

Heisig, Jan Paul

Article — Published Version

Measuring the signaling value of educational degrees: secondary education systems and the internal homogeneity of educational groups

Large-scale Assessments in Education

Provided in Cooperation with:

WZB Berlin Social Science Center

Suggested Citation: Heisig, Jan Paul (2018) : Measuring the signaling value of educational degrees: secondary education systems and the internal homogeneity of educational groups, Large-scale Assessments in Education, ISSN 2196-0739 Berlin, Springer Nature, Berlin, Vol. 6, pp. 1-35, <https://doi.org/10.1186/s40536-018-0062-1>

This Version is available at:

<https://hdl.handle.net/10419/182012>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

RESEARCH

Open Access



Measuring the signaling value of educational degrees: secondary education systems and the internal homogeneity of educational groups

Jan Paul Heisig*

*Correspondence:
jan.heisig@wzb.eu
WZB Berlin Social Science
Center, Reichpietschufer 50,
10785 Berlin, Germany

Abstract

Background: By providing high-quality, internationally comparable data on the cognitive skills of working-age adults, the Programme for the International Assessment of Adult Competencies (PIAAC) offers great potential for illuminating the complex interplay of formal qualifications and skills in shaping labor market attainment as well as social inequalities more broadly. I argue that PIAAC can be used to construct direct, country-level measures of the ‘skill transparency’ or ‘signaling value’ of formal qualifications, that is, of how informative the latter are about a person’s actual skills. The primary goal of the analysis is to extend previous work on skills gaps by educational attainment and map cross-national variation in the internal skills homogeneity of educational groups as a second dimension shaping the signaling value of educational degrees. I also explore whether the internal homogeneity of educational groups is related to national (secondary) education systems.

Methods: I use a sample of 30,646 20-to-34-year-olds in 21 countries that participated in the first round of PIAAC. The internal homogeneity of educational groups is measured using the residual standard deviation of literacy and numeracy skills after adjusting for sex, age, and foreign-birth/foreign-language status. Residual standard deviations for the different educational groups are subjected to a factor analysis to construct a one-dimensional measure of internal homogeneity for each country. This index of internal homogeneity is then related to education system characteristics in a series of country-level regressions.

Results: The internal homogeneity of educational groups with respect to literacy and numeracy skills varies considerably across countries and is highly correlated across both skill domains and educational groups. Educational groups tend to be more homogeneous in countries with stronger (ability-related) tracking in secondary education. In addition, there is some evidence that internal homogeneity declines when instructional resources such as computer hardware and lab equipment are distributed more unequally across schools. An unexpected finding is that internal homogeneity is negatively associated with standardization of input (e.g., curricula, textbooks).

Conclusions: The signaling value of educational degrees varies substantially across advanced economies, not only in terms of skills gaps among educational groups, but also in terms of their internal homogeneity. Some features of secondary education systems appear to be systematically related to the extent of internal homogeneity. The

findings lend empirical support to so far untested assumptions about the relationship between formal qualifications and skills in cross-national research on labor market inequalities.

Keywords: PIAAC, Education systems, Educational credentials, Labor market attainment, Signaling, Screening, Human capital theory

Background

For a long time, empirical studies of educational success and its importance for labor market attainment have largely defined education in terms of formal qualifications (i.e., in terms of educational degrees). Only recently has it become possible to also consider educational success in terms of an individual's actual competencies: Especially from the 1990s onwards, an ever-growing number of (international) large-scale assessment studies have begun to collect high-quality data on the actual skills of individuals by administering carefully designed test items to representative samples. Most large-scale assessment studies focus on school-aged children and adolescents, but a few have also surveyed working-age adults. The first cycle of the *Programme for the International Assessment of Adult Competencies* (PIAAC) is the so far most ambitious effort of the latter type.

Large-scale assessment data on adults offer numerous analytic possibilities. One of the most exciting ones is that they enable us to better understand the complex relationships among educational attainment, actual competencies, and labor market outcomes. For example, previous studies have found that adults with higher formal qualifications have higher (average) skills, but that the magnitude of skill differentials among educational groups varies considerably across countries and is related to (secondary) education systems (Heisig and Solga 2015; Park and Kyei 2011).

In this article, I extend this line of research by studying another crucial dimension of the qualification-skill nexus: the internal homogeneity of educational groups. Using PIAAC data covering 21 advanced economies, I seek to answer two primary research questions: (1) How homogeneous are educational groups with respect to the actual skills of their members across a diverse set of advanced economies? (2) Are country differences in the extent of homogeneity related to key features of secondary education systems such as stratification (tracking) and standardization?

The answers to these questions are interesting because they will contribute to a more comprehensive picture of the relationship between qualifications and skills by moving 'beyond the mean' (see also Spörlein and Schlüter 2018, recent study of within-group variation in competencies among immigrant and native-born adolescents). More importantly, investigating cross-national variation in the internal homogeneity of educational groups will enhance our understanding of the role that formal qualifications and actual skills play for labor market inequalities. Several studies have shown that the relationship between formal qualifications and labor market outcomes is stronger in some countries than in others and that the strength of the association is related to secondary education systems (e.g., Andersen and Van de Werfhorst 2010; Bol and Van de Werfhorst 2011; Shavit and Müller 1998; Van de Werfhorst 2011). One explanation for this pattern is that some education systems are more 'skill transparent' than others (see, in particular,

Andersen and Van de Werfhorst 2010; Bol and Van de Werfhorst 2011). In a more skill transparent system, the argument goes, formal qualifications are more informative about the actual skills a person has—in the terminology of Spence (1973), they are a stronger ‘signal’ of productivity. Employers should therefore attach greater weight to formal qualifications in more skill transparent contexts, which in turn should amplify the importance of formal qualifications for labor market attainment.

While plausible, empirical tests of the skill transparency hypothesis have so far relied on untested assumptions about the relationship between certain education system characteristics and the extent of skill transparency. In particular, scholars have argued that the skill transparency of educational credentials increases with the extent of tracking and with the emphasis on vocational training in secondary education (Andersen and Van de Werfhorst 2010; Bol and Van de Werfhorst 2011). Attempts to measure skill transparency more directly remain rare. This is where the contribution of the present study lies. As I argue below, formal qualifications are more informative about the skills a person has when, (a), there are large skill differentials among educational groups and when, (b), groups are internally homogeneous. Whereas recent work on ‘skills gaps’ (Heisig and Solga 2015; Park and Kyei 2011) has begun to investigate cross-national variation in skill differentials, cross-national variation in the internal homogeneity of educational groups has not been studied so far. The following analysis addresses this gap by quantifying the extent of internal homogeneity for a set of 21 advanced economies and by investigating whether it is systematically related to the way secondary education is organized.

