IZA Journal of Labor Economics

ORIGINAL ARTICLE

Open Access

CrossMark

A new measure of skill mismatch: theory and evidence from PIAAC

Michele Pellizzari [iD][*] and Anne Fichen

*Correspondence:
michele.pellizzari@unige.ch
[1]Geneva School of Economics and
Management, University of Geneva,
Uni-Mail, 40 Bd. du Pont d'Arve,
CH-1211 Geneva 4, Geneva,
Switzerland

Abstract

This paper proposes a new measure of skill mismatch to be applied to the recent OECD Survey of Adult Skills (PIAAC). The measure is derived from a formal theory and combines information about skill proficiency, self-reported mismatch and skill use. The theoretical foundations underling this measure allow identifying minimum and maximum skill requirements for each occupation and to classify workers into three groups: the well-matched, the under-skilled and the over-skilled. The availability of skill use data further permits the computation of the degree of under- and over-usage of skills in the economy. The empirical analysis is carried out using the first round of the PIAAC data, allowing comparisons across skill domains, labour market statuses and countries.

JEL Classification: J0, J20, J24

Keywords: Mismatch, Skills

1   Introduction

A large number of studies investigate the nature and consequences of mismatch, generally defined as some sort of discrepancy between the characteristics of employed workers and the requirements of the jobs that they occupy (Quintini 2011a). For example, several papers compare the formal education qualifications held by employed workers with the requirements of their jobs, commonly finding large numbers of workers being more qualified than necessary (Chevalier 2003; Dolton and Vignoles 2000; Groot and Maassen van den Brink 2000; Quintini 2011b; Rubb 2003; Sicherman 1991; Sloane et al. 1999). This finding can be rationalized by arguing, for example, that over-qualified workers may not have benefited from formal education as much as they could have and that their actual competencies are less advanced than those one would normally expect them to possess based on their formal educational qualifications. At the same time, workers who are found to be under-qualified for their jobs may have acquired the necessary skills to perform satisfactorily outside formal schooling, through experience, on-the-job learning and adult education (Green and McIntosh 2007; Chevalier and Lindley 2009). Hence, it is interesting to contrast qualification mismatch with skill mismatch, namely the discrepancy between the skills possessed by a workers and those required to perform his/her job (Allen and van der Velden 2001; Desjardins and Rubenson 2011). Over-skilled workers are those who are more skilled than required by their jobs, the opposite for under-skilled workers.

Springer Open

Unfortunately, measuring skill mismatch is particularly challenging, mostly due to the lack of direct information about workers' skills and job requirements. A large literature has now emerged proposing various methodologies to measure mismatch in skills (Allen and van der Velden 2001; Green and McIntosh 2007; Quintini 2011a; Flisi et al. 2014; Desjardins and Rubenson 2011; CEDEFOP 2010; van der Velden and Bijlsma 2016) and the comparison and assessment of these many methodologies is the subject of a sometimes heated debate, centred around the definition of the skill requirements of jobs or the appropriateness of direct comparisons between skill endowments and skill use (Levels et al. 2013). Our view of this debate is that it suffers from a serious lack of theory. The typical paper in this area addresses the measurement problem without really providing a formal definition of the underlying theoretical notion that is meant to be measured, thus making it very difficult to compare the many proposed indicators. In most cases, they simply measure different underlying concepts.

In this paper, we develop a simple theory that guides our use of the data from *OECD Survey of Adult Skills* (*PIAAC*) to construct a new indicator of mismatch. Our model is closely anchored to the specific data that we use and cannot be seen as a general theory of mismatch. Nevertheless, the approach to measurement of skill mismatch that we derive can be easily generalized to any other dataset sharing the same key features.

The *OECD Survey of Adult Skills* (*PIAAC*) includes a rich battery of questions on skill use at work and direct indicators of workers' skill proficiency derived from a purposely designed assessment exercise. The survey covers a large number of countries and guarantees a high degree of comparability across borders thanks to the harmonized sampling procedures and the common questionnaire (OECD 2013a).[1]

In summary, the proposed methodology uses the simple theoretical framework to overcome the fundamental problem of defining the skill requirements of jobs from a survey of workers. Specifically, for each available skill domain and each occupation, minimum and maximum requirements are defined as the minimum and the maximum proficiency of self-reported well-matched workers.[2] Within this framework, workers are classified as well-matched in a skill domain if their proficiency score in that domain is between the minimum and maximum requirements of her occupation. Workers are over-skilled or under-skilled in a domain if their score is above the maximum or below the minimum requirement.

Three additional features of the approach described in this paper are worth mentioning. First, alternative measures of the minimum and maximum skill requirements can be produced by comparing the extremes of the distributions of assessed competencies for the under- and over-skilled and the well-matched. Such comparison allows assessing the relevance of misreporting in the estimated requirements. Second, exploiting the rich background questionnaire of the PIAAC survey, it is possible to compare the utilization of skills in the workplace by similarly proficient workers who are well-matched or mismatched in their jobs, thus constructing indicators of the degree of under- and over-utilization of skills associated with mismatch. Finally, our approach allows designing simple reassignment algorithms that, far from solving the problem of optimally allocating workers to jobs, can be used to compare the distribution of skill mismatch across alternative allocations and thus measure their relative efficiency.

In addition, we also develop a general procedure to construct standard errors for estimates of skill-mismatch derived from surveys like PIAAC, where the sampling frames can

differ substantially across countries and where the test scores are derived from imputation models. Such a procedure can be computationally intense, but it is very general and it can be easily applied to any non-linear estimator constructed with the PIAAC data. The procedure is described in the Appendix.

The results of our analysis show that on average, across the entire survey, approximately 75% of dependent employees are well-matched in the literacy domain, about 9% are under-skilled and 16% are over-skilled. The overlap between literacy and numeracy mismatch is substantial: 90% of the workers who are well-matched in literacy are also well-matched in numeracy. Men are more likely to be over-skilled than women, whereas gender differences in under-skilling are minor. Tertiary graduates are less likely to be under-skilled than less educated workers, and they are also more likely to be over-skilled. Foreign workers are more than twice more likely to be under-skilled than natives and substantially less likely to be over-skilled. Differences emerge also when looking across age groups.

The rest of the paper is organized as follows. Section 2 briefly summarises the relevant literature. Section 3 lays out the proposed methodology to measure skill mismatch, starting from its theoretical underpinnings and including a discussion of the empirical implementation and of the impact of misreporting. Section 4 briefly describes the PIAAC data and provides some descriptive statistics. Section 5 reports comparable estimates of skill mismatch and skill under- and over-utilization across the countries covered in PIAAC, for the entire population and for various subgroups. Section 6 presents an extension of the approach to construct measures of the over- and under-utilization of workers' skills. Section 7 compares the distribution of skill mismatch observed in the data with those resulting from a variety of reassignment procedures. Finally, Section 8 concludes by highlighting the importance of this analysis for both academic research and policy making.

2 Measuring mismatch: a brief review of the literature

The term mismatch is often used to refer to rather different concepts in economics, thus creating a certain confusion in an area that is attracting more and more policy attention and that, therefore, would benefit a lot from more accurate definitions and measurement.

It is useful to distinguish two broad notions of mismatch, a macro and a micro one. In this paper, we focus on the latter, but to avoid confusion, it is important to mention that it also exists a macro concept of mismatch that is common to a rich strand of studies (Jovanovic 1979; Farber 1999; Robin et al. 2009; Sattinger 1993). In very general terms, in models with heterogeneous jobs and workers, aggregate mismatch is defined as the existence of an allocation of workers to jobs that could improve the realized equilibrium in terms of either employment levels or output. For example, vacancies and jobseekers could be heterogeneous in their locations and mismatch would be present when reallocating them across locations could improve the efficiency of matching (Shimer 2007; Şahin et al. 2012). The same definition could be applied to other (or multiple) dimensions of heterogeneity, such as workers' skills and jobs' requirements. A somewhat dated but still very valid review of models in this area is provided by (Sattinger 1993), who labels them assignment models. Regardless of the nature of the heterogeneity, the aggregate notion of mismatch is a feature of the joint distribution of workers' and jobs' characteristics and, as such, it is an intrinsically macro concept. In this perspective, it is impossible to say whether a single job-worker pair is a mismatch in isolation from the others.

The micro notion of mismatch is very different, as it really refers to each single pair of workers and jobs. Unfortunately, the theoretical foundations of such micro concept are much less clear than for its macro analogue. The entire literature on qualification and skill mismatch, which clearly refers to the micro notion, is exclusively empirical and various measurements have been proposed, but in the absence of a formal definition, it is extremely difficult to compare them and assess their advantages and disadvantages.

In very general terms, skill (or qualification) mismatch is constructed by comparing the skills (or qualifications) of an employed worker with the skill (or qualification) requirements of her job (hence, the non-employed and the vacant jobs are completely disregarded). Then, any given job-worker pair can be classified as a good match if the skills (or qualifications) of the employee are compatible with the requirements of the job. If the worker is more skilled (or qualified) than required, she is classified as over-skilled (or over-qualified) and under-skilled (or under-qualified) in the opposite case.

This measurement exercise is usually carried out using data collected from surveys of workers, so that direct information on the demand side is lacking and the job requirements need to be inferred. Various approached have been proposed to address this problem.

Regarding qualification mismatch, many surveys now include questions on the educational qualifications required by the employer for the job occupied by the respondent. The question may ask about the current requirements or those at the time when the person was hired (or both). This is a reasonable approach but, given that skills are acquired (or lost) also outside formal schooling, under-qualified workers may have acquired the necessary skills to carry out their jobs through experience or training. Similarly, over-qualified workers may have failed to acquire skills in school or may have lost them over time.

