Measurements of skill and skill-use using PIAAC

Kawaguchi, Daiji; Toriyabe, Takahiro

Labour Economics 78 (2022) 102197

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Measurements of skill and skill-use using PIAAC☆

Daiji Kawaguchi[a,b,*], Takahiro Toriyabe[a,c]

[a] The University of Tokyo, Tokyo, Japan
[b] Research Institute of Economy, Trade and Industry, Tokyo, Japan
[c] University College London, London, United Kingdom

A R T I C L E I N F O

JEL classification:
D12
H24
J16
J12
J13
J16
J24

Keywords:
Measurement
Skill use
Gender gap
Parental leave
Immigrant wage penalty

A B S T R A C T

We develop new indices of skill and skill use, drawing on the alley of skill and skill-use questions in the Programme for the International Assessment of Adult Competencies (PIAAC). We demonstrate that the proposed skill and skill use indices explain the wage gap between males and females, as well as the gap between immigrants and natives. We also show that the skill use index captures the side effect of parental-leave policies on females that conventional labor-market outcomes fail to capture. We discuss how the newly developed indices can be merged to conventional survey data.

1. Introduction

Labor economists have made tremendous efforts to measure human capital since the inception of this concept. Years of schooling and years of potential experience were originally used as proxies for human capital, and the subsequent development of the literature expands the scope of the measurement to include health status, cognitive skill, non-cognitive skill, and social skill (Bartel and Taubman, 1979; Deming, 2017; Griliches, 1977; Heckman et al., 2006; Mincer, 1974). In contrast to the attention given to measuring human capital, efforts to measure the usage of human capital have been limited, perhaps because a perfectly competitive labor market entails an efficient use of human capital through the price mechanism. This is Say's law in the labor market: An abundance of skill lowers the skill price and skill demanded increases along with the skill demand curve. Rejections of the efficient market, however, are paramount. For instance, Hsieh et al. (2019) report that 20–40% of US economic growth between 1960 and 2010 was induced by resolving the inefficient allocation of talents caused by the division of the labor market by race and gender. Accumulated evidence points to the prevalence of monopsony in the labor market (Ashenfelter et al., 2010; Dube et al., 2020; Manning, 2013). Recent studies, furthermore, have investigated the role of gender norms in the apparent underutilization of human capital to explain the persistent gender wage gap (Bertrand et al., 2010; 2015). Thus, direct measurement of skill use is indispensable to shed light on skill underutilization and skill mismatch in the labor market. This paper proposes a succinct measure of skill use based on detailed information about tasks implemented on the job, and demonstrates its benefit through examples.

The literature on over-education has attempted to measure skill use. In particular, this strand of literature attempts to measure the degree of skill under-utilization by the gap between skill and an occupation's skill requirement, typically approximated by the minimum requirement of educational attainment or the average educational attainment of workers in the same job. In this exercise, taxi drivers with a university degree are labeled as over-educated, because they have higher educational attainment than the minimum educational requirement, which is typically

☆ This work was supported by the Research Institute of Economy, Trade and Industry and Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research Grant Numbers 15H05692, 16H03630, and 16H06322. For valuable comments, we thank Taiyo Fukai, Maria Guadalupe, Rasmus Lenz, Shiko Maruyama, Mai Seki, Michael Waldman, and seminar participants at the University of Tokyo, the Japan Economics Association Meeting, the Kansai labor workshop, Seoul National University, National Taiwan University, RIETI, Musashi University, Chuo University, Australian National University, the Asian and Australasian Society of Labor Economics and University of Technology, Sydney.

* Corresponding author.
E-mail address: kawaguchi@e.u-tokyo.ac.jp (D. Kawaguchi).

https://doi.org/10.1016/j.labeco.2022.102197
Received 28 March 2021; Received in revised form 30 March 2022; Accepted 6 June 2022
Available online 11 June 2022

a high-school degree. Typical studies validate the measurement of over-education by showing that a negative correlation between the degree of over-education and wages. Leuven and Oosterbeek (2011) comprehensively surveyed the literature and criticized this approach, however, because selection into jobs with a low educational requirement is a mere reflection of the lower unobserved ability conditional on educational attainment, and the lower residual wage just reflects workers' lower ability. This critical survey article provides two important lessons for an attempt to measure skill use. First, skill use should be precisely measured based on detailed information about what tasks are implemented on the job to avoid a mere reflection of the endogenous selection into jobs based on ability. Second, the constructed measure of skill use should be validated against a measurement that is not a trivial reflection of ability.

To construct the skill-use measurement, we rely on microdata from the Programme for the International Assessment of Adult Competencies (PIAAC) compiled by the Organisation for Economic Co-operation and Development (OECD), covering more than 30 countries (of which we use 24 countries). The PIAAC is the best-suited micro data set for our purpose, because it includes measurements on both skills and skill uses: The PIAAC measures the literacy and numeracy skills of adults based on an on-site test, as well as the frequencies of implementing certain tasks requiring a specific skill, such as reading manuals/reference sources or calculating prices/costs. Based on detailed information about the frequency of engaging in various kinds of tasks, we construct measures of literacy and numeracy use drawing on Item Response Theory (IRT), which is widely used to identify the ability of students from their responses to examination items. Applying IRT enables us to construct objective measures of skill use on the job for each individual. The application of a uniform method across countries renders an internationally comparable measure. Furthermore, the proposed skill-use measures of literacy and numeracy exactly correspond to the skill measures of literacy and numeracy, constructed based on IRT, which enables an examination of the determinants of skill use conditional on the skill level.

We demonstrate the validity of the newly developed measure of skill use with three applications: The gender wage gap, the impact of parental leave on gender gap in skill use, and the wage gap between natives and immigrants.

The first application estimates the role of skill use in explaining the gender wage gap. Literature has demonstrated that the accurate measurement of skill is indispensable to estimate the residual gender wage gap. On the other hand, a strand of literature points to an under-utilization of skill due to the work hours constraint faced by females that causes the gender wage gap, particularly at the top end of the wage distribution (Goldin, 2014). To directly examine if the difference in skill use explains the observed gender wage gap, we examine how much the residual gender wage gap narrows by including skill use measures in addition to the conventional covariates that presumably capture workers' skill. We find that skills are roughly the same across genders, but there are substantial gender gaps in skill use in some countries. We also find that the gender gap in skill use, along with the gender gap in tenure, explains a substantial part of the gender wage gap in countries where observed gender wage gap is large such as Korea and Japan.

The second application focuses on assessing the impact of parental-leave policy on the gender gap in skill use. Motivated by the observation that generous parental-leave policies and the glass ceiling limit female career advancement in Nordic countries, some studies point to the backlash effect of parental-leave policies on female career advancement. A few studies indeed report adverse effects of parental-leave policies on women's career advancement by exploiting international differences in the generosity of the policy. We collected parental-leave policies in 2011 from the relevant laws in each country, as well as the Working Conditions Laws Database of the International Labour Organization (ILO) and the OECD family database. Exploiting this database, we estimate the impact of the length of parental leave on the gender gap in skill use, paying attention to heterogeneous impacts by skill level. We find that the relationships between the length of parental leave and the gender gaps in skill use differ across skill levels. Among workers in the lowest-skilled group, the length of parental leave is positively correlated with the country-specific gender gap in literacy use. In contrast, among workers with higher skill levels, we find a negative association between the length of parental leave and skill use among higher-skilled women. We show that the relationships still hold after partialling out the effects of other institutions. Our findings suggest that expanding parental-leave policies entails a trade-off: On one hand, parental leave promotes the employment of the least-skilled women, who would otherwise drop from the labor market, while on the other hand, the policy hinders the intensive skill use of moderately skilled women. To highlight the benefit of using the skill-use index, we estimate the policy impact using conventional market outcomes, the employment, hours worked, and hourly wage. We find that the estimated effect of the parental-leave policy is not as clear cut as the results based on the skill-use indexes, assuring the benefit of the skill-use measure.

The third application analyzes the source of the wage gap between natives and immigrants. Numerous studies examining the wage penalty of immigrants point to skill mismatch as a source. To directly examine the role of skill mismatch as a source of the wage penalty of immigrants, we examine how much of the residual wage gap is explained by the skill and skill use measures. In 6 of the 16 countries where we found an immigrant wage penalty, the wage gap narrows substantially after controlling for observed characteristics. In these 6 countries, a non-negligible part of the reduction in the immigrant-native wage gaps is attributable to the difference in literacy utilization. We further demonstrate that the our skill use index is useful for shedding light on the mechanism behind the dynamic change of the wage gap between immigrants and natives. We show that the wage gap between immigrants and natives narrows as the number of years since immigration increases. Then, the decomposition shows that the gap explained by the gap in literacy use narrows as the number of years since migration increases. Overall, we demonstrate that the gap in skill use explains the variation in the wage penalty among immigrants relative to natives.

Through these three examples, we demonstrate the strength of the proposed skill-use measure. The test batteries to measure skill and skill use in the PIAAC, however, are lengthy and incorporating such test batteries in the conventional labor force survey or national census is not realistic. With this backdrop, we propose to merge skill and skill use measures based on common variables in the PIAAC and the conventional dataset. By using occupation codes, we can merge the skill use index with the conventional data set. This exercise is similar to assigning tasks to each occupation using O*net, as done by Autor et al. (2003). We indeed show that the literacy and numeracy use intensities in each occupation measured by the PIAAC and O*net are similar in the US. By using demographic variables such as education and potential experience, we can merge skill index with the conventional survey data. Merging skill and skill use indices with the conventional survey data allows researchers to shed light on the issues beyond the coverage of the PIAAC data set.

2. Measurements of skill and SKill use

2.1. PIAAC

We draw on the Programme for the International Assessment of Adult Competencies (PIAAC) by the OECD, which aims to measure adults' cognitive and workplace skills. Twenty-four countries participated in the PIAAC Round 1 (2008–2013), and 9 countries participated in Round 2 (2012–2016); participating countries in each round are tabulated in Table 1. Our analysis sample consists of all participating countries in Rounds 1 and 2, except for Australia and Indonesia, whose data sets are not provided for public use, and Russia, whose data set does not include Moscow residents. Accordingly, our analysis sample includes individuals from 30 countries, but 6 countries are excluded because they

Table 1
Participating countries in PIAAC.

Round 1 (2008–2013)	Australia, **Austria, Belgium**, Canada, Cyprus, **Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland**, Russian Federation, **Slovak Republic, Spain, Sweden, United Kingdom, United States**
Round 2 (2012–2016)	**Chile, Greece**, Indonesia, Israel, Lithuania, **New Zealand**, Singapore, **Slovenia**, Turkey
Round 3 (2016–2019)	Ecuador, Hungary, Kazakhstan, Mexico, Peru, United States

Note: The countries in bold text are used in our study.

lack some social-institution indices (See Section 3.2 for those indices).[1] Hence, our main analysis sample consists of the remaining 24 countries. The survey targets individuals ages 16–65 and collects basic background information, such as age, sex, and educational attainment.

A distinguishing feature of the PIAAC is that it tests literacy, numeracy, and problem-solving skills in technology-rich environments. None of the respondents completed all three test sections; rather, they completed two at most, where the sections are randomly assigned; possible combinations are "literacy and numeracy," "literacy and problem solving 1," "literacy and problem solving 2," "numeracy and problem solving 1," "numeracy and problem solving 2," and "problem solving 1 and problem solving 2." The fraction of respondents taking the problem-solving section tends to be small, because its assigning probability is lower than those of the other two test sections and because some countries opted out of it (including France, Italy, and Spain). We thus decided not to use the problem-solving section.

The PIAAC data set contains plausible values (PV), which are computed based on the test results and background information, such as sex and educational attainment (OECD, 2013). Since sex, which is the variable of interest in our analysis, is used to impute the PVs, we do not rely on those PVs, and instead calculate test scores based on Item Response Theory (IRT) by ourselves, as described in detail in the next subsection.

We restrict the sample to prime-age adults, those between 25 and 59 at the time of the survey, while all individuals taking the computer-based assessment are used to estimate the skill and skill-use indices.[2] We restrict the age range so that the sample construction is relatively free from school enrollment and retirement decisions. We exclude full-time students and the permanently disabled from the sample. Also, we exclude observations with missing values in the variables necessary for our analysis.[3]

2.2. Calculation of skill index

Obtaining a skill score from the PIAAC test module is tricky for two reasons. First, scores are not pre-assigned to each test item. Second, there are several testlets and one of them is randomly assigned to each respondent, where the assignment probability depends on one's demographics, so that those with high (low) educational attainment are likely to be assigned to a difficult (easy) testlet. As a result, the standard grading scheme, such as counting the number of correct responses, is not applicable. Instead, the PIAAC test module was designed to apply the IRT, which "assigns" scores to each test item in a data-driven way, and the IRT is the standard method to estimate latent ability in the educational psychology literature. Since some test items are good at discriminating high-ability respondents from low-ability respondents while other test items are not, test takers' response patterns are informative about how good a test item is to measure ability. Although the PIAAC has several testlets with different extents of difficulty, they have some common test items to make it possible to compare respondents with different testlets.

