Lack of skills or formal qualifications? New evidence on cross-country differences in the labor market disadvantage of less-educated adults

Jan Paul Heisig[*], Maurice Gesthuizen[+], and Heike Solga[*~]

* WZB Berlin Social Science Center
+ Radboud University, Department of Sociology
~ Freie Universität Berlin, Institute of Sociology

DRAFT VERSION: JANUARY 25, 2019

PLEASE DO NOT CITE OR DISTRIBUTE WITHOUT PERMISSION

Abstract

Less-educated adults (i.e., those who have not completed upper secondary education) bear high risks of labor market marginalization, but the extent of their disadvantage differs considerably across countries. Using PIAAC data on the actual literacy and numeracy skills of 59,067 adults in 27 countries, we examine three explanations for this cross-country variation, focusing on the occupational status gap between less-educated adults and those with a degree at the upper secondary level: Does the disadvantage of less-educated workers vary because of variation in the (individual-level) skills of less-educated adults, because of the "skills transparency" of educational degrees (i.e., the strength of the association between formal qualifications and actual skills), or because of differences in the vocational orientation of upper secondary education? Our analysis supports especially the first two explanations: Country-specific decompositions show that individual-level skills differences account for a substantial portion of the occupational status gap in most countries. Using two novel country-level measures of skills transparency (the "skills gap" and the "index of internal homogeneity"), we further find that higher skills transparency exacerbates the occupational status gap and that it is an important mediating channel for the effects of tracking in upper secondary education found in previous studies.

Keywords: Education; Inequality; Social Stratification; International Comparison; Skills; PIAAC

1 INTRODUCTION

It is well-established that educational degrees are positively associated with labor market outcomes such as employment rates, occupational status, or wages. Less-educated adults, that is, adults who did not complete upper-secondary education, bear particularly high risks of labor market marginalization (e.g., Abrassart, 2013; Gesthuizen et al., 2011). While the less educated are facing difficulties throughout the industrialized world, the extent of their labor market disadvantage varies considerably across countries (Abrassart, 2013; Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011; Gesthuizen et al., 2011; Shavit and Müller, 1998).

In this paper, we examine three explanations for cross-national variation in the labor market disadvantage of less-educated adults. A first explanation suggests that the levels of skills achieved by less- and more-educated workers vary across countries. If employers reward skills, these differences should more or less directly translate into differences in labor market attainment (e.g., Gesthuizen et al., 2011; Murnane et al., 1995). A second explanation, closely related to signaling and screening theories of labor market inequalities (Spence, 1973; Weiss, 1995), emphasizes cross-national differences in the "skills transparency" of educational certificates (e.g., Andersen and van de Werfhorst, 2010). The central argument is that the relationship between (easily observable) formal qualifications and (hard-to-observe) skills is closer—that is, formal qualifications are more "skills transparent"—in some countries than in others. When skills transparency is high, formal qualifications will be better proxies for actual skills. This might then affect labor market inequalities by exacerbating statistical discrimination against less-educated workers because of their (lack of) formal qualifications. A third explanation argues that especially in countries with a strong vocational orientation of upper secondary education, a shortage of occupational skills or credentials produces larger labor market disadvantages for less-educated workers (e.g., Shavit and Müller, 1998; Solga, 2008).

Our overarching research question in this paper is: how relevant are these different explanations? While they are widely referenced, empirical evidence on their validity remains limited. This is partly because very few cross-national data sets contain good measures of the actual skills of working-age adults. Most previous studies, therefore, cannot account for skills differentials between less- and more-educated workers. The few exceptions (Gesthuizen et al., 2011; van de Werfhorst, 2011) mostly use data from the mid-1990s *International Adult Literacy Survey* (IALS) and must be viewed with caution because severe problems with this data set have been detected in recent years (see Solga, 2014, for further details). More importantly, previous studies have tested the skills transparency explanation rather indirectly by exploring the role of education system characteristics. This line of research has found that the labor market disadvantage of less-educated adults increases with the extent of external differentiation (i.e., ability-related tracking) and vocational orientation (i.e., emphasis on occupation-specific skills) in secondary education (e.g., Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011; Shavit and Müller, 1998). The prevailing interpretation of these results is that both external differentiation and vocational orientation strengthen the skills transparency of educational certificates (Andersen and van de Werfhorst, 2010), but this is assumed rather than demonstrated empirically.

We provide novel and up-to-date evidence on the empirical validity of the three aforementioned explanations using recent data on 27 countries from the first and second rounds of the *Programme for the International Assessment of Adult Competencies* (PIAAC). PIAAC is a unique data set that provides high-quality, internationally comparable data on labor market outcomes and, crucially, also on the literacy and numeracy skills of working-age adults. We use occupational status as our measure of labor market attainment and compare less-educated adults (who have less than upper secondary education) with intermediate-educated adults (who have a degree at the upper-secondary or non-tertiary postsecondary level), that is, we focus on the "occupational status gap" between the two groups. We exclude adults with tertiary education because they are unlikely to be direct competitors of the less educated on the labor market.

The first major contribution of our paper is to test the skills transparency explanation more directly by using two *novel country-level measures of skills transparency*, at least with respect to general literacy and numeracy skills: the *skills gap* (the adjusted differential in mean literacy and numeracy competencies between adults with low and intermediate formal qualifications) and the *index of internal homogeneity* (which measures the residual skills variation within these educational groups). In countries where the skills gap is large and where educational groups are internally homogeneous, educational credentials are highly informative about an individual's actual skills (Heisig, 2018). In such settings, formal qualifications should play a particularly important role as screening devices on the labor market. Our novel measures enable us to provide more direct evidence on the empirical validity of the skills transparency explanation than previous studies. In addition, we shed new light on the well-documented relationship between tracking (or "external differentiation") in secondary education and labor market returns to formal qualifications (e.g., Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst 2011). Unlike previous studies, we can *empirically* assess the claim that it is an important mediating channel for the role of tracking in upper secondary education.

As our second contribution, we examine the importance of the *individual-level relationship between literacy/numeracy skills and labor market attainment.* Importantly, and in contrast to many previous studies, this also means that we account for possible individual-level effects of skills when exploring the role of country-level explanatory variables. Finally, we revisit the role of the *vocational orientation of a country's system of upper secondary education,* after accounting for country differences in skills transparency and for individual-level differences in skills.

2 EXPLAINING THE LABOR MARKET DISADVANTAGE OF LESS-EDUCATED ADULTS

We now review previous research on labor market returns to education and the labor market disadvantage of less-educated adults, with a particular focus on comparative work. We begin with the three explanations outlined above: individual-level differences in skills, country-level differences in skills transparency and vocational orientation. We close the section with a discussion of prominent alternative explanations.

Before turning to the different explanations, we note that skills are not homogeneous, but comprise a diverse set of capabilities that differ in their transferability across different types of jobs (Becker, 1964). One important distinction in this respect is between general and occupation-specific skills (e.g., Müller and Jacob, 2008). Whereas general skills such as literacy and mathematical skills are useful in a wide variety of jobs, occupation-specific skills (e.g., an auto mechanic's understanding of how to repair a car engine) are, by definition, valuable only in a narrow set of particular occupations. The PIAAC data used in our empirical analysis provide measures of individuals' general (literacy and numeracy) skills, but no direct measures of occupation-specific skills. This limitation is important to keep in mind as we present our hypotheses and empirical approach.

2.1 Individual-level differences in skills

A common explanation for the labor market disadvantage of less-educated workers is based on human capital theory (Becker, 1964). It is argued that skills enhance productivity and are therefore rewarded by employers (e.g., by higher wages or job placements). Accordingly, because less-educated adults achieve on average lower levels of skills than intermediate-educated workers (Heisig and Solga, 2015a; Park and Kyei, 2011), they should have poorer occupational attainment (Bills, 1990, 2003; Solga, 2002, 2008). While this argument is widely used, empirical tests with direct measures of cognitive skills remain rare. A major reason for this has been a shortage of direct skills measures, especially in cross-national surveys. PIAAC's high-quality measures of general skills allow us to assess the empirical relevance of this argument for a large set of advanced economies. We therefore expect to find:

Hypothesis 1: In all countries, less-educated workers attain lower occupational status than intermediate-educated workers, and this disadvantage of less-educated adults is partly explained by differences in individual literacy and numeracy skills.

Importantly, this argument about the importance of *individual* skills also suggests a first and straightforward explanation for cross-national variation in the labor market disadvantage of the less educated. Figure 1 in Heisig and Solga's (2015a) analysis of PIAAC data shows that less-educated workers have relatively high numeracy skills in Finland, Japan, and Norway. Moreover, the difference in mean skills levels between less- and intermediate-educated workers is quite small in these countries. By contrast, less-educated workers tend to have very low numeracy skills in the United States, Belgium, and Germany and skills differentials are quite large in these countries. Concerning country differences in the labor market disadvantage of less-educated workers, these findings suggest that, in some countries, less-educated adults might attain higher occupational status simply because they are, on average, better equipped with skills. We therefore expect that:

Hypothesis 2: Accounting for differences in literacy and numeracy skills at the individual level reduces cross-national variation in the occupational status gap between less- and intermediate-educated adults.

2.2 Skills transparency

Previous research (e.g., Abrassart, 2013; Andersen and van de Werfhorst, 2010; Solga, 2002, 2008) has already stated that the aggregate (country-level) relationship between formal qualifications and skills might affect the labor market disadvantage of less-educated workers—above and beyond the direct individual-level effect of skills emphasized in the previous section. This research has mainly relied on signaling and screening accounts for theoretical justification.[1] In their weak versions, these accounts do not dispute the aforementioned argument that higher qualifications are rewarded by employers because qualifications are positively related to skills (Bills, 2003). However, the signaling approach emphasizes that skills are very difficult to observe and that employers therefore heavily rely on more readily observable proxies for skills and "trainability" in hiring, job placement, and promotion decisions (Spence, 1973; Thurow, 1979). Degrees and other indicators of educational attainment such as grades therefore serve as crucial sources of information (Arrow, 1973; Hirsch, 1977; Thurow, 1979; Weiss, 1995).

When employers assess the skills of applicants based on beliefs about how well educational certificates indicate (i.e., "signal") an applicant's skill level, they effectively apply so-called statistical discrimination (Aigner and Cain, 1977; Phelps, 1972). Employers should be particularly likely to statistically discriminate on the basis of educational credentials when the latter are strongly predictive of an individual's actual skills—in other words, when skills transparency is high (Andersen and van de Werfhorst, 2010). Hence, even after accounting for skills at the individual level, the labor market disadvantage of less-educated adults should still increase with a country's level of skills transparency—reflecting stronger statistical discrimination against all less-educated adults, independent of their individual skills, in more skill-transparent contexts.

[1] We treat signaling and screening theories as one general approach in this article. While some scholars view the two as distinct approaches, we concur with Bills' (2003) reading of Weiss (1995) that the two approaches are conceptually very similar and that the primary "difference between screening and signaling models is that, in the former, firms move first and, in the latter, students move first" (Bills, 2003, p. 447).

Measuring skills transparency is not trivial, however, and previous studies have mostly proxied it using education system indicators. Andersen and van de Werfhorst (2010), for example, tried to capture a country's level of skills transparency using an index based on several education system characteristics, including the extent of tracking, the prevalence of vocational enrollment, and participation in tertiary education. Based on this operationalization, and not accounting for skills at the individual level, they concluded that skills transparency seems to be "the primary moderator" explaining country differences in the relationship between educational degrees and occupational status (Andersen and van de Werfhorst, 2010, p. 336).

Bol and van de Werfhorst (2011) included self-reported years of schooling as a proxy for individuals' skills and found similar results. Instead of using a summary index, they analyzed the moderating role of external differentiation and vocational enrollment separately. Their main findings were that higher levels of external differentiation and vocational orientation are both associated with higher returns to formal qualifications in terms of occupational status. In line with the above argument, they speculated that this was due to the signaling value of educational degrees being higher in countries with stronger tracking and vocational orientation.

Some studies have used data from the mid-1990s *International Adult Literacy Survey* (IALS), the most important cross-national survey with direct measures of individual skills before PIAAC. Van de Werfhorst (2011) found that earnings returns to educational degrees are positively related to external differentiation and vocational orientation, even after controlling for individuals' skills. Abrassart (2013) employed a more direct measure of skills transparency: the skills differential (or "skills gap") between less- and intermediate-educated workers at the country level. He found that the labor market disadvantage of less-educated adults (with respect to employment rates) increases with the aggregate skills differential. However, he did not include skills at the individual level, so it remains unclear if the effect of the aggregate skills gap simply picks up direct, individual-level effect of skills. Finally, Gesthuizen, Solga, and Künster (2011) found that, net of individual general skills, the skills mean of the less-educated group is positively related to their average occupational status. Yet, the skills mean alone is a poor measure of skills transparency. This is because an educational degree can only function as a useful signal to the extent that it indicates differences in the likely skills of a person relative to another one with a different degree. In this sense, the notion of skills transparency involves a comparative or relational element that cannot be captured by the skill mean of a single educational group.

Taken together, the aforementioned studies provide meaningful hints that country differences in skills transparency might be an important part of the explanation why less-educated adults face greater labor market disadvantages in some countries than in others. But they leave important questions unanswered. Some studies only look at the moderating role of education system characteristics and argue on theoretical grounds that the latter are related to the skills transparency of educational degrees (Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011). Other studies attempt to measure skills transparency more directly, but do not control for skills differences at the individual level (Abrassart, 2011) or use a suboptimal measure of skills transparency (Gesthuizen et al., 2011).

In the present paper, we use a more sophisticated approach to measuring skills transparency. We understand skills transparency as the extent to which formal qualifications are predictive of actual skills and focus on two aspects of the distribution of skills conditional on formal qualifications (Heisig, 2018). The first is the difference in the average skills levels of different educational groups, adjusted for other readily observable factors such as age or gender, which we refer to as the *skills gap*. Formal qualifications should become more informative about the actual skills a person has (i.e., they should become more skills transparent) as the skills gap increases. The less educated should therefore face stronger statistical discrimination and consequently also greater labor market disadvantages in countries where the skills gap is large. The second aspect of skills transparency is the *internal skills homogeneity* of educational groups: Other things being equal, including the skills gap, degrees are a less noisy proxy of actual individual skills (and therefore send a stronger signal about them) when educational groups are internally more homogeneous (Aigner and Cain, 1977). Based on these considerations, we formulate the following hypotheses, both of which we expect to hold after controlling for literacy and numeracy skills at the individual level:

Hypothesis 3: The occupational status gap between less-educated and intermediate-educated adults is larger in countries where the skills gap (with respect to numeracy and literacy skills) between the two groups is larger.

Hypothesis 4: The occupational status gap between less-educated and intermediate-educated adults is larger in countries where the distribution of literacy and numeracy skills within the two groups is more homogeneous (i.e., has lower variance).

Moreover, previous studies (Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011) have argued that the effect of *external differentiation of secondary education* on labor market returns to formal qualifications is due to differences in the skills transparency of educational certificates. If this interpretation is correct, we should find support for the following hypothesis:

Hypothesis 5: Accounting for the direct measures of skills transparency (i.e., skills gap and internal homogeneity of educational groups) decreases the effect of the index of external differentiation of secondary education on the occupational status gap between less-educated and intermediate-educated adults.

2.3 Vocational orientation

As noted above, the studies by Bol and van de Werfhorst (2011) and van de Werfhorst (2011) both find that the labor market disadvantage of less-educated adults is larger in countries with a stronger vocational orientation of upper secondary education. Earlier work by Shavit and Müller (1998) reached similar conclusions, and Andersen and Werfhorst (2010) also included indicators of vocational orientation in their summary index of skills transparency.

There are several possible explanations for these results. Theories of credentialism, for example, suggest that the benefits of holding a vocational certificate might derive from occupational licensing and closure (see Section 2.4 for further discussion). Another argument builds on the two explanations (differences in individual skills and in skills transparency) that we have focused on so far. In countries with a strong vocational orientation, most adults with an upper secondary degree have completed a program that focuses on occupation-specific skills, which likely ensures that "more job-relevant skills are acquired that are directly applicable in the workplace" (van de Werfhorst, 2011, p. 1080). This suggests that these countries are characterized by greater inequalities in occupational skills between less- and more-educated workers. These differences should translate into greater labor market inequalities, either because of the direct relationship between (occupational) skills and labor market attainment emphasized by human capital theory, or because formal qualifications are more transparent with respect to occupational skills when the education system emphasizes vocational programs.[2]

For our purposes, it is important to acknowledge that, unlike for (general) literacy and numeracy skills, we cannot disentangle these explanations. To do so, we would need direct, individual-level measures of occupation-specific skills which were not assessed in PIAAC. Similar to previous studies, we can therefore only look at the role of vocational orientation of upper secondary education, an institutional characteristic varying at the country level. However, we can at least control for general skills at the individual level, which are known to influence educational decisions and thereby also the probability of entering and completing educational programs at the upper secondary level (e.g., Lleras, 2008). Thus, our final hypothesis is:

Hypothesis 6: The higher the prevalence of enrollment in vocational upper secondary education programs, the larger the occupational status difference between less- and intermediate-educated adults.

2.4 Alternative explanations

Another prominent explanation why less-educated workers are disadvantaged on labor markets is *credentialism* (e.g., Collins, 1979). In its strong version, it states that higher educational certificates are rewarded on labor markets, "apart from anything they have learned in schools. Education is thus more a selector, sorter, and allocator than it is a socializer" (Meyer, 1977, p. 59). Weaker versions "merely […] argue that the relation between education and productivity is smaller than that between education and rewards" (Bills, 2003, p. 452). Such excess returns to educational credentials are often attributed to credentials functioning as devices of social closure that artificially restrict access to advantageous positions (Bol and Weeden, 2014; see also Sørensen's, 2000, theory of rent generation). In short, strong versions of credentialism dispute any meaningful relationship between educational credentials and job-relevant skills; weak versions allow for such a relationship but argue that labor market returns to educational degrees at least partly reflect entitlements and occupational rents that are unrelated to worker productivity. Empirical support for hypotheses 1 to 4 would be difficult to reconcile with strong

[2] In fact, some studies discuss the moderating effect of vocational orientation on labor market returns to education primarily in terms of human capital theory (van de Werfhorst, 2011), while others (Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011) emphasize the signaling/skills transparency explanation.

versions of credentialism because it would indicate that labor market returns to educational degrees are at least partly due to the fact that the latter are related to general (literacy and numeracy) skills. Weak versions of credentialism, however, could clearly be compatible with support for our hypotheses and might help account for any portion of the labor market disadvantage of less-educated adults that we are unable to explain.

Theories of *labor market segmentation* provide yet another framework for understanding labor market inequalities. While heterogeneous in their details, segmentalist explanations generally view the labor market as divided into a small number of segments, with many positing an essentially dualistic structure consisting of a primary and a secondary sector (e.g., Doeringer and Piore, 1971, Piore, 1994). Jobs in the primary sector are characterized by good career opportunities (on internal labor markets), high levels of job security, high remuneration, and good overall job quality, whereas jobs in the secondary sector tend to be rather low skilled, insecure, badly paid, and unattractive in other respects. Segmentation theory breaks with the supply-side orientation of mainstream economic theories, and of human capital theory in particular, and sees demand-side factors as the primary determinants of job quality: "Industrial organisation, product market and technological conditions, managerial control strategies and systems of labour market regulation are all recognised as having an influence on the structure of jobs and in contrast to the orthodox theory of the labour market, the distinction between good and bad jobs is not based on individual differences in productivity" (Leontaridi, 1998, p. 64).

When it comes to explaining cross-national variation in the labor market disadvantage of less-educated adults, the segmentation perspective suggests that we should consider demand-side factors as a possible explanation. The relative position of less-educated adults in a country might be less a matter of relative skill endowments but rather a function of the availability of "good" and "bad" jobs, which in turn is related to a country's industrial and broader economic structure. To explore this possibility, we will investigate whether our focal country-level relationships are robust to controlling for cross-national differences in the industrial structure.

A final concern could be that our data were collected during the first half of the 2010s, when the countries in our sample were characterized by very *different labor market conditions*. Some countries such as Spain and Greece were still in the midst of the deep recessions that unfolded in the years after the 2007 financial crisis. Other countries such as Austria or Germany were faring much better. Because the labor market prospects of less-educated adults are particularly sensitive to the business cycle (Farber, 1997), these cross-national differences might confound our results. We account for this possibility in two ways. First, we measure labor market attainment in terms of occupational status in the *current or last* job (up to five years before the interview). Thus, we also observe the outcome for respondents who lost their job in the wake of the financial crisis. Second, we explore whether our focal country-level relationships are robust to controlling for the unemployment rate.