Figure 1 illustrates some of the problems associated with measures of mismatch based on educational qualifications. The figure reports the distributions of numeracy skills—as measured in PIAAC (see Section 4 for more information about the data)—for two groups of graduates, namely those employed in jobs requiring a graduated degree (the matched)
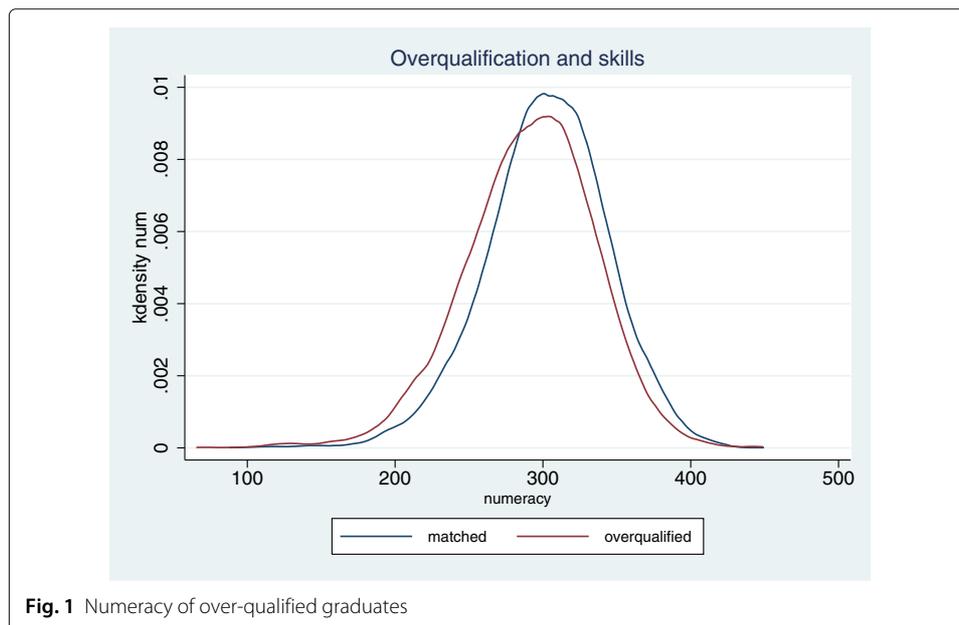

Fig. 1 Numeracy of over-qualified graduates

and those employed in jobs that do not require a graduate degree (the over-qualified).[3] The distribution of the over-qualified is clearly shifted to the left, indicating that these workers have lower numeracy scores than the well-matched (results for literacy are very similar). One possible interpretation of this result is that the reason why some graduates end up in jobs that do not require a graduate degree is that their skills are not exactly those that one would expect from someone who has attended college.

There are many possible explanations for this phenomenon. For example, for some people, the investment in tertiary education might not have been particularly successful or they could have been particularly unlucky and found jobs that did not contribute to maintaining and developing their competencies. There could even be an issue of reverse causality, as graduates employed in non-graduate jobs may see their skills deteriorate rapidly. Whatever the reasons underlying the result in Fig. 1, it is clear that an indicator of mismatch based on direct measures of skills would provide a much more precise description of the phenomenon.

For these reasons, skill mismatch is commonly regarded as a more informative indicator and several studies measure it using data with direct information on workers' skill proficiency. A variety of techniques to identify the skill requirements of the jobs can be found in the literature. One approach makes use of information from surveys asking employed workers whether they have the skills to do a more demanding job than the one they currently do or whether they feel the need of additional training to carry out their job tasks satisfactorily (Allen and van der Velden 2001; Green and McIntosh 2007). Unfortunately, answers to such questions are likely subject to various forms of misreporting, the most obvious being people's overconfidence.

Alternative approaches can be implemented when data on actual skill proficiency and skill usage are available, as in a number of datasets like PIAAC, IALS, TIMMs, PIRLS, ALL and a number of national surveys.[4] For example, using these data, one can compare individual proficiency with the average or the median in the occupation and classify as over-(under-)skilled those workers whose skills are significantly (usually one or two standard deviations) above (below) the centrality measure (Quintini 2011a; Flisi et al. 2014; Montt 2016).

When information on both proficiency and skill use is available, the two can also be compared directly, thus considering over-skilled those workers who do not make full use of their competencies on the job (Desjardins and Rubenson 2011; CEDEFOP 2010). Such an alternative approach is also subject to a number of serious problems. First of all, it implicitly assumes that skill use, which is either self-reported by the worker or derived from occupational titles, can be interpreted as a measure of job requirements, whereas it rather is the outcome of both the matching process and endogenous effort choices. Second, proficiency and use are very different theoretical concepts, and they can hardly be represented along the same metrics. In fact, they are derived from structurally different pieces of information: indicators of skill use normally exploit survey questions about the frequency (and/or the importance) with which specific tasks are carried out in a certain job, whereas skill proficiency is usually measured through cognitive tests.[5]

The methodology proposed in this paper is meant to address these difficult issues, and it rests on a very simple theoretical framework that allows us to formally define mismatch and to provide guidance about its empirical implementation. Obviously, our new indicator also suffers from various important limitations that will be discussed at length in the

next section. For example, it does still use self-reported information by the workers, but it does so in a way that reduces the potential distortions induced by overconfidence or by misinterpretations of the need for training. Our indicator does not use the median skill in an occupation to define job requirements; however, we still use some moments of the distribution of skills within occupations to define them. Our indicator does not require making direct comparisons between measured skill proficiency and skill use, but we do need to impose strong assumptions about the process of skill deployment. Overall, we believe that our indicator improves on the existing ones in many dimensions, but we do acknowledge that it is also subject to a number of important limitations.

3 Deriving the *OECD measure of skill mismatch*

The micro version of skill mismatch considered in this paper is a feature of the single job-worker pair, and it measures whether the skills possessed by the worker are adequate to carry out the tasks required by the job. A worker whose skills are below the level required by the job is classified as under-skilled, a worker whose skills are above those required by the job is classified as over-skilled.

The key difficulty in formalizing the notion of skill mismatch concerns the identification of the job requirements, as most of the times, the data used for this type of analysis are collected through surveys of workers and do not contain direct information on the structure of the production process.

In this section, we develop a simple theoretical framework that is helpful to define job requirements more formally and to spell out explicitly the assumptions imposed on the data to estimate them. One crucial feature of the theory is the treatment of skill use as an endogenous choice of the worker, similar to the choice of effort in standard principal-agent models. By explicitly modelling the choice to deploy skills, our model provides guidance not only for the measurement of skill mismatch but also for the interpretation of the questions regarding the use of skills at work. We see this as an important contribution because, as we discuss more in details in Section 6, it allows constructing meaningful indicators of the degree of skill under-utilization or over-utilization that can be associated with over- and under-skilling. In the absence of some theoretical guidance about skill deployment, it would be very difficult to link empirical measures of skill endowment and skill use.

It is also worth emphasizing that the theoretical framework described in this section serves the simple purpose of providing guidance to the measurement of skill mismatch with the empirical variables available in PIAAC (see Section 4 for a description of the data). Hence, it is very limited in two dimensions. First, it does not aim at formalising an explanation for the existence of mismatch as an equilibrium outcome. A direct implication of this first limitation is that the model assumes an existing allocation of workers to jobs and discusses how the degree of mismatch in such allocation could be measured. The model does not attempt to explain why such an allocation might be observed. In this sense, our theoretical exercise is very different from the so-called assignment models that instead focus specifically on the process by which workers and jobs are matched to one another (Sattinger 1993). Of course, there is a connection between our theory and the assignment models because the efficiency of the assignment process determines the degree of mismatch in the resulting allocation of workers to jobs. Hence, one can view our exercise as complementary to (some)

assignment models, as we provide an approach to measuring their efficiency with some real data.[6]

The second limitation is that our model is specifically designed to be implemented with the PIAAC data and it cannot be seen as a general theory of mismatch measurement. It should however be noticed also that our methodology can presumably be applied to any dataset where direct indicators of skills are available, together with the more common information on employment status and occupations. There are now many datasets in which this type of information is available, such as IALS and ALL, the predecessors of PIAAC, but also TIMMS, PIRLS, and a number of national skill surveys (e.g. UK Employer Skill Survey).

Despite these limitations, we believe that our theory still constitutes a nice contribution to the literature, at a minimum because it allows making explicit the assumptions underlying the proposed measure of skill mismatch. Other indicators of skill mismatch that have been used in the literature are obviously also based on a number of assumptions, but these are rarely made explicit and are often more restrictive than the ones discussed here. For example, the assumption that jobs are homogeneous within occupations or that the production function is kinked are common to virtually all studies.

3.1 Theoretical foundations

For presentational ease, the model in this section rests on a number of simplifying assumptions, many of which can be relaxed without affecting the qualitative implications of the theory in a major way (see Section 3.5).

Building blocks. Consider an economy with heterogeneous workers and heterogeneous jobs. Workers, indexed by i, differ in their endowment of skills, labelled η_i, and they endogenously decide how much skills to deploy in their jobs. For simplicity, η_i is assumed to be a simple uni-dimensional skill, and Section 3.5 discusses how this framework can be extended to multiple skills.[7]

Deploying skills is costless within the limit of one's endowment, and it is subject to a constant marginal cost for any skill level beyond one's endowment, as in Fig. 2. In other words, workers are allowed to deploy a level of skills that goes beyond their endowments provided they pay a utility cost. This is necessary in order to rationalize the existence of under-skilled workers in the economy.

Fig. 2 The cost of deploying skills
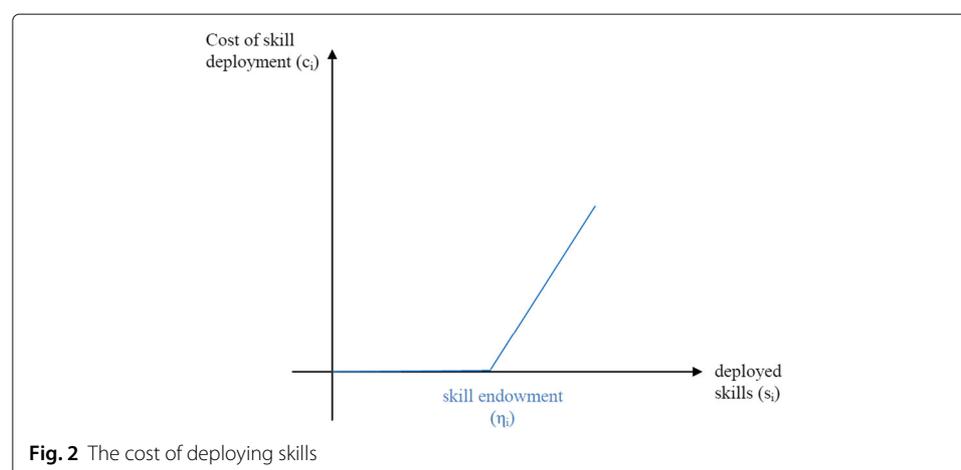
Jobs are defined as production functions, with skills being the only input. Each job employs one worker and is independent of other jobs. Different jobs have different production functions, which are characterized by three key features: (i) local linearity, (ii) fixed operational costs and (iii) discontinuously declining marginal productivity.

More specifically, assume that output y_{ij} of job j filled with worker i is a function of the amount of skills that the worker endogenously chooses to deploy on the job, s_i. Further, assume that there are fixed costs k_j to operate the job and that the marginal product of deployed skills is locally constant and decreases above a certain threshold. For simplicity, we will assume that the marginal product of skills is equal to zero beyond such threshold. Under this set of assumptions, the production function for a generic job looks as in Fig. 3.

The combination of the fixed costs and the discontinuously declining marginal product generates two critical values in the distributions of skills that lead to a very natural definition of skill mismatch. Workers with skill endowments below min_j are under-skilled, workers with skill endowments between min_j and max_j are well-matched and workers with skill endowments above max_j are over-skilled.

We do not allow firms to change their production technologies. In particular, they cannot adapt the technological characteristics of the job to the skill composition of available workers nor to the skills of the specific workers they are matched with. Of course, if such adjustment could take place frictionlessly and instantaneously, no mismatch would be observed in equilibrium. More reasonably, it is plausible to assume that some frictions exist preventing immediate and costless technological adaptation. In this model, we take this assumption to the extreme and impose that the parameters of the production function are fixed. As a consequence, the skill mismatch that we measure should be interpreted as a short-run phenomenon that could disappear over time if employers adjust the requirements of their jobs to the skills of their employees.