While the test scores (PVs) estimated by the IRT are included in the PIAAC, the PVs provided by OECD do not serve our purpose, because the PVs are calculated based on demographic variables in addition to the responses to the skill and skill-use questions. Thus, the OECD's PVs already reflect gender differences and using this as a dependent variable is tautological in certain applications. Therefore, we construct our own test scores that depend only on the responses to test items.

Following OECD (2013), we employ a two-parameter logistic model, which characterizes each test item by its "difficulty" and "discrimination." It specifies the probability of selecting the correct response as

$$\Pr(y_{ij} = 1 \mid a_j, b_j, \theta_i) \equiv \frac{\exp\left(a_j(\theta_i - b_j)\right)}{1 + \exp\left(a_j(\theta_i - b_j)\right)}, \tag{1}$$

where y_{ij} takes one if respondent i correctly answers test item j, and zero otherwise, and θ_i is the latent trait of respondent i. Each test item j is characterized by two parameters: a_j, the "discrimination" parameter of item j, which represents the sensitivity of being correct to the ability; and b_j, which represents the "difficulty" that shifts the probability of being correct irrespective of the ability. This specification assumes that test items measure the uni-dimensional latent trait summarized by θ_i, and that observed item responses are independent, conditional on the latent trait, θ_i. In fact, test items in the PIAAC are designed to apply this model, such that each question is independent of the others. Letting $y_i = (y_{i1}, \cdots, y_{iJ})$ and $B = (a_1, \cdots, a_J, b_1, \cdots, b_J)$, the conditional distribution for respondent i is denoted as

$$f(y_i \mid B, \theta_i) = \prod_{j=1}^{J} \left[\Pr(Y_{ij} = 1 \mid a_j, b_j, \theta_i)\right]^{y_{ij}} \left[1 - \Pr(Y_{ij} = 1 \mid a_j, b_j, \theta_i)\right]^{1-y_{ij}}. \tag{2}$$

Given the prior distribution of the latent trait θ_i, which is assumed to follow the standard normal distribution, $\hat{B}$ is chosen to maximize the log-likelihood,

$$\ln L(B) = \sum_{i=1}^{N} \ln\left(\int f(y_i \mid B, \theta) d\Phi(\theta)\right), \tag{3}$$

where Φ is the standard normal distribution function.

Finally, the latent trait parameter θ_i, is estimated using Bayes' theorem; its immediate application gives the posterior distribution of the latent trait, θ_i, conditional on the estimated parameters and response patterns. Then, the empirical Bayes mean (or posterior mean) of θ_i is

$$\tilde{\theta}_i = \int_{-\infty}^{\infty} \theta\phi(\theta \mid y_i, \hat{B}) d\theta = \int_{-\infty}^{\infty} \theta \frac{f(y_i \mid \hat{B}, \theta)\phi(\theta)}{\int f(y_i \mid \hat{B}, \theta)\phi(\theta) d\theta} d\theta. \tag{4}$$

We estimate the latent parameters for each country, allowing discrimination and difficulty parameters to differ across countries. To facilitate the interpretation, we normalize the estimated skill indices so that they each have exactly zero mean and one standard deviation. A set of 49 test items is used to estimate the literacy skill score, and another set of 49 test items is used to estimate the numeracy skill score.

[1] We obtained the German scientific-use file from GESIS.

[2] If a respondent does not have the basic ability to use a computer or if he/she refuses to use a computer, he/she takes a paper-based assessment (PBA). OECD (2013) suggests that the computer-based tests and the paper-based tests are comparable. Our results are robust to the use of the PBA sample.

[3] Item non-response is relatively rare, and the fraction of observations dropped due to some missing values is limited to 3.3 percent.

2.3. Calculation of skill-use index

In addition to skill possession, respondents in the PIAAC report their skill use at work with well-defined responses, which enable us to compute the latent traits for skill use. For example, they are asked, "In your job, how often do you usually read directions or introductions?" for use of literacy skill, and "In your job, how often do you usually calculate prices, costs or budgets?" for use of numeracy skill. Respondents answer these questions using a five-point frequency scale: (1) Never, (2) Less than once a month, (3) Less than once a week but at least once a month, (4) At least once a week but not every day, or (5) Every day. There are 8 items for literacy use and 6 items for numeracy use. (See Appendix A for details.) These responses are more objective than responses such as "often" and "rare," because the measurement units are well defined. Note that the workers who work short hours mechanically use the skill less frequently. Thus, the skill use measure captures the effect through the length of hours worked.

Using this information, we apply the general partial credit model (GPCM; Muraki, 1992), which is an extension of the two-parameter logistic model to the ordered responses to each set of skill-use items. Then, we obtain two skill-use indices for each respondent as the empirical Bayes means of the posterior distribution of latent skill-use intensity; i.e., skill use of literacy and skill use of numeracy. The skill-use indices are normalized to have a zero mean and one standard deviation.

2.4. Validation of skill and skill-use indices

Before conducting a detailed analysis using these skill and skill-use indices, we check their validity by examining whether they are correlated with conventional proxy variables for each worker's productivity or career advancement. We restrict the analysis sample to men to abstract gender issues away and to mitigate possible selection biases. Figs. 1 and 2 illustrate the relationship between the occupation-average hourly wages and literacy skill and skill use in each country, where the size of the circles indicates the number of observations in each occupation. The figures demonstrate the positive correlation in all countries, suggesting that occupations with skilled workers or intensive skill use are associated with higher wages. This positive correlation between wages and skill and skill use ensures that skill and skill-use measures carry substantive information correlated with wages, the conventional proxy for productivity.

We further demonstrate that literacy use is closely related to occupation, by examining the literacy-use distribution across occupations. Fig. 3 shows that professionals stochastically dominate other occupations, followed by managers, armed forces, and technicians, whereas the elementary occupations use literacy skill the least frequently. Hence, our literacy-use score well reflects across-occupational differences.

Given the high correlation of literacy skill and occupations, how much does skill use explain the wage variation within an occupation? To address this question, we estimate the following equation:

$$\ln(wage)_{ij} = \beta^s Skill_{ij} + \beta^{su} SkillUse_{ij} + X_{ij}\beta^x_j + \lambda_{s(i),j} + u_{ij}, \tag{5}$$

where i and j indicate each individual and country; X_{ij} include age indicators, years of education, and dummy variables, indicating that the test language is the same as the respondent's native language and that parents are immigrants; and $\lambda_{s(i),j}$ is country-occupation fixed effects, with $s(i)$ indicating individual i's occupation in country j. We estimate the model with and without $\lambda_{s(i),j}$. The estimates demonstrate, for example, that one-standard-deviation increases in literacy skill and skill use are associated with 6.0% and 9.8% increases in hourly wages, respectively. Since both the skill and skill-use measures are normalized, the estimated coefficients are comparable and it is notable that skill-use explains more wage variation than skill. The estimated impacts are dampened to 4.8% and 6.6% after controlling for occupation dummy variables as shown in Column 2 of Table 2. The change of the results shows that both skill and skill-use measures capture substantial wage variation within occupations. We repeat the same exercise using numeracy and numeracy use and obtain similar results to those reported in Columns 3 and 4 of Table 2.[4] Finally, the joint distribution of skill use and wage demonstrates that the density is high at the diagonal part and thus, the rank of the wage is closely related to the rank of the literacy use (Fig. 4). In particular, top (bottom) wage earners are likely to be those who use literacy skill most (least) frequently.

Table 2
Regression estimates of hourly wages on skill and skill use.

Dep.Var. $ln(wage)$	Skill: Literacy		Skill: Numeracy	
	(1)	(2)	(3)	(4)
Skill	0.060***	0.048***	0.057***	0.042***
	(0.005)	(0.005)	(0.005)	(0.004)
Skill-use	0.098***	0.066***	0.071***	0.040***
	(0.005)	(0.006)	(0.004)	(0.004)
Occupation		X		X
Observations	12,790	12,790	12,642	12,642
Countries	21	21	21	21

Note: This table shows the estimation results of Eq. (5). We did not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language was the same as the native language of the respondent, or that parents were immigrants. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Overall, both skill and skill-use indices constructed by IRT using the PIAAC are strongly associated with the conventional labor market variable, namely, hourly wage. First, both indices are strongly associated with occupations. Second, these indices are associated with wages within occupations. Third, the skill-use index explains wage variation even after conditioning on the skill index. The last finding suggests the importance of skill-use in addition to skill as a wage determinant.

3. Applications

This section demonstrates the usefulness of the newly developed skill-use index through three applications: the gender wage gap, the effect of parental leave, and the wage gap between immigrants and natives.

3.1. Gender wage gap

As the first application of the newly developed skill-use measure, we consider explaining the gender wage gap. Literature has demonstrated that an accurate measurement of skill is indispensable to credibly identify the gender wage gap not explained by the gender gap in skill. For example, Weichselbaumer and Winter-Ebmer (2005) reviewed the studies around the world and showed that the lack of actual years of labor-market experience results in an over-estimation of the gender wage gap. Furthermore, they point out that the lack of skill variables, such as on-the-job training, biases the gender wage gap. On the other hand, Blau and Kahn (2017) examined the evolution of gender wage gap in the US and reported that the gender wage gap remains large at the top end of the wage distribution and pointed to the importance of women's work force interruptions and shorter hours. This claim arguably suggests that the skill of high-skilled women is not fully utilized on the job. Since both skill and skill use are suggested as important determinants of the gender wage gap, examining gender wage gap conditional on skill and skill-use measures using PIAAC sheds new light on the mechanism behind the observed gender wage gap.

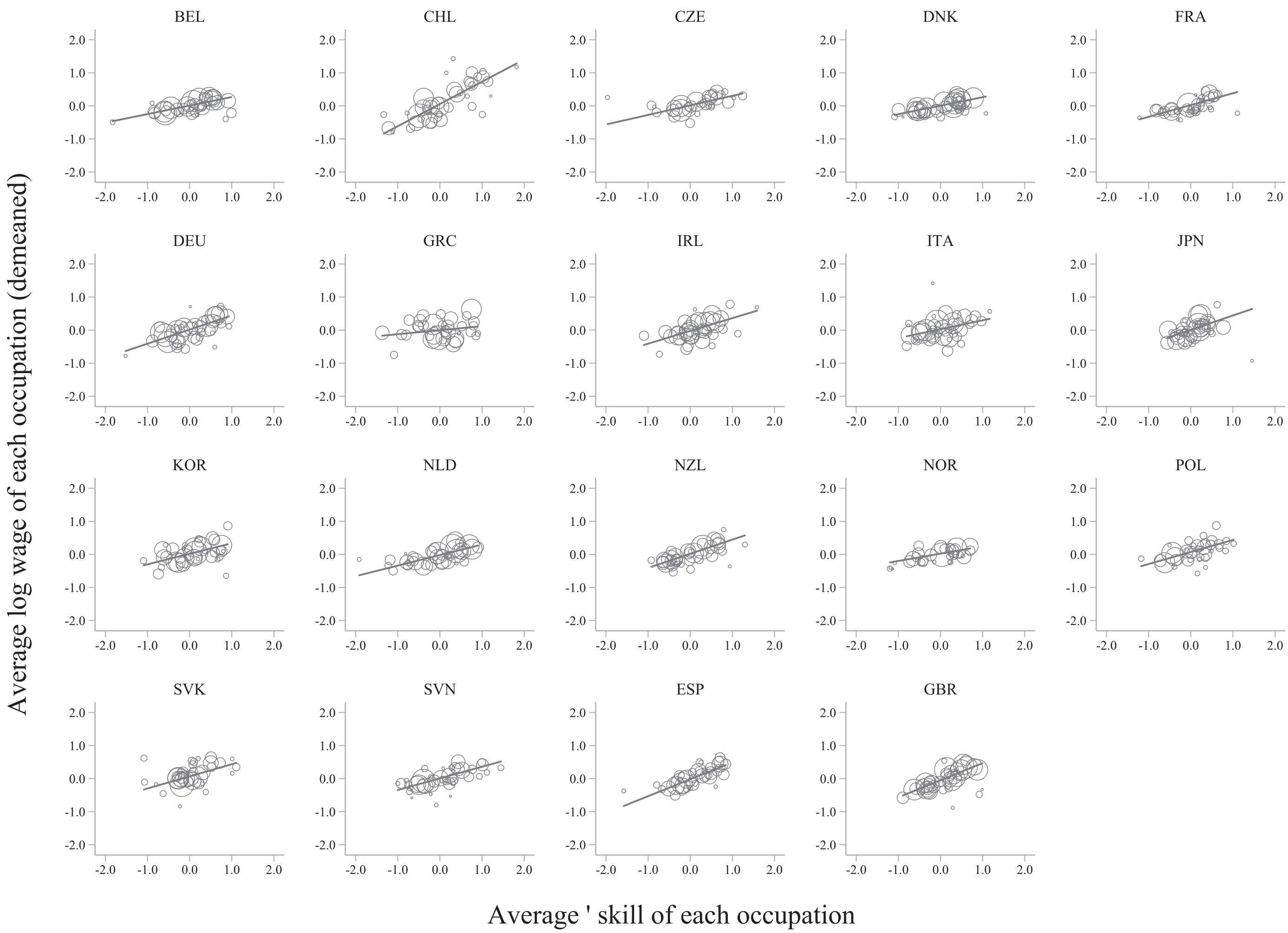


Fig. 1. Occupation-average wage rates and literacy skill. Note: This figure shows the correlation between occupation-average wage rates and average literacy skill. The size of each circle indicates the number of observations engaging in each occupation. The line is the fitted value by the weighted least squares, where the number of observations in each occupation is used as a weight.