The remainder of the paper is structured as follows. In the next section, I review prominent explanations of the relationship between educational degrees and labor market attainment, with particular emphasis on how the different approaches conceive of the role of actual competencies. In the ensuing section, I argue that the signaling value of educational degrees not only depends on skills gaps among educational groups, but also on their internal homogeneity. I then go on to review some related studies and motivate the main research questions of the present article. I also formulate hypotheses about how education system characteristics might be related to the internal homogeneity of educational groups. After describing the PIAAC data and methods of analysis, I present the main empirical results. I first construct a country-level index of the internal homogeneity of educational groups and then test my hypotheses by regressing it on indicators of education system characteristics. The last part of the empirical section reproduces key findings from a related study (Heisig et al. 2016) to show that cross-national variation in the skills gap and in the extent of internal homogeneity help to account for variation in the labor market disadvantage of less-educated adults, even after controlling for cognitive skills at the individual level. The final section draws conclusions and discusses some limitations as well as directions for future research.

Education and skills in theories of labor market attainment

Numerous studies show that educational attainment in the sense of formal qualifications is positively related to labor market outcomes. In all advanced economies, individuals with higher educational degrees have higher employment rates, occupational status, and earnings than their less-educated counterparts. At the same time, labor market returns to formal qualifications vary widely across countries (see, for example, Shavit and Müller

1998; Van de Werfhorst 2011). Social scientists have proposed several explanations for these empirical regularities (for overviews, see Bills 2003; Bills et al. 2017). Three broad classes of theoretical accounts have been particularly influential: human capital theory, signaling/screening explanations, and credentialism.

According to the human capital explanation (e.g., Mincer 1970) the advantages of more educated workers are largely due to their higher levels of skills and productivity: ‘Schooling provides marketable skills and abilities relevant to job performance. This makes the more highly schooled applicants more valuable to employers, thus raising their incomes and their opportunities for securing jobs’ (Bills 2003, 444). The simple human capital model can be refined considerably, for example, to differentiate between different types of general and specific skills (e.g., industry- or occupation-specific skills; Becker 1962). However, such extensions do not alter the central themes of human capital theory: that education serves the development of productive skills, that skills in turn are the main driver of the association between educational attainment and labor market outcomes, and that these relationships are rather straightforward and direct.

Like human capital theory, signaling (Spence 1973) and screening (Arrow 1973; Stiglitz 1975) approaches¹ generally subscribe to the notion that there is a positive link between skills and productivity. They do, however, emphasize a crucial aspect that may complicate the link between individual skills and labor market outcomes, namely that actual skills are very difficult to observe. The central idea of signaling explanations is that employers will therefore rely on more readily available proxies (i.e., on signals) in forming beliefs about the actual skills (or trainability) of a person. In addition to the direct link emphasized by human capital theory, the signaling story thus suggests a second pathway through which the association between formal qualifications and skills might influence labor market inequalities: by making (easy-to-observe) qualifications a useful signal of (hard-to-observe) skills. From this perspective, the advantages of higher educated workers at least partly stem from employers’ assumptions about group-level differences in productivity and from concomitant (positive) statistical discrimination (Aigner and Cain 1977; Arrow 1973)—rather than from direct employer responses to individual-level variation in skills.

While quite heterogeneous in their details, credentialist perspectives (Berg 1971; Collins 1979) generally break with the assumption that skills and productivity differentials are the primary reason why individuals with higher formal qualifications tend to be more successful on the labor market. The roots of credentialism can be traced back to Weber to whom ‘educational credentials were essentially cultural-political constructions of competence and organizational loyalty that bore little relationship to the technical demands of modern work’ (Brown 2001, 21). In its strongest forms, credentialism disputes any meaningful relationship between formal qualifications and job performance. Weaker versions ‘merely [...] argue that the ratio between education and productivity is smaller than that between education and rewards’ (Bills 2003, 452).

¹ I concur with (Bills 2003, 447) that screening and signaling theories are very closely related and primarily differ in that ‘in the former, firms move first and, in the latter students move first’, as he remarks in his discussion of Weiss (1995). Hence, I treat them as a single perspective in this paper.

A prominent theme of credentialism is that educational degrees are used to restrict access to advantageous positions (e.g., via occupational licensing), thereby reducing the supply of certain types of workers and generating monopoly rents (Sørensen 2000). From this perspective, credentials are instruments of social closure that generate, maintain, and legitimize social inequalities (Collins 1979). Another (alleged) phenomenon emphasized by credentialists is ‘credential inflation’, a trend toward ever-increasing educational attainment that is viewed as unrelated to any real changes in work demands (Berg 1971; Collins 1979). If employers nevertheless look to formal qualifications in hiring decisions, such a trend may become self-sustaining because individuals seek ever-higher qualifications in order to stick out from the pool of applicants and to be ranked ahead of their competitors in the ‘labor queues’ emphasized in models of job competition (Thurow 1979).²

These different explanations of labor market inequalities are not mutually exclusive and it is not straightforward to disentangle them empirically, but quite some progress has been made in recent decades (for reviews, see Bills 2003; Bills et al. 2017). One promising line of research has begun to investigate how the relative importance of the different mechanisms depends on education systems and other macro-structural conditions (e.g., Bol and Van de Werfhorst 2011; Di Stasio et al. 2016; Van de Werfhorst 2011). A crucial prerequisite for advancing this agenda is to conceptualize and measure potentially relevant contextual factors. The main goal of the present study therefore is to further our understanding of cross-national differences in the ‘signaling value’—or ‘skill transparency’ (Andersen and Van de Werfhorst 2010)—of formal qualifications, that is, of how informative the latter are about a person’s actual skills. If such differences exist and if they can be measured, this may help us to assess the importance of the signaling explanation and to better understand why we find greater labor market inequalities according to formal qualifications in some countries than in others. As a first step towards these goals, I now elaborate how the signaling value of educational degrees can be conceptualized in terms of the distribution of actual skills conditional on formal qualifications.