Workers are assigned to jobs according to some assignment mechanism that we do not model and, conditional on the characteristics of their jobs, they choose

Fig. 3 The production function
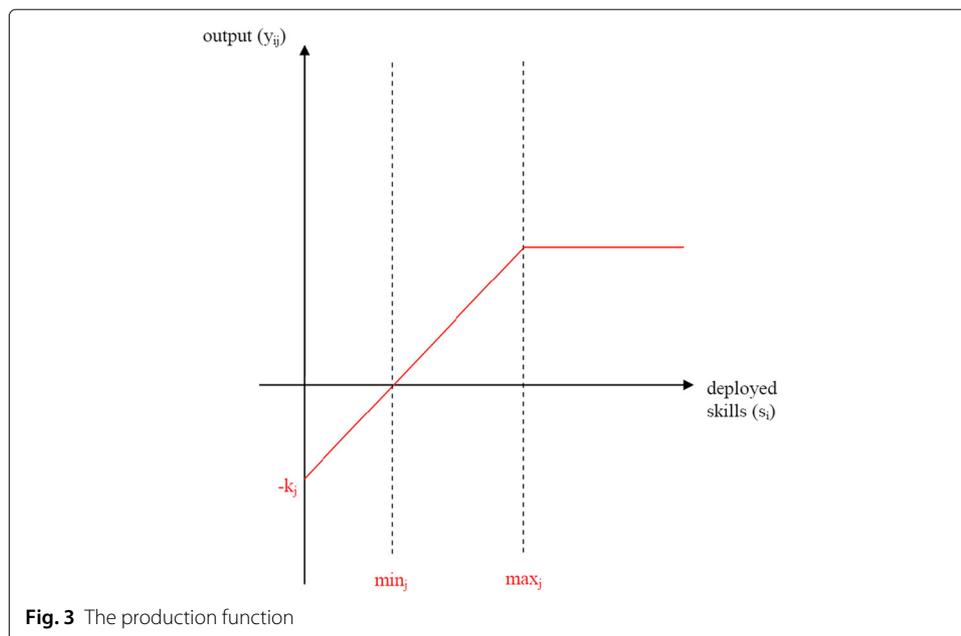
how much of their skills to deploy in order to maximize the following utility function:

$$U_{ij} = w_{ij} - 1(y_{ij} < 0)F - c_i(s_i) \qquad (1)$$

where w_{ij} is the wage worker i paid in job j, F is a utility cost associated with producing negative output (e.g. the cost of being fired and suffering a spell of unemployment) and $c_i(s_i)$ is the cost of deploying skills (Fig. 2):[8]

$$c_i = \begin{cases} 0 & \text{if } s \leq \eta_i \\ \delta s_i & \text{if } s > \eta_i \end{cases} \qquad (2)$$

with $\delta \geq 0$.

Assume wages are proportional to productivity:[9]

$$w_{ij} = \gamma_i y_i j \qquad (3)$$

where for simplicity, γ_i is allowed to vary only across workers and output is defined as[10]

$$y_{ij} = \begin{cases} \beta_j s_i - k_j & \text{if } s_i \leq min_j \\ \beta_j max_j - k_j & \text{if } s_i > max_j \end{cases} \qquad (4)$$

with $\beta_j \geq 0$ and $k_j \geq 0$ for all j.

Optimal skill deployment. Consider the following three cases.

1. Worker i is a good skill match with job j, i.e. $min_j \leq \eta_i \leq max_j$. Given the above assumptions, workers in this condition would obviously find it optimal to deploy their entire endowment of skills on the job, $s_i^* = \eta_i$.
2. Worker i is under-skilled for job j, i.e. $\eta_i < min_j$. Assuming that F is large enough to make the decision to deploy skills below min_j always suboptimal, under-skilled workers choose to deploy the minimum level of skills that allows them not to incur in the cost F: $s_i^* = min_j$.
3. Worker i is over-skilled for job j, i.e. $\eta_i > max_j$. Workers in this condition are indifferent between any level of skill deployment in the interval $[max_j, \eta_i]$.

It is now possible to look more formally at the meaning of skill mismatch. In order to do so, the optimal skill deployment of over- and under-skilled workers should be compared to the counterfactual of their being well-matched. Importantly, such comparison should be independent of other matches. In other words, the counterfactual should be viewed as a move of the mismatched worker to a previously vacant or even non-existent job or, equivalently, as a transformation of the production function of the job held by the mismatched worker. The alternative counterfactual, whereby the mismatched worker takes a job previously held by someone else, requires considering the effect of such a transition on the latter worker, thus making it impossible to define skill mismatch as a feature of the job-worker pair and bringing it nearer to the macro notion of mismatch.

In the simple theory spelled out in this section, jobs are characterized by three parameters: the operational costs (k_j), the returns to deployed skills (β_j) and the maximum skill level (max_j).[11] Hence, in order to become well-matched, any mismatched worker needs to move to a job with a different combination of these three parameters.

Consider the over-skilled first. In order to be well-matched, they need to find a job h such that $max_h > max_j$ (j indicating their current jobs), where they would deploy more skills, as their optimal skill deployment increases from max_j to η_i. Unless the new job is

also characterized by lower returns to skills ($\beta_h < \beta_j$), such a transition would also result into higher output.

As regards the under-skilled, in order to become well-matched, they need to be in a job h characterized either by lower operational costs ($k_h < k_j$) or by higher returns to skills ($\beta_h > \beta_j$) or both. In any event, where they are well-matched, they would deploy less skills but output would be unambiguously higher.

Hence, based on the definitions above, over- and under-skilled workers are mismatched in the sense that their skills could be more productively used if the structural features of their jobs were different and such that they would be well-matched.

3.2 Empirical implementation

Having access to data that include observable measures of the skills possessed by employed workers, as in PIAAC, it is possible to identify and estimate the parameters min_j and max_j for each job, where jobs are defined as occupations or, depending on the size and quality of the data, as the combination of occupation and industry classes. In other words, all the jobs in the same class are assumed to be homogeneous, i.e. using the same production technology.

The identification of job requirements rests on two questions that are asked to employed respondents in PIAAC but that are also common to other surveys, sometimes with variations (Allen and van der Velden 2001; Mavromaras et al. 2007; Green and McIntosh 2007). The first question asks about whether one feels to have the skills to do a more demanding job. The exact phrasing is the following: *"Do you feel that you have the skills to cope with more demanding duties than those you are required to perform in your current job?"*. The second question is about the need of training and reads as follows: *"Do you feel that you need further training in order to cope well with your present duties?"*. We assume that respondents who answer negatively to both questions are neither over-skilled nor under-skilled; hence, they are well-matched. According to our simple theory, well-matched workers deploy their entire endowment of skills and we can then estimate min_j and max_j as the minimum and the maximum of their tested skills, respectively:

- $\widehat{min}_j$ = minimum level of assessed skills of workers who neither feel they could do a more demanding job nor feel the need of further training
- $\widehat{max}_j$ = maximum level of assessed skills of workers who neither feel they could do a more demanding job nor feel the need of further training

For the moment, the assumption that selecting workers who answer negatively to both questions correctly identifies good matches, i.e. job-worker pairs such that $\eta_i \in [min_j, max_j]$, is maintained. The obvious concerns about misreporting in such questions are the object of the next section (Section 3.3).

Now, it is possible to classify under-skilled workers as those whose skill endowments are below min_j and, similarly, over-skilled workers are those whose skill endowments are above max_j. In Section 5, we produce empirical estimates of such categorization.

Next, an optimal level of skill use can be defined for every worker in the economy as the skill use observed for workers with a similar level of skill endowments who are well-matched. Such a comparison is informative about the amount of skills that are under- or over-utilized. We perform this analysis on the PIAAC data in Section 6.

Finally, it is possible to use this theoretical framework to assess the efficiency of the observed allocation of workers to jobs, the efficiency of the assignment mechanism. In Section 7, we compare the observed degree of skill mismatch with what would be observed in an alternative allocation generated by an assignment procedure designed to minimize over- and under-skilling.

3.3 Misreporting

The use of self-reported information about one's ability to perform one's current job and one's need for training may question the validity of the estimates of the job requirements. Despite not being immune to measurement error and misreporting, the methodology described in Section 3.2 allows the derivation of alternative estimators of the job requirements and, by comparing such alternative n, it also allows producing evidence that is informative about the extent of the problem.

Specifically, in addition to the estimators described in Section 3.2, min_j could alternatively be estimated as the maximum skill endowment of workers who report feeling the need of further training and not feeling able to do a more demanding job. Similarly, max_j could be estimated as the minimum skill endowment of workers who report feeling able to do a more demanding job and not feeling the need for further training.

It is useful to define these alternative estimators as follows:

- $\widetilde{min}_j$ = maximum skill endowment of workers feeling the need of further training
- $\widetilde{max}_j$ = minimum skill endowment of workers feeling able to do more demanding jobs

Figure 4 visually summarizes the intuition behind these estimators, each of which is affected differently by the most cumbersome sources of mismeasurement, namely overconfidence and the generalized need for training.

Fig. 4 Alternative estimators of job requirements
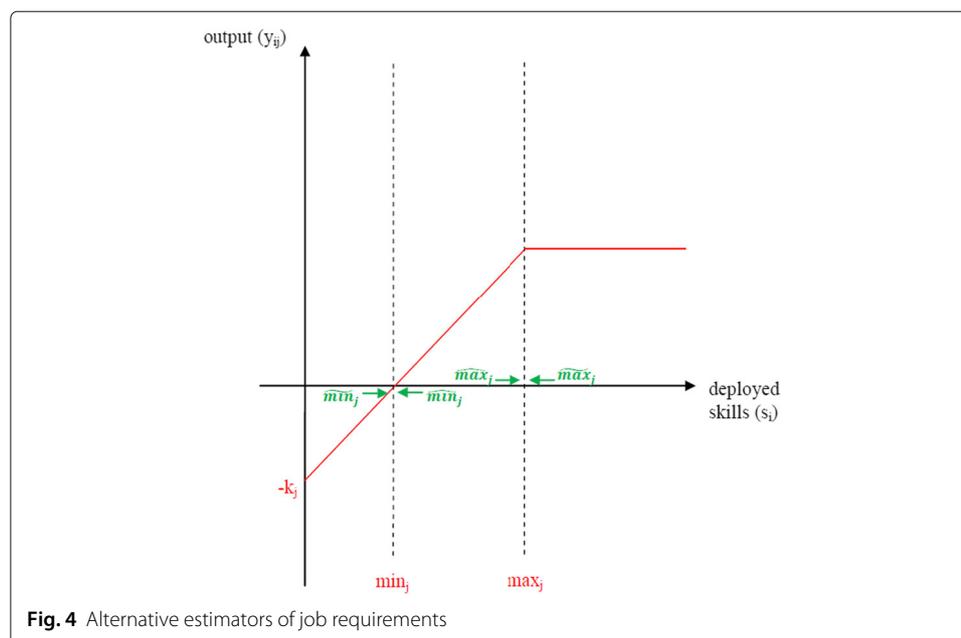
Overconfident respondents might report being capable of doing more demanding jobs even when they are indeed well-matched or even under-skilled in their current employment. Interestingly, overconfidence is much more likely to bias $\widetilde{max}_j$ than $\widehat{max}_j$. In fact, a single (truly) well-matched worker who is overconfident and consequently reports being over-skilled crucially changes $\widetilde{max}_j$. On the other hand, only if the most skilled worker among the (truly) well-matched is overconfident $\widehat{max}_j$ changes. In practice, one can look at the magnitude of the difference between $\widehat{max}_j$ and $\widetilde{max}_j$ to assess the importance of overconfidence in the data.