Fig. 5 summarizes the gender differences in skill and skill use, where each point is the gender gap of skill or skill use and the bars indicate the 95% confidence intervals. Literacy scores are roughly the same across genders, though women's scores tend to be slightly lower than men's in some countries. In contrast, the gender gaps in literacy use scores are substantially different across countries: Women use literacy more in Poland and Slovenia and use it less in Japan, Korea, Netherlands, and Norway. In terms of numeracy, women tend to score lower and use it less at work than men. From casual observation, gender gaps in skill use tend to be small or reversed in ex-communist countries, such as Poland and Slovakia.[5] de Haan (2012) documents that these countries encourage women to participate in the labor market by providing opportunities for education and training to meet the demands of labor-intensive industries under socialist regimes.

While the international variation in gender skill-use gaps is notable, the gender gaps in skill use in this figure should be interpreted with caution, because the index is not separable from the labor supply at both the extensive and intensive margins. At the extensive margin, only labor market participants are asked about their skill use at work. Therefore, the gender difference in skill use reflects the gender difference in selection into the labor market. In particular, if the female labor force participation rate is low and working women are positively selected, it will reduce the observed gender gap in skill use (unconditional on skill level). At the intensive margin, the difference in skill use partially reflects the difference in the hours worked. Often the jobs requiring intensive skill use, such as management or professional jobs, entail long working hours or inflexible work schedule (Blau and Kahn, 2013; Goldin, 2014), and thus the short working hours can well be a fundamental source of skill under-utilization. Therefore, the gender gap in skill use conditional on the gap in hours worked is *not* necessarily a better measure of the gender gap in skill use than the unconditional measure. With this caveat, we calculate the gender gap in skill use, conditional on hours worked, by regressing the raw skill use measure on weekly hours worked and recording the residuals. Fig. 6 illustrates the hours adjusted skill use. The figure shows that gender gaps in literacy and numeracy use narrow substantially from raw gender gaps in skill use. This result suggests that short working hours are an important source of the gender gap in skill utilization.

For brevity of exposition, the following analysis focuses on literacy skill and its utilization, instead of those of numeracy. The choice of literacy over numeracy is partially based on the concern that numeracy skill is acquired by taking labor-market prospects into consideration. The usage of numeracy seems limited to market production, in comparison with the usage of literacy, which applies to both market and household

[4] Note that we cannot include both literacy and numeracy indices in the same model as the respondents for literacy and numeracy tests are different.

[5] We define ex-communist countries as including Czech, Estonia, Poland, and Slovakia.

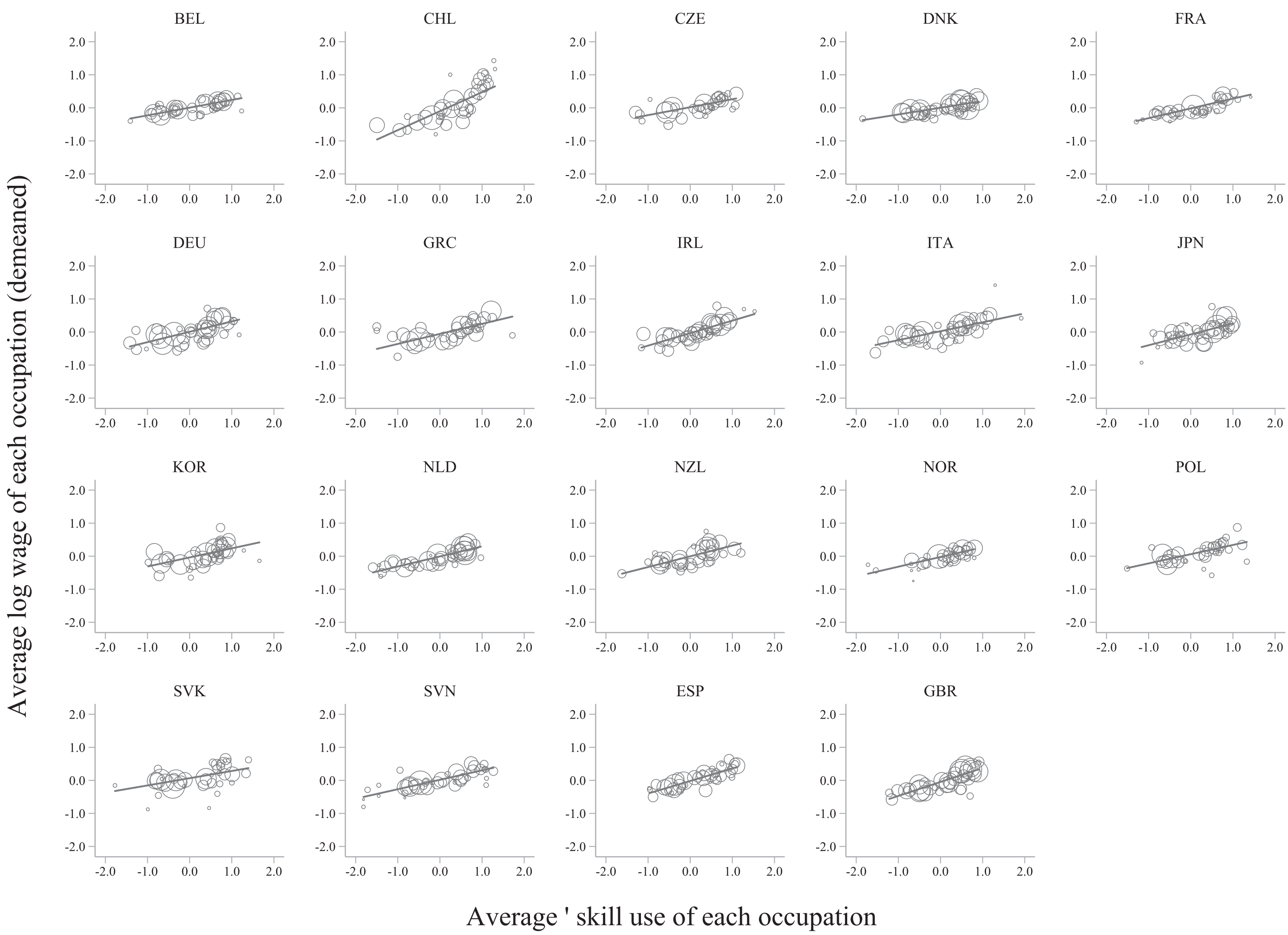


Fig. 2. Occupation-average wage rates and literacy skill use. Note: This figure shows the correlation between occupation-average wage rates and average literacy skill use. The size of each circle indicates the number of observations engaging in each occupation. The line is the fitted value by the weighted least squares, where the number of observations in each occupation is used as a weight.

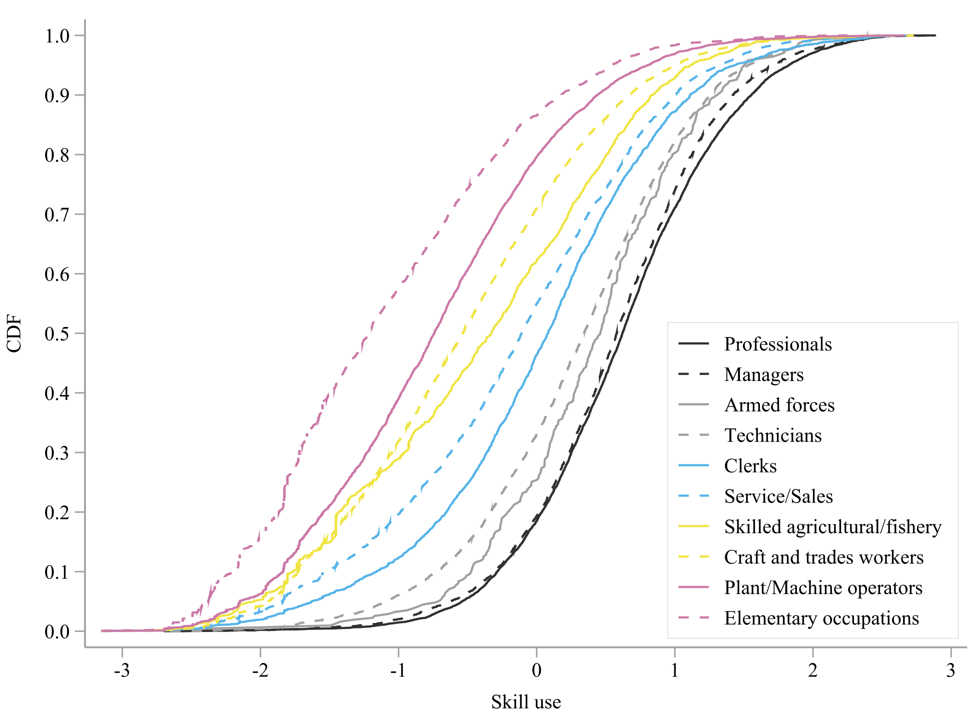


Fig. 3. CDF of literacy use at work by occupation. Note: This figure shows the CDF of literacy use at work by each occupation.

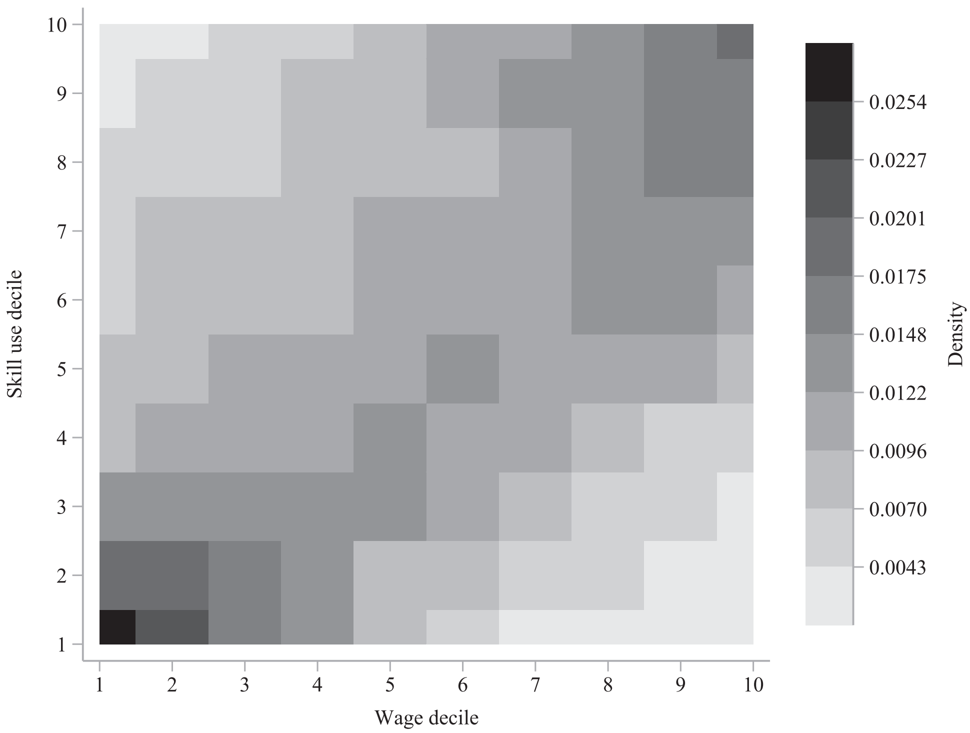


Fig. 4. Joint distribution of literacy use at work and wage. Note: This figure shows the joint density of literacy use at work and wage.

production. As a result, women with high numeracy skill might differ from other women in unobserved ways, such as attitudes toward work (Fryer and Levitt, 2010; Guiso et al., 2008; Nollenberger et al., 2016). Furthermore, items that measure numeracy skill use (e.g., use of algebra) do not seem to be as general as items that measure literacy skill use. In fact, Table 2 shows that literacy skill use is more closely correlated with wage rates than numeracy skill use. Note that we did not choose to use both literacy and numeracy tests because only a portion of respondents take both literacy and numeracy tests and thus the effective sample size decreases significantly.