Two dimensions of skill transparency: skills differentials and internal homogeneity

The importance that employers attach to formal qualifications should depend on (at least) two aspects of the conditional distribution of actual skills. The first is the extent of skills differentials or ‘skills gaps’ among different educational groups (Heisig and Solga 2015; Park and Kyei 2011). Other things being equal, formal qualifications are more informative about the actual skills that people with different qualifications have when the skills differential between their respective educational groups is large. The second crucial dimension is internal homogeneity. Other things, including the skills gap, being equal, formal qualifications send a stronger signal about an individual’s actual skills when the educational group that the individual belongs to is internally more homogeneous.

² It is worth noting that job competition models have a strong theoretical affinity with signaling and screening explanations as well and that the latter might similarly give rise to an ‘educational arms race’ (Bills 2016, 69) where individuals seek ever higher qualifications to distinguish themselves from their peers (see, for example, Di Stasio et al. 2016).

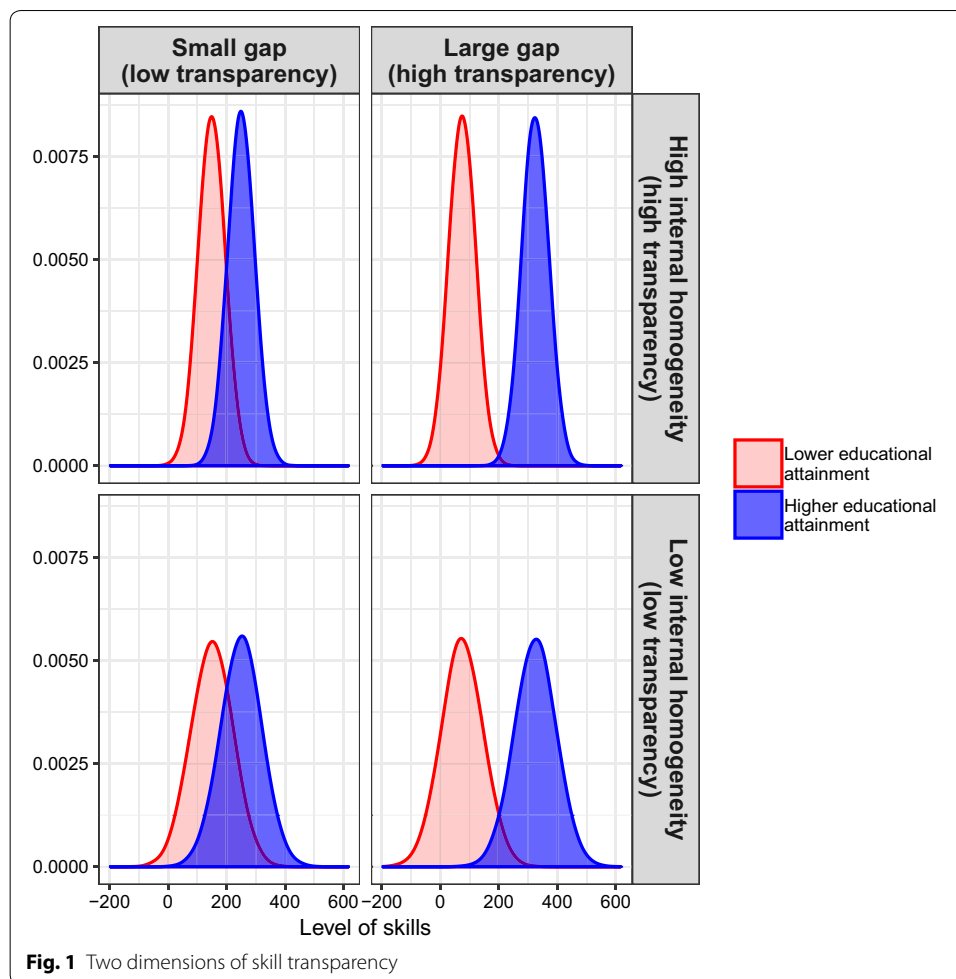


Figure 1 illustrates these ideas graphically. The density curves represent skill distributions for two hypothetical groups, a lower-educated one represented by the light (red) curves and a higher educated one represented by the dark (blue) curves. Skill transparency is lowest in the lower left graph. Here, the skill gap between the two groups is small, that is, the skill means are quite similar across the two groups, and both groups are internally heterogeneous. Members of the higher-educated group tend to have higher skills, but there clearly is considerable overlap among the two groups. In this situation, a hypothetical employer can learn comparatively little from observing the formal qualifications of applicants. His/her best guess would be that an individual belonging to the higher-educated group has higher skill than a person from the lower-educated group. However, the expected difference between any such pair of applicants would be quite small and there would always be a good chance that, for a given pair of applicants, the difference might even be reversed. In such a situation, an employer would likely pay greater attention to other easily observable characteristics that are correlated with skills or invest resources into learning more about the actual skills of the competing applicants (e.g., by inviting both rather than only the higher-educated applicant for a job interview, or by hiring both for a limited screening period).

In the bottom right graph, the skills gap is considerably larger than in the bottom left graph (but within-group variability is the same). Clearly, this reduces overlap between the two groups and renders group membership a stronger predictor of skills. The predictive power of formal qualifications also increases as one moves from the bottom to the top row, but here the reason is that both groups become internally more homogeneous. Skill transparency is highest in the top right graph where the skills gap is large and members of the same group tend to be very similar in terms of the actual skills that they have. In this hypothetical situation, an employer could be almost certain that he/she would hire a more skilled employee by choosing an applicant from the higher-educated rather than the lower-educated group. Moreover, the expected difference between applicants from the two groups would be quite large.

In sum, this discussion suggests that direct (country-level) measures of the signaling value (or skill transparency) of educational degrees should capture two crucial dimensions of the distribution of skills conditional on educational attainment: the size of skills differentials among educational groups and their internal homogeneity.

Previous research

Despite the pervasiveness of (implicit) assumptions about the relationship between formal qualifications and skills in research on labor market inequalities, there is very little robust empirical knowledge about what this relationship actually looks like and whether it differs across countries—partly due to a shortage of data on the skills of adults. Many cross-national studies of labor market inequalities by educational attainment refer to signaling explanations (e.g., Andersen and Van de Werfhorst 2010; Bol and Van de Werfhorst 2011), but they do not include direct measures of skill transparency based on empirical information about skill distributions. Abrassart (2013) uses data on 14 countries from the *International Adult Literacy Survey* (IALS) and finds that the labor market disadvantage of less-educated relative to intermediate-educated adults (measured as the adjusted difference in employment rates) is related to the skills gap at the country level. He also speculates that this relationship might be attributable to the signaling mechanism (i.e., statistical discrimination related to formal qualifications) being stronger in countries with a larger skills gap. However, his analysis does not control for skills at the individual level, so it is unclear to what extent the aggregate-level association picks up the direct, individual-level effects of skills (as emphasized by the human capital approach). Moreover, Abrassart (2013) does not consider the internal homogeneity of the different educational groups.

