realized matches (mean). RM (mode) = Realized matches (mode).

Specific estimates: **a -** Natives **b -** Immigrants **c -** First job **d -** Current job **e -** Men **f -** Women **g -** Last job **h -** Refined measure **i -** Immigrantes (EU countries) **j -** Immigrantes (other countries) **k -** Six months after graduation **l -** Five years after graduation

n.a. = Not applicable

Table 3: **Variation in phrasing across surveys with questions of self-assessment mismatch.**

Panel A: Direct Self-Assessment definitions.

Question	**Dataset**
Do you consider you have skills or qualifications to do a more demanding job than you have now?; Have you had formal training or education that has given you skills needed for your present type of work?	ECHP
Which of the following alternatives would best describe your skills in your own work? “My duties correspond well with my present skills”, “I need further training to cope well with my duties”, “I have the skills to cope with more demanding duties”	ESWC
Do you think that the job is/was appropriate to your qualifications and/or experience? "Reasonably appropriate", "Higher that my qualifications", "Lower than my qualifications".	IVIE
Do you consider that the job you do is correct according to the qualification you have? "It is the correct", "It is lower than my qualification", "It is above my qualification", "I would need a different qualification to which I have".	LWCS
Thinking about whether this qualification is necessary for doing your job satisfactorily, which of the following statements would be most true? “This level is necessary”, “A lower level would be sufficient” or “A higher level would be needed”.	PIAAC
Would you say that your qualifications are/were more than sufficient, sufficient or insufficient for your current/ last job?	SCCB

Panel B: Indirect Self-Assessment definitions.

I) Requirements to *get* the job (hiring criterion).

Question	Dataset
If someone was applying nowadays for the job you do now, would they need any education of vocational schooling beyond compulsory education?	EES
If you had to advise someone to occupy your job, what level of qualification would you advise to pursuit?	IVIE
Still talking about your current job: If applying today, what would be the usual qualifications, if any, that someone would need to get this type of job? [answer categories: ISCED-97 categories]	PIAAC
How much formal education is required to get a job like yours?	PSID

II) Requirements to *do* the job (performance criterion).

Question	Dataset
What kind of education does a person need in order to perform your job?	LWCS
What type of education do you feel was most appropriate for your work? "PhD", "Other postgraduate qualification", "Master", "Bachelor" or "Lower than higher education".	REFLEX
What level of education do you think is currently the most suitable for the work you do/did?	SCCB

Data: **ECHP** = European Community Household Panel. **ESS** = European Social Survey. **ESWC** = European Survey of Working Conditions. **IVIE** = Observatorio de Inserción Laboral de los Jóvenes (IVIE-Bancaja). **LWCS** = Living and Working Conditions Survey. **PIAAC** = Programme for the International Assessment of Adult Competencies. **PSID** = Panel Study of Income Dynamics. **REFLEX** = The Flexible Professional in the Knowledge Society. **SCCB** = Structure, Consciousness and Class Biography Survey.

Table 4: Covariates included in the estimations.

Covariates and categories	Models		
	(1)	**(2)**	**(3)**
10.5**Gender** (Male = 1, Female=0).	X	X	X
Living with spouse or partner (With partner = 1, without partner = 0).	X		
Children (With children = 1, without children = 0).	X		
Work experience (in years).	X	X	X
Age (in years).	X	X	X
Country of birth (Non native = 1, native = 0).	X		
Participated in non-formal education in 12 months preceding survey (1 = Yes, 0 = No)	X	X	X
Firm size (1 = 1 to 10 people, 2 = 11 to 50 people, 3 = More than 50 people).	X	X	X
Type of contract (1 = indefinite, 0 = temporary, fix term, appreticeship or no contract).	X	X	X
Hours worked per week (1 = more than 35 hours, 0 = less that 35 hours).	X		X
Highest of mother or father's level of education (1 = At least one parent has attained tertiary education, 0 = Neither parent has attained tertiary education).	X		X
Graduate (1 = higher education, 0 = without higher education).	X	X	
Numeracy scale scores (10 plausible values, ranging from 0 to 500).	X	X	X
Years in country (1 = Less than 5 years in country, 2= Between 5 and 10 years in country, 3 = More than 10 years in country) [only for non-natives].		X	
Region of birth (1 = North America and Western Europe, 2 = Arab States, South and West Asia, 3 = Latin America and the Caribbean, 4 = Sub-Saharan Africa, 5 = East Asia and the Pacific (poorer countries), 6 = Central Asia, 7 = East Asia and the Pacific (richer countries), 8 = Central and Eastern Europe) [only for non-natives].		X	
Field of study (1 = General programmes, 2 = Teacher training and education science, 3 = Humanities, languages and arts, 4 = Social sciences, business and law, 5 = Science, mathematics and computing, 6 = Engineering, manufacturing and construction, 7 = Agriculture and veterinary, 8 = Health and welfare, 9 = Services).			X
Interactions			
Gender x Age.	X		
Graduate x Numeracy scores.	X	X	