3 DATA AND METHODS

3.1 Individual-level data and sample

Our individual-level data are from the first and second round of PIAAC, conducted in 33 countries in 2011/12 and 2014/15, respectively (OECD, 2013, 2016). The PIAAC data are representative of the noninstitutionalized 16-to-65-year-old population. The OECD requested a minimum sample size of 4,500 or 5,000 cases per country[3] and a minimum response rate of 50 percent. All countries were required to provide a non-response bias analysis after data collection, and the results of this analysis were taken into account in the construction of the final survey weights, which were used in all analyses reported in this article. We also use the replicate weights provided by PIAAC to correct the standard errors for the complex survey design (for further details, see OECD, 2016).

PIAAC was conducted in 33 countries. Two of these, Australia and Indonesia, provide no public use files. We decided to exclude two further cases, Cyprus and Russia, because of concerns about data quality.[4] Influence diagnostics for the remaining 29 countries revealed that the inclusion of Israel and Slovenia has a dramatic impact on the main regression results reported below (as indicated by the DFBETA and Cook’s D statistics; see Fox, 1991). We therefore chose to drop these two cases, resulting in a sample of 27 countries for the main analysis (see Table 1 below for the individual countries). We provide a detailed account of the influence diagnostics in

[3] The higher sample size was required if respondents were also tested in the optional “problem solving in technology-rich environments” (PS-TRE), in addition to the (mandatory) literacy and numeracy domains. We ignore PS-TRE skills because they are not available for all countries.

[4] Cyprus has a very high share (almost 18%) of so-called literacy-related non-respondents, that is, of sampled respondents who did not complete the survey because of language difficulties (OECD, 2013). The country with the second-highest share is Belgium (5.2%). Among several concerns about the quality of the Russian data, a major one is that the Moscow municipal region was not included in the survey (OECD, 2016, p.21).

Section D of the Online Supplement, including the main regression results when the Israel and Slovenia are included. A brief summary is provided in Section 4.4 below.

Our goal is to explain the labor market disadvantage of less-educated workers. We therefore compare the occupational status attainment of less-educated workers—defined as those with the highest degree below the upper secondary level—to those with upper secondary education degrees. We exclude respondents with a tertiary degree[5] from the analysis, as they rarely compete for the same kinds of jobs as less-educated adults. We restrict the analysis to men and women aged 16 to 54 who, a), worked for pay at the time of interview or within the last five years before the interview, b), were not enrolled in full-time education at the time of interview, and, c), had obtained their highest educational degree in the country where they were surveyed.

A total of 59,556 cases meet the sample restrictions, after excluding 1,463 so-called literacy-related non-respondents (OECD, 2013) and 171 cases with missing values on at least one of the variables defining the sample.[6] The only variables with non-negligible proportions of missing data are parental education and occupational status, which are unavailable for 4,223 and 811 cases, respectively. We use multiple imputation via chained equations to fill in missing values on these two measures. All other variables have very low proportions of missing data. To simplify the imputation procedure, we drop the 489 cases that are incomplete with respect to these variables. We generate ten imputations, one for each of the so-called plausible values for the skills measures (see Section 3.2). The final sample comprises 59,067 (= 59,556 - 489) respondents, with country-specific sample sizes ranging from 1,216 cases in Singapore to 7,430 cases in Canada (see Table 1 below).

3.2 Individual-level variables

PIAAC provides information on the respondent's highest educational degree in terms of the 1997 revision of the International Standard Classification of Education (ISCED). We differentiate between less-educated (ISCED levels 0–2) and intermediate-educated (ISCED levels 3–4) adults.

[5] That is, those with levels 5 and 6 according to the 1997 revision of the International Standard Classification of Education (ISCED).

[6] These case numbers refer to the sample of 27 countries used in the main analysis.

This corresponds to the highest degree being at the lower secondary level or below and at the upper secondary or non-tertiary post-secondary level, respectively.

We operationalize the labor market disadvantage of less-educated adults by the *occupational status gap* between less- and intermediate-educated adults. *Occupational status* is measured using the International Socio-Economic Index of Occupational Status (ISEI). The ISEI scores are "weighted averages of standardized measures of the income and education of incumbents of each occupation" (Ganzeboom and Treiman, 1996, p. 204)—based on relative weights for (standardized) education and earnings, "such that the direct effect of education on earnings is minimized. […] The resulting index was then projected onto a 10 … 90 range using linear transformation" (Ganzeboom and Treiman, 2010, p. 13). The ISEI score thus indicates the *relative* position of occupations in the hierarchical occupational stratification system. We assign scores based on one-digit 2008 International Standard Classification of Occupation (ISCO-08) codes. For respondents who worked at the time of interview, occupation codes refer to the current job. For those who did not work (but stopped working no more than five years ago) codes refer to the respondent's last job.

The one-digit ISCO-08 groups workers into ten broad occupational categories. It would be preferable to assign occupational status using occupational categories at the two- or higher-digit level, but four countries in our sample (Austria, Canada, Estonia, and Finland) only provide one-digit codes in their PIAAC public use file. To ensure consistency we use the one-digit version of ISCO-08 for all countries. Reassuringly, ISEI gaps based on more detailed occupational categories are almost identical to those based on one-digit groups for the countries where the former are available. For the 23 countries that provide two-digit ISCO-08 codes in their public use files, the Pearson correlation between ISEI gaps based on one-digit and two-digit codes is .99, after adjusting for literacy and numeracy skills and additional controls (see the discussion of "fully adjusted ISEI gaps" in Section 3.4). Even ISEI gaps based on four-digit occupation codes (which, in addition to the previously mentioned countries, are unavailable for Germany, Ireland, Sweden, and the United States) still show a Pearson correlation of .97 with the gaps based on one-digit occupation codes.

PIAAC also provides information on other labor market outcomes, most importantly on respondents' employment status and earnings. The primary reason why we do not analyze (un)employment is that, as noted above, the countries in our sample were facing very different macroeconomic conditions in the early 2010s. The employment rates of less-educated workers in particular have been found to be highly sensitive to overall labor market conditions (Farber, 1997). Adequately controlling for country differences macroeconomic conditions would thus be crucial, but doing so is difficult given limited degrees of freedom at the country level and uncertainty about the precise functional form of the relationship. Education-related differentials in occupational status should be less sensitive to macroeconomic context, especially since we also observe the occupation in the last job for respondents who were not employed at the time of interview. Nevertheless, some of our specifications additionally control for the unemployment rate (see Section 4.4).

We have two main reasons for not analyzing wages or earnings in the main article. First, the estimated country-specific wage/earnings gaps between less- and intermediate-educated workers are noisier than the gaps in occupational status. We investigated this issue by computing I^2 statistics for the occupational status and various earnings/wage gaps after adjusting for literacy and numeracy skills and the additional controls. I^2 is commonly used in meta-analysis to distinguish "true" between-study variability in effect sizes from variability that is due to sampling error, that is, to the fact that the effect size for each individual study is subject to statistical uncertainty. In the present context, the statistic can be interpreted as the proportion of overall between-country variation in the estimated labor market outcome gap that is attributable to true between-country differences rather than to sampling error; in other words: to signal rather than noise. Formally, I^2 is calculated as $\hat{\tau}^2 / (\hat{\tau}^2 + \hat{\sigma}^2)$, where $\hat{\tau}^2$ denotes estimated between-country variance and $\hat{\sigma}^2$ the estimated (average) statistical error of the country-specific estimates (for details on the underlying random effects model and its estimation, see Viechtbauer, 2010). For the ISEI gap between less- and intermediate-educated adults, a reasonable 74.1 percent of the between-country variance reflects true variation according to the I^2 statistic.[7] For log hourly earnings[8], this proportion is only 69.7 percent.[9] The second reason why we prefer to focus on

[7] All estimates of I^2 reported here are based on the so-called Empirical Bayes estimator as implemented in the R package *metafor* (Viechtbauer, 2010). Other estimation approaches such as (restricted) maximum likelihood yield very similar values.

[8] For confidentiality reasons, some countries do not provide the exact hourly earnings of respondents in the PIAAC public use files. For these countries, only the respondent's decile rank in the distribution of hourly earnings is available. For consistency, we therefore used the median wage within a respondent's wage decile for all respondents, just like we generally used the average score for one-digit ISCO groups to assign ISEI scores. The decile medians were kindly provided by (NAME OMITTED). In a previous analysis of the PIAAC data, Hanushek et al. (2015) found that using decile medians instead of exact wages had only a very limited impact on the results.

occupational status is that an analysis of wage gaps would also need to account for several country-level factors that influence overall wage inequality (e.g., collective bargaining arrangements and minimum wage legislation; see Koeniger et al., 2007). Such factors are difficult to control due to imperfect measurement and limited degrees of freedom at the country level. In supplementary analyses, we reran the main sequence of regression models with the gap in hourly earnings as the dependent variable. The results provide less support for our hypotheses than those for occupational status, but we are inclined to attribute this to the abovementioned complications (see Section 4.4 for further details).

The unique feature of PIAAC is the availability of high-quality measures of respondents' *actual skills*. All PIAAC-participating countries administered test items to assess the reading and text comprehension skills (*literacy*) and practical mathematical skills (*numeracy*) of participants (OECD, 2013, 2016). To limit respondent burden, each participant received only a relatively small number of test items, rendering individual competence estimates quite uncertain. PIAAC therefore provides ten plausible values rather than a single competence score for each case. To appropriately handle the plausible values (as well as the multiply imputed values for parental education and occupational status), we run all analyses ten times and apply the appropriate rules for multiply imputed data to obtain final point estimates, standard errors, and p-values (Little and Rubin, 2002).

We include several individual-level control variables: *sex*; *potential work experience* (linear splines with knots at 10, 20, and 30 years); *foreign-birth/foreign-language status* (four categories; see Table 1); *parental educational attainment* (low = no parent has completed upper secondary education; intermediate = at least one parent has completed upper secondary education; high = at least one parent has completed tertiary education); *self-employment* in last/current job (dummy variable). Table 1 provides descriptive statistics for the individual-level variables.

Our sample includes both currently and formerly employed respondents, but we do not control for current employment status because it is endogenous to the outcome variable (people with lower occupational status have higher risks of unemployment). As a robustness check, we reran the analysis using only respondents who worked at the time of interview and results were similar (see Section 4.4).

[9] For two alternative earnings measures we considered, the individual's decile rank in the distribution of hourly wages and in the distribution monthly earnings, I^2 estimates are only 45.6 and 31.3 percent, respectively.

Table 1. Individual-level descriptive statistics by country

							Foreign-birth/foreign-language status				Parental education				
	Mean ISEI score	% less-educated (ISCED 0-2)	% intermediate-educated (ISCED 3-4)	Mean literacy score	Mean numeracy score	Mean potential work experience (years)	% native-born, test language is first language	% native-born, test language is not first language	% foreign-born, test language is first language	% foreign-born, test language is not first language	% with low parental education	% with intermediate parental education	% with high parental education	% self-employed	N
Austria	40.6	17.5	82.5	270.9	276.4	21.3	92.7	2.5	1.7	3.1	26.5	59.3	14.2	10.0	2,225
Belgium	36.1	18.9	81.1	264.7	270.1	22.1	90.2	2.9	2.6	4.2	46.0	39.3	14.7	11.3	1,663
Canada	40.3	19.3	80.7	264.1	254.0	20.5	84.8	5.6	4.1	5.4	29.0	42.5	28.5	12.3	7,430
Chile*	28.7	36.6	63.4	207.5	192.8	19.8	95.3	0.6	3.8	0.3	57.5	33.9	8.7	23.2	2,128
Czech Rep.	36.0	9.1	90.9	268.8	270.5	20.2	96.2	0.0	1.7	2.1	10.7	81.0	8.4	16.0	2,402
Denmark	36.4	29.2	70.8	263.6	271.6	18.5	89.1	0.7	1.4	8.8	33.3	45.7	21.0	9.6	1,851
Estonia	34.4	20.6	79.4	265.9	263.4	18.7	88.4	2.1	8.2	1.3	29.4	44.2	26.4	9.2	2,511
Finland	32.6	15.1	84.9	284.4	277.1	17.2	93.9	1.6	1.4	3.1	38.0	47.1	15.0	11.6	1,422
France	34.2	26.9	73.1	253.1	243.7	20.2	85.5	2.3	5.1	7.1	50.9	40.1	9.0	9.0	2,515
Germany	35.0	15.0	85.0	261.0	262.1	21.4	83.1	2.3	3.4	11.2	13.6	63.4	23.0	7.0	2,060
Greece*	31.0	36.0	64.0	244.1	243.9	22.7	87.4	0.6	5.3	6.8	72.5	21.3	6.1	28.5	1,853
Ireland	34.7	31.2	68.8	258.7	247.1	19.4	79.5	0.7	10.2	9.6	60.4	27.9	11.7	14.7	1,976
Italy	34.1	53.5	46.5	248.3	248.4	23.3	85.9	1.8	2.1	10.2	79.2	18.6	2.2	18.3	2,131
Japan	35.2	15.3	84.7	292.3	281.1	21.1	99.4	0.1	0.3	0.2	25.0	54.6	20.4	8.7	1,287
Korea	32.4	19.1	80.9	260.6	250.8	23.7	97.4	0.3	1.2	1.1	65.6	26.2	8.2	23.6	1,985
Lithuania*	31.2	9.9	90.1	258.2	257.3	21.2	87.4	9.4	1.6	1.6	35.8	28.4	35.8	8.6	1,794
Netherlands	41.8	36.2	63.8	277.5	273.7	20.0	87.5	1.0	3.2	8.2	58.1	26.7	15.2	11.9	1,733
New Zealand*	38.2	37.9	62.1	270.2	259.8	18.4	83.2	3.0	7.4	6.5	44.4	28.0	27.6	11.9	1,797
Norway	36.6	32.1	67.9	270.1	268.7	17.5	84.4	1.4	0.7	13.5	29.3	45.0	25.8	7.9	1,559
Poland	31.6	10.5	89.5	254.9	250.2	19.8	99.0	1.0	0.0	0.0	29.1	66.0	4.9	16.7	3,001
Singapore*	39.1	36.5	63.5	228.2	222.7	26.1	18.5	70.4	1.0	10.0	69.2	25.0	5.8	14.6	1,216
Slovak Rep.	36.6	11.4	88.6	272.9	275.7	20.3	92.8	5.5	0.9	0.8	26.4	67.4	6.2	15.1	2,482
Spain	31.4	65.6	34.4	242.2	237.1	22.8	79.1	2.2	11.9	6.8	81.5	12.8	5.7	13.5	2,256
Sweden	37.2	20.2	79.8	276.1	275.2	18.1	82.2	2.6	1.7	13.6	40.2	29.0	30.8	9.0	1,504
Turkey*	32.3	70.5	29.5	225.7	222.0	21.1	95.4	4.1	0.4	0.0	92.9	5.8	1.3	21.0	1,837
United Kingdom	35.9	34.8	65.2	263.7	252.0	18.1	88.5	1.6	5.2	4.7	36.5	49.9	13.6	15.6	2,824
United States	36.1	16.9	83.1	253.6	235.4	20.5	80.3	2.6	4.0	13.1	23.1	48.8	28.1	13.0	1,625

Notes: * Second PIAAC round. Values for ISEI score, literacy, numeracy, and parental education are averages across 10 imputations. ISEI=International Socio-Economic Index of Occupational Status; ISCED=International Standard Classification of Education. Low parental education: no parent has completed upper secondary education; intermediate parental education: at least one parent has completed upper secondary education, but no parent has completed teriary education; high parental education: at least one parent has completed tertiaryeducation.

Sources: PIAAC (rounds 1 and 2), authors' calculations.

3.3 Country-level predictors

A key innovation of our study is to measure *skills transparency* directly using the *skills gap* between less- and intermediate educated adults and the *internal homogeneity* of these groups. In constructing the measures, we closely follow the work of Heisig and Solga (2015a) and Heisig, 2018.

The *skills gap* is the adjusted mean skills difference between less- and intermediate-educated adults. We construct this measure by running country-specific regressions of literacy and numeracy skills on a dummy variable for highest educational attainment, with sex, potential experience, foreign-birth/foreign-language status, and parental education as controls. We adjust the skills gap for these characteristics because they are readily observable and because we want to isolate the additional information conveyed by an individual's educational degree.[10] The skills gap for a given country is the coefficient estimate on having intermediate rather than low formal qualifications in the country-specific regression. Note that this coding is the opposite of that used in the regression models for occupational status, so larger (i.e., more positive) values correspond to a larger skills gap. Our final measure is the unweighted average of the estimated literacy and numeracy gaps for each country.

The index of internal homogeneity measures how homogenous the skills distribution within educational groups is, independent of their levels of skills. To compute the index, we first obtain the residuals from the country-specific regressions used in constructing the skills gap measure. For each educational group and for both literacy and numeracy, we then calculate the standard deviation of the residuals as a straightforward measure of within-group heterogeneity. The resulting four standard deviations (i.e., of the residual literacy and numeracy scores for less- and intermediate-educated adults, respectively) turn out to be strongly positively correlated (Heisig, 2018). To reduce the dimensionality, we run a principal factor analysis of the four standard deviations. The first factor loads positively on all four standard deviations and has an eigenvalue of 2.38 (averaged across the ten plausible values). The internal consistency of the four standard deviations is high, with the value of Cronbach's alpha (standardized) being equal to .82 (again, averaging across the ten plausible values). We reverse-code the factor scores so that higher values on the index indicate greater homogeneity.

[10] Parental education might be more difficult to observe than the other characteristics, but there is evidence that employers infer class background from other worker characteristics such as name, school attended, and leisure activities (Jackson, 2009). We also reran the analyses without adjusting the skills gap (and the index of internal homogeneity, see next paragraph) for parental education and results were similar (see Section 4.4).

For the emphasis on acquiring occupation-specific skills in upper secondary education, we use the *prevalence of vocational enrollment* measured by the percentage of students in upper secondary education who are enrolled in a vocational program. To reduce measurement error, we average the values provided in two sources: OECD (2006: Table C2.5) and UNESCO's online database (http://data.uis.unesco.org/). Values refer to 2004 (OECD) and 2006 (UNESCO) or the closest year available.[11] Our indicator is highly correlated (r = .99) with Bol and van de Werfhorst's (2013) vocational orientation index, which is based the same sources but not available for all countries in our sample.

We measure tracking in secondary education using the *external differentiation index* by Bol and van de Werfhorst (2013). The index is based on a principal factor analysis of three measures: age of first selection into different tracks (reverse coded), number of tracks available at age 15, and length of tracked education as a proportion of the total duration of primary and secondary education. Values for these variables refer to 2003 (age of first selection and number of tracks at age 15) and 2002 (length of tracked curriculum) or the closest year available (for details, see Bol and van de Werfhorst, 2013). The index is not available for three of the countries in our main analysis sample: Estonia, Lithuania, and Singapore.

Table 2 reports the values of the focal country-level predictors and of the unadjusted and fully adjusted ISEI gaps (see Section 3.4). Table 3 shows the pairwise correlations among them. In supplementary analyses (see Section 4.4), we include the unemployment rate and the employment shares of different economic sectors and labor market segments in the country-level regressions. We provide further details on these measures in Section 4.4 and in Section A of the Online Supplement.

[11] For a few countries, the OECD measure is unavailable. We simply use the UNESCO measure in these cases. Neither the OECD nor UNESCO provide data for Singapore, so we had to use the enrollment data from the World Bank available at http://datatopics.worldbank.org/education/.