Two studies have examined cross-national differences in skills gaps and related them to various country-level explanatory variables. Using data on 19 countries from IALS, Park and Kyei (2011) study cross-national variation in skills gaps by educational attainment, differentiating among adults with low (highest degree below upper secondary level), intermediate (highest degree at upper secondary level), and high education (highest degree at tertiary level). They find substantial country differences in skills differentials among the educational groups, particularly between the low educated and the two higher-educated groups. They further show that the skills gaps between the low educated and the other groups are larger in countries where educational resources (such as instructional resources, teacher experience, or class size) are more unequally distributed

across schools. A likely explanation that Park and Kyei do not investigate empirically is that low-achieving students tend to cluster in disadvantaged schools, resulting in a vicious cycle of cumulative disadvantage.

In a more recent study using data on 18 countries from PIAAC, Heisig and Solga (2015) investigate the link between secondary education systems and skills, focusing on the skills gap between adults with low and intermediate formal qualifications. They confirm Park and Kyei's (2011) result of substantial cross-national variation in the skills gap and find that the latter increases with the extent of external differentiation in lower and upper secondary education and decreases with the extent of vocational orientation of upper secondary education. External differentiation refers to the extent of tracking in secondary education, that is, to the extent to which students are allocated to different programs depending on their academic abilities and to how early this kind of separation occurs (Bol and Van de Werfhorst 2016). Vocational orientation refers to the prevalence of vocational/occupation-specific—as opposed to general academic—programs in upper secondary education (*ibid.*).

One should be cautious in attaching a causal interpretation to the cross-sectional country-level relationships uncovered by Heisig and Solga (2015). That said, the authors discuss several pathways through which secondary education systems might (causally) affect skills differentials among educational groups. As for external differentiation, they stress the importance of selection by external gatekeepers that may negatively affect low-achieving students' opportunities for participation in upper secondary education. A further possibility is that tracked systems deprive low-achieving students of stimulating interactions with higher-achieving peers and thereby reinforce preexisting inequalities (Gamoran 2000). As for vocational orientation, Heisig and Solga (2015) adopt an argument by Soskice (1994) and suggest that vocational options might reduce inequalities by providing incentives for low-achieving students to work hard and stay in school (see also Green and Pensiero 2016).

Research questions and contributions of the present study

Quantifying cross-national differences in internal homogeneity

In this paper, I extend previous work on cross-national variation in skills gaps by examining variation in a second crucial dimension of skill transparency, the internal homogeneity of educational groups with respect to literacy and numeracy skills. As cross-national variation in the internal homogeneity of educational groups is largely uncharted territory, the first part of the analysis is primarily exploratory. The questions addressed in this part are:

1. Does the internal homogeneity of educational groups (in terms of literacy and numeracy skills) vary across countries?
2. Does the extent of internal homogeneity differ according to the skill domain (literacy or numeracy)?
3. Does the internal homogeneity of different educational groups vary independently or is it highly correlated?

National education systems and internal homogeneity

In the second part of the analysis, I relate the internal homogeneity of educational groups to key features of national education systems. In particular, I explore the roles of between-school resource inequality (Park and Kyei 2011), external differentiation/tracking and vocational orientation (Heisig and Solga 2015), and standardization of input and output (Bol and Van de Werfhorst 2016).

For most of these education system characteristics, hypotheses concerning their relationship with the internal homogeneity of educational groups suggest themselves quite naturally. As noted above, Park and Kyei (2011) found that greater between-school inequalities in instructional resources are associated with larger skills differentials among educational groups. It seems plausible that, by creating more diverse learning environments and experiences, between-school inequality also reduces the internal homogeneity of educational groups—unless resource inequalities are specifically targeted towards the reduction of inequalities (e.g., by giving greater resources to schools with high shares of disadvantaged students). Park and Kyei's (2011) findings are indirect evidence that such compensatory targeting is not the predominant reason for resource inequalities, however.

Stronger tracking can be expected to increase the internal homogeneity of educational groups. In externally differentiated systems, 'gatekeepers' such as teachers and school principals (and in apprenticeship systems also employers) tend to have considerable control over who is admitted to which programs at the lower and upper secondary levels. To the extent that such selection takes academic abilities into account (as 'meritocratic' selection procedures typically require) this external selection should result in more homogeneous student populations relative to 'choice-driven' (Jackson et al. 2012) comprehensive education systems.

It is more difficult to formulate clear expectations concerning the relationship between vocational orientation and internal homogeneity. I therefore mainly include this predictor because of its prominent role in past research on labor market returns to education (e.g., Bol and Van de Werfhorst 2011; Shavit and Müller 1998) and skills gaps among educational groups (Heisig and Solga 2015).

With respect to standardization (Allmendinger 1989), Bol and Van de Werfhorst (2016) distinguish between standardization of input and output. 'Standardisation of input refers to the extent to which schools have limited control over the input in education' (Bol and Van de Werfhorst 2016, 75), in particular over the content of, and instruments (e.g., textbooks) used in, teaching. Standardization of output, by contrast, is defined as the extent to which educational output is benchmarked against unified external standards. The most widely studied instrument for achieving this latter type of standardization are central exit examinations (Bol and Van de Werfhorst 2016).

As for the relationship between standardization and the internal homogeneity of educational groups, one would expect higher standardization of input to result in more homogeneous skills distributions by homogenizing the learning experiences of students. Similarly, standardization of output might also increase internal homogeneity by setting uniform and clearly defined goals for education.

Internal homogeneity and labor market inequalities

In the last step of the empirical analysis, I reproduce key findings from Heisig et al.'s (2016) analysis of cross-national variation in the labor market disadvantage of less-educated adults to illustrate that direct measures of skill transparency help to account for cross-national variation in labor market inequalities by educational attainment.