Models: 1 = All employees; 2 = Only non-natives workers; 3 = Only graduate workers.

Table 5: **Multinomial logit estimates of the probability of mismatch with different definitions. Pooled models (Model 1. All employees).**

Panel A. Probability of over-qualification.

	Y_{JA}	Y_{Me}	Y_{Mo}	Y_{DSA}	Y_{ISA}
Gender (Male)	0.189	-0.0699	-0.5478*	0.4026*	-0.4032
Living with spouse or partner	-0.1008	0.0074	-0.1464*	-0.0435	-0.113
With children	0.0358	-0.0895	-0.0214	-0.065	-0.0855
Work experience	-0.0262***	-0.0556***	-0.0453***	0.0059	-0.0138**
Age	-0.0062	0.0386***	0.0032	-0.0091	-0.0005
Gender x Age	-0.0037	0.0054	0.0137*	-0.0099*	0.0083
Non-native	0.3561***	0.9472***	0.4490***	0.2171*	0.4813***
Non-formal education	-0.1787**	-0.1463	0.1088	-0.1118*	-0.2068**
Firm size					
- 11 to 50 workers	-0.4054***	-0.5502***	-0.1249	-0.0595	-0.2345***
- More than 50 workers	-0.3400***	-0.2256**	0.065	-0.0192	-0.2080***
Indefinite contract	0.0951	-0.2190*	0.4189***	-0.0339	-0.1618*
Full time	0.3935***	0.7468***	0.2880***	-0.0806	-0.0933
One parent with tertiary education	-0.3607***	0.059	0.0303	-0.1111	-0.2029**
Skills (numeracy scores)	0.0055***	-0.0024	0.0161***	-0.0028**	0.0038***
Constant	0.9515***	-4.8978***	-0.3892*	-0.8203***	0.1142
N	21,164	21,164	21,164	21,164	21,164

Panel B. Probability of under-qualification.

		Y_{JA}	Y_{Me}	Y_{Mo}	Y_{DSA}
Y_{ISA}					
Gender (Male)	0.0844	0.7932**	.7932**	0.3877	0.3206
Living with spouse or partner	0.0399	-0.0061	-0.0061	-0.0128	-0.0061
With children	0.0454	0.0819	0.0819	-0.049	0.1793
Work experience	0.0103	0.0287***	.0287***	0.0031	0.0384***
Age	0.0072	0.0180**	.0180**	0.011	0.009
Non-native	-0.1679	-0.0044	-0.0044	-0.0695	-0.234
Non-formal education	0.4183***	0.0023	0.0023	0.1175	0.3783***
Firm size					
- 11 to 50 workers	0.0329	0.0565	0.0565	-0.0108	-0.0778
- More than 50 workers	0.1886*	-0.0267	-0.0267	0.0378	0.0715
Indefinite contract	0.0057	-0.125	-0.125	-0.049	0.1574
Full time	0.0922	0.0303	0.0303	-0.087	0.2429**
One parent with tertiary education	0.0372	0.0412	0.0412	0.1452	-0.1348
Skills (numeracy scores)	1.8598***	-0.8601	-0.8601	1.9028***	0.1139
Constant	0.9281***	-1.0348	-1.0348	-1.0746***	-0.6661**
N	21,164	21,164	21,164	21,164	21,164

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Models include country fixed effects and controls for level of education and it interaction term with scores. They are run for each plausible value in the numeracy score and replicate weights.

Mismatch definitions used as dependent variables for each model: Y_{JA} = Job analysis; Y_{Me} = Mean procedure; Y_{Mo} = Mode definition; Y_{DSA} = Direct self-assessment question; Y_{ISA} = Indirect self-assessment question.