Table 2. Values of focal country-level predictors

	Country code	Unadjusted ISEI gap (1)	Fully adjusted ISEI gap (2)	Skills gap (3)	Index of internal homogeneity (4)	Prevalence of vocational enrolment (5)	Index of external differentiation (6)
Austria	AT	-11.1	-7.9	23.8	1.09	78.3	1.82
Belgium	BE	-7.8	-5.0	21.9	0.33	61.8	1.02
Canada	CA	-7.8	-4.1	36.5	-1.23	2.8	-1.32
Chile*	CL	-6.3	-4.5	35.1	-0.05	37.0	0.32
Czech Rep.	CZ	-10.8	-8.1	24.6	1.49	79.2	1.62
Denmark	DK	-6.9	-4.7	22.7	-0.69	50.6	-0.87
Estonia	EE	-7.7	-5.4	26.8	0.19	31.0	Not available
Finland	FI	-3.3	-2.8	12.6	-0.25	57.1	-0.87
France	FR	-5.3	-4.0	27.6	-0.91	49.6	-0.47
Germany	DE	-9.4	-4.9	35.3	-0.83	60.3	1.86
Greece*	GR	-7.5	-5.6	21.4	-0.09	33.9	-0.47
Ireland	IE	-5.7	-4.7	28.9	-0.23	32.9	-0.30
Italy	IT	-11.8	-10.1	28.2	0.31	61.7	0.17
Japan	JP	-6.1	-4.7	21.8	1.54	24.6	-0.47
Korea	KR	-7.4	-5.3	24.7	1.62	28.6	0.07
Lithuania*	LT	-6.9	-5.9	12.2	0.30	28.2	Not available
Netherlands	NL	-9.1	-6.6	27.6	0.33	68.5	0.94
New Zealand*	NZ	-5.9	-4.6	29.0	-0.45	24.3	-0.42
Norway	NO	-3.4	-2.0	16.4	-0.47	60.2	-1.04
Poland	PL	-7.3	-4.9	18.6	-0.93	47.3	-0.08
Singapore*	SG	-12.4	-7.5	51.6	-2.37	11.3	Not available
Slovak Rep.	SK	-13.0	-8.2	31.1	1.16	73.6	1.62
Spain	ES	-8.7	-6.9	24.6	0.76	40.6	-1.02
Sweden	SE	-6.7	-4.2	22.3	0.37	55.8	-0.87
Turkey*	TR	-8.5	-7.4	27.1	0.36	37.6	1.20
United Kingdom	UK	-7.9	-5.2	26.2	-0.76	51.6	-1.04
United States	US	-10.2	-5.6	30.3	-0.57	0.0	-1.32
Mean		-8.0	-5.6	26.3	-0.00	44.0	0.00
Standard deviation		2.4	1.8	7.9	0.93	21.4	1.05

Notes: * Second round of PIAAC. For the country-level regressions, all predictors were (re-)standardized to have a mean of 0 and a standard deviation of 1 within the sample of 27 countries included in the analysis.

Sources(1) & (2): PIAAC, rounds 1 and 2, authors' calculations; (3): OECD (2006, Table C2.5), UNESCO online database (http://data.uis.unesco.org/) and World Bank online database (http://datatopics.worldbank.org/education); (4): Educational Systems Database, Version 4 (Bol and Van de Werfhorst, 2013).

Table 3. Pairwise Pearson correlations between focal country-level predictors

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Unadjusted ISEI gap	1					
(2) Fully adjusted ISEI gap	0.880^{**}	1				
(3) Skills gap	-0.540^{**}	-0.310	1			
(4) Index of internal homogeneity	-0.118	-0.327^{+}	-0.439^{*}	1		
(5) Prevalence of vocational enrolment	-0.148	-0.235	-0.340^{+}	0.393^{*}	1	
(6) Index of external differentiation	-0.579^{**}	-0.562^{**}	0.231	0.459^{*}	0.619^{**}	1

Notes: N=27. For pairwise correlations involving index of external differentiation N=24 because the index is not available for Estonia, Lithuania, and Singapore. + p < 0.1, * p < 0.05, ** p < 0.01.

Sources: See Table 2.

3.4 Analytical strategy and estimation

In the first step of the analysis, we seek to test hypotheses 1 and 2, which state that individual-level differences in literacy and numeracy skills can partly account for the labor market disadvantage of less-educated adults (H1) as well as for cross-national variation in its magnitude (H2).[12] We use the decomposition technique pioneered by Kitagawa (1955), and commonly referred to as the Oaxaca-Blinder decomposition, to assess these hypotheses. Following the notation of Fortin, Firpo, and Lemieux (2011), the variant of the decomposition that we use takes the following form:

$$\hat{\Delta}_O^{\mu} = \underbrace{(\bar{X}_B - \bar{X}_A)\hat{\beta}^*}_{\hat{\Delta}_X^{\mu}} + \underbrace{[\bar{X}_B(\hat{\beta}_B - \hat{\beta}^*) + \bar{X}_A(\hat{\beta}^* - \hat{\beta}_A)]}_{\hat{\Delta}_S^{\mu}}, \tag{1}$$

where the subscripts *A* and *B* index the two groups being compared and $\hat{\Delta}_O^{\mu} = \mu_B - \mu_A$ is the observed difference in the group means of the outcome variable (i.e., the unadjusted ISEI gap). In our case, the less-educated are group *B* and the intermediate educated group *A*. $\hat{\Delta}_O^{\mu}$ is decomposed into an explained part $\hat{\Delta}_X^{\mu}$ and an unexplained part $\hat{\Delta}_S^{\mu}$. The explained part is the sum of the differences in the group means for a set of *k* explanatory variables (i.e., $\bar{X}_B - \bar{X}_A$), with the mean difference for each variable weighted (or "priced") according to the corresponding coefficient estimate from the vector $\hat{\beta}^*$. This vector is estimated by running a regression of the ISEI score on the explanatory variables (the skill measures and the individual-level controls) using the data of both educational groups.[13] The terms $\hat{\beta}_A$ and $\hat{\beta}_B$ in the unexplained part represent the coefficient vectors estimated using only the data from group *A* or *B* (for further details, see Fortin et al. 2011).

Due to the linear additive nature of the decomposition, it is possible to calculate the contributions of individual variables or subsets of variables, often referred to as a "detailed decomposition" (Jann, 2008; Fortin et al., 2011). Given our research questions, we are particularly interested in the combined contribution of group differences in literacy and numeracy skills to the ISEI gap. To assess H1, we examine whether the skill measures explain

[12] For simplicity, the following presentation of our empirical approach abstracts from the fact that we have to run each analysis step multiple times to account for the multiply imputed/plausible values.

[13] As recommended in the literature, this regression also includes a group indicator (i.e., a dummy for having intermediate education; see Fortin et al., 2011).

a substantial and statistically significant portion of the ISEI gap in the majority or even all countries. To assess H2, we investigate whether adjusting for differences in literacy and numeracy skills reduces the cross-country variation of the ISEI gaps. That is, we investigate whether the cross-country variance of the unexplained portion of the gap remaining after adjusting for group differences in literacy and numeracy skills, the "skills-adjusted" ISEI gap, is smaller than the cross-country variance of the observed (unadjusted) ISEI gap.[14]

In the second step of the analysis, we test hypotheses 3 to 5 using country-level regressions. The dependent variable in these regressions is the "fully adjusted" ISEI gap, that is, the ISEI gap after adjusting not only for differences in literacy and numeracy skills but also for compositional differences with respect to the socio-demographic controls. To estimate it, we run country-specific regressions of the ISEI score on the skill measures, the individual-level controls, and an indicator for belonging to the less-educated group, with the coefficient on the latter variable providing the estimate of the fully adjusted ISEI gap. The full results of these country-specific regressions are reported in Table C1 in the Online Supplement. While we use pooled regressions with a group dummy to estimate the fully adjusted gap, it is worth noting that it is conceptually equivalent to the unexplained component of the wage gap (i.e., $\hat{\Delta}_S^u$) in Equation 1 above (Elder et al., 2010).[15]

The independent variables in the country-level regressions are the focal explanatory variables (i.e., the skills transparency measures, the prevalence of vocational enrollment, and the index of external differentiation) and the additional country-level controls (i.e., the unemployment rate and sectoral composition). The regressions are estimated using a Feasible Generalized Least Squares (FGLS) approach that accounts for the fact that the dependent variable is estimated rather than observed and therefore subject to sampling error (i.e., the regressand is a set of coefficient *estimates* from the first-step regressions rather than the unobservable true coefficients). By accounting for country differences in the precision of the first-step estimates, FGLS addresses the resulting heteroskedasticity and achieves greater

[14] Note that while the skills-adjusted ISEI gap is adjusted only for group differences in average literacy and numeracy skills (and not for differences in the socio-demographic controls), the "skill prices" (i.e., the coefficient estimates) used in calculating the adjustment are "net" skill prices from a (pooled) regression that does include the controls in addition to the skill measures (and a group dummy).

[15] In the present case, the two pooled regression and decomposition methods of obtaining the ("adjusted or unexplained") gap are even mathematically equivalent, at least with respect to the point estimate, because we use a pooled regression with a group membership dummy to estimate the coefficient vector for the decomposition (this mathematical equivalence would not hold if the model for the reference coefficient were estimated using only one of the groups or if we did not include a dummy in the pooled model; see Elder et al., 2010). The two methods produce somewhat different standard error estimates, however, and we prefer the ones from the pooled regression, which tend to be more conservative.

efficiency than OLS estimation of country-level relationships (Heisig et al., 2017; Lewis and Linzer, 2005).

4 RESULTS

We start this section with results for the role of individual skills based on country-specific decompositions. We then turn to a country-level analysis to examine the roles of skills transparency (skills gaps and internal homogeneity of educational groups) and vocational enrollment. The next step of the analysis investigates whether the effect of external differentiation (tracking) in secondary education on labor market inequalities is mediated by skills transparency. The section concludes with several robustness checks (e.g., for the role of macroeconomic context and sectoral composition).

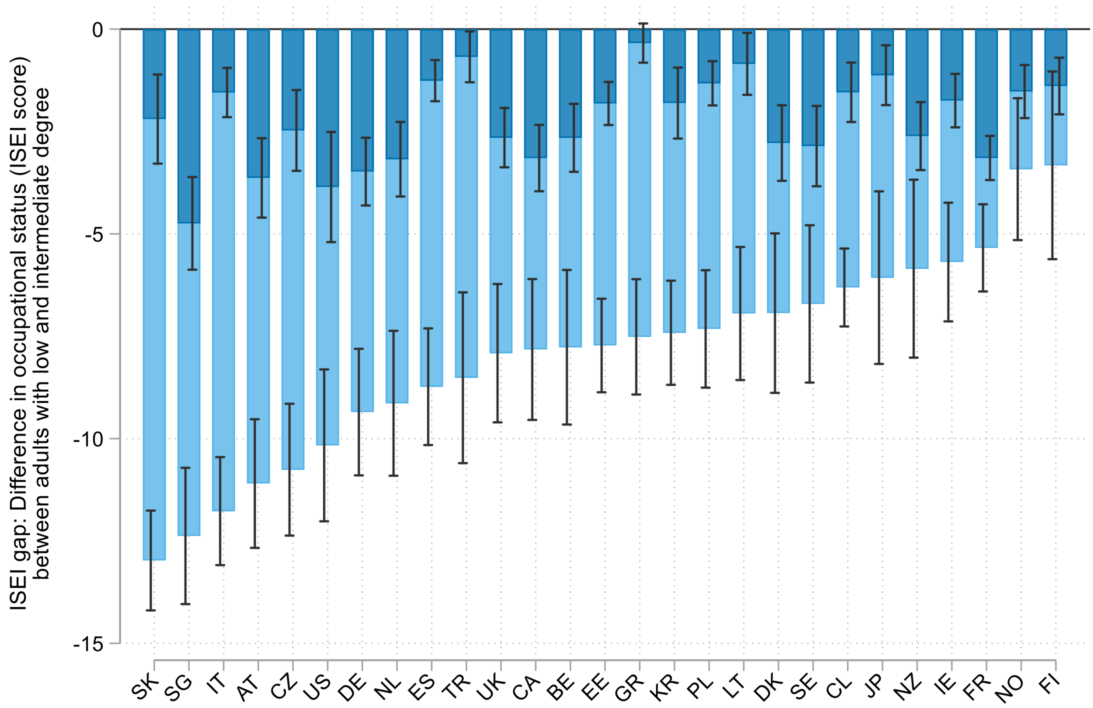
4.1 The role of individual skills

Figure 1 summarizes the country-specific Kitagawa-Oaxaca-Blinder decompositions of the occupational status (ISEI) gap between adults with low (ISCED 0-2) and intermediate (ISCED 3-4) formal qualifications. As throughout this paper, the gap is calculated as the average ISEI score for less-educated adults minus the average score of intermediate-educated adults. “More negative” values thus correspond to a greater labor market disadvantage for the less educated.

For each of the 27 countries in our main analysis sample, Figure 1 shows the unadjusted ISEI gap between less-and intermediate-educated adults as well as the part of the gap that is attributable to group differences in literacy and numeracy skills. The unadjusted gap is represented by the overall length of the bars. It is negative and statistically significant ($p < .05$) in all countries. Cross-national variation is considerable, with the gap ranging from -13.0 points in Slovakia and -12.4 points in Singapore to only -3.4 and -3.3 points in Norway and Finland, respectively (see also Table 2 above). The average unadjusted gap equals -8.0 points, with a cross-country standard deviation of 2.4 points.

To what extent can the labor market disadvantage of less-educated adults be attributed to differences in literacy and numeracy skills between less- and intermediate-educated adults? This question is answered by the darker segments of the bars in Figure 1, which represent the part of the gap that is explained by differences in literacy and numeracy skills according to the decomposition results. In most countries, differences in literacy and numeracy skills account for a substantial portion of the ISEI gap (and this portion is statistically significant at the five percent level for all countries except Greece). The explained part of the occupational status gap averages -2.2 ISEI points across the 27 countries, somewhat less than 30 per cent of the total gap of -8.0 points. We thus find ample support for hypothesis 1. The average unexplained part, represented by the lighter segments of the bars in Figure 1, is approximately -5.8 (= -8.0 – [-2.2]) ISEI points.

Figure 1. The ISEI gap between less- and intermediate-educated adults in 27 countries



Notes: See Table 2 for country codes. The bars represent the occupational status gap, measured in ISEI points, between less- and intermediate-educated adults. The darker segment indicates the part of the gap that is attributable to differences in literacy and numeracy skills according to the decomposition results (see text for details). The capped lines indicate 95% confidence intervals for the overall gap and for thr portion attributable to skills.

Sources: PIAAC (rounds 1 and 2), authors' calculations.

Hypothesis 2 states that accounting for individual-level differences with respect to literacy and numeracy skills will reduce cross-national variation in the ISEI gap. Consistent with this prediction, we find that the unexplained portion of the ISEI gap that remains after accounting for skills exhibits less cross-country variation than the unadjusted gap (i.e., the length of the lighter segments of the bars in Figure 1 is less variable than the overall length). Whereas the cross-country variance is 6.0 in the unadjusted case, it is only 5.0 after accounting for literacy and numeracy skills—a reduction of approximately 17 percent.

The overall cross-country variation of the estimates for both the unadjusted gap and the unexplained portion that remains after adjusting for skills comprises both "true" variation in the ISEI gap and variation due to sampling error. When we estimate a random effects model to separate the two components (see the discussion in Section 3.2 above), we find that the true variation (τ^2) declines from 5.3 (95% confidence limits: 3.0; 10.5) to 4.2 (95% confidence limits: 2.3; 8.5)—a reduction of approximately 20 percent.[16] As expected by hypothesis 2, accounting for individual-level differences in skills thus appears to reduce cross-country variation in the estimated ISEI gap, but the considerable overlap between the confidence intervals indicates that this result must be viewed as suggestive.

We also investigated to what extent the ISEI gap can be explained by the control variables (sex, foreign-birth/foreign-language status, parental education, and potential experience). As discussed above (see Section 3.4), the "fully adjusted gap" (i.e., the dependent variable in the country-level regressions presented below) is effectively the unadjusted gap minus the contributions of literacy/numeracy skills and of the lower-level controls. Figure B1 in the Online Supplement shows that the combined contribution of the control variables is ambiguous. In most countries, compositional differences with respect to the controls contribute to the ISEI gap, but in quite a few countries they also appear to reduce it, meaning that the observed ISEI gap would be even larger in the absence of these compositional differences. In terms of the size of the contributions, the controls tend to play a smaller role than literacy and numeracy skills, although there are also a few countries (e.g., Slovakia, Greece, and France) where their impact is quite substantial.

4.2 The roles of skills transparency and vocational orientation

How can the remaining cross-country variation in the ISEI gap be explained and what, in particular, is the role of skills transparency? We make a first attempt to answer these questions in Figure 2 where we visually explore the relationships between the skills transparency measures, vocational enrollment, and the fully adjusted ISEI gap (i.e., the unexplained gap that remains after adjusting the ISEI gap for literacy and numeracy skills as well as the additional

[16] For technical reasons, the standard errors for the unexplained part of the ISEI gap cannot be corrected for the complex sampling design using the jackknife replication weights. To address this issue, we multiplied the uncorrected standard errors for the unexplained part by the ratio of the corrected and the uncorrected errors for the unadjusted gap.

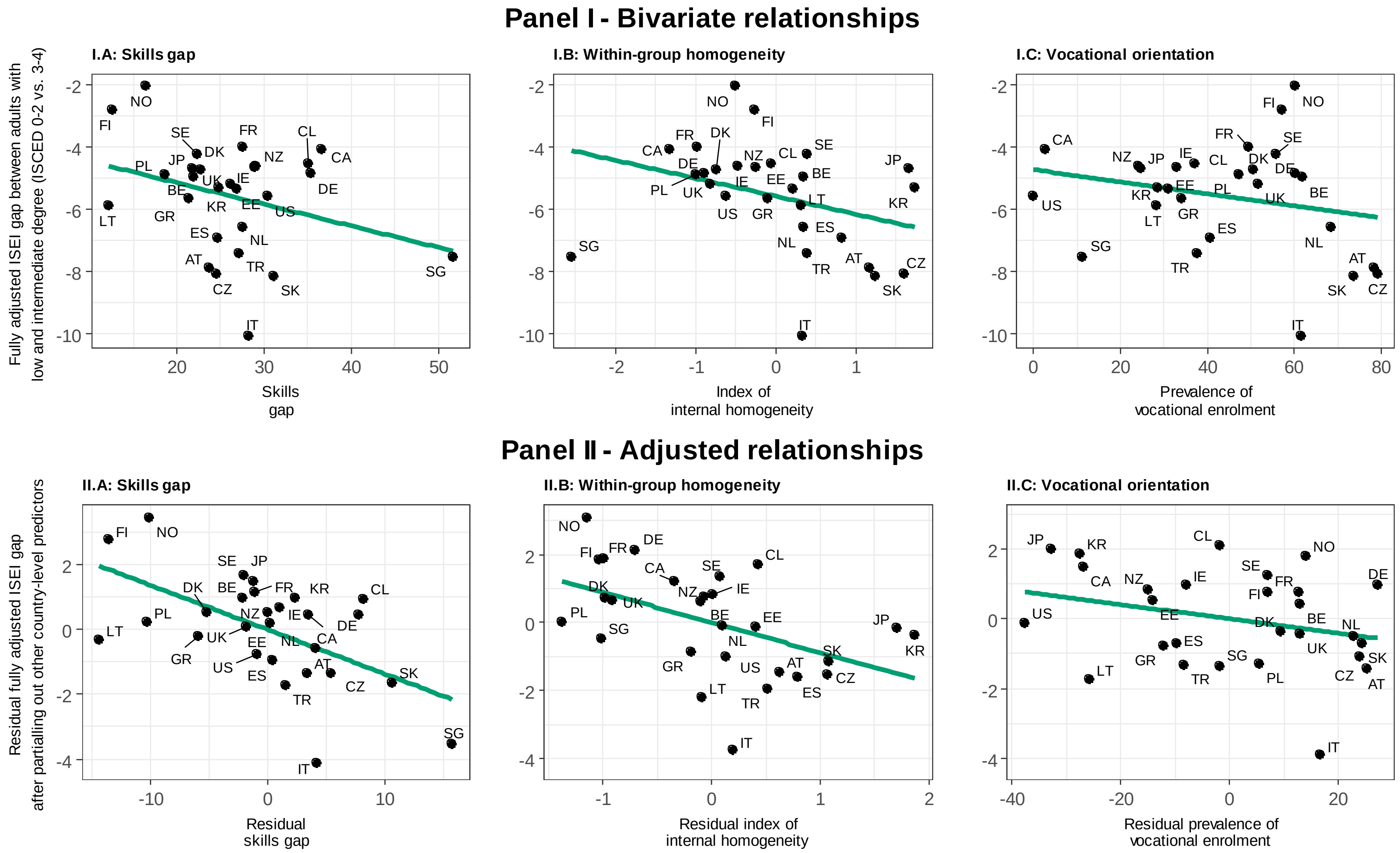
control variables). The three graphs in Panel I (top row) depict simple bivariate relationships, with the lines representing simple linear fits estimated by OLS.[17]

Consistent with hypotheses 3, 4, and 6, we find that the ISEI gap between less- and intermediate-educated adults tends to increase with the skills gap between less- and intermediate-educated adults, with the internal homogeneity of these groups, and with the prevalence of vocational enrollment in upper secondary education: The ISEI gap becomes "more negative" as these country-level characteristics increase. None of the relationships seems to be driven by single countries, although there are clearly some potentially influential cases. Finland and Norway, the two countries where the ISEI gap is smallest, also have very small skills gaps. Singapore stands out as a country with a very large skills gap and very low levels of internal homogeneity. We further check for potential outlier issues below (see Section 4.4 and Section D in Online Supplement).