Overconfidence is less of an issue for the estimation of min_j, as the question about having the skills to cope with a more demanding job is not used for this purpose. It does however remain possible that some overconfident workers who are truly under-skilled end up being classified as well-matched because they believe that their skills are appropriate for their jobs and thus report both not feeling able to do a more demanding job and not needing training. Hence, our methodology is not completely immune from mismeasurement induced by overconfidence. Nevertheless, we believe that it is a limited problem given that workers answering positively to the specific question about being able to do more demanding jobs are not used by our procedure.

Beside overconfidence, another source of misreporting might affect the respondents' answers to the question about the need for training, which is the basis for estimating min_j. Such a question specifically asks whether the respondent feels the need of additional training to "cope well" with her present duties and people may attach different interpretations to the notion of "coping well," given that the quality of how tasks are performed can vary substantially. Hence, some people might answer that they do feel the need of additional training, under the assumption that with more training, they could carry out their current tasks better (e.g. more rapidly, less expensively) even though they already do so at an acceptable level or, in the terminology of our simple theory, they already deploy skills above min_j.

It seems reasonable to argue that the bias in $\widehat{min}_j$ is likely to be smaller than in $\widetilde{min}_j$. This is because any (truly) well-matched or over-skilled worker who misinterprets the question and reports needing training would crucially affect $\widetilde{min}_j$. On the other hand, $\widehat{min}_j$ is biased only if the least skilled among the (truly) well-matched reports being in need of training.

An additional, although less worrisome, source of mismeasurement is the heterogeneity of jobs within occupations (or occupation-industry cells). In fact, despite the theoretical assumption that all jobs are identical within occupations, some heterogeneity necessarily exists in practice. Hence, in order to reduce its implications on the definition of the job requirements, it is useful to consider some bottom and top percentiles of the within-job distributions of workers' skills rather than the actual minimum and the maximum. In Section 5, the 95th and 5th percentiles of the within-occupation distribution of skill endowments among workers who neither feel the need for further training nor feel capable of doing more demanding jobs are used as estimators of max_j and min_j, respectively. In Section 5.2, we show that our results are robust to the choice of the percentile.

3.4 Skill-specific mismatch

So far, the skill endowment of workers, η_i, has been assumed to be a simple unidimensional variable. However, one major advantage of PIAAC is the availability of measures of proficiency in three important skill domains, namely numeracy, literacy and

problem-solving. Hence, it allows producing measures of mismatch that are specific to each skill, as workers could use all their skills in some domains and be over-skilled or under-skilled along other dimensions.

In fact, the methodological framework presented in this section can be readily reinterpreted in the context of multiple skills. Simply allow η_i to be a vector of several skills and, similarly, also the job requirements, min_j and max_j will be multidimensional vectors. Then, assume workers who report being over/under-skilled do so whenever any of their skills is above/below the corresponding minimum/maximum requirement, even if they are well-matched with regard to all the other skill dimensions. Under this additional assumption, minimum and maximum requirements for each skill type can still be estimated as discussed in the section above and workers can be classified as under- or over-skilled by each skill domain.

Of course, the survey cannot cover the entire set of skills that are needed at work so that some individuals may still be mismatched along some dimensions that are not observed in the data.

3.5 Extensions

The theoretical framework described above clearly rests on a number of simplifying assumptions and, although some of them are crucial for the purpose of constructing measures of skill mismatch that can be implemented empirically, some serve the more modest purpose of simplifying the model.

For example, in order to make sense of the notions of minimum and maximum requirements, it is crucial to define production functions with either kinks or negative intercepts or both. Similarly, in order to conceptualize separately the endowment of skills and their deployment, one needs to introduce some costs of deploying one's endowment into the job.

However, the sharp assumptions about the return to skills dropping all the way to zero above max_j and the cost of skill deployment being exactly zero up to one's endowment can be relaxed. Specifically, the production and cost functions could very well look as in Fig. 5 without compromising any of the implications that we derived from the model.

Fig. 5 Alternative production and cost functions
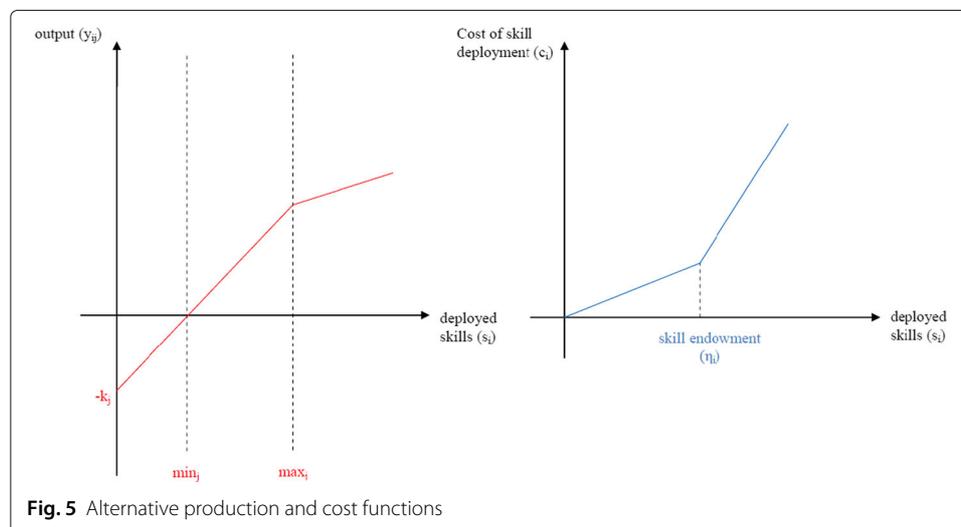
Provided the marginal cost of skill deployment increases above η_i and the returns to skills decline beyond max_j, nothing would change substantially in our framework. Only one additional assumption would be needed regarding the relative ratio of the returns to skills above and below max_j and the marginal costs above and below η_i to avoid unreasonable and uninteresting equilibria in which, for example, the under-skilled find it optimal to deploy skills above max_j.

Other assumptions that are worth mentioning here are the lack of complementarity of workers in the production process, the random assignment of workers to jobs and the limited variation in the sharing parameter γ which is constrained to be constant within workers across jobs.

Regarding complementarities, it is important to note that skill complementarity can be very easily incorporated in the model of Section 3.1. The linearity of output with respect to each specific skill is what makes the identification of job requirements particularly simple. However, it is still possible to allow the production function in Fig. 3 to shift vertically in reaction to changing inputs of other skills. The model would still require some additional assumptions to avoid the minimum and maximum requirements for each skills to be affected by changes in the inputs of the others, a situation that would make the very definition of requirements extremely unclear. Hence, skill complementarity does not need to be totally ruled out, but only some specific forms of complementarity can be incorporated in the model. In any event, incorporating them would necessarily complicate the model and make it empirically less tractable.

A similar argument can be made for complementarity across workers, which could be taken into account, provided it takes forms that still allow defining worker job-specific requirements. In the current "one worker/one job" formulation, requirements indifferently refer either to the total input of skills in the production function or to the input provided by the single worker. With multiple workers contributing to the same production function, these two notions of requirements do not coincide and they need to be defined separately.

Finally, allowing the sharing parameter, γ, to vary both across workers and across jobs is possible, but it complicates the interpretation of mismatch. One convenient feature of the current formulation that would be lost if γ varied by job is the very sharp implications for optimal skill deployment. This is, in part, the result of having jobs and workers being defined by structural features that do not overlap with one another: workers are characterized by skill endowments (η_i) and jobs by the parameters of the production function (β_j, k_j and max_j). A sharing parameter that varies across both i and j would break this useful separation and make the derivation of both optimal deployment and the implications of mismatch much less clear.

4 The *Survey of Adult Skills* (*PIAAC*)

The *Survey of Adult Skills* is the main output of the *Programme for the International Assessment of Adult Competencies* run by the OECD in collaboration with national governments and a consortium of experts supporting the implementation of the survey and the preparation of the data.

The survey is a collection of country-specific samples designed to be representative of the adult population aged between 16 and 65 years. The samples are constructed from potentially very different sampling frames but according to harmonized statistical

procedures aimed at guaranteeing comparability across countries. The same background questionnaire is administered to all sampled individuals in all the countries, merely translated in the local language.[12]

There currently are two rounds of PIIAC data. The first round covers 23 countries (round 1) and was collected between 2008 and 2013.[13] The second round covers 9 countries (that were not in round 1) and was collected between 2012 and 2016. In this paper, we only use the data from round 1 and the descriptive statistics of some key socio-economic variables are presentedin Table 1.

One key element of PIAAC is the skill assessment exercise that all respondents are asked to take as part of the interview process. The exercise consists of a set of test questions organized into three domains: numeracy, literacy and problem-solving. By default, all three tests are carried out on computers but literacy and numeracy can also be done on paper for those who prefer to do so and for those who lack basic IT literacy. Problem-solving can only be taken on computers and those who refuse or cannot use a PC are simply routed out. As a consequence, the number of missing values in problem-solving is relatively high in many countries (on average about 10% across all participating countries but up to over 35% in some). For this reason, the analysis of problem-solving skills is excluded from this paper.

As it is customary in the design of competency tests (OECD 2012; 2013), not all respondents are administered all the questions and a purposely designed routing algorithm guides each respondent through a subset of the test items. This procedure allows reducing the time required to complete the assessment, thus maximising participation. Then, the entirety of the answers for all respondents in all countries is used to estimate a psychometric model based on *Item Response Theory* (*IRT*) that produces a skill proficiency measure for each participant in the survey with completed information from the background questionnaire (Ackerman 2010; Jakubowski 2013; Jacob and Rothstein 2016).