Methods

Data and sample restrictions

I use data from the first round of PIAAC (OECD 2013a, b), collected in 2011/2012 in a total of 24 countries. I exclude Cyprus and Russia because of concerns about data quality and Australia because it provides no public use file. The PIAAC target population for each country is the non-institutionalized population aged 16–65 residing in the country when the survey was conducted. All samples are probability samples of the target population. All individual-level analyses are weighted using the final sampling weights to account for unequal selection probabilities.

The analysis includes all respondents aged 20–34 who were not enrolled in full-time education at the time of interview and who obtained their highest degree in the country where they participated in the survey.³ In combination with the lower age bound, the restriction to respondents who were not enrolled in education ensures that sample members have completed their main educational biographies. The upper age threshold ensures a good match with the education system measures, which generally refer to the mid-1990s to mid-2000s (see below).

Another reason for focusing on adults in their 20s and early 30s is that signaling theory is most compelling as an account of how employers assess the likely productivity of young and inexperienced workers. For more experienced workers, employers can draw on additional (work history) information (Altonji and Pierret 2001). A drawback of restricting the analysis to a relatively narrow age range is that sample sizes become quite small, especially for adults with low levels of education who are a relatively small group in many countries. Table 1 shows that the less-educated group accounts for less than ten percent of the population under study in many countries. At about 3%, the group is particularly small in South Korea, which has seen rapid educational expansion during recent decades (Park 2007). Reassuringly, however, the South Korean case does not have a major impact on the results of the analysis, as further discussed in the “Robustness checks” section below.

As another means of addressing the issue of small sample sizes and assessing the robustness of the findings, I reran the analysis with respondents aged 16–54 at the time of interview (after enforcing the sample restrictions concerning participation in education, foreign-degree status, and literacy-related nonresponse). This more generous age restriction also matches that of Heisig et al.'s (2016) study of labor market inequalities. The original results of the present study remain similar when using the broader age restriction (again, see the “Robustness checks” section for details).

³ I exclude literacy-related non-respondents who did not complete the survey for literacy-related reasons and for whom very little information is available (in most cases, only age and sex; OECD 2013a; Van de Kerckhove et al. 2013).

Table 1 Descriptive statistics for individual-level measures by country

Country	Mean literacy score	Mean numeracy score	% less-educated (ISCED 0-2)	% intermediate-educated (ISCED 3-4)	% high-educated (ISCED 5-6)	% male	Mean age	SD age	% native-born and tested in first language	% native-born and not tested in first language	% foreign-born and tested in first language	% foreign-born and not tested in first language	N
Austria	279.5	282.0	13.0	72.7	14.3	51.0	27.9	4.0	87.2	3.8	2.0	7.0	986
Belgium	290.1	292.5	8.5	48.4	43.1	49.9	28.7	3.9	91.5	3.1	2.0	3.4	1006
Canada	283.7	274.8	9.6	40.8	49.6	52.9	28.5	3.9	80.6	5.2	5.1	9.1	4508
Czech Rep.	282.4	283.4	8.6	67.6	23.8	53.6	29.1	3.8	95.8	0.0	1.1	3.2	1520
Denmark	278.5	282.7	17.6	41.4	41.0	53.2	28.8	3.8	83.0	1.1	2.4	13.5	970
Estonia	280.9	277.4	18.5	44.0	37.5	53.6	28.4	3.9	93.1	2.7	3.8	0.4	1572
Finland	306.2	297.2	8.2	53.3	38.5	50.9	28.6	3.9	94.7	1.7	1.1	2.6	974
France	275.1	266.0	15.2	48.7	36.1	49.9	28.3	3.9	89.0	2.2	3.4	5.4	1456
Germany	274.8	274.6	14.6	57.5	28.0	52.5	28.5	3.8	80.9	2.6	3.2	13.4	1057
Ireland	272.9	262.2	13.3	42.9	43.8	47.4	28.9	3.9	73.6	0.5	11.1	14.9	1480
Italy	254.2	254.7	35.0	46.8	18.2	53.4	28.9	3.8	83.8	1.7	2.2	12.3	855
Japan	306.8	293.6	8.6	39.2	52.2	49.6	28.4	3.9	99.5	0.0	0.4	0.1	1127
Korea	287.6	277.2	3.1	37.7	59.1	50.3	29.0	3.7	97.1	0.3	0.8	1.8	1407
Netherlands	294.4	288.6	19.7	44.4	36.0	51.8	28.6	3.9	87.1	2.3	3.4	7.2	846
Norway	283.7	280.4	20.2	41.6	38.2	52.6	28.6	3.9	80.9	1.6	1.1	16.4	958
Poland	272.9	265.7	7.2	54.1	38.6	51.9	28.6	3.6	98.6	1.2	0.0	0.1	3240
Slovak Rep.	274.8	275.0	14.0	63.8	22.3	53.0	28.7	3.8	92.7	6.3	0.6	0.4	1462
Spain	256.5	251.3	41.3	25.0	33.7	51.4	29.2	3.6	77.8	2.2	11.4	8.6	1226
Sweden	289.3	287.3	14.3	56.8	28.9	53.9	28.0	4.1	82.9	2.6	1.6	12.9	858
United Kingdom	274.9	262.5	19.5	37.6	42.8	50.1	27.9	4.0	80.7	1.8	7.1	10.5	2053
United States	271.8	255.1	11.8	52.3	36.0	50.1	28.1	3.9	81.3	3.8	3.5	11.4	1085

Regressions include age using 5-year categories. Means and standard deviations of age reported in this table were calculated by assigning the midpoints of the categories. Age is included in the regressions using 5-year groups. To compute the mean of standard deviation of age reported here cases were assigned the midpoints of their age categories

ISCED international standard classification of education, SD standard deviation

Handling of plausible values and missing data

To control respondent burden, PIAAC administered only a limited number of test items to each individual participant. To accurately reflect statistical uncertainty about individual competence levels, the data therefore provide ten plausible values rather than a unique competence score for each respondent. All results presented below are based on running the respective analysis ten times, once for each plausible value, and combining the resulting estimates using the appropriate rules (Little and Rubin 2002).

The prevalence of missing data is low in PIAAC. 30,705 cases meet the sample restrictions, after excluding 17 cases with missing information on whether they were in full-time education at the time of interview and/or obtained their highest degree in a foreign country. 59 of these cases are excluded because of missing information for at least one of the variables included in the analysis, resulting in a final sample size of 30,646 cases. Country-specific sample sizes range from 846 in the Netherlands to 4508 in Canada.