Panel II of Figure 2 displays the partial relationships between the ISEI gap and three country-level characteristics. The graphs relate residual variation in the ISEI gap to residual variation in the focal country-level predictor, after accounting for the effects of the respective other two predictors. For example, we regressed the ISEI gap and the skills gap on the indices of internal homogeneity and prevalence of vocational enrollment to compute the residuals depicted in Panel II.A. According to the Frisch-Waugh-Lovell theorem, this has the same effect as controlling for the other two characteristics in conventional multiple regression (Davidson and MacKinnon, 2004, Chapter 2). Thus, the bottom row of graphs displays the adjusted (partial) relationships between the ISEI gap and the country-level predictors. We see that the adjusted relationships continue to go in the expected negative direction. Especially for the skills gap and the index of internal homogeneity, the adjusted relationships appear clearer than the simple bivariate associations. Several countries that look like potential outliers in the bivariate case no longer appear problematic when the respective other two characteristics are taken into account. The Singaporean case in particular gives less reason for concern in Panels II.A and II.B than in Panels I.A and I.B. Again, we investigate potential outlier issues more systematically in Section 4.4 below.

[17] The slopes of these lines differ somewhat from those estimated in the formal country-level regression analysis (see Table 4 below) because they are based on the final point estimates (rather than running the country-level regressions on each of the ten imputed data sets). Moreover, they are based on unweighted regressions, whereas the formal country-level regression analysis uses an FGLS approach that gives greater weight to more precise estimates of the ISEI gap (see Section 3.4).

Figure 2. Country-level relationships between fully adjusted ISEI gap and measures of skills transparency



Notes: See Table 2 for country codes. Lines are linear fits estimated using ordinary least squares. Panel II shows relationships after partialing out the effects of the respective other two characteristics.

Table 4 displays the results of the more formal country-level analysis based on FGLS regressions. We present seven models that cover all possible combinations of the skills gap, the index of internal homogeneity, and the prevalence of vocational enrollment in upper secondary education. Models 1 to 3 include the three characteristics one at a time. Models 4 to 6 show the three possible two-way combinations between them, and Model 7 includes all three variables simultaneously. All three predictors are z-standardized, so the coefficient estimates can be interpreted as the predicted change in the fully adjusted ISEI gap associated with a standard deviation increase in the respective characteristic.

The signs of the coefficient estimates in Table 4 are generally consistent with hypotheses 3, 4, and 6: The coefficient estimates for all three country-level characteristics are negative throughout the table, indicating that an increase in the respective predictor is associated with a greater labor market disadvantage of less-educated adults. The bivariate associations in Models 1 to 3 do not reach statistical significance, but several coeffcient estimates become statistically significant once we include at least two of the predictors simultanously in Models 4 to 6. Our preferred specification, Model 7, includes all predictors simultaneously.

Table 4. Country-level regressions of ISEI gap on measures of skills transparency and vocational orientation

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Skills gap	-0.513			-0.906*	-0.726+		-0.992*
	(0.366)			(0.352)	(0.370)		(0.354)
Index of internal homogeneity		-0.630+		-0.990*		-0.529	-0.859*
		(0.343)		(0.352)		(0.371)	(0.358)
Prevalence of vocational enrollment			-0.480		-0.707+	-0.279	-0.469
			(0.361)		(0.363)	(0.381)	(0.349)
Intercept	-5.622***	-5.608***	-5.604***	-5.615***	-5.603***	-5.601***	-5.600***
	(0.341)	(0.331)	(0.341)	(0.301)	(0.324)	(0.334)	(0.296)
N	27	27	27	27	27	27	27
R2	0.08	0.12	0.07	0.33	0.21	0.14	0.38
Adjusted R2	0.04	0.09	0.03	0.27	0.14	0.07	0.30

Notes: Feasible Generalized Least Squares (FGLS) estimates, based on 10 imputations/plausible values. Dependent variable: the fully adjusted ISEI gap between less-educated and intermediate-educated adults aged 16-54 (see Figure 1 above and Section A in Online Supplement). All country-level variables are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Sources: PIAAC (rounds 1 and 2), authors' calculations.

According to this specification, a standard deviation increase in the skills gap is associated with a -.992 point change in the ISEI gap between less- and intermediate-educated adults ($p < .05$). This effect size is substantial, given that the average fully adjusted ISEI gap across the 27 countries in our sample is -5.58 points, with a cross-country standard deviation of 1.77

points—it corresponds to about 56 percent of the standard deviation. This result supports hypothesis 3, which expects the aggregate skills differential between less- and intermediate-educated adults to have an independent effect on the labor market disadvantage of less-educated adults above and beyond the direct individual-level effect of skills (which is accounted for in the first-step regressions).

Hypothesis 4, which posits that higher internal homogeneity of the educational groups increases the ISEI gap between less- and intermediate-educated adults, is supported as well. The ISEI gap is larger in countries where the less- and intermediate-educated groups are internally more homogeneous in terms of literacy und numeracy skills. According to Model 7, the ISEI gap grows by -.859 points ($p < .05$) with each standard deviation increase in the index of internal homogeneity. This equates to approximately 49 percent of the cross-country standard deviation of the fully adjusted ISEI gap.

These results support the notion that skills do not only matter at the individual level. The skills transparency of educational certificates (as captured by the skills gap and internal homogeneity of educational groups) appears to be an additional source of cross-national variation in the labor market disadvantage of less-educated workers.

Turning to the role of vocational education and training systems, the direction of the coefficient on the vocational enrollment measure is consistent with previous findings (e.g., Bol and van de Werfhorst, 2011; Shavit and Müller, 1998; van de Werfhorst, 2011). As expected in hypothesis 6, Table 4 suggests that the ISEI gap between less- and intermediate-educated adults is larger in countries where upper secondary education puts greater emphasis on occupation-specific skills. That being said, the empirical evidence for this relationship is weaker than one might expect given its central place in the extant literature: Although the coefficient estimates on the prevalence of vocational enrollment are quite sizable, their (absolute) magnitude is smaller than for the two direct measures of skills transparency. Moreover, the coefficient on vocational enrollment is mostly far from attaining significance, except in Model 5 ($b = -.707$, $p = .065$).

4.3 Does skills transparency mediate the effect of external differentiation?

Can our novel measures of skills transparency help us make better sense of findings in the existing literature? In particular, can we provide more direct evidence that the moderating role of external differentiation (tracking) in secondary education is due to a positive relationship between external differentiation and skills transparency (Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011)? We address these questions with an additional sequence of regression model in Table 5. As noted above (see Section 3.3), the external differentiation index is unavailable for Estonia, Lithuania, and Singapore. We therefore have to exclude these countries from this step of the analysis, which reduces the country sample to 24 cases.

Model 1 in Table 5 regresses the fully adjusted ISEI gap on the external differentiation index. Consistent with previous research (e.g., Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011; van de Werfhorst, 2011), the coefficient estimate is negative and statistically significant, indicating that the labor market disadvantage of less-educated adults increases with the extent of tracking in secondary education. At -1.034 ($p < .01$) the size of the coefficient is quite substantial and broadly comparable to those of the direct skills transparency measures in Table 4 above. Model 2 adds the prevalence of vocational enrollment. The coefficient on vocational enrollment is negative, but rather small and statistically insignificant. The coefficient on the external differentiation index is only slightly weaker than in Model 1 and remains significant at the five percent level. Thus we are able to reproduce the finding that stronger external differentiation is associated with greater labor market disadvantages for less-educated adults.

Adding the direct measures of skills transparency (Model 3) leads to a dramatic attenuation of the effect of external differentiation, as the coefficient estimate declines by more than 80 percent (in absolute size) relative to Model 2 (from -0.995 to -.172). By contrast, the coefficients on the direct measures of skills transparency—the skills gap and the index of internal homogeneity—are broadly similar to those from the previous step of the analysis (see Table 4 above). Only the coefficient of internal homogeneity reaches statistical significance ($p < .05$), but a comparison with Table 4 shows that the loss of statistical significance for the skills gap measure is primarily due to decline in precision. At -.935 the coefficient estimate for the skills gap measure is actually very similar to Model 7 in Table 4 (-.992). The standard error, however, is considerably larger: .615 rather than .354. The reason is multicollinearity: External differentiation is strongly related to both internal homogeneity (Heisig, 2018) and especially to the size of the skills gap, at least when vocational orientation is controlled (Heisig and Solga, 2015a)—as it needs to be if these two dimensions of skills transparency are to be plausible mediators of the effect of tracking. The final column of Model 4 underlines this point by showing that the coefficients on both the skills gap and the index of internal homogeneity do not change much when the external differentiation index is dropped (while maintaining the reduced sample of 24 countries). The absolute size of the coefficients on the skills transparency measures increases only slightly, but their standard errors decline considerably. The vocational enrollment measure also comes close to reaching statistical significance in this specification (b = -.646; p = .099).

Table 5. Country-level regressions of ISEI gap on measures of external differentiation, skills transparency, and vocational orientation

	Model 1	Model 2	Model 3	Model 4
Index of external differentiation	-1.034**	-0.995*	-0.172	
	(0.345)	(0.443)	(0.570)	
Prevalence of vocational enrollment		-0.065	-0.532	-0.646+
		(0.463)	(0.547)	(0.370)
Skills gap			-0.935	-1.035*
			(0.615)	(0.471)
Index of internal homogeneity			-0.917*	-0.975*
			(0.423)	(0.368)
Intercept	-5.546***	-5.534***	-5.440***	-5.427***
	(0.331)	(0.345)	(0.325)	(0.314)
N	24	24	24	24
R2	0.30	0.30	0.47	0.47
Adjusted R2	0.26	0.23	0.36	0.39

Note: Estonia, Lithuania, and Singapore are excluded because of missing information for the index of external differentiation. All variables are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001 (two-tailed tests). See text and note to Table 4 for further information.

Sources: PIAAC (rounds 1 and 2), authors' calculations.

In sum, these findings provide substantial support for hypothesis 5 which expects the well-document effect of external differentiation on the labor market disadvantage of less-educated adults to be attenuated substantially once direct measures of skills transparency are included in the regression. In more substantive terms, these findings indicate that skills transparency is an important channel through which external differentiation is related to labor market inequalities among educational groups.

4.4 Alternative explanations and robustness checks

In Table 6, we present a series of further analyses to assess potential alternative explanations and explore the robustness of our findings. The first two models control for the unemployment rate to address the concern that our results might be driven by country differences in macroeconomic conditions. Data come from the World Bank.[18] We use the

[18] https://data.worldbank.org/, downloaded on September 3, 2018.

mean of the 2011 and 2012 values for the first-round and the mean of 2014 and 2015 unemployment rates for the second-round countries.

Model 1 includes the unemployment rate linearly and Model 2 adds a squared term. The linear term is z-standardized, while the squared term is the square of the standardized variable. The untransformed values of the unemployment rate can be found in Table A3 in the Online Supplement. Model 1 suggests that less-educated workers tend to face greater disadvantages when unemployment is high, but the coefficient on the unemployment rate is not statistically significant. More importantly, the effects of our three focal country-level predictors—the skills gap, the index of internal homogeneity, and vocational enrollment measure—are very similar to Model 7 in Table 4. This does not change when we add the square of the unemployment rate to allow for a non-linear effect in Model 2.

Models 3 to 6 account for demand-side factors by including the prevalence of employment in different sectors and labor market segments, expressed as the share of overall employment. The employment shares are based on International Labour Office (ILO) data on employment by industry, defined according to the fourth revision of the International Standard Industrial Classification (ISIC).[19] As with the unemployment rate, we use the mean of the 2011 and 2012 shares for first-round and of the 2014 and 2015 shares for second-round countries.[20]

Models 3 and 4 include the share of employment in the primary and secondary (manufacturing) sector. The employment share of the tertiary (service) sector is omitted because it is perfectly collinear with the employment shares of the other two sectors. Models 5 and 6 use a more fine-grained typology that groups industries into six labor market segments. It builds on the work of Stinchcombe (1979), as implemented in Carroll and Mayer (1986), and distinguishes among the following segments: traditional primary; small competitive; competitive; large-scale engineering; professional; bureaucratic. [21] We omit the share of the traditional primary segment from the regressions to avoid perfect multicollinearity. We provide further information on the different segment measures in Section A in the Online Supplement, including the values of the employment shares for each country (see Table A3). For the regression models, the measures were again z-standardized.

[19] Data were obtained from https://www.ilo.org/global/statistics-and-databases/lang--en/index.htm on October 8, 2018. For two countries, Canada and Chile, industries are classified according to the third revision of the ISIC.

[20] The one exception is Canada where we have to use the values for 2016, the only year covered by the ILO data.

[21] Carroll and Mayer (1986) identify a seventh segment, "classical capitalist", but the industry classification provided by the World Bank is not fine-grained enough to differentiate it from the "small competitive" segment.

Table 6. Additional country-level regressions

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Skills gap	-1.068**	-1.076**	-0.958*		-0.663	
	(0.347)	(0.348)	(0.383)		(0.538)	
Index of internal homogeneity	-0.821*	-0.858*	-0.690+		-0.528	
	(0.356)	(0.356)	(0.368)		(0.534)	
Prevalence of vocational enrollment	-0.557	-0.526	-0.460		-0.500	
	(0.341)	(0.342)	(0.365)		(0.470)	
Unemployment rate	-0.456	-0.812				
	(0.297)	(0.481)				
Unemployment rate (squared)		0.194				
		(0.216)				
Employment shares of broader sectors (Ref.: tertiary sector)						
Primary sector			-0.367	-0.132		
			(0.334)	(0.330)		
Manufacturing			-0.395	-0.865*		
			(0.338)	(0.319)		
Employment shares of detailed segments (Ref.: traditional primary segment)						
Small competitive					-0.156	-0.691
					(0.722)	(0.581)
Competitive					0.248	0.182
					(0.512)	(0.387)
Large-scale engineering					-0.246	-1.050+
					(0.738)	(0.592)
Professional					0.650	0.315
					(0.621)	(0.606)
Bureaucratic					-0.207	-0.443
					(0.387)	(0.319)
Intercept	-5.571***	-5.752***	-5.597***	-5.616***	-5.569***	-5.576***
	(0.287)	(0.353)	(0.288)	(0.315)	(0.296)	(0.304)
N	27	27	27	27	27	27
R2	0.45	0.47	0.47	0.24	0.51	0.39
Adjusted R2	0.35	0.35	0.34	0.18	0.30	0.24

Note: All variables, except square of unemployment rate, are z-standardized (mean of 0, standard deviation of 1). Square of unemployment rate is the square of the z-standardized unemployment rate. Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests). See text and note to Table 4 for further information.

Sources: PIAAC (rounds 1 and 2), authors' calculations.

The coefficient of the skills gap proves very robust to the inclusion of the broader sector share measures in Table 6. It remains statistically significant at the five percent level and the effect size is very similar to Model 7 in Table 4. The effect of internal homogeneity is noticeably weaker when the sector shares are controlled (b = -.690 as opposed to -.859 in Model 7 in Table 4) it no longer meets conventional standards for statistical significance (p = .078). However, this finding must be seen in the context of the limited degrees of freedom available in this country-level analysis. It is also worth noting that effect sizes for the sectoral share measures are rather small in Model 3 and that both fail to attain statistical significance. This is underlined by Model 4, which omits the skills transparency and vocational enrolment measures, to examine the effects of sectoral composition with a more parsimonious specification. The coefficient on the primary share remains small and insignificant in this specification, while manufacturing share shows a negative and statistically significant coefficient ($b = -.865$; $p < .05$). While this provides some evidence for the relevance of demand-side explanations, it does not indicate that sectoral composition—at least at this level of aggregation—is a major driver of cross-national differences in the labor market disadvantage of less-educated adults. In particular, the effects of both skills transparency measures appear more robust to the inclusion of measures of sectoral composition than *vice versa*.

Model 5 in Table 6 includes the segment shares in addition to the skills transparency and vocational enrollment measures. It fails to show any statistically significant effects. The effect of the skills gap again appears somewhat more robust to the inclusion of the segment shares than the effect of internal homogeneity, but even the former is far from reaching statistical significance. However, the segment measures themselves are not strongly predictive of the ISEI gap either. When the skills transparency and vocational enrolment measures are omitted in Model 6, we do find the share of employment in large-scale engineering comes close to reaching statistical significance ($p = .092$). With a coefficient estimate of -1.050 it is also substantially sized, but the comparison with Model 5 makes very clear that it cannot be viewed as more robust than the effects of the skills transparency measures that have been the focus of our analysis.

In summary, the supplementary analyses presented in Table 6 yield two main conclusions. First, country differences in labor market conditions, while potentially of some relevance, do not seem to drive the relationships between our focal predictors and the labor market disadvantage of less-educated adults. Second, country differences in the structure of employment may likewise play some role for cross-national variation in the labor market disadvantage of less-educated adults, especially when it comes to the relative size of the manufacturing sector or the related large-scale engineering segment. However, we find no clear evidence that the relationships between our focal predictors and the ISEI gap are spurious and ultimately attributable to cross-national differences in demand-side factors.

We conducted several further robustness checks, which we report in the Online Supplement. In a first set of analyses, reported in Section D of the Online Supplement, we examined the influence individual country cases and pairs of countries on the country-level regression results, focusing on our preferred specification from the main sequence of regression models, Model 7 in Table 4 above. Initial analyses showed that Israel and Slovenia jointly have a dramatic impact on the regression results, as measured by the DFBETA and

Cook’s D influence statistics. This led us to exclude these to countries from the main analysis. In general, support for H4 (internal homogeneity) is weaker and support for H6 (vocational orientation) stronger when Israel and Slovenia are included. Additional influence diagnostics for the main analysis sample of 27 countries revealed no further cases with extreme influence.

In a second set of robustness checks, we explored the impact of changing the individual-level sample restrictions. In particular, we reeaximined the main sequence of country-level regressions in Table 4 above after excluding respondents who were not employed at the time of interview (see Section E in the Online Supplement) and after exluding respondents who were self-employed (see Section F in the Online Supplement). Results were very similar to the main analysis, with the most noteworthy difference being that the coefficient on the index of internal homogeneity just misses statistical significance in Model 7 when the self-employed are excluded (b = -.765; se = .390; p = .064).

In a third robustness check (see Section G in the Online Supplement), we omitted we omitted parental education from the control variables used in constructing the skills transparency indicators (the skills gap and index of internal homogeneity) because parental education may be less readily visible to employers than the other characteristics that we adjusted for in constructing these measures (sex, age, and foreign-birth/foreign-language status). Results were very similar to the main analysis.

In a fourth check (see Section H in the Online Supplement), we reran the regressions in Table 4 with the gap in log hourly wages rather than occupational status as the dependent variable. This analysis provides essentially no clear for either Hypothesis 3, 4, or 6. As noted above (see Section 3.2), we suspect that these inconclusive results are due to a combination of more noise in the measured wage gaps and unmodeled confounding by contextual factors such as collective-bargaining institutions and minimum wage regulations.

5 CONCLUSIONS

The main goal of our paper was to further our understanding of cross-national differences in the labor market disadvantage of less-educated adults. We used the recent PIAAC data, which provide higher quality measures of the actual (literacy and numeracy) skills of adults than previous cross-national data sets.