The purpose of the IRT model is the estimation of the unobservable respondents' ability in each domain (literacy, numeracy) using information about their observed performance in tasks that are associated to such domains. The number of tasks that could be associated

Table 1 Descriptive statistics

Variable	Mean	St. dev.	Min	Max
1 = woman	0.480	0.500	0.000	1.000
1 = graduate[a]	0.431	0.495	0.000	1.000
1 = foreign born	0.102	0.302	0.000	1.000
Age distribution				
24 or less	13.75%			
25–34	22.67%			
35–44	24.17%			
45–54	23.72%			
55 plus	15.69%			
	Total	Mean	Min	Max
Sample size[b]	78,422	3565	1666	15,060

Source: OECD Survey of Adult Skills (PIAAC)
[a]Individuals with some tertiary education
[b]Total is the total size of the sample used for the analysis, across all countries. Mean is the mean sample size across countries. Min and max are the minimum and maximum sample size across countries

with each skill is potentially infinite, and only a subset of them can be tested in practice. In PIAAC, each respondent answers on average 20 questions in literacy and about the same in numeracy, taking approximately 1 min for each item.

A number of arbitrary assumptions necessarily need to be made in this context. First, the association of tasks to skills is entirely discretionary and, while reading a text is clearly a literacy task and summing numbers is clearly a numeracy test, there are numerous examples of test items that could be associated with several skills.[14] Additionally, the theory does not provide guidance about the specific formulation of the IRT model in terms of both functional form and explanatory variables, and the choice is usually made on the basis of computational convenience and data quality. PIAAC adopts a logistic model with two parameters, one reflecting the difficulty of the task and one measuring how well the task discriminates among respondents along the underlying skill. The resulting estimates are used to impute an indicator of skill proficiency for each respondent with completed information on the variables used in the IRT model.

For ease of use and interpretation, the skill indicators are transformed into a scale ranging from 0 to 500. The first two lines of Table 2 report some basic descriptive statistics for the indicators of proficiency in literacy and numeracy for the pooled sample of all PIAAC participating countries. The average proficiency is around 277 for literacy and slightly lower (270) for numeracy. In both cases, the median is higher than the mean, suggesting that the distribution is skewed to the left due to a tail of individuals with very low scores. Additionally, the distribution of numeracy proficiency appears to be slightly more dispersed than that of literacy.

The background questionnaire of PIAAC also includes a very detailed section about the use of skills at work. Participants are asked about the frequency with which they perform specific tasks, such as reading documents or making calculations, in the course of their work activities. This paper focuses on a limited set of such questions to construct indicators of the use of literacy and numeracy at work.[15]

The original frequency questions allow respondents to answer on a discrete scale of 5 values: never (1), less than once a month (2), less than once a week but at least once a month (3), at least once a week but not every day (4) and every day (5). The set of tasks considered to construct the indicator of literacy use includes reading and writing of a very wide set of documents.[16] The tasks considered for numeracy are also numerous and very detailed, including making various types of calculations and using calculators.[17]

This large number of questions is averaged to construct skill use indicators for literacy and numeracy. This simple procedure remains agnostic about the relative importance of each task and maintains a rather intuitive interpretation of the resulting scales, where a value of zero signifies that none of the tasks considered is ever performed and a value of 5 corresponds to performing each of the tasks every day. Basic descriptive statistics for

Table 2 Proficiency and use of literacy and numeracy

Variable	Mean	Median	Std. dev.
Proficiency in literacy	277.735	281.319	46.500
Proficiency in numeracy	270.169	273.998	51.575
Use of literacy	2.663	2.750	0.943
Use of numeracy	2.328	2.167	1.080

Source: OECD Survey of Adult Skills (PIAAC)

these indicators are shown in the bottom two lines of Table 2. The mean use of literacy is around 2.7, which is very close to the median (2.7). Numeracy tasks seem to be performed slightly less frequently, with a mean use around 2.3.

5 Empirical results

The methodology described in Section 3 is applied to the PIAAC survey, and the main results are in Tables 3 and 4 for literacy and numeracy, respectively.

Jobs are defined separately for each country on the basis of 2-digit occupational codes (ISCO 2-digit).[18] Due to the small sample sizes, armed forces (ISCO code 0) are dropped. Furthermore, possible observations with missing two- codes have been recoded according to their 1-digit occupation. Finally, occupations with fewer than 50 observations (about 3% of the overall sample) are also dropped. In the end, we have 492 country-occupation cells, with a median of 25 occupations per country (mean is 22). The final working sample is restricted to dependent employees holding only one job.

The computation of the standard errors for the estimates presented in this section needs to take into account both the differences in the sampling frames across countries and the variation induced by the imputation of the ability scores. The Appendix discusses in details how this is done. Tables 3 and 4 present our main results disaggregated by country. For brevity, all the following results will be reported pooling all countries together.[19]

Considering literacy proficiency, approximately 75% of dependent employees are classified as well-matched across all the countries covered by the survey, about 16% are over-skilled and 9% are under-skilled (Table 3). These average results mask a large heterogeneity across countries. For example, over-skilling can affect as many as 25% of workers in Spain and as few as 5.9% in France. Under-skilling is lowest in Austria (2.2%) and Canada (2.4%) and is highest in Spain (17.1%). The results for numeracy (Table 4) are broadly similar to those for literacy, and the ranking of countries is also similar. The Spearman rank correlation between the incidence of mismatch—i.e. the sum of the under- and over-skilled—in literacy and in numeracy is equal to 0.55.

In fact, Table 5 shows that 90% of the workers who are well-matched in literacy are also well-matched in numeracy. The overlap is less strong but still very important among the under- and the over-skilled.

Table 6 describes the incidence of under- and over-skilling across socio-demographic groups. Men appear to be affected by over-skilling more frequently than women, both with regard to literacy and numeracy, whereas gender differences in under-skilling are minor. This result is not obvious, as one may think that women, who often find employment more difficultly than men, might be more willing to take jobs that do not necessarily match their skills perfectly. On the other hand, (OECD 2013a) shows that women use their skills less frequently than men, mostly because of the jobs in which they are occupied. Being in jobs where skills are not often used, they might also be less likely to be mismatched.

As one might expect, graduate workers are less likely to be under-skilled than non-graduates. They are also more likely to be over-skilled (Quintini 2011a; 2011b; OECD 2013a). Literacy and numeracy follow similar patterns. All these differences are statistically significant at the 5% level.

Table 3 Skill mismatch by country—literacy

Country	Under-skilled	Well-matched	Over-skilled	Country	Under-skilled	Well-matched	Over-skilled
Austria	0.022	0.878	0.100	Korea	0.057	0.758	0.185
	(0.041)	(0.540)	(0.041)		(0.025)	(0.032)	(0.029)
Belgium[a]	0.058	0.855	0.087	The Netherlands	0.040	0.893	0.067
	(0.016)	(0.022)	(0.019)		(0.012)	(0.019)	(0.016)
Canada	0.024	0.880	0.095	Norway	0.058	0.833	0.109
	(0.006)	(0.013)	(0.013)		(0.022)	(0.028)	(0.023)
Czech Republic	0.042	0.818	0.140	Poland	0.103	0.778	0.119
	(0.021)	(0.032)	(0.029)		(0.026)	(0.037)	(0.030)
Denmark	0.047	0.839	0.114	Russia	0.100	0.793	0.107
	(0.011)	(0.018)	(0.014)		(0.042)	(0.047)	(0.036)
Estonia	0.042	0.892	0.066	Slovak Republic	0.102	0.777	0.120
	(0.013)	(0.023)	(0.019)		(0.025)	(0.030)	(0.026)
Finland	0.030	0.900	0.070	Spain	0.171	0.580	0.249
	(0.008)	(0.019)	(0.018)		(0.036)	(0.038)	(0.039)
France	0.044	0.897	0.059	Sweden	0.091	0.810	0.099
	(0.012)	(0.021)	(0.015)		(0.025)	(0.033)	(0.022)
Germany	0.097	0.670	0.233	UK[b]	0.067	0.825	0.108
	(0.036)	(0.045)	(0.041)		(0.017)	(0.024)	(0.020)
Ireland	0.117	0.740	0.142	USA	0.144	0.612	0.244
	(0.032)	(0.042)	(0.033)		(0.039)	(0.037)	(0.042)
Italy	0.102	0.766	0.132	Total	0.093	0.746	0.162
	(0.030)	(0.042)	(0.032)		(0.072)	(0.133)	(0.073)
Japan	0.031	0.876	0.093				
	(0.011)	(0.021)	(0.018)				

Bootstrapped standard errors in parentheses. See Appendix for details on the bootstrap procedure. *Source:* OECD Survey of Adult Skills (PIAAC)

[a]Only the Flanders region is included in the survey

[b]Only England and Northern Ireland are included

Table 4 Skill mismatch by country—numeracy

Country	Under-skilled	Well-matched	Over-skilled	Country	Under-skilled	Well-matched	Over-skilled
Austria	0.018	0.834	0.148	Korea	0.056	0.792	0.152
	(0.043)	(0.060)	(0.041)		(0.024)	(0.035)	(0.031)
Belgium[a]	0.059	0.859	0.082	The Netherlands	0.038	0.904	0.058
	(0.017)	(0.021)	(0.019)		(0.013)	(0.021)	(0.015)
Canada	0.028	0.874	0.098	Norway	0.074	0.848	0.078
	(0.007)	(0.015)	(0.012)		(0.018)	(0.028)	(0.021)
Czech Republic	0.038	0.838	0.124	Poland	0.107	0.739	0.155
	(0.023)	(0.034)	(0.028)		(0.025)	(0.037)	(0.027)
Denmark	0.062	0.842	0.096	Russia	0.065	0.816	0.119
	(0.013)	(0.019)	(0.015)		(0.044)	(0.050)	(0.038)
Estonia	0.031	0.910	0.059	Slovak Republic	0.043	0.782	0.176
	(0.010)	(0.019)	(0.017)		(0.027)	(0.033)	(0.031)
Finland	0.040	0.897	0.063	Spain	0.151	0.599	0.250
	(0.008)	(0.018)	(0.016)		(0.033)	(0.038)	(0.035)
France	0.043	0.892	0.065	Sweden	0.075	0.843	0.081
	(0.010)	(0.019)	(0.016)		(0.025)	(0.031)	(0.022)
Germany	0.105	0.653	0.243	UK[b]	0.069	0.822	0.108
	(0.033)	(0.041)	(0.040)		(0.017)	(0.028)	(0.024)
Ireland	0.121	0.726	0.153	USA	0.139	0.598	0.263
	(0.034)	(0.043)	(0.038)		(0.038)	(0.038)	(0.042)
Italy	0.088	0.771	0.141	Total	0.087	0.746	0.167
	(0.030)	(0.039)	(0.030)		(0.070)	(0.137)	(0.079)
Japan	0.035	0.894	0.071				
	(0.011)	(0.020)	(0.018)				

Source: OECD Survey of Adult Skills (PIAAC)

[a]Only the Flanders region is included in the survey

[b]Only England and Northern Ireland are included

Table 5 Overlapping of skill mismatch in literacy and numeracy

		Numeracy		
		Under-skilled	Well-matched	Over-skilled
Literacy	Under-skilled	0.652	0.305	0.043
		(0.017)	(0.016)	(0.008)
	Well-matched	0.033	0.894	0.073
		(0.003)	(0.005)	(0.004)
	Over-skilled	0.016	0.313	0.671
		(0.005)	(0.015)	(0.014)

Shares of workers who are under-skilled, well-matched or over-skilled in numeracy by mismatch status in literacy. The numbers sum to 1 by rows. Bootstrapped standard errors in parentheses. See Appendix for details on the bootstrap procedure. *Source:* OECD Survey of Adult Skills (PIAAC)

Consistent with the higher educational achievement of the younger generations, older workers are more likely to be under-skilled and less likely to be over-skilled, in both literacy and numeracy. This result also conforms with the idea that younger workers need time to experiment and move across jobs in search of what fits their skills well (Topel and Ward 1992). As for older workers, the presence of a non-negligible share of over-skilled might be interpreted as an encouraging finding, especially for those countries facing rapidly ageing populations, as it suggests that improving the matching of older workers may help mitigate the impact of population ageing on productivity.