Individual-level measures

Cognitive skills are measured in terms of *literacy* and *numeracy* skills. According to (OECD 2013a, 59; emphasis in original), ‘Literacy is defined as the ability to understand, evaluate, use and engage with *written texts* to participate in society, to achieve one’s goals and to develop one’s knowledge and potential’ whereas ‘Numeracy is defined as the ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life.’

Educational attainment in terms of the highest educational degree is measured using a coarsened, three-category version of the 1997 revision of the International Standard Classification of Education (ISCED). I differentiate between workers with low (ISCED levels 0–2), intermediate (ISCED levels 3–4), and high education (ISCED levels 5–6). This is equivalent to the highest degree being below the upper secondary level, at the upper secondary (including the non-tertiary post-secondary) level, and at the tertiary level, respectively. In a sensitivity analysis, I explored the consequences of using a more fine-grained education measure with five categories. Findings were reassuring (see the “[Robustness checks](#)” section below).

Further individual-level measures used in the main analysis are *sex*, *age* (5-year groups), and *foreign-birth/foreign-language status*, a four-category variable indicating if the respondent was born in the country where he/she participated in the survey and if the language of the assessment was his/her first language. Table 1 displays country-specific means and proportions for the individual-level measures.

Country-level measures

The first country-level measure is the *between-school inequality of instructional resources*. The variable is based on information from the eighth-grade (middle) school principal questionnaire of the 1995 Trends in International Mathematics and Science Study (TIMSS).⁴ Principals were asked to what extent (four-point scale) their school’s

⁴ Finland did not participate in the 1995 round of TIMSS, so I use data from the 1999 round.

capacity for instruction was affected by shortages and inadequacies in 17 different domains such as heating/lighting, computer hardware, and lab equipment (for a full list, see Park and Kyei 2011, 887). Following Park and Kyei (2011), I average all 17 items to construct a school-level measure of resources and then compute the Theil index to measure inequalities among schools.

The four additional country-level predictors all come from version 4 of the Educational Systems Database by Bol and Van de Werfhorst (2016).⁵ The *external differentiation index* captures differences in the extent and timing of tracking in lower and upper secondary education. It is based on a principal factor analysis of three indicators: 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 measures 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.

The *vocational orientation index* is based on a principal factor analysis of the proportion of students at the upper secondary level who are enrolled in a vocational program, as provided in two sources: OECD (2006, Table C2.5) and UNESCO's online database (<http://data.uis.unesco.org/>), with values referring to 2004 (OECD) and 2006 (UNESCO) or the closest year available (further details are provided in Bol and Van de Werfhorst 2016).

Standardization of input is measured using three items administered to school principals in the 2006 round of PISA. The variables refer to the extent to which schools can autonomously choose textbooks, course content, and the types of courses being offered. Bol and Van de Werfhorst (2016) first computed, for each country and item, the proportion of school principals reporting that their school can make autonomous decisions. They then constructed a summary index by running a principal component factor analysis (Bol and Van de Werfhorst 2016, 81). Higher values on the index correspond to greater standardization (i.e., fewer principals reporting autonomy).

The measure of *standardization of output* is based on the existence of centralized exit exams, with the value one indicating that such exams exist and the value zero indicating that they do not. Bol and van de Werfhorst scored countries based on several sources (for details, see Bol and Van de Werfhorst 2016, 81f.). For three of the countries analyzed here (Canada, Germany, and the United States) the value lies between zero and one and corresponds to the proportion of regions with central examinations.

The education system measures generally refer to the state of education systems at some point between the mid-1990s and the mid-2000s, with the precise reference year varying across measures. Respondents included in the main analysis sample were 20–34 years old in 2011/2012. Thus they were born between 1977 and 1992. As secondary school usually starts between ages 10 and 12 and ends between ages 16 and 18, these cohorts attended secondary-level programs between 1987 and 2010, which ensures a rather good match with the education system measures.

⁵ Data are available for download at <http://thijsbol.com/data/> (last accessed: December 2, 2016).

Table 2 Values of country-level predictors

Country	Between-school inequality of instructional resources	External differentiation index (tracking)	Vocational orientation index	Standardization of input	Standardization of output
Austria	0.0130	1.817	1.701	− 0.095	0
Belgium	0.0093	1.018	0.945	0.033	0
Canada	0.0240	− 1.321	− 1.723	0.905	0.51
Czech Rep.	0.0105	1.621	1.744	− 0.595	1
Denmark	0.0332	− 0.870	0.455	− 0.298	1
Estonia	0.0177	n/a	− 0.441	− 0.647	1
Finland	0.0153	− 0.870	0.737	− 0.614	1
France	n/a	− 0.474	0.393	− 0.008	n/a
Germany	0.0185	1.862	0.887	0.018	0.44
Ireland	0.0309	− 0.302	− 0.354	− 0.236	1
Italy	0.0341	0.166	0.948	− 0.340	1
Japan	0.0190	− 0.474	− 0.729	− 1.243	1
Korea	0.0240	0.072	− 0.550	− 1.151	1
Netherlands	0.0104	0.937	1.260	− 0.701	1
Norway	0.0147	− 1.043	0.885	0.459	1
Poland	n/a	− 0.083	0.296	− 0.498	n/a
Slovakia	0.0180	1.621	1.492	0.817	1
Spain	0.0213	− 1.020	− 0.001	− 0.841	0
Sweden	0.0212	− 0.870	0.686	− 0.127	0
United Kingdom	0.0185	− 1.043	0.467	− 0.592	1
United States	0.0280	− 1.321	− 1.844	− 1.089	0.09
Mean	0.0201	− 0.029	0.345	− 0.326	0.69
Std. dev.	0.0075	1.111	0.995	0.581	0.44

See text for sources

n/a not available

Table 2 shows the values of the country-level predictors for the 21 countries included in the analysis. Not all measures are available for all countries. The main analysis of the country-level relationships between internal homogeneity and the education system measures will therefore focus on the 18 countries with complete data.

Table 4 in the [Appendix](#) shows pairwise correlations between the country-level predictors for the 18 countries without missing information. There is a strong positive correlation ($r = 0.681$) between the external differentiation and vocational orientation indices. Despite this relatively high correlation, there is substantial variation in the extent of tracking among countries with similar levels of vocational orientation and vice versa. Figure 6 in the [Appendix](#) shows a scatterplot of the two measures. While there are no countries with a strong vocational orientation and very low levels of external differentiation, several countries (notably the Scandinavian ones) combine moderate levels of external differentiation with a relatively strong vocational orientation. In countries such as Germany or Austria, both measures take high values as tracking occurs very early and vocational programs play an important role in upper secondary education. There are two further instances of relatively high correlations: the measure of between-school inequality of instructional resources correlates

quite strongly with the external differentiation (-0.477) and vocational orientation (-0.521) indices. All remaining correlations are quite low.