Previous research has shown that formal qualifications are more important for occupational attainment in countries with extensive ability-related external differentiation or “tracking” in secondary education (e.g., Andersen and van de Werfhorst, 2010; Bol and van de Werfhorst, 2011; Shavit and Müller, 1998). This research has assumed that skills transparency of educational certificates—that is, the degree to which formal qualifications are indicative of actual skills—is the major mediating channel behind this effect of external differentiation. Our study is the first that examines this assumption more directly by employing two direct measures of skills transparency, the mean skills difference between less- and intermediate-educated adults (which we refer to as the “skills gap”) and an index that captures the internal homogeneity educational groups in terms of literacy und numeracy skills. Our analysis confirms that stronger external differentiation of upper secondary education systems is associated with larger labor market disadvantages for less-educated adults, measured as the adjusted difference in occupational status between less- and intermediate-educated workers. More importantly, we also find that the country-level association between external differentiation and the labor market disadvantage of less-educated adults is largely accounted for by our two direct measures of skills transparencyNotably, these findings hold after controlling for differences in literacy and numeracy skills at the individual level.

The finding that direct measures of the skills transparency explain labor market inequalities between educational groups is consistent with theories of labor market signaling and screening. In countries where skills transparency is high—and a person’s formal qualifications therefore send a stronger signal about his/her actual skills—employers seem to be more likely to use these qualifications as a basis for statistical discrimination. Taken together, our findings also speak against a strong version of credentialism, which claims that labor market returns to educational certificates—in our study, returns to completing upper secondary education—are unrelated to skills and productivity.

Compared to previous studies (Bol and van de Werfhorst, 2011; Shavit and Müller, 1998; van de Werfhorst, 2011), we found only relatively weak support for the notion that a stronger vocational orientation of upper secondary education exacerbates the labor market disadvantage of less-educated adults. This suggests a need for further research and replication to ascertain the robustness of this finding.

A crucial improvement over most previous research is that we also examined how differences in actual literacy and numeracy skills are related to occupational status at the individual level. Our results show that the actual skills of less-educated workers are an important predictor of their occupational status attainment, as suggested by human capital theory. Hence, country differences in the (relative) level of skills achieved by less-educated adults appear to be another important source of country variation in their labor market disadvantage. An obvious policy implication of this last finding is to improve the education system and adult training in order to raise the skills of less-educated adults (Heisig and Solga, 2015a; Park and Kyei, 2011).

On a more subtle note, our findings concerning the role of skills transparency point to a possible trade-off that may need to be taken into account when designing policies to improve the labor market prospects of less-educated adults. On the one hand, skills transparency (in the sense of a stronger association between formal qualifications and skills) should facilitate labor market matching and may contribute to merit-based hiring and promotion decisions. Low skills transparency might undermine trust in educational degrees and, thus, employers might pay greater attention to social origin, ethnicity, or gender when assessing applicants, raising inequalities by these (ascriptive) characteristics.

On the other hand, a potential downside of high levels of skills transparency is that it may reinforce the disadvantages of less-educated adults who are perceived to have low skills, possibly even leading to a larger (statistical) discrimination of the group (Solga, 2002). Even in "skills transparent" countries where the less educated are relatively homogeneous, we still find substantial within-group variation in literacy and numeracy skills (Heisig, 2018). Yet if less-educated workers are facing statistical discrimination based on their formal qualification, the more skilled members of the group might find it difficult to translate their higher skills into better labor market outcomes (e.g., because they are screened out during the early stages of the hiring process). This suggests that labor market returns to skills may be particularly low for the less educated. Moreover, skills transparency might moderate this individual-level interaction between formal qualification and skills because statistical discrimination on the basis of formal qualification should be when skills transparency is high (i.e., there may be a three-way interaction between skills and educational degrees at the individual and skills transparency at the country level). Future research should investigate this possibility in greater depth, although much larger sample sizes than provided by PIAAC are probably needed to identify such complex relationships (for some suggestive evidence, see Heisig and Solga, 2015b).

We conclude with some limitations of our study and with related questions for future research. A first limitation is that our direct measures of skills remain incomplete. While PIAAC is the richest and most advanced cross-national survey of adult skills to date, general literacy and numeracy skills likely are not the only skills that matter for labor market attainment. Other important types of skills emphasized in the literature include occupation-specific (e.g., van de Werfhorst, 2011) and non-cognitive skills (e.g., Heckman et al., 2006).[22] Information on these skills would be a valuable resource for extending and refining our analysis. The next round of PIAAC (planned for 2021/22), which is set to collect also information on noncognitive skills, will be an important step forward in this respect.

A second limitation is that we cannot rule out that our results are confounded by unobserved third variables. While this concern applies to any observational study, it looms particularly large in cross-national comparisons, due to small sample sizes and a lack of data on potentially relevant country characteristics. That being said, we did control for country differences in overall labor market conditions and in the industrial structure—and found little evidence that it is these factors, rather than the extent of skills transparency, which drives cross-national variation in the labor market disadvantage of less-educated adults.

Third, while employer perceptions and employer behavior play a central role in theoretical explanations for the labor market disadvantage of less-educated workers, we cannot observe them directly with survey data like PIAAC. A few studies have recently begun to use innovative designs such as correspondence studies and factorial surveys to better understand employer decision making (e.g., Di Stasio, 2015; Di Stasio and van de Werfhorst, 2016; Protsch and Solga, 2015), and at least one of these studies has included a country-comparative element (Di Stasio and van de Werfhorst, 2016). This line of research nicely complements studies such as ours that approach the process of labor market attainment from the employee side.

Finally, we focused on the less educated as a group that faces particularly high labor market risks. One obvious extension would be to study the labor market advantage of higher-education graduates. However, the literature on statistical discrimination argues that employers will also look to characteristics other than education when they want to infer an individual's actual level of skills. Hence, the approach taken in this paper—to explain labor market inequalities not only with the skills of an individual herself, but also with the "skill profile" of the groups that she belongs—might be useful for understanding inequalities by race, sex, or age as well.

[22] However, it should be noted that our skill measures will partly pick up the effects of other types of skills if the latter are correlated with them.

REFERENCES

Abrassart, A. 2013. Cognitive Skills Matter: The Employment Disadvantage of Low-Educated Workers in Comparative Perspective. *European Sociological Review* 29(4), 707–719.

Aigner, D.J., Glen G.C. 1977. Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review* 30(2), 175–187.

Andersen, R., van de Werfhorst, H.G. 2010. Education and Occupational Status in 14 Countries: The Role of Educational Institutions and Labour Market Coordination. *British Journal of Sociology* 61(2), 336–355.

Arrow, K.J. 1973. Higher Education as a Filter. *Journal of Public Economics* 2(3), 193–216.

Becker, G.S. 1964. *Human Capital*. New York, NY: Columbia University Press.

Bills, D.B. 1990. Employers' Use of Job History Data for Making Hiring Decisions. *Sociological Quarterly* 31(1), 23–35.

Bills, D.B. 2003. Credentials, Signals, and Screens. Explaining the Relationship between Schooling and Job Assignment. *Review of Educational Research* 73(4), 441–469.

Bol, T., van de Werfhorst, H.G. 2011. Signals and Closure by Degrees: The Education Effect across 15 European Countries. *Research in Social Stratification and Mobility* 29, 119–132.

Bol, T., van de Werfhorst, H.G. 2013. Educational Systems and the Trade-Off between Labor Market Allocation and Equality of Educational Opportunity. *Comparative Education Review* 57(2), 285–308.

Bol, T., Weeden, K. 2014. Occupational Closure and Wage Inequality in Germany and the United Kingdom. *European Sociological Review* 31(1), 354–369.

Carroll, G.R., Mayer, K.U. 1986. Job-Shift Patterns in the Federal Republic of Germany: The Effects of Social Class, Industrial Sector, and Organizational Size. *American Sociological Review* 51(3), 323-341.

Collins, R. 1979. *The credential society*. New York: Academic Press.

Davidson, R., MacKinnon, J.G. 2004. *Econometric Theory and Methods*. Oxford: Oxford University Press.

Di Stasio, V. 2015. Education as a signal of trainability: results from a vignette study with Italian employers. *European Sociological Review* 30 (6), 796—809.

Di Stasio, V., van de Werfhorst, H.G. 2016. Why does education matter to employers? A vignette study in England and Netherlands. *Social Forces* 95(1), 77–106.

Doeringer, P., Piore, M.J. 1971. *Internal labor markets and manpower analysis*. Lexington, MA: Heath.

Elder, T.E., Goddeeris, J.H., Haider, S.J. 2010. Unexplained gaps and Oaxaca-Blinder decompositions. *Labour Economics* 17, 284–290.

Farber, H. 1997. The Changing Face of Job Loss in the United States, 1981-1995. *Brookings Papers on Economic Activity: Microeconomics*, 55–128.

Fortin, N.M., Lemieux, T., Firpo, S. 2011. Decomposition Methods in Economics. In *Handbook of Labor Economics, Volume 4A*, edited by Ashenfelter, O.C., Card, D., pp. 1–102. Amsterdam/San Diego: North Holland/Elsevier.

Fox, J. 1991. *Regression Diagnostics: An Introduction*. Newbury Park: Sage.

Ganzeboom, H.B.G. 2010. *Occupational Status Measures for the New International Standard Classification of Occupations ISCO-08.* Last revised: May 24, 2010. Online: http://www.harryganzeboom.nl/isol/isol2010c2-ganzeboom.pdf (retrieved July 18, 2018).

Ganzeboom, H.B.G and Treiman, D.J. 1996. Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research* 25, 201–239.

Gesthuizen, M., Solga, H., Künster, R. 2011. Context Matters: Economic Marginalisation of Low-Educated Workers in Cross-National Perspective. *European Sociological Review* 27(2), 264–280.

Hanushek, E.A., Schwerdt, G., Wiederhold, S., Woessmann, L. 2015. Returns to Skills around the World: Evidence from PIAAC. *European Economic Review* 73, 103–130.

Heckman, J.J., Stixrud, J., Urzua, S. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3), 411–482.

Heisig, J.P. 2018. Measuring the Signaling Value of Educaional Degrees: Secondary Education Systems and the Internal Homogeneity of Educational Groups. *Large-scale Assessments in Education*. 6(9).

Heisig, J.P., Schaeffer, M., Giesecke, J. 2017. The Costs of Simplicity: Why Multilevel Modles May Benefit from Accounting for Cross-Cluster Differences in the Effects of Controls. *American Sociological Review* 82(4), 796–827.

Heisig, J.P., Solga, H. 2015a. Secondary Education Systems and the General Skills of Less- and Intermediate-educated Adults: A Comparison of 18 Countries. *Sociology of Education* 88(3), 202–225.

Heisig, J.P., Solga, H. 2015b. Ohne Abschluss keine Chance. *WZBrief Arbeit 19*. Berlin: WZB Berlin Social Science Center.

Hirsch, F. 1977. *Social Limits to Growth*. Cambridge, MA: Harvard University Press.

Jackson, M. 2009. Disadvantaged through Discrimination. The Role of Employers in Social Stratification. *British Journal of Sociology* 60(4), 669–692.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453–479.

Kitagawa, E.M. 1955. Components of a Differences between Two Rates. *Journal of the American Statistical Association* 50(272), 1168–1194.

Koeniger, W., Leonardi, M., Nunziata, L. 2007. Labor Market Institutions and Wage Inequality. *Industrial and Labor Relations Review* 60(3), 340–356.

Leontaridi, M.R. 1998. Segmented Labour Markets. *Journal of Economic Surveys* 12 (1), 63–101.

Lewis, J.B., Linzer, D.A. 2005. Estimating Regression Models in Which the Dependent Variable Is Based on Estimates. *Political Analysis* 13(4), 345–364.

Little, R.J.A., Rubin, D.B. 2002. *Statistical Analysis with Missing Data*. New York: Wiley.

Lleras, C. 2008. Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research* 37(3), 888–902, Meyer, J.W. 1997. The Effects of Education as an Institution. *American Journal of Sociology* 83(1), 55–77.

Müller, W., Jacob, M. 2008. Qualifications and the Returns to Training Across the Life Course. In *Skill Formation*, edited by Mayer, K.U., Solga, H., pp. 126–172. New York: Cambridge University Press.

Murnane, R., Willett, J., Levy, F. 1995. The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics* 77(2), 251–266.

OECD. 2006. *Education at a Glance*. Paris: OECD.

OECD. 2013. *Skills Outlook 2013: First Results from the Survey of Adult Skills*. Paris: OECD.

OECD. 2016. *Technical Report of the Survey of Adult Skills (PIAAC)* (second edition). Paris: OECD.

Park, H., Kyei, P. 2011. Literacy Gaps by Educational Attainment. *Social Forces* 89(3), 879–904.

Phelps, E.S. 1972. The Statistical Theory of Racism and Sexism. *American Economic Review* 62(4), 659–661.

Piore, M.J. 1994. The dual labor market. In *Social stratification: Class, race, and gender in sociological perspective*, edited by Grusky, D.B. pp. 359–361. Boulder, CO: Westview Press.

Protsch, P., Solga, H. 2015. How Employers Use Signals of Cognitive and Noncognitive Skills at Labour Market Entry: Insights from Field Experiments. *European Sociological Review*, 31(5), 521–532.

Shavit, Y., Müller, W. (Eds.). 1998. *From School to Work*. Oxford: Clarendon Press.

Sørenson, A.B. 2000. Toward a Sounder Basis for Class Analysis. *American Journal of Sociology* 105(6), 1523–1558.

Solga, H. 2002. “Stigmatization by Negative Selection”: Explaining Less-Educated People’s Decreasing Employment Opportunities. *European Sociological Review* 18(2), 159–178.

Solga, H. 2008. Lack of Training – The Employment Opportunities of Low-Skilled Persons from a Sociological and Micro-Economic Perspective. In *Skill Formation*, edited by Mayer, K.U., Solga, H., pp. 173–204. New York: Cambridge University Press.

Solga, H. 2014. Education, economic inequality, and the promises of the social investment state. *Socio-Economic Review* 12(2), 269–297.

Spence, M. 1973. Job Market Signaling. *Quarterly Journal of Economics* 87(3), 355–374.

Stinchcombe, A. 1979. Social Mobility in Industrial Labor Markets. *Acta Sociologica* 22(3), 217–245.

Thurow, L.C. 1979. A Job Competition Model. In *Unemployment and Inflation*, edited by Piore M.J., pp. 17–32. New York: M.E. Sharpe.

van de Werfhorst, H.G. 2011. Skill and Education Effects on Earnings in 18 Countries. *Social Science Research* 40(4), 1078–1090.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 36(3), 1–48.

Weiss, A. 1995. Human Capital vs. Signalling Explanations of Wages. *The Journal of Economic Perspectives* 9(4), 133–154.

Online supplement to

"Lack of skills or formal qualifications? New evidence on cross-country differences in the labor market disadvantage of less-educated adults"

Contents

A. FURTHER DETAILS ON THE ADDITIONAL COUNTRY-LEVEL PREDICTORS: UNEMPLOYMENT RATE AND SECTOR/SEGMENT SHARE MEASURES

In Table 6 in the main article, we present supplementary country-level regressions account for country differences in labor market conditions by controlling for the unemployment rate and for demand-side explanations of labor market inequalities by controlling for the proportion of employment in different sectors (primary and secondary) and labor market segments. Here we provide additional information on the additional country-level predictors.

The unemployment rate data come from the World Bank and are based on data from the International Labour Office.[1] The rate is number of unemployed people, expressed as a percentage of the total labor force. We use the mean of the 2011 and 2012 values for the first-round and the mean of 2014 and 2015 unemployment rates for the second-round countries.

Both the sector and the segment shares are based on data employment by industry from the International Labor Office.[2] For most countries data are provided in terms of the fourth revision of the International Standard Industrial Classification of All Economic Activities (ISIC). The exceptions are Canada and Chile where they are based on the third revision of the ISIC. On assigning industries to the primary, secondary, and tertiary sector we followed standard conventions, as applied, for example, in the OECD's Annual Labor Force Statistics. Specifically, we assigned ISIC Rev. 3 categories A-B and ISIC Rev. 4 category A to the primary sector, ISIC Rev. 3 categories C-F and ISIC Rev. 4 categories B-F to the secondary sector, ISIC Rev. 3 categories G-Q and ISIC Rev. 4 categories G-U to the tertiary sector. The share of employment in the residual category "X" (not elsewhere classified) was distributed over the different sectors and labor market segments in proportion to the shares calculated on the basis of the other categories valid categories to ensure that the sector/segment shares sum to 1 for all countries. This is equivalent to assuming that information on sector is missing completely at random.

In assigning industries to labor market segments, we followed the classification of Carroll and Mayer (1986), which is based on conceptual and empirical work by Stinchcombe (1979). The mapping used by Carroll and Mayer is documented in Appendix A on p. 338 of their 1986 article.

[1] https://data.worldbank.org/, downloaded on September 3, 2018.

[2] https://www.ilo.org/global/statistics-and-databases/lang--en/index.htm; downloaded on October 8, 2018:

We had to make some adaptations because the ISIC classifications do not fully align with the industry classification used by Carroll and Mayer. Table A1 shows how we assigned of ISIC Rev. 4 industry categories to the different labor market segment. Table A2 shows the mapping for ISIC 3 categories. As also noted in the main article, the ISIC classifications were not fine-grained enough to identify the share of employment located in the “classical capitalist” segment. We generally assigned groups of industries that included industries treated as “classical capitalist” by Carroll and Mayer to the “small competitive” segment. We thus differentiate only six (rather than) seven segments.

As noted in the main article, we use the mean of the 2011/2012 sector/segment shares for the PIAAC first-round and the mean 2014/2015 sector/segment share for the PIAAC second-round countries. For Japan, a first-round country, we only have valid values for 2012. For Canada, also a first-round country, we use the 2016 values, as this is the closest year to 2011/2012 for which employment shares are available.

Table A3 reports the values of the unemployment rate variable and of the sector and segment share measures. All measures were z-standardized before including them in the regression models reported in Table 6 in the main article (the square of the unemployment rate in Model 2 in Table 6 in the main article is the square of the z-standardized rate).

Table A1. Assignment of industry categories (ISIC Rev. 4) to labor market segments

ISIC Rev. 4 category	Labor market segment
A. Agriculture, forestry and fishing	Traditional primary
B. Mining and quarrying	Large-scale engineering
C. Manufacturing	Large-scale engineering
D. Electricity, gas, steam and air conditioning supply	Large-scale engineering
E. Water supply; sewerage, waste management and remediation activities	Large-scale engineering
F. Construction	Competitive
G. Wholesale and retail trade; repair of motor vehicles and motorcycles	Small competitive
H. Transportation and storage	Large-scale engineering
I. Accommodation and food service activities	Small competitive
J. Information and communication	Large-scale engineering
K. Financial and insurance activities	Bureaucratic
L. Real estate activities	Professional
M. Professional, scientific and technical activities	Professional
N. Administrative and support service activities	Small competitive
O. Public administration and defence; compulsory social security	Bureaucratic
P. Education	Professional
Q. Human health and social work activities	Professional
R. Arts, entertainment and recreation	Professional
S. Other service activities	Competitive
T. Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use	Small competitive
U. Activities of extraterritorial organizations and bodies	Bureaucratic

Sources: Authors' adaptation of classification in Carroll and Mayer (1986), which is based on Stinchcombe (1979).

Table A2. Assignment of industry categories (ISIC Rev. 3) to labor market segments

ISIC Rev. 3 category	Labor market segment
A. Agriculture, hunting and forestry	Traditional primary
B. Fishing	Traditional primary
C. Mining and quarrying	Large-scale engineering
D. Manufacturing	Large-scale engineering
E. Electricity, gas and water supply	Large-scale engineering
F. Construction	Competitive
G. Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods	Small competitive
H. Hotels and restaurants	Small competitive
I. Transport, storage and communications	Large-scale engineering
J. Financial intermediation	Bureaucratic
K. Real estate, renting and business activities	Competitive
L. Public administration and defence; compulsory social security	Bureaucratic
M. Education	Professional
N. Health and social work	Professional
O. Other community, service and personal service activities	Small competitive
P. Activities of households as employers and undifferentiated production activities of private households	Small competitive
Q. Extraterritorial organizations and bodies	Bureaucratic

Sources: Authors' adaptation of classification in Carroll and Mayer (1986), which is based on Stinchcombe (1979).