The analysis of the skill and skill-use indices indicates that there are substantial gender gaps in skill use in some countries. How much does skill under-utilization explain the gender wage differentials that are also known to be different across countries? To address this question, we examine how much the gender wage gap narrows by adjusting for the literacy-utilization index along with the standard demographic variables and the literacy-skill index.

The unadjusted (raw) gender wage gap is estimated as the coefficient for the female dummy variable F_i in the following model:

$$\ln(wage_i) = \beta_0^{base} + \beta_1^{base} F_i + u_i. \tag{6}$$

In contrast, with the vector of control variables including literacy and literacy-use indices, $X_i = [X_{1i} X_{2i} \dots X_{ki}]$, the adjusted gender wage gap is estimated as the coefficient for F_i in the following model:

$$\ln(wage_i) = \beta_0^{full} + \beta_1^{full} F_i + \sum_{j=1}^{k} X_{ji}\gamma_j + v_i. \tag{7}$$

Gelbach (2016) articulates that $\hat{\beta}_1^{full} - \hat{\beta}_1^{base} = -\sum_{j=1}^{k} \hat{\delta}_j \hat{\gamma}_j$ where $\hat{\delta}_j$ is the regression coefficient of X_{ji} on F_i due to the standard omitted variable bias formula. He proposed to use this formula to decompose the difference of the unadjusted and adjusted wage gaps. The control variables include age, immigration status, educational background, job tenure, literacy skill, and literacy skill use.

The estimates of the unadjusted and adjusted gender wage differentials are reported in Fig. 7. Panel (a) compares the unadjusted and adjusted gender wage differentials. The adjusted gender wage gaps are smaller than the unadjusted ones in 11 countries, implying that the gender gap in covariates explains a part of the unadjusted gender wage gap. Among these countries, the difference between the unadjusted and adjusted wage gap is particularly notable in Japan and Korea. In Japan, the unadjusted gap is 45 percent, and it shrinks to 24 percent after conditioning on demographics, skill, and skill use. In Korea, the unadjusted gap is 26 percent, and it shrinks to 14 percent after conditioning on demographics, skill and skill use. In other words, about one half of the raw gender gap is explained by the gender difference in demographics, skill, and skill use in these countries.

Panel (b) of Fig. 7 shows the result of Gelbach (2016)'s decomposition. In most countries of our sample, education contributes to expand the gender wage gap. Since the educational attainment of women exceeds that of men in many developed countries, conditioning on the level of education rather expands the gender gap. Exceptions are Japan and Korea, however, where conditioning on education shrinks the gender wage gap but its contribution is relatively minor, while the contribution of tenure and literacy use is large. Interestingly, the decomposition results of the gender wage gap in Japan and Korea are quite similar. About 50 percent of the explained gender gap is attributable to tenure and about 25 percent is attributable to literacy use. In other countries, the magnitude of the explained gap itself tends to be small, but the contribution of literacy use to the explained gap is not necessarily small. For instance, the contribution of literacy use is around 60 percent in Belgium and the UK, and almost all of the explained gap is attributed to literacy use in Norway, while in some countries such as, Finland and Italy, literacy use plays a relatively minor role. Overall, our decomposition analysis suggests that acquired human capital is not necessarily utilized in the market, and indeed, the gender gap in skill utilization is reflected in the gender wage gap at least to some extent. Therefore, this exercise suggests that measuring human capital is not sufficient, but we need to consider how the human capital is used to better understand wage determinants in the labor market.

3.2. Parental leave and women's skill utilization

We next demonstrate the strength of the skill-utilization measure in the context of the assessing the impact of parental-leave policy on

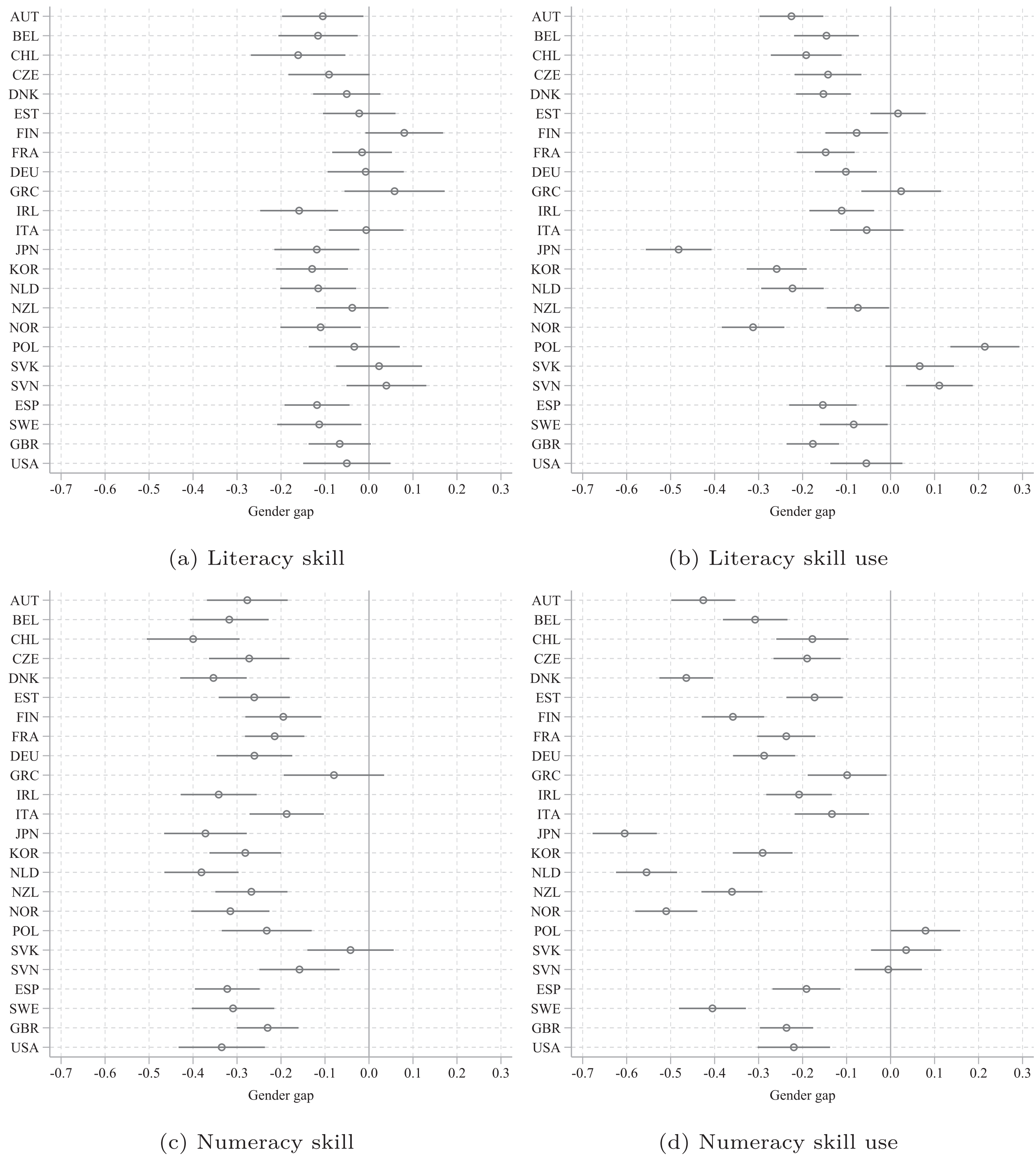
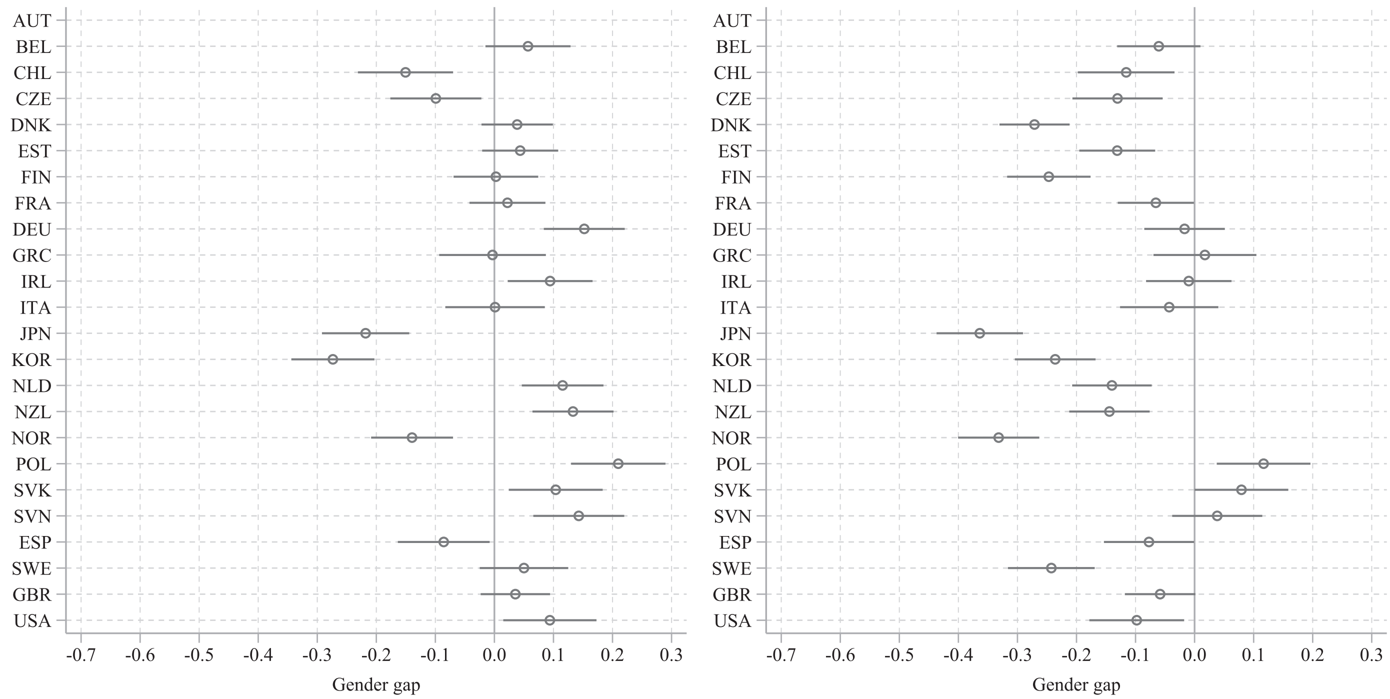


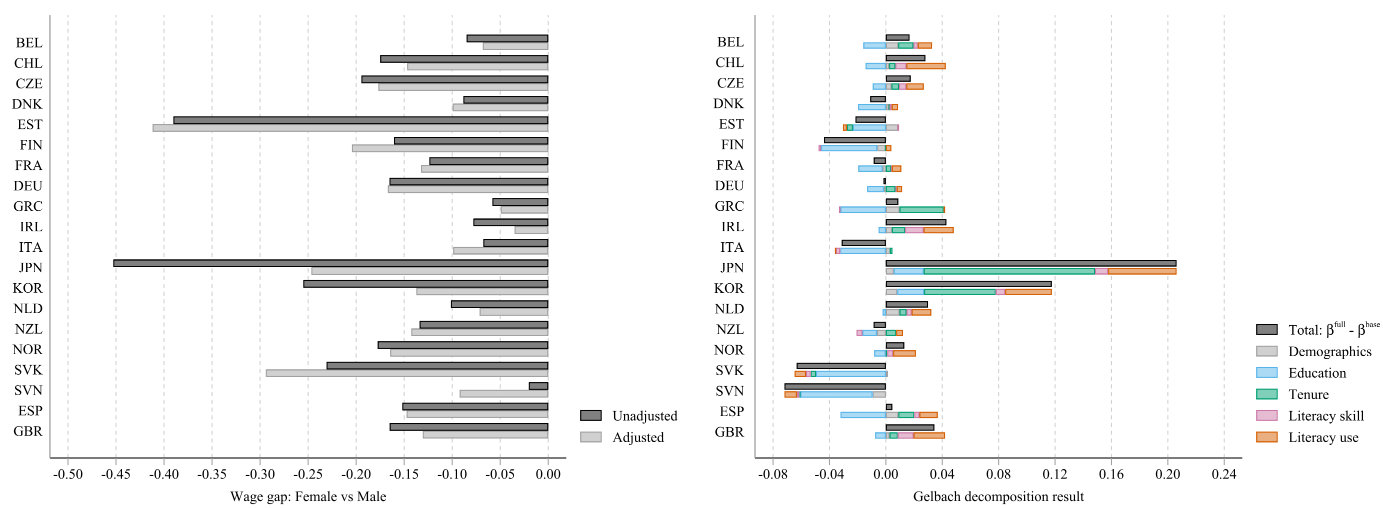
Fig. 5. Gender gaps in skill and skill use: Females vs Males. Note: This figure shows the unconditional gender gaps in skill and skill use. Each point represents the gender gap, and the bars indicate its 95 percent confidence interval.