Statistical analysis

Measurement of internal homogeneity

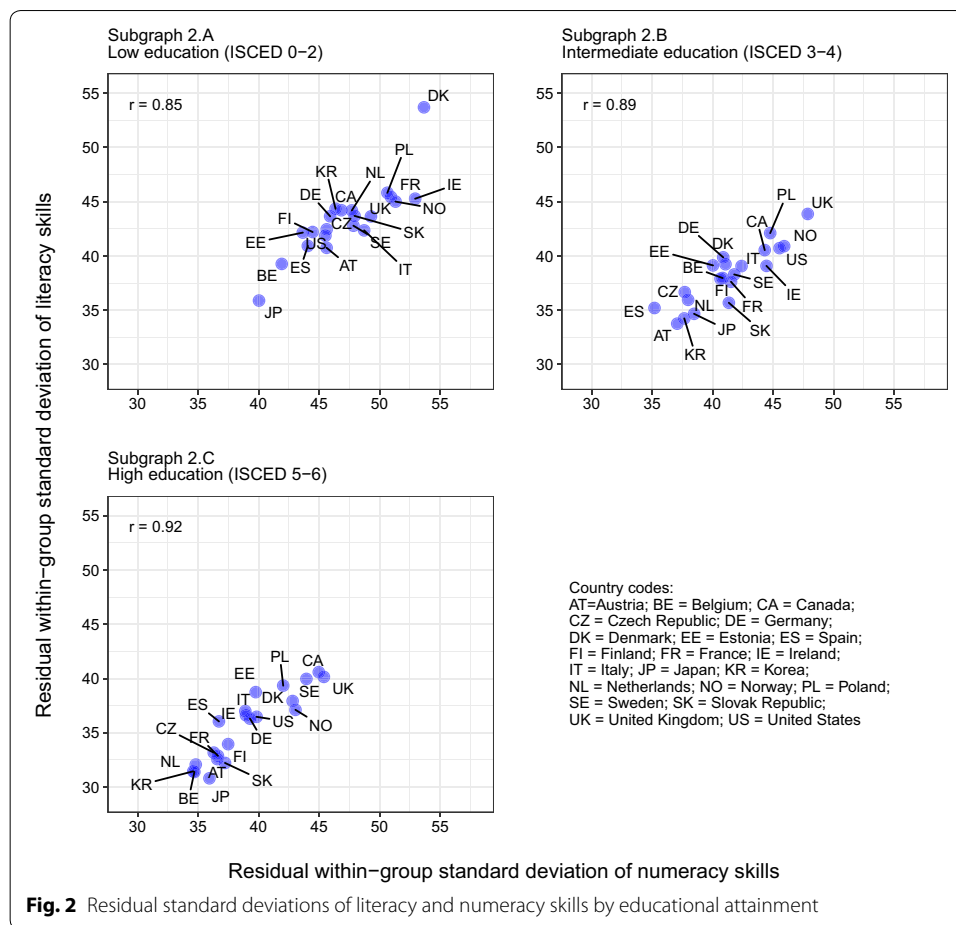
To measure the *internal homogeneity* of educational groups, I begin with a measure of internal *heterogeneity*, namely the *within-group standard deviation* of literacy and numeracy skills. Before calculating these standard deviations, I first adjust the competence scores for compositional differences with respect to basic socio-demographics. More concretely, I run country-specific linear regressions of the competence scores on the educational group variables and sex, age, and foreign-birth/foreign-language status and then calculate the standard deviations of the residuals from these regressions within the educational groups.

Not only should adjusting for these socio-demographics improve the comparability of country-specific estimates, it also makes sense for theoretical reasons: sex, age, and immigrant background are readily observable characteristics that employers likely use as further signals of skills—i.e., in addition to educational attainment—a possibility that is also emphasized in the rich literatures on (statistical) discrimination according to these characteristics (see, for example, the review article on racial discrimination by Pager and Shepherd 2008).

As the within-group standard deviations turn out to be highly correlated, both across skill domains and across educational groups, I use a principal factor analysis to reduce the dimensionality and create a summary index. I reverse-code the factor scores from this analysis to arrive at a measure of internal homogeneity, that is, a measure where higher values correspond to greater skill transparency (i.e., a stronger signaling value) of educational degrees.

There are two related objections to this empirical approach to measuring internal homogeneity. Both have to do with the fact that employers arguably observe more information about individuals than is used in constructing the measures of internal homogeneity. First, they observe more detailed levels of educational attainment than the relatively coarse three-category scheme used in the analysis. This suggests that internal homogeneity should likewise be measured at more detailed levels, especially since research based on PIAAC documents meaningful competence differentials by detailed educational attainment (Massing and Schneider 2017). Second, in addition to sex, age, and foreign-birth/foreign-language status, employers can presumably observe further characteristics that are related to skills and might therefore factor into their assessment of an individual's likely level of skills. It could be argued that these characteristics, too, should be taken into account before calculating the internal homogeneity of educational groups. On the other hand, it is difficult to verify if and at what point of the hiring process or employment relationship employers learn about different worker characteristics, so some uncertainty about the appropriate list of covariates is probably inevitable.

To address these concerns within the constraints of the data, I conducted two supplementary analyses, with the first using more finely grained educational categories and the second adjusting competence scores for a richer set of covariates. Results were reassuring (see the “[Robustness checks](#)” section for details).