Table A3. Values of additional country-level predictors used in Table 6 in main article

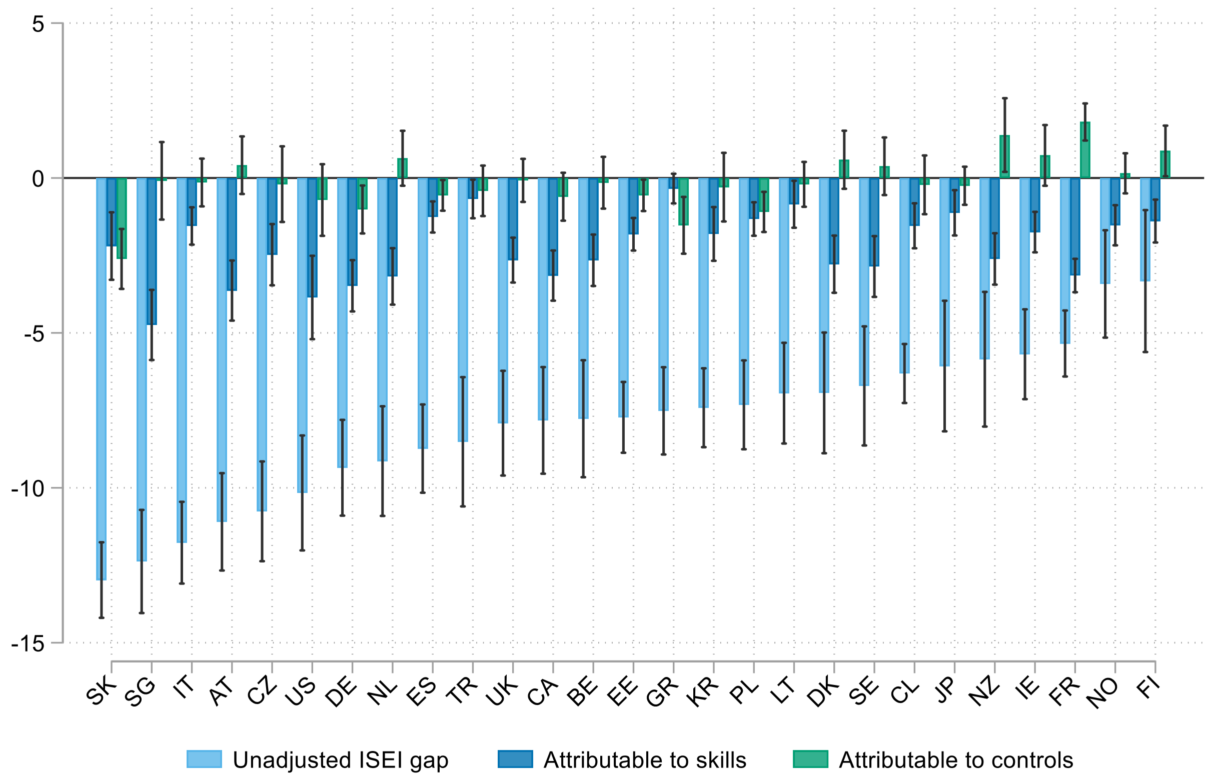
			Employment shares of broad sectors (%)			Employment shares of detailed labor market segments (%)					
	Country code	Unemployment rate (%) (1)	Primary sector (2)	Secondary sector (3)	Tertiary sector (4)	Traditional primary (5)	Small competitive (6)	Competitive (7)	Large-scale engineering (8)	Professional (9)	Bureaucratic (10)
Austria	AT	4.7	4.8	26.0	69.2	4.8	24.9	12.4	24.7	22.8	10.4
Belgium	BE	7.3	1.2	22.5	76.2	1.2	21.8	10.0	24.2	29.4	13.4
Canada	CA	7.4	2.0	19.6	78.4	2.0	29.2	21.3	18.0	19.9	9.6
Chile	CL	6.6	9.4	23.1	67.5	9.4	32.9	15.2	21.8	13.1	7.5
Czech Rep.	CZ	6.8	3.0	38.3	58.7	3.0	18.8	11.5	38.7	18.9	9.0
Denmark	DK	7.6	2.5	19.7	77.8	2.5	21.3	9.4	22.7	35.2	8.9
Estonia	EE	11.2	4.4	31.8	63.8	4.4	19.4	13.2	33.2	21.6	8.2
Finland	FI	7.7	4.2	22.9	73.0	4.2	20.0	11.0	25.7	32.2	6.8
France	FR	9.1	2.9	21.9	75.2	2.9	22.7	11.2	22.8	27.3	13.1
Germany	DE	5.6	1.6	28.1	70.3	1.6	23.1	10.2	29.5	25.1	10.5
Greece	GR	25.7	13.2	15.0	71.8	13.2	30.4	6.3	17.7	21.0	11.3
Ireland	IE	14.6	5.8	17.2	77.1	5.8	25.7	7.7	22.0	29.0	9.9
Italy	IT	9.5	3.7	27.9	68.4	3.7	26.7	11.5	27.3	21.6	9.2
Japan	JP	4.4	3.8	26.2	70.0	3.8	28.9	12.8	27.0	20.7	6.8
Korea	KR	3.3	6.2	24.7	69.0	6.2	27.7	14.4	25.9	18.3	7.4
Lithuania	LT	9.9	9.1	24.9	65.9	9.1	23.5	10.8	26.7	22.3	7.4
Netherlands	NL	5.4	2.8	17.2	80.0	2.8	23.5	9.0	20.4	34.4	10.0
New Zealand	NZ	5.4	6.2	21.4	72.4	6.2	24.9	12.9	20.6	27.5	8.0
Norway	NO	3.2	2.3	20.3	77.4	2.3	20.2	10.2	22.3	37.1	8.0
Poland	PL	9.9	12.7	30.5	56.7	12.7	19.7	10.6	30.1	17.8	9.1
Singapore	SG	3.8	0.0	28.5	71.5	0.0	25.9	25.6	24.8	11.7	12.0
Slovak Rep.	SK	13.8	3.2	37.5	59.4	3.2	19.8	12.4	36.2	18.1	10.3
Spain	ES	23.1	4.1	21.2	74.7	4.1	32.5	9.9	21.9	21.2	10.3
Sweden	SE	7.9	2.1	19.8	78.2	2.1	19.9	10.7	22.3	36.8	8.2
Turkey	TR	10.1	20.7	27.6	51.7	20.7	24.4	10.5	25.5	12.3	6.5
United Kingdom	UK	8.0	1.2	19.1	79.6	1.2	23.6	11.1	20.4	33.1	10.6
United States	US	8.5	1.4	19.8	78.8	1.4	25.8	11.3	22.1	30.6	8.8
Mean		8.9	5.0	24.2	70.8	5.0	24.3	12.0	25.0	24.4	9.3
Standard deviation		5.3	4.6	5.8	7.5	4.6	4.0	3.8	5.1	7.4	1.8

Note: Table reports the untransformed values of the predictors, constructed as described in the text. For the country-level regressions all predictors were standardized to have a mean of 0 and a standard deviation of 1 within the sample of 27 countries included in the analysis.
Sources: (1): International Labour Office, retrieved from the World Bank at https://data.worldbank.org/ on September 3, 2018. (2)-(10): International Labour Office, retrieved from https://www.ilo.org/global/statistics-and-databases/lang--en/index.htm on October 8, 2018.

B. FULL DECOMPOSITION RESULTS

Figure 1 in the main article shows the unadjusted ISEI gaps as well as the portion of the gap explained by differences in literacy and numeracy skills between less- and intermediate educated adults. For brevity, it does not show the estimated contributions of the lower-level control variables (sex, potential work experience, foreign-birth/foreign-language status, parental educational attainment, and self-employment status). In Figure B1 we show, for each country, the combined contribution of control variables (rightmost bar), next to the unadjusted gap (leftmost bar) and the combined contribution of literacy and numeracy skills (middle bar).

Figure B1. Full decomposition results



Notes: See Table 2 in main article for country codes. The leftmost (light blue) bars represent the occupational gap, measured in ISEI points, between less- and intermediate-educated adults. The middle (dark blue) bars show the portion of the gap that os attributable to differences in literacy and numeracy skills. The rightmost bars (turquoise) show the portion of the gap that is attributable to differences of the lower-level control variables (sex, potential work experience, foreign-birth/foreign-language status, parental educational attainment, and self-employment status). The unadjusted gap and the portion attributable to skills are also shown in Figure 1 in the main article (as stacked bars). The capped lines indicate 95% confidence.

Sources: PIAAC (rounds 1 and 2), authors' calculations.

C. COUNTRY-SPECIFIC REGRESSION RESULTS (FIRST-STEP REGRESSIONS)

Table C1 shows the results of the country-specific regressions for estimating the fully adjusted ISEI gap. The fully adjusted ISEI gap is the coefficient estimate on the indicator variable for having low education in these regressions. It is the dependent variable in the country-level regressions in Tables 4, 5, and 6 in the main article and in the additional country-level regressions reported in this online supplement.

The literacy and numeracy scores can range from 0 to 500 on the original scales. For easier interpretation, Table C1 is based on z-standardized scores that have a mean of 0 and a standard deviation of 1 in our analytic sample. The standard deviations of the original variables are 46.5 for the literacy and 51.4 for the numeracy score.

Results for Israel and Slovenia are presented here, although they are not included in our main analysis (for details on why we excluded these two countries, see Section D below).

Table C1. Country-specific regressions of ISEI score on individual-level predictors

	Austria	Belgium	Canada	Chile	Czech Republic	Denmark	Estonia
Education (highest degree)							
Low (ISCED 0-2)	-7.88***	-4.96***	-4.07***	-4.55***	-8.09***	-4.74***	-5.35***
	(0.87)	(1.03)	(1.07)	(0.82)	(1.07)	(1.10)	(0.64)
Intermediate (ISCED 3-4)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Skills							
Literacy	4.07***	2.38*	2.91*	0.15	1.44	3.09**	-0.44
	(0.95)	(1.01)	(1.03)	(0.55)	(1.05)	(1.03)	(0.72)
Numeracy	2.05+	2.14+	0.92	1.78+	2.59*	1.60	3.76***
	(1.01)	(1.09)	(1.05)	(0.79)	(1.26)	(1.04)	(0.80)
Gender							
Female	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Male	-1.53*	1.12	-6.06***	0.72	-3.88***	-3.13**	-1.58*
	(0.71)	(0.85)	(0.75)	(0.73)	(0.87)	(0.92)	(0.64)
Foreign-birth/foreign-language status							
Native-born, test language is first language	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Native-born, test language	2.27	5.53***	3.92**	-0.85	9.70	14.87+	2.01
is not first language	(2.16)	(1.57)	(1.23)	(1.20)	(24.17)	(8.38)	(1.82)
Foreign-born, test	7.10+	-1.44	2.11	-2.48**	3.99	-1.42	-3.50**
language is first language	(3.85)	(2.50)	(2.45)	(0.70)	(4.56)	(2.63)	(1.03)
Foreign-born, test	-1.60	2.40	1.17	5.63	-2.93	-1.39	-5.83*
language is not first language	(2.03)	(2.25)	(1.64)	(13.14)	(2.93)	(1.08)	(2.59)
Parental education							
Neither parent completed upper secondary education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
At least one parent	3.22**	3.08**	2.14*	2.21*	3.90***	0.68	1.93*
completed upper secondary education	(0.98)	(0.94)	(0.88)	(0.88)	(0.97)	(1.05)	(0.75)
At least one parent	5.84***	8.07***	3.32**	1.66	7.10***	2.26	5.38***
completed tertiary education	(1.27)	(1.33)	(1.02)	(1.24)	(1.62)	(1.39)	(0.92)
Potential work experience (linear splines)							
0-10	0.05	-0.07	0.82***	0.20	0.80**	0.36+	0.60***
	(0.22)	(0.27)	(0.17)	(0.27)	(0.24)	(0.20)	(0.16)
10-20	0.25	0.53**	0.36*	0.17	-0.39*	0.33+	-0.08
	(0.16)	(0.17)	(0.17)	(0.14)	(0.18)	(0.19)	(0.13)
20-30	-0.05	0.03	0.07	0.00	0.15	0.46*	-0.02
	(0.14)	(0.15)	(0.15)	(0.13)	(0.19)	(0.18)	(0.13)
30+	0.41+	0.09	-0.12	-0.18	0.00	-0.50	0.23
	(0.21)	(0.19)	(0.20)	(0.18)	(0.30)	(0.31)	(0.24)
Self-employed							
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	0.66	4.03**	4.13**	1.38	3.82**	-2.64+	10.63***
	(1.54)	(1.46)	(1.31)	(0.99)	(1.22)	(1.51)	(1.48)
Constant	35.43***	28.88***	31.53***	28.47***	27.92***	31.95***	28.02***
	(1.82)	(2.18)	(1.47)	(2.34)	(2.05)	(1.68)	(1.15)
N	2225	1663	7430	2128	2402	1851	2511
R2	0.16	0.15	0.14	0.13	0.11	0.14	0.15

Table continues on next page (countries in alphabetic order).

Table C1. Country-specific regressions of ISEI score on individual-level predictors (continued)

	Finland	France	Germany	Greece	Ireland	Israel	Italy
Education (highest degree)							
Low (ISCED 0-2)	-2.81*	-4.00***	-4.86***	-5.65***	-4.67***	-6.27***	-10.08***
	(1.29)	(0.58)	(0.98)	(0.87)	(0.79)	(1.19)	(0.71)
Intermediate (ISCED 3/4)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Skills							
Literacy	1.22	-0.64	3.02***	-0.29	1.45	1.07	0.44
	(0.83)	(0.63)	(0.77)	(0.77)	(1.19)	(0.85)	(0.77)
Numeracy	2.48**	4.66***	0.89	0.98	1.58	2.63**	1.84*
	(0.85)	(0.57)	(0.77)	(0.79)	(1.14)	(0.77)	(0.75)
Gender							
Female	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Male	-0.09	2.31***	-2.86***	0.59	-3.12***	-4.41***	-0.22
	(0.82)	(0.48)	(0.63)	(0.70)	(0.70)	(0.95)	(0.69)
Foreign-birth/foreign-language status							
Native-born, test language is first language	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Native-born, test language is not first language	5.18+	1.14	0.71	-0.60	-1.49	-0.41	-3.42
	(2.85)	(1.61)	(2.40)	(3.08)	(3.53)	(1.62)	(2.15)
Foreign-born, test language is first language	-0.26	-1.59	-1.03	-2.69	-1.84	-2.36	0.31
	(2.84)	(1.24)	(1.40)	(1.66)	(1.40)	(3.14)	(2.62)
Foreign-born, test language is not first language	-1.89	1.03	-1.34	-6.90***	-2.89*	-1.43	-5.85***
	(2.13)	(0.96)	(1.16)	(1.00)	(1.39)	(1.57)	(0.85)
Parental education							
Neither parent completed upper secondary education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
At least one parent completed upper secondary education	1.18	1.31*	1.82	1.25	0.83	3.67**	3.32**
	(0.92)	(0.63)	(1.12)	(0.87)	(0.81)	(1.20)	(0.98)
At least one parent completed tertiary education	3.81**	3.56**	3.88**	5.34**	1.39	5.62***	8.15*
	(1.25)	(1.19)	(1.34)	(1.77)	(1.21)	(1.43)	(3.38)
Potential work experience (linear splines)							
0-10	0.24	0.37*	0.61**	0.54+	0.28	0.99***	0.37
	(0.18)	(0.16)	(0.19)	(0.31)	(0.19)	(0.23)	(0.29)
10-20	0.19	0.33**	-0.28*	-0.23	0.16	0.18	0.31*
	(0.16)	(0.12)	(0.14)	(0.16)	(0.16)	(0.17)	(0.14)
20-30	0.37*	0.22*	0.12	0.17	0.07	0.15	0.19
	(0.17)	(0.09)	(0.13)	(0.14)	(0.16)	(0.15)	(0.12)
30+	-0.42+	0.16	0.03	-0.23+	0.02	0.62**	-0.16
	(0.23)	(0.15)	(0.20)	(0.13)	(0.23)	(0.22)	(0.16)
Self-employed							
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	-0.60	-7.19***	2.47	-2.70***	-1.73	3.85*	2.79**
	(1.24)	(0.86)	(1.72)	(0.72)	(1.22)	(1.58)	(0.98)
Constant	26.37***	28.11***	30.57***	29.66***	34.69***	30.19***	32.66***
	(1.40)	(1.22)	(1.73)	(2.43)	(1.39)	(1.88)	(2.46)
N	1422	2515	2060	1853	1976	1484	2131
R2	0.09	0.13	0.15	0.13	0.08	0.18	0.23

Table continues on next page (countries in alphabetic order).

Table C1. Country-specific regressions of ISEI score on individual-level predictors (continued)

	Japan	Korea	Lithuania	Netherlands	New Zealand	Norway	Poland
Education (highest degree)							
Low (ISCED 0-2)	-4.69***	-5.31***	-5.89***	-6.60***	-4.62***	-2.05*	-4.90***
	(1.17)	(0.87)	(0.90)	(0.90)	(1.18)	(0.90)	(0.82)
Intermediate (ISCED 3/4)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Skills							
Literacy	0.43	1.48	-0.67	3.07**	1.45	3.70***	1.68*
	(1.59)	(1.14)	(1.07)	(1.06)	(1.25)	(0.88)	(0.76)
Numeracy	1.86	1.17	4.32***	1.53	2.69+	0.21	1.36
	(1.48)	(1.06)	(1.00)	(1.15)	(1.34)	(0.94)	(0.82)
Gender							
Female	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Male	-0.55	-0.95+	-2.01**	-0.40	-2.32*	1.59+	-1.96**
	(0.85)	(0.55)	(0.67)	(0.86)	(0.99)	(0.82)	(0.64)
Foreign-birth/foreign-language status							
Native-born, test language is first language	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Native-born, test language is not first language	-9.07***	0.38	0.37	5.15	-2.24	-3.15	-0.96
	(1.46)	(4.22)	(1.48)	(5.23)	(2.53)	(2.86)	(2.05)
Foreign-born, test language is first language	5.00	2.58	-2.21	-1.63	3.38*	8.43	3.42+
	(7.36)	(2.59)	(1.60)	(3.02)	(1.54)	(6.48)	(1.95)
Foreign-born, test language is not first language	-12.21+	0.39	3.52	-1.99	1.26	-3.56**	(dropped)
	(6.82)	(3.69)	(6.61)	(1.83)	(2.03)	(1.31)	
Parental education							
Neither parent completed upper secondary education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
At least one parent completed upper secondary education	0.75	1.05	-0.05	0.38	0.60	0.17	4.18***
	(1.13)	(0.83)	(0.88)	(1.18)	(1.31)	(0.95)	(0.95)
At least one parent completed tertiary education	3.76*	6.94***	2.98*	2.52+	2.13	1.26	7.37***
	(1.57)	(1.54)	(1.26)	(1.47)	(1.37)	(1.18)	(1.91)
Potential work experience (linear splines)							
0-10	0.38	1.04***	0.66**	0.69**	1.33***	0.46*	0.35*
	(0.25)	(0.25)	(0.22)	(0.25)	(0.20)	(0.18)	(0.17)
10-20	0.35*	-0.33*	0.11	0.24	-0.17	0.26+	0.10
	(0.16)	(0.14)	(0.18)	(0.19)	(0.21)	(0.14)	(0.14)
20-30	0.29+	-0.12	-0.10	-0.27	0.37	0.22	-0.07
	(0.17)	(0.11)	(0.20)	(0.22)	(0.25)	(0.14)	(0.13)
30+	-0.37	0.22+	-0.12	0.42	-0.42+	-0.50+	-0.11
	(0.28)	(0.12)	(0.30)	(0.27)	(0.25)	(0.26)	(0.22)
Self-employed							
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	1.01	2.60**	0.49	6.40***	5.10**	-4.36*	-1.69+
	(1.99)	(0.81)	(1.47)	(1.55)	(1.91)	(1.87)	(0.98)
Constant	26.83***	25.34***	25.10***	34.30***	27.94***	30.01***	27.11***
	(2.41)	(2.13)	(1.68)	(1.82)	(1.60)	(1.62)	(1.30)
N	1287	1985	1794	1733	1797	1559	3001
R2	0.08	0.10	0.10	0.15	0.15	0.13	0.10

Table continues on next page (countries in alphabetic order).