the gender gap in skill use. Motivated by the observation that generous parental-leave policies and the glass ceiling limit female career advancement in Nordic countries, some studies point to the backlash of parental-leave policies on female career advancement, because females are put on "mommy track" (Albrecht et al., 2003; 2015; Datta Gupta et al., 2008). A few studies indeed report adverse effects of parental-leave policies on women's career advancements by exploiting international differences in the generosity of the policy. Blau and Kahn (2013) report that a generous parental-leave policy increases female labor-force participation but decreases full-time employment among women, as well as the fraction of managers and professionals, based on cross-country data from 22 OECD countries. Olivetti and Petrongolo (2017) report that parental-



(a) Literacy use (b) Numeracy use

Fig. 6. Skill use gap: Females vs Males (with skill use residualized by work hours). Note: This figure shows the unconditional gender gaps in skill and skill use. The skill use score is residualized by work hours by each country. Each point represents the gender gap, and the bars indicate its 95 percent confidence interval.



(a) Unadjusted and adjusted log wage gap (b) Decompsotion result

Fig. 7. Decomposition of the wage gap: Females vs Males. Note: This figure shows the (un)adjusted wage gap between males and females, and demonstrates Gelbach (2016)'s decomposition result. Demographics includes age and immigration status.

leave policies increase the employment of low-educated women but decrease the wages of high-educated women, based on cross-country data from 30 OECD countries.

Thomas (2018) further investigates the mechanism with which the parental leave backlashes the maternal career through the lens of statistical discrimination. In the presence of asymmetric information on a worker's preference, an employer provides training to (or promotes) a female worker only if her production output (which is a noisy signal about her preference) exceeds a certain threshold, and since the parental-leave policy increases the employer's risk of training/promoting non-career-oriented female workers, the employer raises the threshold for training/promotion. Therefore, her model suggests that the parental-leave policy has a negative impact, particularly on female workers around the training/promotion threshold, and she provides some empirical evidence in line with the model prediction.

We push forward the analysis using the newly developed skill-use measure and demonstrate that the impacts are more clearly detected based on the new measure than on traditional measures of labor-market outcomes, such as employment or wages. Furthermore, the skill measure in our dataset is useful to detect heterogeneity across the skill distribution. On the one hand, the glass ceiling literature suggests that the negative impact of the parental leave is severe particularly among very high skilled women, but on the other hand, the statistical discrimination model provided by Thomas (2018) suggests that the negative impact is severe among moderately skilled women.

We collected parental-leave policies in 2011 from the relevant laws in each country, as well as the Working Conditions Laws Database of the International Labour Organization (ILO) and the OECD family database. See Appendix B for a full description of the data sources. We define the duration of parental leave as the sum of maternity and parental leave duration, in years, in a particular country. To be sure, these two policies are distinct, in the sense that maternity leave is given only to women, while parental leave is gender neutral; in reality, however, the parental leave is most likely to be taken by women in many countries.[6]

Since parental-leave policies have two functions, job protection and income compensation, we measure these aspects by the duration of paid leave and the duration of job protection. Figure C1 summarizes the duration of parental leave in each country in 2011 in terms of the paid parental-leave period and the job-protection period. We confirm that the paid-leave policy has sufficient variation across countries, and many countries support substantially long job-protection periods that extend more than three years, while some of them, such as Finland, France, and Spain, provide cash benefits for less than one year.

Since we implement cross-country comparisons that associate the length of parental leave with women's skill utilization, the correlation may be driven by gender norms or other market institutions that affect both the policy and the outcome. To control for those institutions, we construct a quantitative measure of the strength of traditional gender norms using internationally comparable social surveys: the World Values Survey Wave 6 and the European Values Study 2008.[7] We further collect other quantitative indicators for social institutions, such as tax policy[8], child care policy, the strength of employment protection, and the unionization rate from the OECD database. In addition, following Blau and Kahn (2013), we construct the indicators for the right to part-time work and equal treatment of part-time workers from OECD (2010). Since the industrial structure could affect both the policy and the outcome, we control for the fractions of public-sector employment and service sector employment, respectively, which are calculated using the PIAAC. Table C1 provides summary statistics for these indices.

Fig. 8 displays the relationship between the literacy-use gender gap and the paid-leave policy by each literacy quartile group, where these quartile groups are defined by each country.[9] We exclude ex-communist countries in these figures because their social institutions are different from those of other countries (de Haan, 2012), but include these countries in the regression analysis, which controls for various aspects of social institutions. The gender gap is measured by subtracting the average skill-use levels of men from women's; Thus, the negative value indicates that women tend to use less skill than men. Among workers in the lowest-skilled group, the length of paid parental leave is positively correlated with the country-specific gender gap in literacy use. In contrast, among workers with higher levels of literacy (literacy levels: Q3 and Q4), the longer the length of paid parental leave, the larger are the gender gaps in literacy use.

We next investigate whether the relationships observed in Fig. 8 hold after partialling out the effects of other institutions. We also incorporate the non-working population in the analysis. Since those who are not working use no skills for market production, their skill-use scores are lower than the lowest values observed among those in the labor force. Note that we do not attempt to measure the *potential* literacy use of non-participants that would be realized if they worked in the market. We instead measure *actual* skill use in the labor market. Hence, the skill-use indices are considered to be left-censored, where the threshold, the minimum value of literacy use among labor-force participants, varies across countries. Since the censored Tobit model takes into account non-utilized skill due to non-participation as well as skill use within the market, it captures both the extensive and intensive margins.

Using the Tobit model, we estimate the effect of the literacy score on literacy use by regressing the literacy-use score on the dummy variables, indicating the literacy-score quartile. We examine the difference of the relationship between the literacy score and literacy use by gender and the length of parental leave by interacting the female dummy variable and the length of leave with the dummy variables for the literacy-score quartiles. Specifically, we estimate the following model, pooling all individuals from the sample countries:

$$\begin{aligned} y^*_{ijs} = & \sum_{q=1}^{4} 1\{q=s\}\cdot(\beta_{0q} + \beta_{1q} Female_i + \beta_{2q} Female_i \times PL_j \\ & + \beta_{3q} Female_i \times Inst_j + x_i'\beta_{4q} + c_{js}) + u_{ijs}, \end{aligned} \tag{8}$$

and the latent skill-use level is observed if it exceeds a certain threshold;

$$y_i = \begin{cases} y^*_{ijs} & \text{if} \quad y^*_{ijs} > y^L_j, \\ y^L_{ijs} & \text{if} \quad y^*_{ijs} \le y^L_j, \end{cases} \quad u_{ijs} \mid Female_i, s, x_i, c_{js} \sim N(0, \sigma^2_j), \tag{9}$$

where i, j, and $s \in \{1,2,3,4\}$ indicate individuals, countries, and skill quartile groups, respectively. The threshold y^L_j is the minimum of skill-use score among those who are employed in country j. The indicator function, $1\{q=s\}$, takes one if individual i's literacy skill belongs to the literacy quartile q; PL_j is the duration of the parental leave of country j measured in two ways: paid-leave length and job-protection period; and $Inst_j$ is the vector of institutional variables of country j, including an ex-communist dummy variable, the childcare center utilization rate, an index of the tax system, public sector size, service sector size, an index of employment protection policy, and union density. The vector x_i is a collection of individual characteristics, which include age indicators, years of education, and immigrant status. Country × skill quartile group fixed effects, c_{js}, captures the country-specific relationship between the literacy-skill quartile and literacy use.[10] We report the clustering robust standard errors by the cell, defined by country times skill-quartile group.

Table 3 shows the estimation results of the Tobit model consisting of Eqs. (8) and (9), using the duration of paid leave as the measurement of parental-leave policies. The basic specification in Column 1 does not control for the gender-specific institutional term (i.e., $Female_i \times Inst_j$) except for parental leave, the ex-communist dummy variable, and the size of the public and service sectors. The result shows that one-year-longer parental leave does not affect the gender gap in skill use at the

[6] As a caveat, we note that some US states, such as California, have more generous parental leave policies, and that some other countries may have similar within country variation. In general, we do not use the variation of the policy within a country because of the absence of a systematic data set. In addition, some private companies offer more generous parental leave than the national requirement as a part of their compensation package to attract workers. In this case, the variation of the generosity among companies facing the same legal requirement is highly endogenous reflecting the unobserved determinants of workers' quality as in the usual empirical work on the compensating wage differentials. Thus it is not trivial to estimate the impact of company-level parental-leave policy.

[7] Both surveys asked "When jobs are scarce, should men have more right to a job than women?" with possible responses "Agree" (= 1), "Neither" (= 0) and "Disagree" (= −1). We defined the index as the average of individual responses within each country.

[8] Since characteristics of the tax system depend on the levels of earnings, the OECD evaluates it at 133% and 200% of the mean earnings of a single household. Although we collect the index evaluated at 200%, the differences associated with this choice are minor and the qualitative argument is unaffected.

[9] See Figure C2 and Table C2 for the analysis in terms of numeracy skill and skill use.

[10] One might argue that the effects of the paid parental leave policy at its introduction and the extension of the period are different. This is probable but our data set does not allow us to estimate such non-linearity because most OECD countries have at least 6 months parental leave policy.

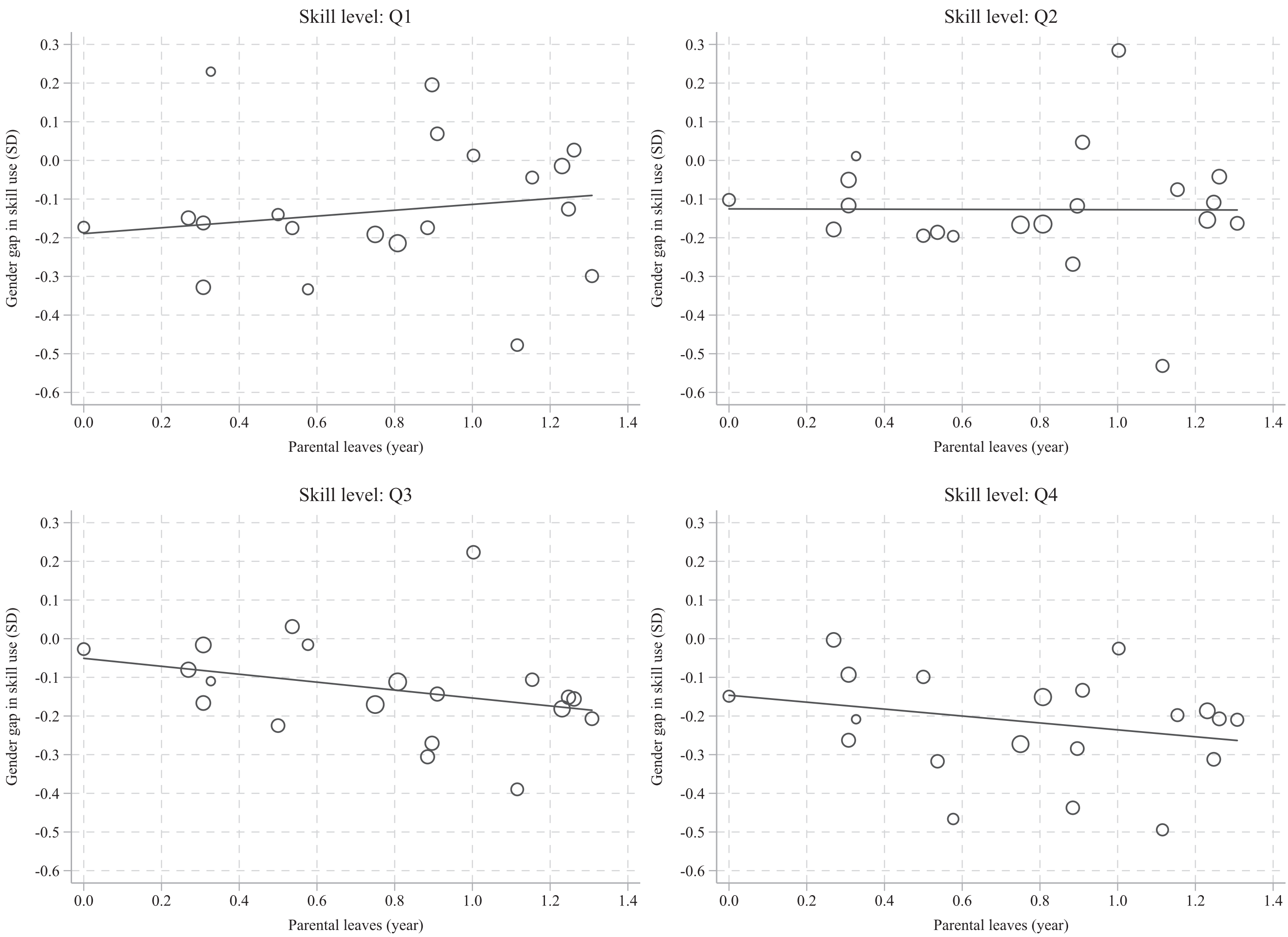


Fig. 8. Unconditional gender gap in literacy skill use and the paid-leave policy. Note: This figure shows relationship between the gender gap in literacy skill use and the paid-leave policy. The gender gap in each country is calculated as a raw difference in average skill-use levels between employed women and men. The line is the fitted value by the weighted least squares, where the number of observations in each country is used as a weight. In this figure, ex-communist countries are excluded, because their social institutions tend to differ from those of other countries (de Haan, 2012).