Finally, foreign workers are twice more likely than natives to be under-skilled in either literacy or numeracy. The incidence of over-skilling in numeracy (literacy) is 70% (40%) larger for foreigners than natives. This result is easy to rationalize for literacy, given that in most cases, the language of the destination country is different from migrants' mother

Table 6 Skill mismatch by socio-demographic groups

|	Literacy		Numeracy	
	Under-skilled	Over-skilled	Under-skilled	Over-skilled
Men	0.095	0.184	0.081	0.201
	(0.005)	(0.006)	(0.005)	(0.006)
Women	0.090	0.138	0.094	0.131
	(0.005)	(0.005)	(0.005)	(0.006)
Non-graduates	0.106	0.132	0.102	0.139
	(0.005)	(0.006)	(0.005)	(0.006)
Graduates	0.076	0.202	0.068	0.204
	(0.006)	(0.006)	(0.006)	(0.008)
Age <45	0.083	0.186	0.080	0.185
	(0.005)	(0.005)	(0.005)	(0.006)
Age ≥45	0.108	0.125	0.098	0.139
	(0.005)	(0.004)	(0.005)	(0.005)
Natives	0.082	0.169	0.078	0.172
	(0.004)	(0.005)	(0.005)	(0.005)
Foreigners	0.183	0.100	0.171	0.121
	(0.008)	(0.006)	(0.008)	(0.006)

Shares of under- and over-skilled workers. Bootstrapped standard errors in parentheses. See Appendix for details on the bootstrap procedure. *Source:* OECD Survey of Adult Skills (PIAAC)

tongues. For numeracy, the lower incidence of over-skilling contrasts with the common finding that immigrants often hold formal educational qualifications that are higher than those required by their jobs. The over-qualification of migrants is often attributed to the difficulties in having educational qualifications officially recognized across countries. However, the results in Table 6 seem to suggest that some of the over-qualified foreigners simply do not have the necessary skills to carry out their jobs satisfactorily, pointing to a large heterogeneity in the quality of schooling across countries.

5.1 Comparison with other measures of skill mismatch

As we already discussed in Section 2, we are certainly not the first to measure skill mismatch and a variety of methodologies have been already proposed in the literature.

In Fig. 6, we show the distribution of skill mismatch for the pooled PIAAC sample obtained using the two most popular approaches to measuring it. The left panel of Fig. 6 shows the percentages of under-skilled, well-matched and over-skilled based on the fully self-reported approach, which only makes use of the self-reported answers to the questions about needing training and feeling capable of doing more demanding jobs. The under-skilled are those who report needing training, the over-skilled are those who report feeling capable of doing more demanding jobs and the well-matched are those answering negatively to both questions. Applying this method to the PIAAC data shows that a large 82% are classified as over-skilled, suggesting that overconfidence might actually be a very common attitude. An additional problem with this method is that a sizeable fraction of workers report both needing training and feeling able to do more demanding tasks. In the pooled PIAAC sample, this group represents a good one fourth of all employed workers. Notice also that self-reported mismatch cannot be attributed to a specific skill (literacy or numeracy).

Fig. 6 Comparison with other measures of skill mismatch
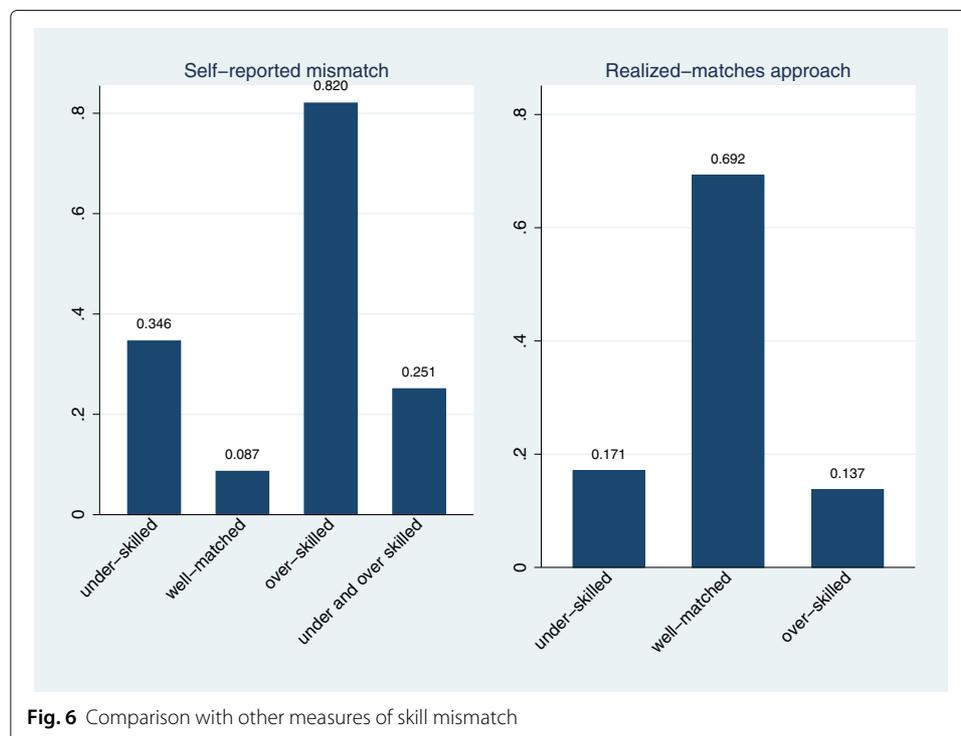
The right panel of Fig. 6 shows results obtained with the statistical or realised-match approach for numeracy mismatch. This approach proceeds by first computing the median observed skill of workers employed in each occupation, and then, it defines minimum and maximum requirements in each occupation by, respectively, adding and subtracting one standard deviation to the median. Workers are classified as well-matched if their observed skills are within a one standard deviation interval around the median, they are under-skilled if their skills are below the median minus one standard deviation and they are over-skilled if they are above the median plus one standard deviation.

Results indicate that according to this method, about two thirds of the workers are well-matched and the remaining one third is rather equally divided between under-skilled and over-skilled. In fact, this result is a direct consequence of the normality of the distribution of the skill scores, which is imposed by item response theory, the methodology used to compute them.

5.2 Robustness checks

Our mismatch indicator is based on the minimum and maximum skill requirements by occupations, which are estimated as the minimum ($\widehat{min}_j$) and maximum ($\widehat{max}_j$) of the country-occupation distribution of proficiency for those workers who report neither feeling the need of training nor feeling to be able to do more demanding jobs. As discussed in Section 3.3, the same requirements could also be estimated as the maximum proficiency level of workers who report feeling the need of training ($\widetilde{min}_j$) and the minimum proficiency of workers who feel they can do a more demanding job ($\widetilde{max}_j$). However, the first set of estimators ($\widehat{min}_j$ and $\widehat{max}_j$) is preferred because it is more robust to the most common sources of misreporting, such as respondents' overconfidence and the misinterpretation of the question about needing training. Comparing these alternative estimators can, therefore, provide an indication of the extent of mismeasurement.

Table 7 performs such a comparison. The table reports the average absolute (columns 1 and 3) and percentage (columns 2 and 4) difference between these alternative estimators across all the country-occupation cells. Results show that the two sets of estimates are massively different, thus emphasizing the importance of deriving indicators of mismatch that take misreporting into careful consideration. On average, across all occupations and countries, $\widetilde{min}_j$ is approximately 67% larger than $\widehat{min}_j$ for literacy and

Table 7 Alternative estimates of the skill requirements

	$\widetilde{min}_j - \widehat{min}_j$[a]	$\frac{\widetilde{min}_j - \widehat{min}_j}{\widehat{min}_j}$	$\widetilde{max}_j - \widehat{max}_j$[b]	$\frac{\widetilde{max}_j - \widehat{max}_j}{\widehat{max}_j}$
Literacy	126.04	0.669	−116.36	−0.365
	(1.533)	(0.013)	(1.235)	(0.004)
Numeracy	140.95	0.852	−111.74	−0.350
	(1.608)	(0.036)	(1.349)	(0.004)

Bootstrapped standard errors in parentheses. See Appendix for details on the bootstrap procedure. All figures are averages over occupational categories and countries. *Source:* OECD Survey of Adult Skills (PIAAC)

[a] $\widehat{min}_j$ = the 5th percentile of the proficiency distribution of workers not feeling able to do more demanding jobs nor feeling the need of training (by occupation). $\widetilde{min}_j$ = the 95th percentile of the proficiency distribution of workers feeling the need of further training

[b] $\widehat{max}_j$ = the 95th percentile of the proficiency distribution of workers not feeling able to do more demanding jobs nor feeling the need of training. $\widetilde{max}_j$ = the 5th percentile of the proficiency distribution of workers feeling able to do more demanding jobs

85% larger for numeracy. $\widetilde{max}_j$ is approximately 35% times smaller than $\widehat{max}_j$ in both skill domains. These findings indicate that using the pure self-reported information to define skill-mismatch would lead to classify workers as over-skilled even if their assessed proficiency levels are very often below those of the self-reported well-matched or even under-skilled.

In Table 8, we investigate the stability of our results to another important parameter of our methodology, namely the specific choice of the estimators of $\widehat{min}_j$ and $\widehat{max}_j$. For our main results, we use the 5th and 95th percentiles of the skill distributions, and in Table 8, we report results obtained under alternative choices, namely the actual minimum and maximum, the 1st and 99th percentiles and the 2nd and 98th percentiles. We find that our main findings are very robust across all these alternatives.

6 The misuse of skills

According to the theoretical framework of Section 3.1, workers who are well-matched are the only ones who fully deploy their skill endowments. The over-skilled are indifferent between deploying any amount of skills between the maximum required by their jobs and their entire endowments. The under-skilled need to stretch the deployment of their skills to reach the minimum required by their jobs. These theoretical implications can now be readily taken to the PIAAC data, where together with information about skill endowments, respondents are also asked about their use of skills at work.