Table C1. Country-specific regressions of ISEI score on individual-level predictors (continued)

	Singapore	Slovak Republic	Slovenia	Spain	Sweden	Turkey	United Kingdom	United States
Education (highest degree)								
Low (ISCED 0-2)	-7.54***	-8.17***	-8.55***	-6.91***	-4.23***	-7.42***	-5.19***	-5.60***
	(1.19)	(0.84)	(0.76)	(0.72)	(0.98)	(1.12)	(0.91)	(1.23)
Intermediate (ISCED 3/4)	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Skills								
Literacy	-0.96	-0.34	1.36+	-0.09	2.75+	-1.75+	1.72	2.60+
	(1.22)	(0.89)	(0.80)	(0.70)	(1.50)	(0.89)	(1.26)	(1.23)
Numeracy	4.61***	2.86*	1.58+	2.49**	1.53	2.62**	2.38+	1.44
	(1.22)	(1.08)	(0.85)	(0.78)	(1.52)	(0.88)	(1.21)	(1.20)
Gender								
Female	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Male	-0.06	-3.44***	-1.90*	-0.91	-0.01	-0.10	-3.44***	-2.82***
	(0.83)	(0.75)	(0.72)	(0.65)	(1.00)	(0.93)	(0.89)	(0.76)
Foreign-birth/foreign-language status								
Native-born, test language is first language	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Native-born, test language is not first language	-1.54	0.53	0.16	1.68	0.88	-4.95**	9.87**	2.11
	(1.02)	(0.96)	(3.13)	(2.46)	(2.68)	(1.54)	(3.66)	(2.44)
Foreign-born, test language is first language	-3.05	0.04	-0.13	-1.85	-3.82	-1.53	5.37*	1.46
	(5.22)	(2.66)	(2.29)	(1.12)	(2.88)	(1.13)	(2.08)	(2.39)
Foreign-born, test language is not first language	0.72	2.59	-1.39	-3.40**	-0.50	-2.20	-1.94	-1.82
	(1.63)	(4.39)	(0.99)	(1.26)	(1.31)	(1.66)	(1.69)	(1.24)
Parental education								
Neither parent completed upper secondary education	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
At least one parent completed upper secondary education	1.54	3.84***	3.30***	3.53**	0.44	3.83+	2.22+	1.32
	(1.15)	(0.96)	(0.70)	(1.11)	(1.12)	(2.02)	(1.10)	(1.12)
At least one parent completed tertiary education	5.70**	9.77***	7.30***	3.07+	3.03*	2.61	2.31	2.37
	(1.96)	(1.94)	(1.43)	(1.59)	(1.20)	(3.14)	(1.54)	(1.41)
Potential work experience (linear splines)								
0-10	0.75**	0.42*	0.06	-0.18	0.48*	0.20	0.09	0.48*
	(0.28)	(0.21)	(0.31)	(0.23)	(0.20)	(0.20)	(0.18)	(0.20)
10-20	-0.01	0.01	0.30*	0.13	0.73***	-0.14	0.50**	0.20
	(0.21)	(0.15)	(0.15)	(0.15)	(0.17)	(0.16)	(0.18)	(0.17)
20-30	0.18	0.19	0.02	-0.03	-0.28+	0.16	-0.31	0.06
	(0.20)	(0.14)	(0.13)	(0.13)	(0.16)	(0.13)	(0.19)	(0.16)
30+	-0.33*	-0.27	-0.06	-0.01	0.30	0.10	0.14	0.20
	(0.15)	(0.21)	(0.16)	(0.14)	(0.29)	(0.16)	(0.25)	(0.24)
Self-employed								
No	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Yes	7.06***	10.11***	2.16+	5.80***	1.50	-1.88+	0.07	-1.35
	(1.44)	(1.18)	(1.21)	(1.16)	(1.79)	(0.97)	(1.31)	(1.21)
Constant	36.27***	29.13***	33.03***	37.03***	27.69***	36.71***	35.02***	32.30***
	(2.51)	(1.62)	(2.93)	(1.88)	(1.54)	(1.61)	(1.62)	(2.28)
N	1216	2482	1859	2256	1504	1837	2824	1625
R2	0.24	0.15	0.17	0.15	0.13	0.10	0.14	0.13

Notes: Multiple imputation estimates (10 imputations/plausible values). Survey weights applied. Standard errors in parentheses. Literacy and numeracy skills are z-standardized. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.
ISEI = International Socio-Economic Index of Occupational Status. ISCED = International Standard Classification of Education. Ref. = Reference category.

Source: PIAAC 2011/12, authors' calculations.

D. OUTLIER ANALYSIS

Influential cases that have an extreme impact on coefficient estimates are a concern in any regression analysis but particularly when working with small samples—as in our country-level analysis in the present paper. We therefore performed extensive and systematic checks to identify potentially influential country cases, focusing on Model 7 in Table 4, that is, the country-level model that simultaneously includes the two skills transparency measures (skills gap and index of internal homogeneity) and the vocational enrollment measure. As noted in the main article, we initially ran these checks on a sample of 29 countries that included Israel (IL) and Slovenia (SI) in addition to the 27 countries included in the final analyses reported in the main article.[3]

A standard approach to assessing potential problems with influential cases is to use "delete-1" diagnostics, that is, to inspect how dropping a given case from the sample affects the parameter estimates or predictions from the model. There are several well-established statistics for quantifying the influence of individual observations within this framework. In the following, we will focus on DFBETA as a measure of the observation's impact on individual coefficient estimates and on Cook's D as a measure of an observation's overall impact. Intuitively, the DFBETA value for the i^{th} unit and the j^{th} coefficient is the difference in the j^{th} coefficient estimate between the full sample and the reduced sample with unit i removed, expressed in terms of the standard error for the coefficient estimate in the reduced sample. A DFBETA value of .5 thus means that inclusion that inclusion of the case in question shifts the estimate of the coefficient upward (i.e., in the positive direction) by half of its (reduced-sample) standard error. A DFBETA value of -.3 means that the inclusion of the unit shifts the estimate downward (i.e., in the negative direction) by 30% of the (reduced-sample) standard error. Cook's D can be understood as a weighted sum of an observation's DFBETA values for all coefficients. For further details, see Fox (2016).

Figures D1 and D2 summarize the results of the influence diagnostics that led us to drop Israel and Slovenia from the main analysis. Figure 1 shows the result of the delete-1 analysis. Panels D1.A to D1.C show DFBETA statistics for each of the three predictors and Panel D1.D shows Cook's D. There are no universally accepted cutoff values for either DFBETA or Cook's D.

[3] As also noted in the main article, we excluded four PIAAC participating countries *a priori* because of data unavailability (Australia, Indonesia) or concerns about data quality (Cyprus, Russia).

Perhaps the most widely used ones are $\pm 2/\sqrt{n}$ for DFBETA and $4/(n-k-1)$ for Cook's D, where *n* refers to the number of observations in the full sample and *k* to the number of predictors. These cutoffs fall on the conservative side of the spectrum. Other common cutoffs that tolerate higher levels of individual influence are ± 1 for DFBETA and 1 or three times the mean value for Cook's D. All of these values are indicated by solid lines in Figure D.1.

The influence statistics for the 29 country sample in Figure D1 highlight several influential cases, with Israel (IL), Norway (NO), Slovenia (SI), and Lithuania (LT) being the most noteworthy ones. Panel IV shows that all of these four countries have values of Cook's D exceeding the intermediate threshold of three times the average distance. Panels I-III reveal how their inclusion impacts the individual coefficient estimates. Lithuania draws the estimate for the skills gap and for vocational enrollment in the positive direction, away from the direction predicted by the corresponding hypotheses. Inclusion of Norway draws the estimates for the skills gap downward (in the direction of the corresponding hypotheses) and the estimate for vocational enrollment upwards (against the corresponding hypotheses). Israel and in particular Slovenia draw the estimate for internal homogeneity (against the corresponding hypothesis) upward and the one for vocational enrollment downward (in the direction of the corresponding hypothesis).

While all of these four country cases thus give reason for concern, there, and that is the fact that the enormous influence of Israel and Slovenia on the estimated effect of internal homogeneity is not counterbalanced by other influential countries drawing the coefficient estimate in the opposite direction. That is, the distribution of DFBETA statistics looks much more asymmetric for internal homogeneity than for the other predictors. In fact, this delete-1 analysis even tends to understate the influence of Israel and Slovenia on the estimated coefficient for internal homogeneity: When one of the two cases is deleted, the other one remains in the sample and exerts considerable influence on the coefficient estimate. As the influence of the remaining works in the same direction as that of the omitted case, the *additional* influence of the omitted case captured by the delete-1 influence statistics is attenuated. To better understand the joint influence of Israel and Slovenia on the regression results, we therefore conducted a "delete-2" influence analysis to examine the influence of all possible pairs of countries on the regression results. With 29 countries, there are 406 (29 choose 2) such country pairs.

Figure D1: Delete-1 influence statistics for 29 country sample

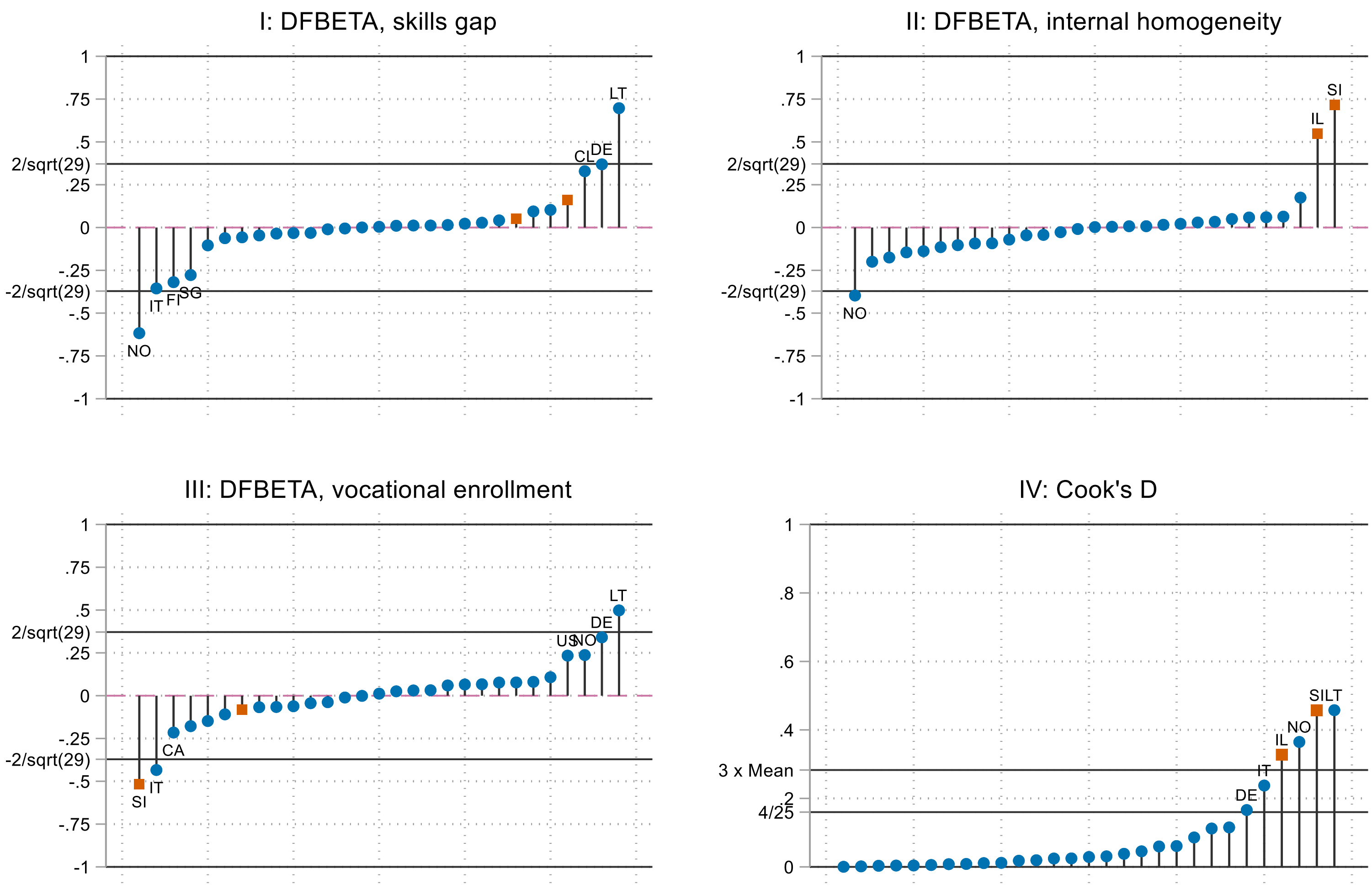


Figure D2 visualizes the DFBETA and Cook's D statistics from the delete-2 analysis. Only extreme values are labeled for better readability. Solid lines again indicate the cutoff values that were also highlighted in Figure D1, multiplied by 2 to adapt them for the delete-2 case (except the "three times the average" threshold for Cook's D). Figure D2 clearly shows that Israel and Slovenia jointly have a dramatic influence on the regression estimates, particularly for the index of internal homogeneity where their exclusion shifts the point estimate by approximately 1.7 standard errors. Their impact on the coefficient of vocational orientation is also noteworthy, but not nearly as dramatic. The Cook's D statistic also underlines the special role played by Israel and Slovenia. While quite a few other country pairs cut the more conservative thresholds as well, Israel/Slovenia is the only pair with a Cook's D above 2 and its value of 3.09 is approximately 2.5 times as high as the second highest of 1.23 (for Germany and Lithuania). Confirming the results from the delete-1 analysis in Figure D1, the case of the internal homogeneity measure is also special in that the distribution of the DFBETA statistic is highly asymmetric. While we also find quite a few influential country pairs for the other two coefficients, their impact generally seems to be counterbalanced by other country pairs drawing the point estimate in the opposite direction. In light of these findings, we decided to omit Israel and Slovenia from the main analysis.

For completeness, we present the main sequence of country-level regressions (cf. Table 4 in the main article) for the 29 country sample in Table D1. As is to be expected given the results of the influence analysis, the effect of internal homogeneity is considerably weaker and that of vocational orientation noticeably stronger when Israel and Slovenia are included.

Figure D2: Delete-2 influence statistics for 29 country sample

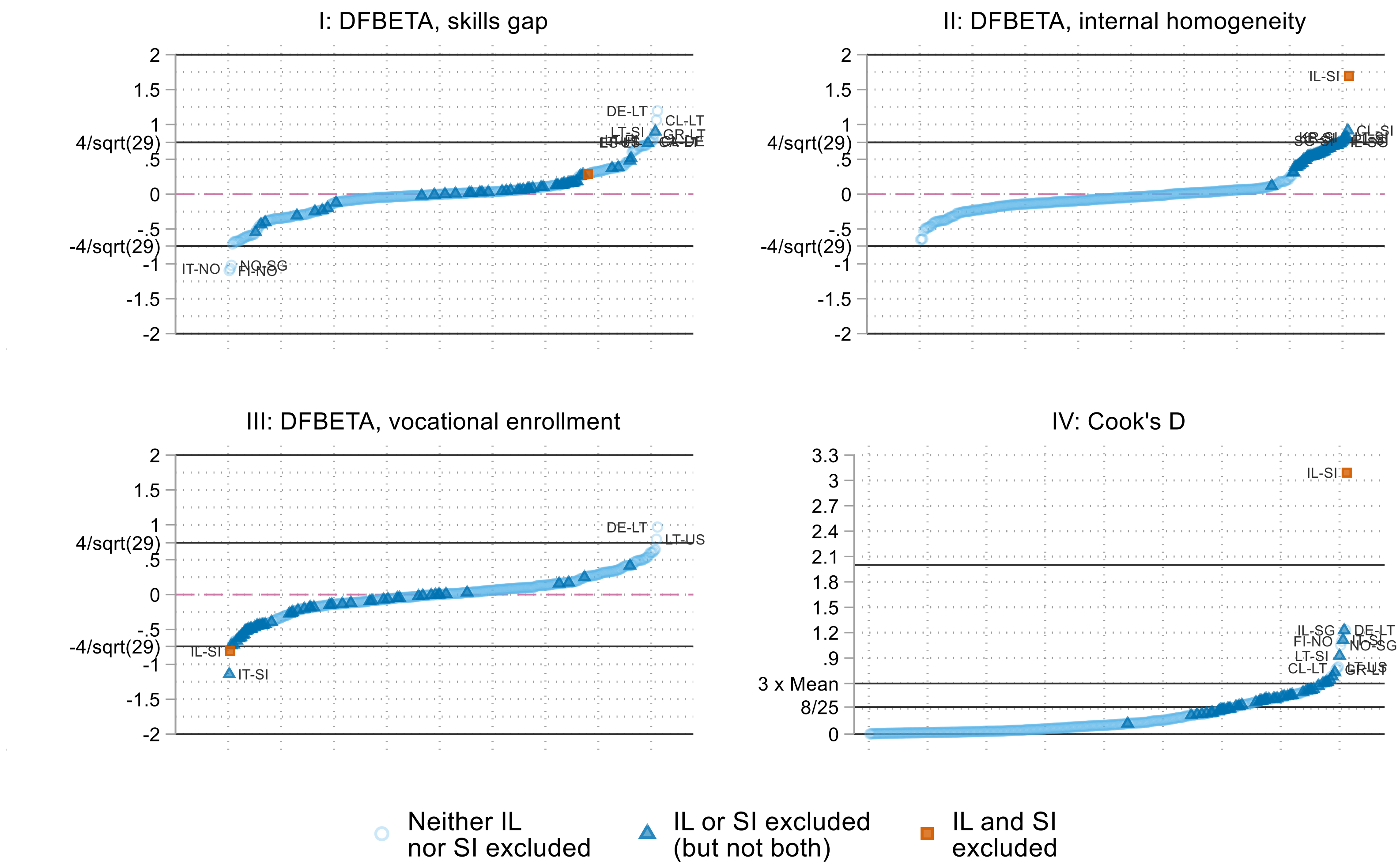


Table D1. Country-level regressions of ISEI gap on measures of skills transparency and vocational orientation (Israel and Slovenia included in the analysis)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Skills gap	-0.497			-0.699+	-0.739+		-0.873*
	(0.359)			(0.381)	(0.358)		(0.376)
Index of internal homogeneity		-0.234		-0.511		-0.094	-0.388
		(0.359)		(0.381)		(0.365)	(0.364)
Prevalence of vocational enrollment			-0.551		-0.787*	-0.522	-0.724+
			(0.349)		(0.350)	(0.368)	(0.354)
Intercept	-5.755***	-5.742***	-5.730***	-5.746***	-5.733***	-5.728***	-5.727***
	(0.335)	(0.343)	(0.331)	(0.330)	(0.313)	(0.337)	(0.313)
N	29	29	29	29	29	29	29
R2	0.07	0.02	0.09	0.13	0.23	0.09	0.26
Adjusted R2	0.04	.	0.05	0.07	0.17	0.02	0.17

Notes: Feasible Generalized Least Squares (FGLS) estimates, based on 10 imputations/plausible values. Dependent variable: the fully adjusted ISEI gap between less-educated and intermediate-educated adults aged 16-54 (see Figure 1 above and Section A in Online Supplement). All country-level variables are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).
Sources: PIAAC (rounds 1 and 2), authors' calculations.

We repeated the influence diagnostics for the reduced sample of 27 countries used in the main analysis. Figures D3 and D4 show the delete-1 and delete-2 results, respectively. The results show that main analysis sample includes several influential cases as well, with Singapore, Lithuania, Norway, and Italy being the most conspicuous ones. Overall, the results give much less reason for concern than those for the 29 country sample, however. In particular, both the DFBETA statistics are distributed much more symmetrically than in the 29 country sample, both in the delete-1 analysis and in the delete-2 analysis. The case that appears most asymmetric is that of the index of internal homogeneity (see Panel II in Figure D3), where we have several countries pulling the point estimate in the upward direction, whereas only the Norwegian case pulls the estimate substantially in the negative direction. If anything, influential cases in our sample therefore pull the point estimate in the positive direction, away from the expected negative effect that we find the main analysis. Hence, the main conclusions concerning the role of internal homogeneity are not called into question by the influence diagnostics.