Table 3
The paid leave policy and utilization of literacy skill at work.

Dep.var. literacy skill use	Full sample				Employed
	(1)	(2)	(3)	(4)	(5)
Female×PL×Literacy skill: Q1	0.072 (0.060)	0.008 (0.034)	0.056 (0.045)	0.034 (0.054)	-0.043 (0.049)
Female×PL×Literacy skill: Q2	-0.173*** (0.064)	-0.196*** (0.066)	-0.171** (0.080)	-0.160*** (0.062)	-0.095* (0.053)
Female×PL×Literacy skill: Q3	-0.238*** (0.064)	-0.292*** (0.064)	-0.250*** (0.070)	-0.301*** (0.083)	-0.099* (0.051)
Female×PL×Literacy skill: Q4	-0.109* (0.058)	-0.095 (0.059)	-0.037 (0.064)	-0.033 (0.059)	0.003 (0.027)
Country×Skill quartile FE	X	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X	X
Female×Skill×Family policies		X	X	X	X
Female×Skill×Gender norm			X	X	X
Female×Skill×Market institutions				X	X
Countries	24	24	24	24	24
Observations	48,970	48,970	48,970	48,970	41,223

Note: This table shows estimation results of the censored Tobit model consisting of equations (8) and (9) for literacy score. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

lowest skill quartile. In contrast, longer parental leave widens the gender gap in literacy usage at the second, third, and fourth skill quartiles. For example, one-year-longer parental leave widens the gender gap of skill use by 0.24 standard deviation at the third quartile of the skill distribution. Adding other institutional variables as control variables does not significantly change the estimated impacts as reported in Columns 2–4. For example, the estimated coefficients for the 2nd and 3rd skill groups range between - 0.16 and - 0.20, and - 0.25 and - 0.30, respectively. Thus, we robustly find that longer parental leave widens the gender gap in skill use among modestly skilled women, suggesting that the longer parental leave reinforces the division of gender roles.[11] As shown in Table C4, our result is robust to use of quintiles instead of quartiles.

The estimation results reported in Columns 1–4 are the Tobit results using the sample that includes those who do not work. Thus, the estimated coefficients capture the mixture of the extensive and intensive margins. To focus on the impact at the intensive margin, we estimate the same model by OLS, using those who work as the analysis sample, and the estimation result is reported in Column 5 of Table 3. The estimated coefficient for the first quartile is negative and not statistically significant. At the second and third quartiles, the estimated coefficients become attenuated, but the estimated coefficients remain statistically significant with about two-thirds and one-third of the full sample estimate in Column 4, respectively. This implies that longer parental leave suppresses the skill use of higher-skilled women not only at the extensive margin but also at the intensive margin. According to the most preferred specification reported in Column 4, a three-month-longer paid leave, for example, decreases literacy-skill use by 0.075 SD. Considering that the gender gap in literacy use is, on average, around 0.1 SD among these skill groups, the size of the impact is substantial.

We discuss the potential mechanisms behind the robust finding that generous parental-leave policies affect the skill use of women heterogeneously across skill levels. Parental-leave policies could affect women's skill utilization through at least two channels: 1) job protection, and 2) statistical discrimination. First, the job protection provided by parental-leave policies allows women who would otherwise drop out of the labor force to continue working. Considering that lower-skilled workers generally have weaker labor-force attachment than higher-skilled workers, the job protection would be effective for lower-skilled workers. To test this hypothesis, we use the length of job protection, instead of the length of paid leave, as the independent variable and report the regression results in Table 4. Indeed, job-protection policies narrow the gender gap of literacy use at the bottom skill quartile, but this effect disappears when we focus on the employed population (Column 5 in Table 4). We observe a similar tendency in terms of other market outcomes. Thus, the positive effect of parental-leave policies on the least-skilled group is arguably driven by job protection, while other mechanisms could also work for those in the market.

Generous parental-leave policies can potentially encourage employers to statistically discriminate against women of a certain type, as discussed earlier, and indeed, our empirical evidence, at least, does not contradict the prediction that parental leave leads those on the verge of promotion threshold to a non-career track, while those in the top of skill distribution is unaffected.[12] Since a source of statistical discrimination against women is the gender-biased feature of parental leave, it is likely to be mitigated when more men take up the parental leave policy.

To highlight the benefit of using the skill-use index as the outcome variable, we estimate the policy impact using conventional labor market outcomes: the employment, hours worked, and hourly wage. Table 5 tabulates the estimation results of the same model using employment status, hours worked, and hourly wage as the dependent variables. Column 1 shows that the paid-leave policy does not have a significant impact on employment in either the economic or statistical sense. Column 2, in contrast, shows that the policy prolongs work hours by about 2 hours among working women with the 1st, 2nd, and 4th quartile literacy skill. This result is consistent with the notion that a generous parental-leave policy enables women to stay in full-time jobs, but the coefficients are only imprecisely estimated. The estimated coefficients for the wage equation reported in Column 3 of Table 5 are negative for all quartiles, but the effects are only imprecisely estimated.[13]

The nuanced results from these conventional outcomes suggest that they are not as informative as the skill-use score in measuring the degree of skill use. For instance, "employment," as a binary variable, does not have any information about the tasks in which the worker engages, and similarly, the hours worked measures the quantity of labor input but not its quality. Although the wage rate could be seen as productivity in a perfectly competitive market, this one-to-one relationship does not hold in reality for various reasons, including discrimination, monopsony, search friction, internal labor-market consideration, collective bargaining, and labor market interventions, such as the minimum wages. Our skill-use score, in contrast, summarizes both quantity and quality inputs; the quantity is measured by the frequency of doing a certain task, and the quality is measured by the content of that task. Furthermore, since the items used to construct the score are directly related to the production process, it is less sensitive to the market structure that distorts the wage from the marginal product of labor.

3.3. Wage gap between natives and immigrants

We next consider applying the skill use measure to analyze the source of the wage gap between natives and immigrants. The large literature examining the wage penalty of immigrants has been attributed to the under-utilization of skill among immigrants. Studies find little or no returns to foreign attained schooling and foreign work experience, suggesting that the immigrant skills acquired in the source country are not fully utilized (Imai et al., 2019; Schaafsma and Sweetman, 2001). As direct evidence of skill under-utilization among immigrants, Chiswick and Miller (2008) report a skill and occupation mismatch. A series of studies attributes the lower return to education or other forms of human capital among immigrants to their limited language proficiency (Chiswick and Miller, 1995; 2003; 2012). The accumulated evidence that immigrants under-utilize their skill warrants a direct examination of the role of skill under-utilization as a source of wage penalty among immigrants.

To shed light on the role of skill utilization among immigrants, we apply Gelbach (2016)'s decomposition technique as we did to analyze the gender wage gap in the Section 3.1. Fig. 9 shows the results of the exercise. Panel (a) reports the unadjusted and adjusted wage differentials of immigrants compared with natives. Of 20 countries, an immigrant wage penalty is reported in 16 of them. Among these 16 countries, the penalty is substantially reduced by controlling for the covariates including skill and skill-use indices in Estonia (14 ppt), Germany (15 ppt), Italy (18 ppt), Korea (23 ppt), Slovenia (17 ppt), and Spain

[11] Career advancement seems to involve not only literacy use but also other non-routine tasks, such as, adapting to new environments or collaborating with others. Related to these tasks, the PIAAC asks about the frequency with which the respondent learns new work-related things from co-workers or through learning-by-doing, as well as the frequency of influencing others via instruction, presentations, advice, or negotiation. In addition, the survey asks for the frequency with which the respondent engages in writing tasks. We confirmed the robustness of our findings by using these tasks (Table C3).

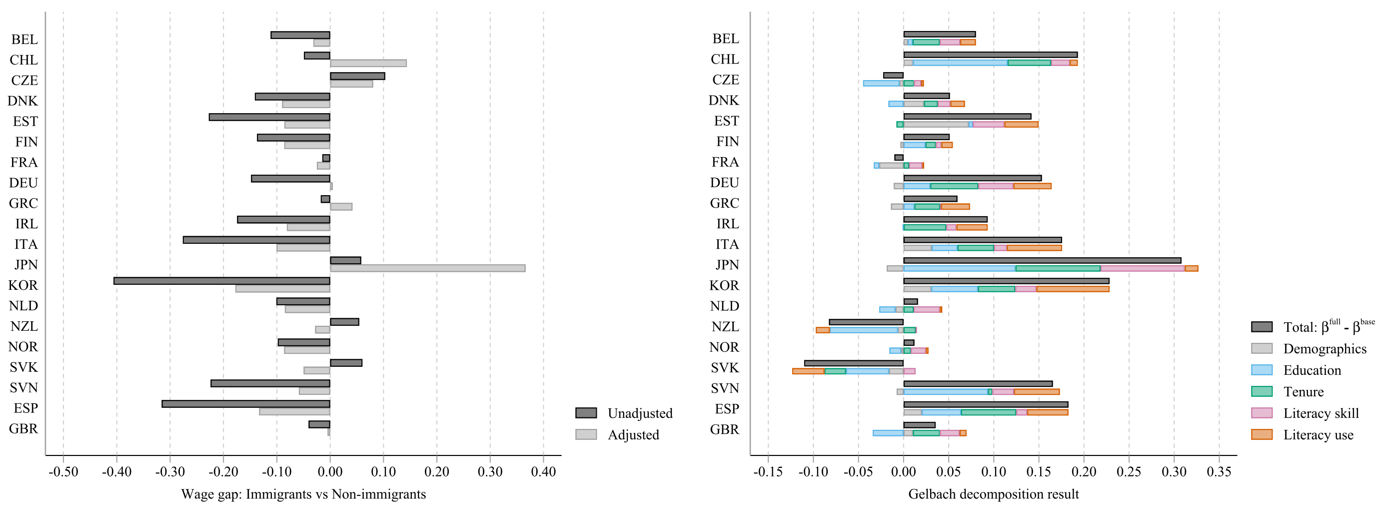
[12] Since we do not rely on panel data or direct measurements of career paths, such as the number of promotion, to regard our skill-use measure as a proxy of career advancement, we need the synthetic cohort assumption or the stationary assumption: namely age profiles estimated by multiple cohorts in a single cross section data represents the life cycle of a cohort.

[13] In this analysis, we used the Heckman sample selection correction method without any variables excluded from the wage equation, and hence, non-random sample selection issues if any, are unlikely to be mitigated. In fact, the resulting estimate is almost identical to the OLS estimate in Panel B.

Table 4
The job protection policy and utilization of literacy skill at work.