For each mismatched worker (either under- or over-skilled), it is possible to compare the use of skills with well-matched workers at their same level of proficiency and in the same country. Table 9 shows that, on average, across countries, the indicator of literacy use at work for individuals who are under-skilled in literacy is about 16.3% higher than the corresponding indicator for similarly proficient workers who are well-matched, suggesting that they do actually over-use their skills. Consistent with the large overlap of mismatch across skill domains (see Table 5), literacy under-skilled workers also appear to over-use their numeracy at work (11.1% more than the well-matched). Notice that the over-usage

Table 8 Skill mismatch under alternative thresholds for job requirements

	Literacy			Numeracy		
	Under-skilled	Well-matched	Over-skilled	Under-skilled	Well-matched	Over-skilled
Min–max[a]	0.087	0.763	0.150	0.082	0.762	0.156
	(0.078)	(0.153)	(0.086)	(0.075)	(0.156)	(0.090)
1st–99th[b]	0.087	0.762	0.151	0.082	0.761	0.157
	(0.077)	(0.152)	(0.085)	(0.075)	(0.155)	(0.090)
2nd–98th[c]	0.088	0.760	0.152	0.082	0.760	0.158
	(0.076)	(0.148)	(0.083)	(0.074)	(0.151)	(0.087)
5th–95th[d]	0.093	0.746	0.162	0.087	0.746	0.167
	(0.072)	(0.133)	(0.073)	(0.070)	(0.137)	(0.079)

Bootstrapped standard errors in parentheses. See Appendix for details on the bootstrap procedure. *Source:* OECD Survey of Adult Skills (PIAAC)
[a] Skill requirements estimated as the minimum and maximum skills of the self-reported well-matched
[b] Skill requirements estimated as the 1st and 99th percentiles of the skill distribution of the self-reported well-matched
[c] Skill requirements estimated as the 2nd and 98th percentiles of the skill distribution of the self-reported well-matched
[d] Skill requirements estimated as the 5th and 95th percentiles of the skill distribution of the self-reported well-matched

Table 9 Skill mismatch and the use of skills at work

	Literacy mismatch		Numeracy mismatch	
	Over-use of literacy[a]	Over-use of numeracy[a]	Over-use of literacy[a]	Over-use of numeracy[a]
Under-skilled	0.163	0.064	0.111	0.097
	(0.012)	(0.015)	(0.008)	(0.019)
Well-matched	–	0.002	0.003	–
		(0.001)	(0.001)	
Over-skilled	−0.053	−0.010	−0.031	−0.017
	(0.009)	(0.007)	(0.005)	(0.011)

Bootstrapped standard errors in parentheses. See Appendix for details on the bootstrap procedure. *Source:* OECD Survey of Adult Skills (PIAAC)
[a]Relative difference between the average usage of the indicated category and the well-matched, conditional on proficiency

of skills by the under-skilled is not necessarily an efficient outcome, since they could be more productive, while at the same time exerting less effort and being less stressed, if they were better matched.[20]

Over-skilling is associated with a substantial waste of skills, as workers who are over skilled in literacy appear to use their skills at work substantially less than similarly proficient workers who are well-matched, namely 5.3% lower usage of literacy and 1% lower usage of numeracy. Looking at mismatch in numeracy shows very similar findings.

A further natural development of the analysis in this section would be the computation of the output loss associated with the misuse of skills. However, such an exercise requires causal estimates of the skill-output gradient, whose identification goes beyond the scope of this paper and is left to future research. A similar and equally interesting analysis could be extended to some indicator of welfare or health so as to incorporate more appropriately the potential negative effects of under-skilling on workers well-being.

7 The efficiency of the assignment mechanism

In this section, we propose a simple empirical exercise to assess the efficiency of the assignment of workers to jobs that is observed in the data. Such an exercise consists in reallocating the individuals in our data to the existing jobs according to an artificial assignment mechanism designed to reduce skill mismatch. We perform this reassignment separately for each skill (numeracy and literacy) and on the basis of the skill endowments of the individuals and the skill requirements of the jobs that are observed in the data (i.e. those filled with an employed worker). Hence, we do not attempt to solve the complex problem of finding the optimal allocation of jobs and workers, most notably because we do not have a measure of output nor causal estimates of the skill-output gradients by literacy and numeracy. Moreover, the exercise we perform in this section takes the current stock of jobs as given and does not consider new jobs that could potentially be created thanks to the more efficient assignment mechanism. Similarly, we also take the skill requirements of the existing jobs as given, and we do not endogenize the potential effect of better assignment on the characteristics of the jobs.

Despite all these limitations, we believe that the results in this section can be useful to show whether and by how much the observed degree of skill mismatch could be reduced by reallocating workers to jobs according to some reasonable and easily implementable procedure. Additionally, these results illustrate the important connection

between our approach and the macro-literature on assignment models. As we already discussed in Section 3.1, our model can be viewed as a framework to produce measures of the efficiency of the assignment mechanism when data about workers' skills are available.

We perform three different reassignment exercises. First, we consider only employed workers and the skill requirements of the jobs they occupy. Then, we reallocate workers to jobs by assigning the least skilled worker to the job with the lowest minimum requirement and the most skilled worker to the job with the highest maximum requirement; next, we assign the second least skilled worker to the job with the second lowest minimum requirement and the second most skilled worker to the job with the second highest maximum requirement and so on until all jobs are filled with a worker. The assignment procedure is carried out country-by-country and replicated separately for literacy and numeracy.

The second row of Table 10 reports the distribution of skill mismatch associated with the resulting assignment. For comparison purposes, the first row of the table reports the distribution of skill mismatch observed in the real data, namely the same estimates reported in Tables 3 and 4.[21]

Focusing on the results for literacy, we find that this relatively simple reassignment procedure increases the share of well-matched workers from 74.6 to 90%, an increase of over 20%. This effect is generated mostly by a reduction of the incidence of over-skilling that goes from 16.2 to 1.6%. The contraction of under-skilling is more modest: from 9.3 to 8.4%. When the reallocation is performed on numeracy, the results are similar to the notable exception that now, under-skilling increases slightly. Overall, these findings suggest that there is not a major lack of highly skilled individuals among the employed but rather a misallocation of them to the the existing jobs. On the other hand, there seems to be a certain lack of skill towards the bottom of the distribution and some jobs with relatively high minimum requirements remain filled with insufficiently skilled individuals.

The second reassignment exercise that we perform is similar to the previous one with the exception that we now consider also the unemployed among the pool of workers to be reallocated to the existing jobs. We then have more workers than jobs, and we start selecting out those with skill endowments above the highest maximum requirement observed in the country or below the lowest minimum requirement. In any possible assignment, these workers would certainly be either under- or over-skilled. Then, we apply our usual reassignment algorithm to allocate the remaining workers to the existing jobs. Results are reported in the third row of Table 10 and indicate that now, both under-skilling and

Table 10 Skill mismatch under alternative assignment mechanisms

	Literacy			Numeracy		
	Under-skilled	Well-matched	Over-skilled	Under-skilled	Well-matched	Over-skilled
Observed[a]	0.093	0.746	0.162	0.087	0.746	0.167
Re-assign E[b]	0.084	0.900	0.016	0.095	0.897	0.008
Re-assign E+U[c]	0.044	0.931	0.024	0.045	0.924	0.032
Re-assign E+U+I[d]	0.021	0.899	0.080	0.032	0.904	0.065

Source: OECD Survey of Adult Skills (PIAAC)
[a]Observed distribution of skill-mismatch (same as in Tables 3 and 4)
[b]Distribution of skill mismatch obtained by reallocating employed workers to the existing jobs. See Section 7 for details
[c]Distribution of skill mismatch obtained by reallocating employed and unemployed workers to the existing jobs. See Section 7 for details
[d]Distribution of skill mismatch obtained by reallocating all individuals (employed, unemployed and inactive) to the existing jobs. See Section 7 for details

over-skilling are reduced compared to the observed allocation and the share of jobs filled with a well-matched worker reaches 93.1% for literacy and 92.4% for numeracy. Notice, however, that compared to re-assigning only the employed workers (the second row of the table), the incidence of over-skilling increases. This result can be easily explained by the fact that there is now a larger pool of individuals that can fill the jobs with low minimum requirements and, therefore, there are more skilled individuals available to be assigned to the more demanding jobs. In addition, there also is a non-negligible share of relatively highly skilled individuals among the unemployed, especially women.

A similar but more pronounced patter can be observed in the last row of Table 10, where we report the results of our last reassignment exercise. In this case, we consider all individuals in our data, namely all the employed, unemployed and inactive. As before, we drop those with endowments below the lowest minimum requirement and above the highest maximum requirement and we apply the reassignment algorithm to the remaining ones. Now, there are enough individuals to fill most jobs with a sufficiently competent worker: the incidence of under-skilling goes down to 2.1% when the reassignment is based on literacy and to 3.2% when done on numeracy. However, this also implies that there are now more skilled workers available to fill more demanding jobs, both because they do not have to be allocated to less demanding jobs and also because there some very skilled individuals among the inactive, again especially women. As a consequence, the share of over-skilling increases compared to the other reassignment exercises (but still decreases compared to the observed data).

Overall, the results of this section indicate that the observed skill mismatch in the pooled PIAAC countries is mostly due to an allocative problem rather than to the shortage of certain skills in the population. Simply reallocating those currently employed to the current existing jobs reduces skill mismatch substantially, to an overall level of around 10% (counting both under- and over-skilling), which can probably be considered a reasonable structural level. In terms of policy implications, it seems thus more effective to focus on policies aimed at improving the quality of the matching process rather than to those aimed at modifying the skill composition of labour supply (educational choices). Of course, our analysis does not capture those jobs that might remain vacant or simply not exist because of mismatch and skill shortages might have an effect along this dimension.

8 Conclusions

This paper proposes a novel measure of skill mismatch for the recent PIAAC data. This new measure allows classifying workers into under-skilled, well-matched and over-skilled along the skill domains of literacy and numeracy. The novelty lies mostly in the development of a theory-based procedure to identify jobs' requirements from data on workers in the absence of direct information about the production process.

On average, across the entire pooled sample, approximately 75% of dependent employees are well-matched in the literacy domain, about 9% are under-skilled and 16% are over-skilled. The overlap between literacy and numeracy mismatch is substantial: 90% of the workers who are well-matched in literacy are also well-matched in numeracy.

Men are more likely to be over-skilled than women, whereas gender differences in under-skilling are minor. Tertiary graduates are substantially less likely to be under-skilled than less educated workers, and they are more likely to be over-skilled. Foreign workers are substantially more likely to be under-skilled and substantially less likely to

be over-skilled. Differences emerge also when looking across age groups. Furthermore, we show that skill mismatch is associated with a substantial degree of skill over- and under-utilization, with potential sizeable implications in terms of output loss and workers' well-being. We also perform a series of reassignment exercises, and we find indications that skill mismatch can be substantially reduced by efficiently reallocating workers to jobs.

Despite being mostly illustrative of the methodology, these findings have important implications for policy. A better match of the workers' skills to the requirements of their jobs can reduce the waste of skills among the over-skilled, improve the efficiency of the under-skilled while, at the same time, potentially reducing their levels of stress and, eventually, lead to important improvements of the overall productivity of the economy and the well-being of individuals.