Figure D3: Delete-1 influence statistics for main analysis sample (27 countries)

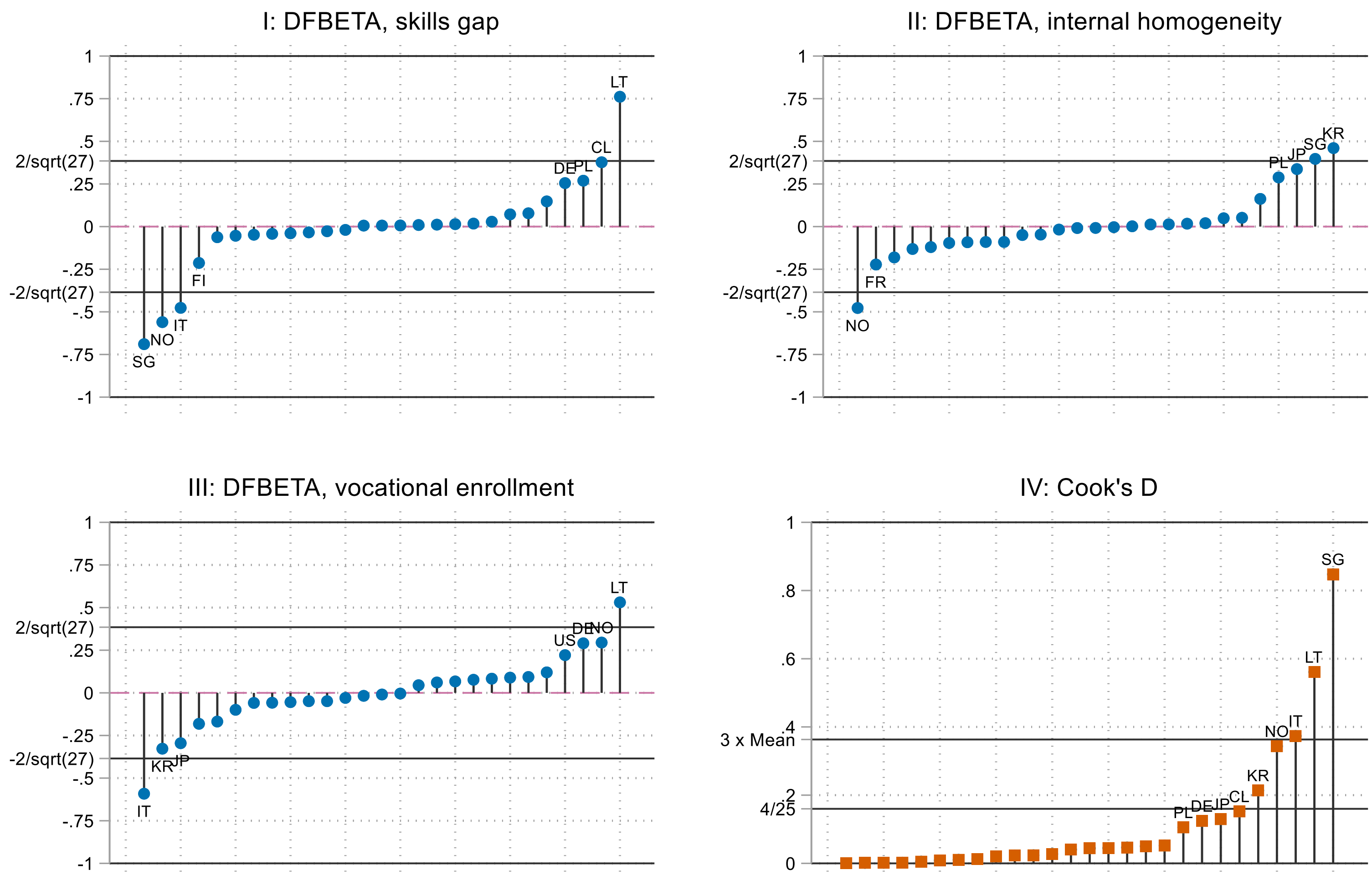
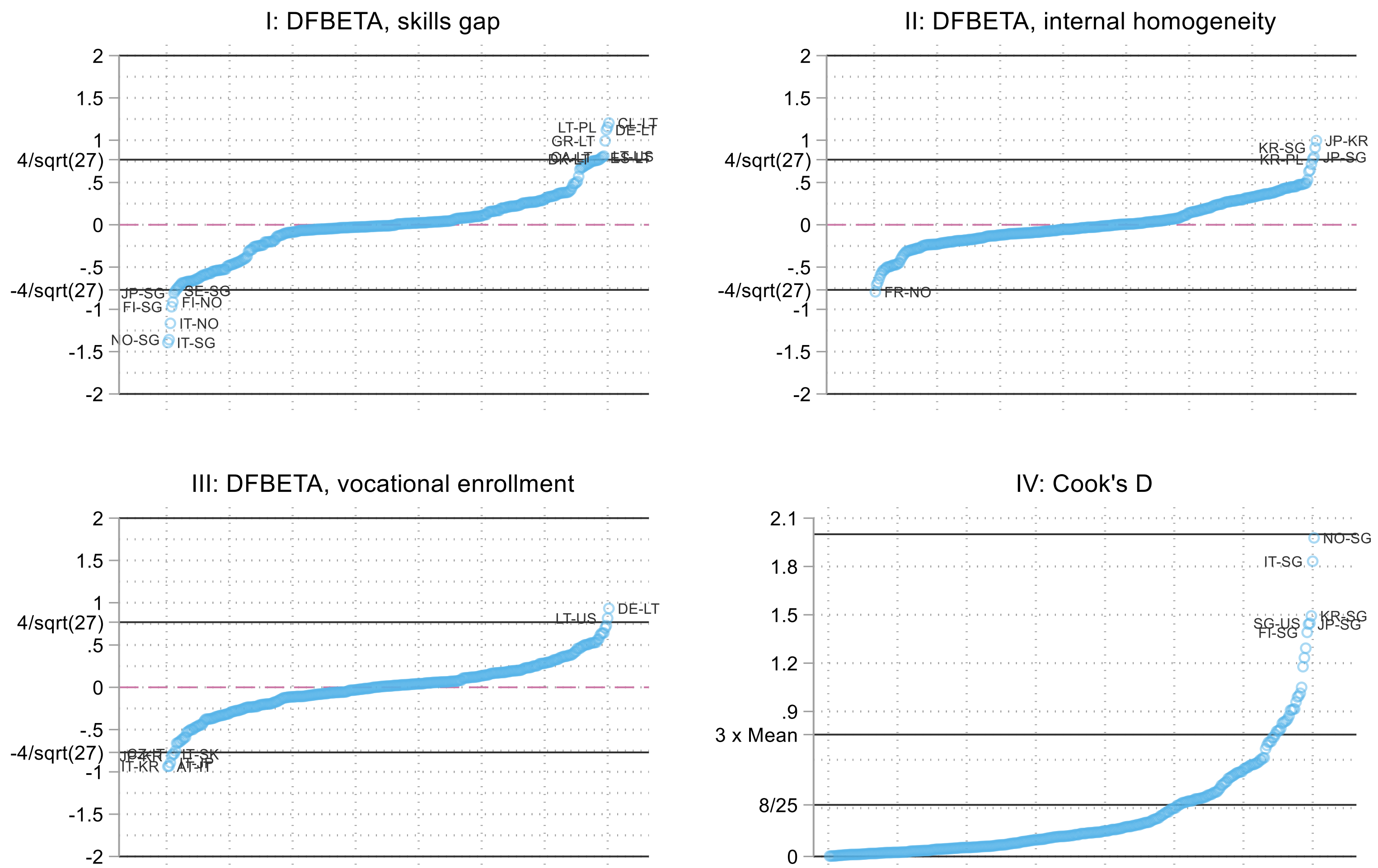


Figure D4: Delete-2 influence statistics for main analysis sample (27 countries)



E. COUNTRY-LEVEL RESULTS WITH NON-EMPLOYED RESPONDENTS DROPPED FROM THE SAMPLE

The sample used in the main analysis includes respondents who were not employed at the time of interview but had left their last job no more than five years ago. For these respondents, occupation codes and hence also ISEI scores refer to the last rather than the current job. As a robustness check, we re-estimated the main sequence of country-level regressions reported in Table 4 in the main article with the sample restricted to respondents who were employed at the time of interview. Table E1 displays the results of these regressions.

Results are similar to the main analysis. Focusing on our preferred specification, Model 7, the coefficients of both skills transparency measures (the skills gap and the index of internal homogeneity) are somewhat larger in absolute size than in Table 4 in the main article. Despite somewhat larger standard errors (which are to be expected because of the reduction of the lower-level sample size) both estimates also remain statistically significant. The coefficient of the vocational orientation also increases somewhat in absolute size (from -.469 in Table 4 in the main article to b=-.537 in Table E1), but it continues to be far from statistical significance.

Table E1. Country-level regressions of ISEI gap on measures of skills transparency and vocational orientation (sample restricted to respondents who were employed at the time of interview)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Skills gap	-0.572			-1.040*	-0.810+		-1.138*
	(0.416)			(0.406)	(0.422)		(0.408)
Index of internal homogeneity		-0.725+		-1.153*		-0.612	-1.015*
		(0.398)		(0.408)		(0.428)	(0.411)
Prevalence of vocational enrolment			-0.549		-0.801+	-0.325	-0.537
			(0.418)		(0.422)	(0.437)	(0.402)
Intercept	-5.559***	-5.545***	-5.543***	-5.558***	-5.548***	-5.539***	-5.548***
	(0.394)	(0.382)	(0.394)	(0.349)	(0.376)	(0.386)	(0.343)
N	27	27	27	27	27	27	27
R2	0.07	0.12	0.07	0.33	0.20	0.14	0.38
Adjusted R2	0.03	0.09	0.03	0.27	0.13	0.07	0.30

Note: Feasible Generalized Least Squares estimates following Lewis and Linzer (2005), based on 10 imputations/plausible values. The dependent variable is the ISEI gap between less-educated (ISCED categories 0-2) and intermediate-educated (ISCED categories 3/4) adults aged 16-54. The ISEI gap measures the occupational status of less-educated adults compared to intermediate educated adults, so more negative values correspond to a larger disadvantage for the less-educated group. ISEI gap is estimated using country-specific individual-level regressions that control for literacy and numeracy skills, sex, potential work experience, foreign-birth/foreign-language status, parental educational attainment, self-employment status (referred to as the "fully adjusted gap" in the text). All predictors are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Sources: PIAAC (rounds 1 and 2), authors' calculations.

F. COUNTRY-LEVEL RESULTS WITH SELF-EMPLOYED RESPONDENTS DROPPED FROM THE SAMPLE

The sample used in the main analysis includes wage and salary workers as well as self-employed respondents. However, signaling and screening theories, which form the theoretical basis for expecting labor market effects of skills transparency, may appear to be most directly applicable to hiring processes for wage and salary workers. That said, there are some reasons why skills transparency might also matter for the self-employed. For example, customers might statistically discriminate against the self-employed on the basis of formal qualifications and the extent of statistical discrimination against wage and salary workers might also indirectly affect selection into self-employment. These considerations suggest that there is no clear best way of treating self-employed respondents. As a straightforward robustness check, we therefore re-estimated the main sequence of country-level regressions reported in Table 4 in the main article with self-employed respondents excluded from the sample.

The results, reported in Table F1, are similar to those from the main analysis. Focusing on our preferred specification, Model 7, the coefficient estimate for the skills gap is marginally stronger than in the main analysis and remains statistically significant. The coefficient estimate for the index of internal homogeneity is somewhat weaker than in the main analysis, but the difference certainly is not dramatic (b = -.765 as opposed -.859 in Table 4 in the main article). The p-value on this estimate is .064, so it is no longer statistically significant acording to conventional standards. The coefficient of vocational orientation index is even weaker than in the main analysis when the self-employed are excluded, falling from -.469 in in Table 4 in the main article to -.341 in Table F1. Further analysis suggests that the attenuation of the coefficient of vocational orientation has two main sources. First, the rate of self-employment among the less educated decreases with the extent of vocational orientation and, second, self-employment is positively associated with occupational status among people with the same level of education. A likely explanation for this pattern is that access to self-employment in certain (relatively skilled) occupations tends to be more strongly regulated in countries with a stronger vocational orientation. In particular, having (at least) intermediate-level formal qualifications is often a prerequisite for establishing a business in certain occupations, effectively barring less-educated people from doing so.

Table F1. Country-level regressions of occupational gap on measures of skills transparency and vocational orientation (respondents who were/are self-employed in their last/current job excluded from the analysis)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Skills gap	-0.588			-0.935*	-0.756+		-1.002*
	(0.381)			(0.382)	(0.395)		(0.391)
Index of internal homogeneity		-0.490		-0.858*		-0.439	-0.765+
		(0.366)		(0.376)		(0.399)	(0.390)
Prevalence of vocational enrolment			-0.306		-0.546	-0.139	-0.341
			(0.388)		(0.390)	(0.414)	(0.385)
Intercept	-5.839***	-5.824***	-5.823***	-5.827***	-5.821***	-5.820***	-5.816***
	(0.356)	(0.359)	(0.367)	(0.330)	(0.350)	(0.365)	(0.331)
N	27	27	27	27	27	27	27
R2	0.09	0.07	0.03	0.27	0.17	0.07	0.30
Adjusted R2	0.05	0.03	0.01	0.21	0.09	0.01	0.20

Note: Feasible Generalized Least Squares estimates following Lewis and Linzer (2005), based on 10 imputations/plausible values. The dependent variable is the ISEI gap between less-educated (ISCED categories 0-2) and intermediate-educated (ISCED categories 3/4) adults aged 16-54. The ISEI gap measures the occupational status of less-educated adults compared to intermediate educated adults, so more negative values correspond to a larger disadvantage for the less-educated group. ISEI gap is estimated using country-specific individual-level regressions that control for literacy and numeracy skills, sex, potential work experience, foreign-birth/foreign-language status, parental educational attainment, self-employment status (referred to as the "fully adjusted gap" in the text). All predictors are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Sources: PIAAC (rounds 1 and 2), authors' calculations.

G. COUNTRY-LEVEL RESULTS WHEN PARENTAL EDUCATION IS IGNORED IN THE CONSTRUCTION OF THE SKILL TRANSPARENCY MEASURES

In constructing the direct country-level measures of skills transparency, we adjusted the skills gap and the within-group standard deviation of skills (which forms the basis of the index of internal homogeneity) for certain characteristics that are readily observable for employers. One of these, parental education, may be less obvious to employers than the others (sex, age, and foreign-birth/foreign-language status), however. We therefore re-estimated the main sequence of country-level regressions reported in Table 4 in the main article with alternative versions of the skills transparency measures that are not adjusted for parental education. Results are very similar to those reported in the main article. Focusing on our preferred specification, Model 7, the coefficients on both skills transparency measures are similar to those from Table 4 in the main article, and both remain statistically significant (with the coefficient estimate for the skills gap now even attaining $p < .01$).

Table G1. Country-level regressions of occupational gap on measures of skills transparency and vocational orientation (no adjustment for parental education in constructing skills gap and index of internal homogeneity)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Skills gap	-0.576			-0.954*	-0.775*		-1.035**
	(0.363)			(0.348)	(0.363)		(0.346)
Index of internal homogeneity		-0.599+		-0.968*		-0.496	-0.839*
		(0.347)		(0.347)		(0.372)	(0.350)
Prevalence of vocational enrolment			-0.480		-0.707+	-0.300	-0.486
			(0.361)		(0.355)	(0.380)	(0.341)
Intercept	-5.622***	-5.608***	-5.604***	-5.614***	-5.602***	-5.600***	-5.598***
	(0.337)	(0.333)	(0.341)	(0.298)	(0.320)	(0.336)	(0.292)
N	27	27	27	27	27	27	27
R2	0.10	0.11	0.07	0.34	0.23	0.14	0.40
Adjusted R2	0.06	0.08	0.03	0.29	0.17	0.06	0.32

Note: Feasible Generalized Least Squares estimates following Lewis and Linzer (2005), based on 10 imputations/plausible values. The dependent variable is the ISEI gap between less-educated (ISCED categories 0-2) and intermediate-educated (ISCED categories 3/4) adults aged 16-54. The ISEI gap measures the occupational status of less-educated adults compared to intermediate educated adults, so more negative values correspond to a larger disadvantage for the less-educated group. ISEI gap is estimated using country-specific individual-level regressions that control for literacy and numeracy skills, sex, potential work experience, foreign-birth/foreign-language status, parental educational attainment, self-employment status (referred to as the "fully adjusted gap" in the text). All predictors are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Sources: PIAAC (rounds 1 and 2), authors' calculations.

H. COUNTRY-LEVEL REGRESSIONS WITH THE WAGE GAP AS THE DEPENDENT VARIABLE

The analysis in the main article uses occupational status, measured in terms of the ISEI score, as the measure of labor market attainment. As discussed in Section 3.2 in the main article, there are two reasons for preferring occupational status to earnings or wages or earnings as the labor market outcome. First, we found that, in the context of our analysis, the “signal to noise ratio” is higher for the occupational status gap between less- and intermediate-educated adults than it is for the corresponding wage gap. More specifically, we used a random effects model to separate true between-country variation in the different labor market gaps from variation due to sampling error. This exercise indicated that 74.1% of the overall between country variance in the occupational status gap represent true variation, compared to only 69.7% for the gap in log hourly wages. The second reason is that wage inequalities are known to depend on a variety of contextual factors such as collective bargaining institutions and minimum wage legislation that are difficult to account for given the limited degrees of freedom at the country level.

Despite these arguments against using wages as the labor market outcome, Table H1 presents the main sequence of country-level regressions with the country-specific wage gaps between less- and intermediate adults as the dependent variable. The restrictions of the individual-level sample underlying these results differ somewhat from those applied in the case of the ISEI gap. In particular, non-employed respondents are excluded (even if they had left their last job no more than five years before the interview), as are the self-employed. This is because information on hourly wages is only available for wage and salary workers who were employed at the time of the interview in PIAAC. The wage gaps are adjusted for the same set of predictors as the fully adjusted ISEI gap in the main article (with the exception of self-employment status because, as just noted, the self-employed are excluded).

The results in Table H1 can be quickly summarized. We find essentially no evidence for either Hypothesis 3, 4, or 6. The only predictor for which we find some hints of the expected negative effect is the skills gap. or the other two predictors, the internal homogeneity and the vocational enrollment measures, the estimated coefficients are very small (both in terms of effect sizes and relative to their standard errors) and even tend to go in the “wrong” direction in that they are positively rather than negatively signed.

Table H1. Country-level regressions of the gap in log hourly wages on measures of skills transparency and vocational orientation

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Skills gap	-0.023+			-0.022+	-0.020		-0.020
	(0.011)			(0.013)	(0.012)		(0.013)
Index of internal homogeneity		0.010		0.002		0.005	0.000
		(0.012)		(0.012)		(0.013)	(0.013)
Prevalence of vocational enrolment			0.015		0.009	0.013	0.009
			(0.011)		(0.011)	(0.012)	(0.012)
Intercept	-0.121***	-0.121***	-0.121***	-0.121***	-0.122***	-0.122***	-0.122***
	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)
N	27	27	27	27	27	27	27
R2	0.14	0.03	0.07	0.15	0.16	0.08	0.17
Adjusted R2	0.11	0.01	0.04	0.07	0.09	0.02	0.06

Note: Feasible Generalized Least Squares estimates following Lewis and Linzer (2005), based on 10 imputations/plausible values. The dependent variable is the gap in log hourly wages between less-educated (ISCED categories 0-2) and intermediate-educated (ISCED categories 3/4) adults aged 16-54. The wage gap measures the log hourly wage of less-educated adults compared to intermediate educated adults, so more negative values correspond to a larger disadvantage for the less-educated group. The wage gaps are estimated using country-specific individual-level regressions that control for literacy and numeracy skills, sex, potential work experience, foreign-birth/foreign-language status, parental educational attainment, self-employment status. All predictors are z-standardized (mean of 0, standard deviation of 1). Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-tailed tests).

Sources: PIAAC (rounds 1 and 2), authors' calculations.

REFERENCES

Carroll, G.R., Mayer, K.U. 1986. Job-Shift Patterns in the Federal Republic of Germany: The Effects of Social Class, Industrial Sector, and Organizational Size. *American Sociological Review* 51(3)*, 323-341*.

Fox, J. 2016. *Applied Regression Analysis and Genralized Linear Models*. Third Edition. Thousand Oaks: Sage.

Stinchcombe, A. 1979. Social Mobility in Industrial Labor Markets. *Acta Sociologica* 22(3), 217-245.