Dep.var. literacy skill use	Full sample				Employed
	(1)	(2)	(3)	(4)	(5)
Female×PL×Literacy skill: Q1	0.094*** (0.027)	0.058*** (0.016)	0.054*** (0.013)	0.054*** (0.013)	-0.009 (0.016)
Female×PL×Literacy skill: Q2	0.006 (0.031)	-0.023 (0.034)	-0.038 (0.031)	-0.041** (0.020)	-0.032** (0.015)
Female×PL×Literacy skill: Q3	-0.036 (0.038)	-0.076*** (0.026)	-0.101*** (0.023)	-0.106*** (0.023)	-0.061*** (0.016)
Female×PL×Literacy skill: Q4	0.003 (0.019)	-0.004 (0.023)	-0.018 (0.020)	-0.020 (0.018)	0.005 (0.017)
Country×Skill quartile FE	X	X	X	X	X
Female×Skill×Industrial structure	X	X	X	X	X
Female×Skill×Family policies		X	X	X	X
Female×Skill×Gender norm			X	X	X
Female×Skill×Market institutions				X	X
Countries	24	24	24	24	24
Observations	48,970	48,970	48,970	48,970	41,223

Note: This table shows estimation results of the censored Tobit model consisting of equations (8) and (9) for literacy score. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates of the coefficients of the interaction terms associated with the literacy skill index and the indicators for social institutions and social norms. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.



(a) Unadjusted and adjusted log wage gap

(b) Decompsotion result

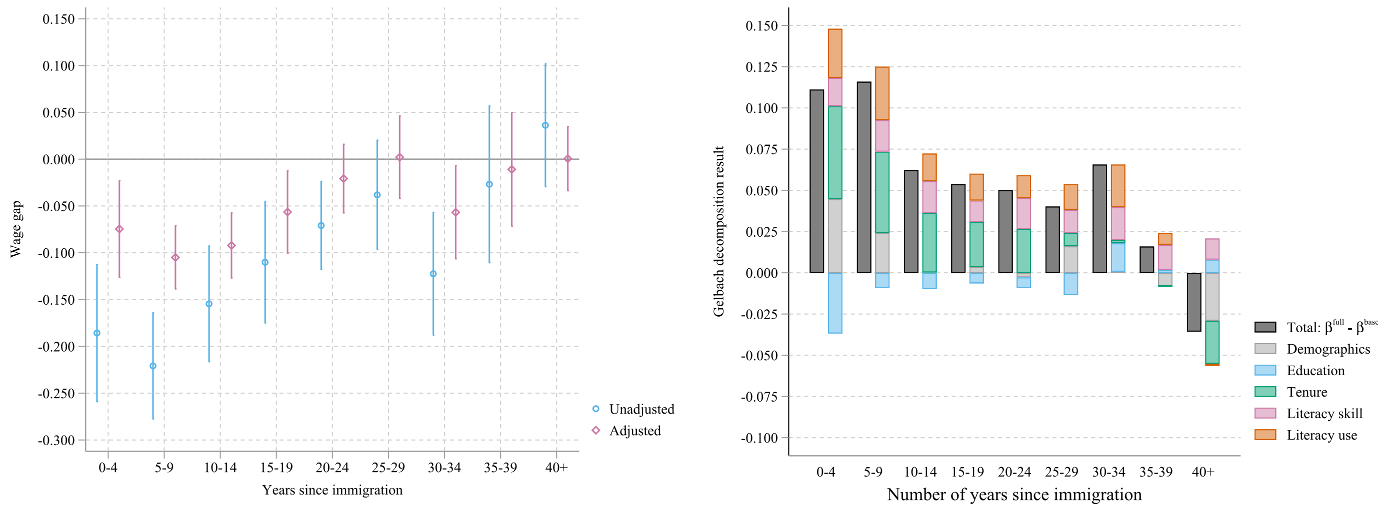
Fig. 9. Decomposition of the wage gap: Immigrants vs Non-immigrants. Note: This figure shows the (un)adjusted wage gap between immigrants and non-immigrants, and demonstrates Gelbach (2016)'s decomposition result. Demographics includes age and gender.

(18 ppt). In particular, the immigrant wage penalty in Germany disappears when we adjust the demographics, literacy skill and literacy use. How much of this reduction in immigrants' wage penalty is attributable to the immigrants-native gaps in skill and skill use? Panel (b) shows the results of Gelbach (2016)'s decomposition. A common feature of the 6 countries (Estonia, Germany, Italy, Korea, Slovenia, and Spain) is that a non-negligible part of the explained the immigrant-native wage gaps is attributable to the difference in literacy utilization. In fact, the contribution of literacy use is 25–36 percent in these countries. At the same time, however, the contribution of literacy skill varies across these countries. The contribution of literacy skill is less than 10 percent of the explained wage gap in Italy and Spain, while it is about 25 percent in Estonia and Germany.[14]

We further demonstrate that the skill-use index is useful to shed light on the mechanism behind the dynamic change of the wage gap between immigrants and natives. Numerous studies on the US have documented the assimilation of immigrants into the US labor market, defined by the wage convergence of immigrants to natives, as the immigrants accumulate years since migration (Borjas, 1985; 1995; Chiswick, 1978; LaLonde and Topel, 1992; 1997). A part of the assimilation process is attributable to language acquisition and this arguably resolves the mismatch between skill and skill use. If this is the case, we would observe a resolution of the gap in skill use between immigrants and natives. With a caveat that the analysis cannot distinguish the effect of assimilation from the effect of selective return migration, we document the wage gap between immigrants and natives by the number of years since immigration.

We first show that the wage gap between immigrants and natives narrows as the number of years since immigration increases. Panel (a)

[14] The heterogeneity of the contribution of literacy skill is possibly due to the fact that some countries have several options for the language used in the survey and skill assessment. For example, Spain has the options of Basque, Castilian, Catalan, Galician, and Valencian while some countries, such as Germany and France, have no options (OECD, 2013). In our sample, about 60 percent of immigrants took the test in languages different from their native languages. If the literacy skill is measured in those languages that are not usually used in the workplace, we would under-estimate the role of literacy skill in the assimilation process.



(a) Unadjusted and adjusted log wage gap (b) Decompsotion result

Fig. 10. Wage gap between immigrants and non-immigrants over the number of years since immigration. Note: This figure shows the (un)adjusted wage gap between immigrants and non-immigrants, and demonstrates Gelbach (2016)'s decomposition result. Demographics includes age and gender.

Table 5
The paid leave policy and conventional labor market outcomes.

Dep.var.	Employment	Work hours	ln($wage$)
	(1)	(2)	(3)
Female×PL×Literacy skill: Q1	0.002	2.971	-0.014
	(0.009)	(1.844)	(0.065)
Female×PL×Literacy skill: Q2	0.003	2.233*	-0.059
	(0.005)	(1.278)	(0.055)
Female×PL×Literacy skill: Q3	-0.009	0.765	-0.030
	(0.007)	(0.758)	(0.057)
Female×PL×Literacy skill: Q4	0.001	1.794**	-0.040
	(0.007)	(0.872)	(0.036)
Mean value among men	0.99	42.19	3.81
Method	OLS	Tobit	Heckit
Country×Skill quartile FE	X	X	X
Female×Skill×Industrial structure	X	X	X
Female×Skill×Family policies	X	X	X
Female×Skill×Gender norm	X	X	X
Female×Skill×Market institutions	X	X	X
Countries	24	23	21
Observations	35,410	33,919	31,515

Note: This table shows estimation results regarding market outcomes. We do not report the estimates of the constant term or the coefficients of age indicators, years of education and dummy variables indicating that the test language is the same as the native language of the respondent, or that parents are immigrants. We also omit some estimates relating to social institutions. Standard errors clustered by each country and skill quartile group are in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

of Fig. 10 shows the raw wage gap and the adjusted wage gap, where the 20 countries used in the previous analysis are pooled, and the raw wage gap is calculated after partialling out country fixed effects. The unadjusted gap is 19% in the 0–4 years since immigration and 22% in 5–9 years. Then the gap narrows as the years since migration increases and becomes close to 0 after 25–29 years. Panel (a) also reports the wage gap adjusting for the gaps in age, gender, education, tenure, literacy skill, and literacy use, where the coefficient for each control variable is allowed to be different across countries to capture different wage structure across those countries. These variables explain more than one half of the raw wage gap in the 0–9 years since immigration, and the difference between the raw wage gap and the adjusted wage gap gradually becomes small. Panel (b) reports what accounts for the reduction of the wage gap by each factor of covariates. The decomposition results show that the differences in tenure, literacy skill, and its use explain the persistent wage gap. Consistent with the prediction, the gap explained by the gap in literacy use narrows as the number of years since immigration increases.

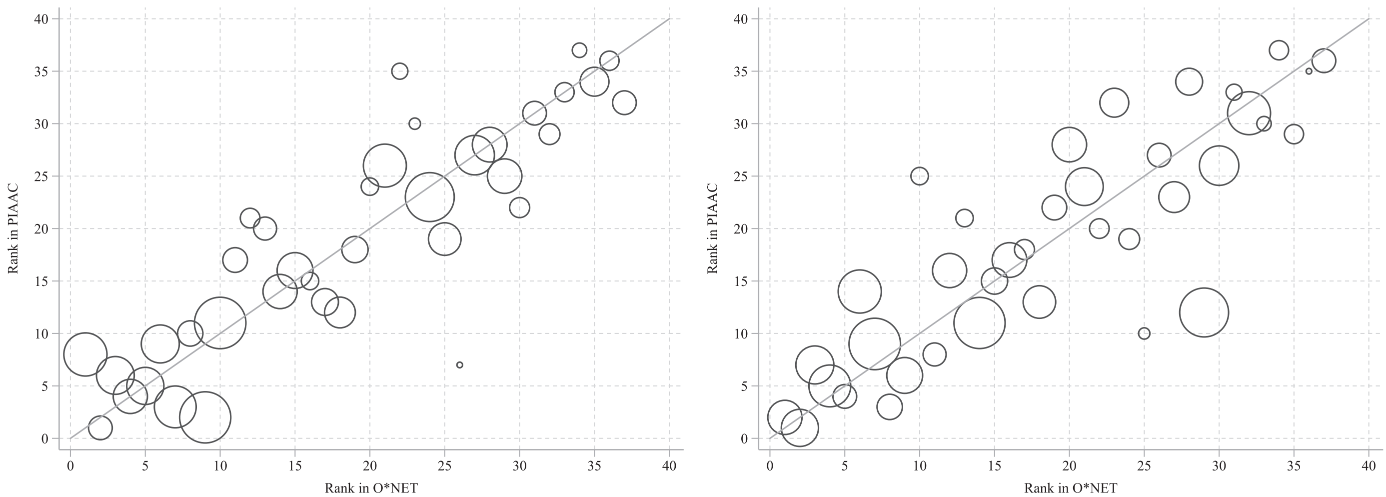
The exercise in this subsection demonstrates that the gap in skill use explains the variation in the wage penalty among immigrants relative to natives.

4. Discussion

The applications of the skill-use measure demonstrate the benefit of this measure in two ways. First, the skill-use measure is useful as an independent variable, as demonstrated by the examples of the gender wage gap and the immigrant wage gap. For both analyses, we showed that the gap in skill use is an important determinant of the wage penalty of females and immigrants in the countries where the unadjusted gaps are substantial. Through these examples, we illustrated that the skill use measure is a useful measure to shed light on the mechanism behind the observed gaps in conventional labor-market outcomes, which are wages in these examples.

Second, the skill-use measure is useful as a dependent variable. The assessment of the parental leave policy demonstrated the potential side effect of the policy that is not captured by conventional labor-market outcomes, such as employment, hours worked, or wages. Labor market outcomes are affected by the many variables other than skill use. Thus, using the skill-use index as a dependent variable is more direct way to capture the policy impact on skill under-utilization or skill mismatch. For the purpose of capturing the skill mismatch, the feature of the PIAAC that enables us to estimate both skill and skill use in the same dimension is particularly useful.

While the skill-use index is a useful measurement, most surveys widely used by labor economists, such as the Labor Force Survey or the population census, do not contain the information on skill use. The skill and skill-use test batteries in the PIAAC is lengthy, and including the batteries in every survey is not realistic. As a remedy, we can use the occupation code and the skill-use index of the PIAAC as a crosswalk to assign the value of skill use to each occupation in conventional data sets. This is analogous to using O*net to assign tasks to each occupation (Autor et al., 2003; Yamaguchi, 2012). The mean and standard errors



(a) Literacy use (b) Numeracy use

Fig. 11. Comparing skill use in PIAAC and O*NET. Note: This figure shows the ranking of literacy use and numeracy use across occupations in the PIAAC and the O*NET. In the O*NET data, we use "Reading comprehension" and"Mathematics" for literacy use and numeracy use, respectively. The size of markers indicates the number of records in the PIAAC, and the solid line is the 45-degree line. To convert the occupation code of O*NET to ISCO-08, we use the code provided by Hardy et al. (2018). This figure includes information from the O*NET 26.1 Database (https://www.onetcenter.org/database.html) by the U.S. Department of Labor, Employment and Training Administration (USDOL/ETA). Used under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/). O*NET®is a trademark of USDOL/ETA. We have modified all or some of this information. USDOL/ETA has not approved, endorsed, or tested these modifications.

of the skill and skill-use scores for each country and occupation are provided in Appendix D.