Endnotes

[1] The indicator of skill mismatch described in this paper is officially adopted by the OECD in the context of the *Programme for the International Assessment of Adult Competencies* (*PIAAC*), of which the *Survey of Adult Skills* is a key element, and hereafter, it will be labelled *OECD measure of skill mismatch*. For simplicity, the acronym PIAAC will be used in this paper to refer to both the overall programme and the survey. Some of the results reported here differ from those in (OECD 2013a) because the latter uses a slightly richer version of the data whose access is restricted. In this paper, we use the publicly available data files and our results are fully replicable.

[2] These are workers who report that they do not feel they *"have the skills to cope with more demanding duties than those they are required to perform in their current job"* and they do not feel they *"need further training in order to cope well with their present duties."*

[3] The distributions are constructed using the same sample of our main analysis in Section 5. The qualification requirements of the jobs are self-reported by the survey respondents.

[4] International Adult Literacy Survey (IALS), Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), Adult Literacy and Lifeskills (ALL).

[5] In their recent work, (van der Velden and Bijlsma 2016) take a more practical approach and investigate how different combinations of skill use and skill endowment indicators correlate with wages.

[6] It is fair to also acknowledge that there are important features of the efficiency of an assignment mechanism that our measure does not necessarily capture very well, such as the efficiency of the allocation of workers between employment and non-employment.

[7] In this framework one might also incorporate an analysis of qualification mismatch by simply defining qualifications as a discretization of skills.

[8] The subscript i to the function $c(\cdot)$ indicates that the function itself varies with η_i, which, in fact, determines the point where the slope of the function changes.

[9] This assumption can be easily justified in the context of search&matching models that have become the standard view of the functioning of the labour market. In the standard version of such models, the equilibrium wage is equal to a fraction of the job's output plus the outside option of the worker. Further, assuming that the worker's outside option is itself a fraction of the wage (as in most unemployment insurance systems) leads precisely to an expression of the equilibrium wage as a fraction of productivity.

[10] Allowing the sharing parameter γ to vary across jobs (or both across workers and jobs) is possible, but it makes it less obvious to formalize a meaningful definition of skill mismatch.

[11] The minimum skill requirement (min_j) can be easily and uniquely derived from the triplet [k_j, β_j, max_j], i.e. for each [k_j, β_j, max_j], there exists one and only one min_j.

[12] In a few countries, the survey is administered in multiple languages.

[13] The data for Australia are not included in the set of public use files and are therefore excluded from our analysis. Hence, we cover only 22 countries.

[14] Notice that the same item can be used to estimate more than one skill measure.

[15] (OECD 2013a) analyses a larger set of skill use indicators.

[16] Directions, instructions, memos, letters, e-mails, articles (in newspapers, magazines, newsletters, professional and scholarly journals), books, manuals and reference materials, bills, invoices, financial statements, diagrams, maps and schematics.

[17] Calculating prices, costs or budgets; calculating fractions, decimals or percentages; using a calculator; preparing charts, graphs or tables; using algebra or formulas; using advanced mathematics (calculus), trigonometry, statistics, regression techniques.

[18] For four countries (Austria, Canada, Estonia and Finland), only 1-digit occupational categories are available in the public use files.

[19] Results by country are available from the authors upon request.

[20] By construction, the degree of over-use of literacy for those who are well-matched in literacy is zero, similar to numeracy.

[21] We do not report standard errors in Table 10 because it is unclear what would be the underlying source of variation when the allocation is generated by an ad hoc assignment procedure.

[22] PIAAC also provides a sequence of *replicate weights* that can be used to assess the sampling variability (OECD 2013). However, it is not obvious how to use them with complex estimation procedures such as the derivation of the skill mismatch indicators and the related statistics. Moreover, additional adjustments would still be needed to take proper account for the imputation of the skill measurements.

[23] To reduce the size of the resulting datasets, all sampling weights have been divided by the minimum weight in the country so that each sampled unit is represented at least once and, at the same time, all relative weights remain unchanged.

[24] Performing correct bootstrapping without expanding the sample would require knowledge of the details of the sampling process in each county, namely stratification units, primary and secondary units, etc. Unfortunately, this information is not provided in PIAAC (and the replicate weights are meant to replace it) for two sets of reasons. First, the sampling structures of the country samples are sometimes quite different. For example, in some cases, the original sampling frame is a standard population register whereas in other instances, data are originally drawn from administrative archives. As a consequence, providing complete information about the sampling structure in a compact and comparable format across all countries is problematic. The second reason is related to the various confidentiality norms present in each participating country, many of which would be breached by the full disclosure of all the sampling information (OECD 2013).

[25] Any plausible value could have been used, and the resulting point estimates would have the exact same asymptotic properties.

Appendix

Inference

In order to make correct inference about the mismatch indicators reported in Section 5, it is necessary to take into proper account both the sample variability and the imputed nature of the skill measures (OECD 2013).

Taking proper account for the sampling variability is a particularly important issue in PIAAC, given the different nature of the country samples. Although harmonized protocols guarantee the comparability of results, PIAAC remains a collection of country surveys, each of which has been constructed independently. Each country sample comes with a sampling weight that indicates the number of units in the target population represented by each sampled unit. Such a weight summarizes all the necessary information to obtain point estimates that are representative of the target population.[22]

In order to compute asymptotically valid standard errors around the estimates presented in the main text, the following procedure has been adopted. First, each sampled unit is identically replicated a number of times equal to its sampling weight (rounded to the closest integer), so as to generate a sample that replicates the full target population.[23] The resulting expanded datasets are fully representative of the target populations and can be used to extract sequences of S bootstrapped samples of the same size of the original country samples. The entire analysis is then repeated on each bootstrapped sample resulting in a sequence of S estimates for each of the statistics presented in Sections 5 and 6. The empirical distribution of such statistics is then used to compute standard errors that are asymptotically valid by construction.[24]

As it is now common practice with IRT-derived measures of psychometric traits, a series of *plausible values* for each trait is provided. In the specific case of the OECD SAS, ten plausible values for each of the three skills considered (literacy, numeracy and problem solving) are available. Each of them is an equally good proxy of the underlying unobservable psychometric construct; however, each of them is a noisy proxy and the dispersion across the plausible values reflect measurement error (OECD 2013; Mislevy 1993a; 1993b). All the point estimates presented in the main text are constructed using one such plausible values for literacy and for numeracy (the first).[25]

In the bootstrapping procedure described above, at each replication, one randomly selected plausible value is used to proxy skills (one plausible value for literacy and one for numeracy), so as to incorporate in the resulting sequence of estimates the additional variability induced by the imputed nature of the measurements.

References

Ackerman TA (2010) The theory and practice of item response theory by de Ayala, r. j. J Educ Measurement 47(4):471–476

Allen J, van der Velden R (2001) Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. Oxford Economic Papers 53(3):434–52

CEDEFOP (2010) The skill matching challenge: analysing skill mismatch and policy implications. Publications Office of the European Union, Luxembourg

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

Chevalier A, Lindley J (2009) Overeducation and the skills of UK graduates. J Royal Stat Soc Ser A 172(2):307–337

Desjardins R, Rubenson K (2011) An analysis of skill mismatch using direct measures of skills. OECD Education Working Papers 63, OECD Publishing, Paris

Dolton P, Vignoles A (2000) The incidence and effects of overeducation in the U.K. graduate labour market. Econ Educ Rev 19(2):179–198

Farber HS (1999) Mobility and stability: the dynamics of job change in labor markets. In: Ashenfelter O, Card D (eds). Handbook of Labor Economics. Handbook of Labor Economics, Vol. 3. Elsevier, New York, pp 2439–2483. Chap. 37

Flisi S, Goglio V, Meroni E, Rodrigues M, Vera-Toscano E (2014) Occupational mismatch in Europe: understanding overeducation and overskilling for policy making. European Commission, JRC Science and Policy Reports, Luxembourg

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

Groot W, Maassen van den Brink H (2000) Overeducation in the labor market: a meta-analysis. Econ Educ Rev 19(2):149–158

Jacob B, Rothstein J (2016) The measurement of student ability in modern assessment systems. J Econ Perspect 30(3):85–108

Jakubowski M (2013) Analysis of the predictive power of pisa test items. OECD Education Working Papers 87, OECD Publishing, Paris

Jovanovic B (1979) Job matching and the theory of turnover. J Polit Economy 87(5):972–990

Levels M, van der Velden RKW, Levels M, Allen JP (2013) Skill mismatch and skill use in developed countries: evidence from the PIAAC study. ROA Research Memorandum 017, Maastricht

Mavromaras K, McGuinness S, Wooden M (2007) Overskilling in the australian labour market. Australian Economic Review 40(3):307–312

Mislevy RJ (1993a) Should "multiple imputations" be treated as "multiple indicators"? Psychometrika 58(1):79–85

Mislevy, RJ (1993b) Some formulas for use with bayesian ability estimates. Educational and Psychological Measurement 53(2):315–328

Montt G (2016) The causes and consequences of field-of-study mismatch. OECD Social, Employment and Migration Working Papers 167, OECD

OECD (2012) PISA 2012 assessment and analytical framework: mathematics, reading, science, problem solving and financial literacy. OECD Publishing, Paris

OECD (2013a) First international report on PIAAC —volume I. OECD Publishing, Paris

OECD (2013) First international report on PIAAC—volume II. OECD Publishing, Paris

Quintini G (2011a) Over-qualified or under-skilled: a review of existing literature. Social, OECD, Employment and Migration Working Papers 121, OECD Publishing, Paris

Quintini, G (2011b) Right for the job: over-qualified or under-skilled?. OECD Social, Employment and Migration Working Papers 120, OECD Publishing, Paris

Robin JM, Meghir C, Lise J (2009) Matching, sorting and wages. 2009 Meeting Papers 180, Society for Economic Dynamics. http://dx.doi.org/10.1016/j.red.2015.11.004

Rubb S (2003) Overeducation: a short or long run phenomenon for individuals? Econ Educ Rev 22(4):389–394

Şahin A, Song J, Topa G, Violante G (2012) Mismatch unemployment. NBER Working Papers 18265, National Bureau of Economic Research, Inc. https://www.aeaweb.org/articles?id=10.1257/aer.104.11.3529

Sattinger M (1993) Assignment models of the distribution of earnings. J Econ Literature 31(2):831–880

Shimer R (2007) Mismatch. Am Econ Rev 97(4):1074–1101

Sicherman N (1991) Overeducation in the labor market. J Labor Econ 9(2):101–22

Sloane PJ, Battu H, Seaman PT (1999) Overeducation, undereducation and the british labour market. Appl Econ 31(11):1437–1453

Topel RH, Ward MP (1992) Job mobility and the careers of young men. Q J Econ 107(2):439–79

van der Velden R, Bijlsma I (2016) Skill, skill use and wages: a new theoretical perspective. Technical report, Research Centre for Education and the Labour Market (ROA), Maastricht