As a validity check, we plot the correlation of the rankings of literacy use and numeracy use from PIAAC and O*net of US occupations in Fig. 11. The figures show that skill-use measures of the PIAAC and O*net are closely related. Although the PIAAC and O*net are largely substitutes in the US context, there are two benefits of drawing on the PIAAC. The first benefit is the availability of the data set outside of the US. The jobs belonging to the same occupation code are not necessarily identical across countries and extrapolating the US occupation structure to other countries fails to capture the specificity of each country. The second benefit is the availability of the skill measure corresponding to the skill-use measure. By regressing literacy or numeracy on workers' covariates such as education and potential experience, and extrapolating the structure to the conventional dataset, we can assign the worker level skill measures to the conventional dataset. Combining this predicted skill measure from the workers' demographic characteristics with the predicted skill-use measure from the occupation code, we can potentially discuss the skill mismatch based on the conventional dataset. Merging skill and skill-use indices with the conventional survey data enables researchers to shed light on issues beyond the coverage of the PIAAC data set. For example, we can describe how the skill mismatch evolves in the long-run or over the short-run business cycle, assuming the stability of the mapping from demographic variables to skill, and from occupation code to skill use.

5. Conclusion

This study constructs objective measures of literacy and numeracy, and their use on the job, drawing on the PIAAC and covering 24 OECD countries. We argue that our proposed measure of skill use complements labor economists' efforts to measure skills on multiple dimensions. Given that skill mismatch is an important concept to explain a wide range of issues such as inefficient allocation of labor, gender wage gap, or assimilation of immigrant, the simultaneous measurement of skill and skill use makes it possible to directly measure the skill mismatch. We demonstrated the benefit of the proposed measurement through three examples: the gender wage gap, the assessment of parental-leave policies on the gender gap in skill use, and the wage convergence between immigrants and natives. Through these three examples, the skill use measure is shown to be useful in both the right- and left-hand sides of the estimation equation.

Finally, we add a remark on the benefit of the continuous effort by the OECD to collect additional waves. This study utilized the data from the 1st cycle that took place between 2011 and 2017. The OECD plans to conduct the 2nd cycle between 2022 and 2023. While the data structure is not panel, the repeated cross-section structure of the dataset will enable researchers to implement cohort-level analyses. Using the quasi-panel structure of the repeated cross section data, we can decompose the wage convergence of immigrants to natives into the selective migration and the causal effect, for example. We can similarly decompose the gender wage convergence into cohort and year effects. Thus, the continuous effort by the OECD to collect the dataset will help labor economists shed new light on skill-related issues in the future.

Acknowledgment

Christian Dustmann, Hyejin Ku, Uta Schoenberg and the Centre for Research and Analysis of Migration, University College London.

Appendix A. Skill use items

A1. Literacy skill use

1. Read directions or instructions
2. Read letters, memos or e-mails
3. Read articles in newspapers, magazines, or newsletters
4. Read articles in professional journals or scholarly publications
5. Read books
6. Read manuals or reference materials
7. Read bills, invoices, bank statements or other financial statements
8. Read diagrams, maps or schematics

A2. Numeracy skill use

1. Calculate prices, costs, or budgets

2. Use or calculate fractions, decimals, or percentages
3. Use a calculator – either hand-held or computer-based
4. Use simple algebra or formulas
5. Use more advanced math or statistics, such as calculus, complex algebra, trigonometry, or regression techniques
6. Prepare charts, graphs, or tables

A3. Learning opportunities

1. In your own job, how often do you learn new work-related things from co-workers or supervisors?
2. How often does your job involve learning-by-doing from the tasks you perform?
3. How often does your job involve keeping up to date with new products or services?

A4. Influencing others

1. How often does your job usually involve instructing, training, or teaching people, individually or in groups?
2. How often does your job usually involve making speeches or giving presentations in front of five or more people?
3. How often does your job usually involve advising people?
4. How often does your job usually involve planning the activities of others?
5. How often does your job usually involve persuading or influencing people?
6. How often does your job usually involve negotiating with people either inside or outside your firm or organization?

A5. Writing skill use

1. Writing skills at work: In your job, how often do you usually write letters, memos, or e-mails?
2. Writing skills at work: In your job, how often do you usually write articles for newspapers, magazines, or newsletters?
3. Writing skills at work: In your job, how often do you usually write reports?
4. Writing skills at work: In your job, how often do you usually fill in forms?

References

Albrecht, J., Bjorklund, A., Vroman, S., 2003. Is there a glass ceiling in sweden? J. Labor Econ. 21 (1), 145–177. https://ideas.repec.org/a/ucp/jlabec/v21y2003i1p145-177.html

Albrecht, J., Skogman Thoursie, P., Vroman, S., 2015. Parental leave and the glass ceiling in Sweden. Res. Labor Econ. 41, 89–114. https://ideas.repec.org/p/hhs/ifauwp/2015_004.html

Ashenfelter, O.C., Farber, H., Ransom, M.R., 2010. Labor market monopsony. J. Labor Econ. 28 (2), 203–210.

Autor, D.H., Levy, F., Murnane, R.J., 2003. The skill content of recent technological change: an empirical exploration. Q. J. Econ. 118 (4), 1279–1333.

Bartel, A., Taubman, P., 1979. Health and labor market success: the role of various diseases. Rev. Econ. Stat. 61 (1), 1–8. http://www.jstor.org/stable/1924825

Bertrand, M., Goldin, C., Katz, L.F., 2010. Dynamics of the gender gap for young professionals in the financial and corporate sectors. Am. Econ. J. 2 (3), 228–255. doi:10.1257/app.2.3.228. https://www.aeaweb.org/articles?id=10.1257/app.2.3.228

Bertrand, M., Kamenica, E., Pan, J., 2015. Gender identity and relative income within households. Q J Econ 130 (2), 571–614. doi:10.1093/qje/qjv001. https://academic.oup.com/qje/article-pdf/130/2/571/30631743/qjv001.pdf

Blau, F.D., Kahn, L.M., 2013. Female labor supply: why is the United States falling behind? Am. Econ. Rev. 103 (3), 251–256.

Blau, F.D., Kahn, L.M., 2017. The gender wage gap: extent, trends, and explanations. J. Econ. Lit. 55 (3), 789–865. doi:10.1257/jel.20160995. https://www.aeaweb.org/articles?id=10.1257/jel.20160995

Borjas, G.J., 1985. Assimilation, changes in cohort quality, and the earnings of immigrants. J. Labor Econ. 3 (4), 463–489.

Borjas, G.J., 1995. Assimilation and changes in cohort quality revisited: what happened to immigrant earnings in the 1980s? J. Labor Econ. 13 (2), 201–245.

Chiswick, B.R., 1978. The effect of americanization on the earnings of foreign-born men. J. Polit. Economy 86 (5), 897–921.

Chiswick, B.R., Miller, P.W., 1995. The endogeneity between language and earnings: international analyses. J. Labor Econ. 13 (2), 246–288.

Chiswick, B.R., Miller, P.W., 2003. The complementarity of language and other human capital: immigrant earnings in canada. Econ. Educ. Rev. 22 (5), 469–480.

Chiswick, B.R., Miller, P.W., 2008. Why is the payoff to schooling smaller for immigrants? Labour Econ. 15 (6), 1317–1340.

Chiswick, B.R., Miller, P.W., 2012. Negative and positive assimilation, skill transferability, and linguistic distance. J. Hum. Cap. 6 (1), 35–55.

Datta Gupta, N., Smith, N., Verner, M., 2008. Child care and parental leave in the nordic countries: a model to aspire to? Rev. Econ. Househ. 6 (1), 65–89.

Deming, D.J., 2017. The growing importance of social skills in the labor market. Q. J. Econ. 132 (4), 1593–1640.

Dube, A., Jacobs, J., Naidu, S., Suri, S., 2020. Monopsony in online labor markets. Am. Econ. Rev. 2 (1), 33–46.

European Values Study 2008, 2016. European Values Study 2008: Integrated Dataset (EVS 2008). GESIS Data Archive, Cologne. ZA4800 Data file. EVS. https://doi.org/10.4232/1.12458.

Fryer, R.G.J., Levitt, S.D., 2010. An empirical analysis of the gender gap in mathematics. Am. Econ. J. 2 (2), 210–240. doi:10.1257/app.2.2.210. http://www.aeaweb.org/articles?id=10.1257/app.2.2.210

Gelbach, J.B., 2016. When do covariates matter? and which ones, and how much? J. Labor Econ. 34 (2), 509–543. doi:10.1086/683668.

Goldin, C., 2014. A grand gender convergence: its last chapter. Am. Econ. Rev. 104 (4), 1091–1119. doi:10.1257/aer.104.4.1091. https://www.aeaweb.org/articles?id=10.1257/aer.104.4.1091

Griliches, Z., 1977. Estimating the returns to schooling: some econometric problems. Econometrica 45 (1), 1–22. http://www.jstor.org/stable/1913285

Guiso, L., Monte, F., Sapienza, P., Zingales, L., 2008. Culture, gender, and math. Science 320 (5880), 1164–1165. doi:10.1126/science.1154094. http://science.sciencemag.org/content/320/5880/1164

de Haan, F., 2012. Women as the motor of modern life. Women Gender Postwar Eur. 87–103.

Hardy, W., Keister, R., Lewandowski, P., 2018. Educational upgrading, structural change and the task composition of jobs in europe. Econ. Trans. Instit. Change 26 (2), 201–231. doi:10.1111/ecot.12145. https://onlinelibrary.wiley.com/doi/abs/10.1111/ecot.12145

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. J. Labor Econ. 24 (3), 411–482. doi:10.1086/504455.

Hsieh, C.-T., Hurst, E., Jones, C.I., Klenow, P.J., 2019. The allocation of talent and US economic growth. Econometrica 87 (5), 1439–1474.

Imai, S., Stacey, D., Warman, C., 2019. From engineer to taxi driver? language proficiency and the occupational skills of immigrants. Can. J. Econ. 52 (3), 914–953.

LaLonde, R.J., Topel, R.H., 1992. The assimilation of immigrants in the us labor market. immigration and the workforce: economic consequences for the united states and source areas. Immig. Workforces.

LaLonde, R.J., Topel, R.H., 1997. Economic impact of international migration and the economic performance of migrants. Handbook Population Family Econ. 1, 799–850.

Leuven, E., Oosterbeek, H., 2011. Overeducation and mismatch in the labor market. In: Hanushek, E.A., Machin, S., Woessmann, L. (Eds.), Handbook of The Economics of Education. In: Handbook of the Economics of Education, Vol. 4. Elsevier, pp. 283–326. doi:10.1016/B978-0-444-53444-6.00003-1. https://www.sciencedirect.com/science/article/pii/B9780444534446000031

Manning, A., 2013. Monopsony in motion. Princeton University Press.

Mincer, J., 1974. Schooling, experience, and earnings. Natl. Bureau Econ. Res..

Muraki, E., 1992. A generalized partial credit model: application of an EM algorithm. ETS Res. Rep. Ser. 1992 (1), i–30.

Nollenberger, N., Rodríguez-Planas, N., Sevilla, A., 2016. The math gender gap: the role of culture. Am. Econ. Rev. 106 (5), 257–261. doi:10.1257/aer.p20161121. http://www.aeaweb.org/articles?id=10.1257/aer.p20161121

OECD, 2010. Oecd employment outlook 2010.

OECD, 2013. Technical Report of the Survey of Adult Skills (PIAAC). Technical Report. OECD.

Olivetti, C., Petrongolo, B., 2017. The economic consequences of family policies: lessons from a century of legislation in high-income countries. J. Econ. Perspect. 31 (1), 205–230. doi:10.1257/jep.31.1.205. http://www.aeaweb.org/articles?id=10.1257/jep.31.1.205

Schaafsma, J., Sweetman, A., 2001. Immigrant earnings: age at immigration matters. Can. J. Econ. 1066–1099.

Thomas, M., 2018. The impact of mandated maternity benefits on the gender differential in promotions: examining the role of adverse selection. Mimeo, Cornell University.

Weichselbaumer, D., Winter-Ebmer, R., 2005. A meta-analysis of the international gender wage gap. J. Econ. Surv. 19 (3), 479–511.

Yamaguchi, S., 2012. Tasks and heterogeneous human capital. J. Labor Econ. 30 (1), 1–53.