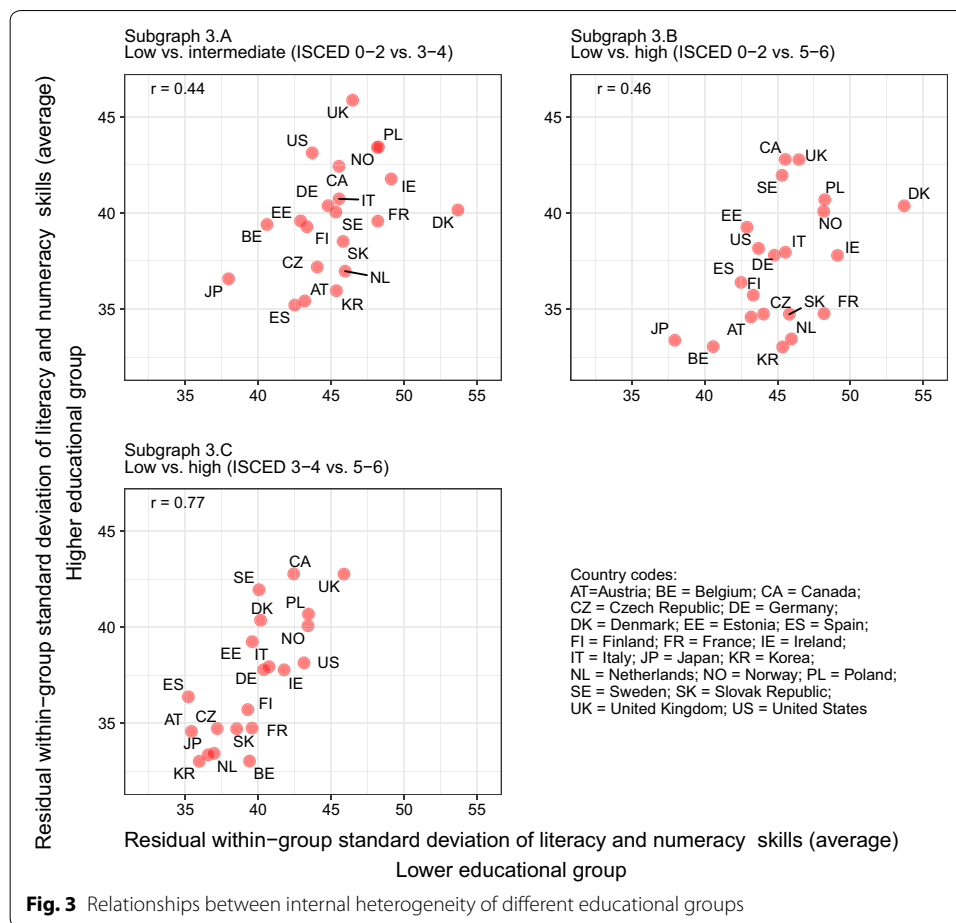
Country-level regressions

To investigate the relationship between internal homogeneity and education systems, I estimate country-level linear regressions with the homogeneity index as the dependent variable and the education system measures as the independent variables. The dependent variable in these regressions is estimated from the PIAAC data and therefore subject to sampling error. The magnitude of sampling error differs across countries (e.g., because of varying sample sizes), making the country-level error term heteroskedastic (Heisig et al. 2017; Lewis and Linzer 2005). I therefore obtain heteroskedasticity-consistent standard errors of the so-called HC3 type, which have been found to have good small-sample properties (Lewis and Linzer 2005; Long and Ervin 2000).

Results

Cross-national variation in the internal homogeneity of educational groups

Figure 2 shows country variation in the internal heterogeneity of the three educational groups. For each group, it plots the residual standard deviation of numeracy skills (y-axis) against the residual standard deviation of literacy skills (x-axis), after adjusting for sex, age, and foreign-birth/foreign-language status.



Three points are worth noting. First, the correlation across the two skill domains is high (r between 0.85 and 0.92) for all three educational groups. Countries that rank high in terms of the heterogeneity of literacy skills also tend to rank high in terms of the heterogeneity of numeracy skills. Second, cross-country differences in internal heterogeneity are substantial for all three groups. For example, the residual within-group standard deviation of numeracy skills among adults with low formal qualifications ranges from less than 42 points in Belgium and Japan to approximately 53 points in Denmark and Ireland (see subgraph 2.A). Thus, the middle 95% of Belgian and Japanese less-educated adults fall into a range that is approximately 40 points narrower than for the middle 95% of their Danish and Irish counterparts.⁶ This is a substantial difference that almost corresponds to the width of one of the four intermediate competence levels distinguished in PIAAC, which span a range of 50 points each (OECD 2013a, 76). A third result in Fig. 2 is that low-educated adults tend to be somewhat more heterogeneous than the higher-educated groups, although this may partly reflect smaller sample sizes (and hence larger sampling error) for the less educated.

In Fig. 3, I explore whether the extent of internal heterogeneity is systematically related across educational groups. That is, I investigate if countries where

⁶ This is because the middle 95% of a normally distributed variable cover a range of approximately 3.92 standard deviations.

low-educated adults are more heterogeneous also tend to have higher levels of heterogeneity among adults with intermediate and high levels of formal qualifications. Because of the high correlations found in Fig. 2, I do not differentiate between the two skill domains and simply use the average of the within-group standard deviations of literacy and numeracy skills. Three subgraphs show the country-level relationships for the pairwise combinations of the three educational groups: low vs. intermediate (Subgraph 3.A), low vs. high (Subgraph 3.B), and intermediate vs. high (Subgraph 3.C), with the internal heterogeneity of the lower educational group on the x-axis and the heterogeneity of the higher group on the y-axis.

Subgraph 3.C in Fig. 3 shows that there is a high country-level correlation ($r = 0.77$) between the within-group standard deviations of literacy and numeracy skills for adults with intermediate and high levels of education. Correlations are lower for the comparisons involving the less-educated group: the correlation with the intermediate-educated group (Subgraph 3.A) is 0.44 and the one with the high-educated group 0.46 (Subgraph 3.B). Nevertheless, the overall picture emerging from Fig. 3 is one of rather strong interrelatedness: Countries where the less educated are very heterogeneous also tend to be countries where the intermediate and high educated are very heterogeneous. The relatively high degree of similarity across educational groups might partly reflect the impact of contextual factors that affect the different groups in similar ways. This possibility will be further pursued in next section where I take a closer look at the role of education systems. Another possible explanation is that some countries have more heterogeneous populations than others, and that these differences in population heterogeneity translate into more heterogeneous educational groups. While the within-group standard deviations in Fig. 3 are calculated after adjusting for differences in sex, age, and foreign-birth/foreign-language status, there clearly are many other individual-level characteristics that might influence the variance of literacy and numeracy skills within a country's population. Potentially relevant factors include detailed adult training participation, immigration history, language proficiency, or childhood conditions. At least some of these factors are included in the more comprehensive set of covariates considered in the "[Robustness checks](#)" section below.

Given the strong interrelatedness of internal heterogeneity, both across skill domains (Fig. 2) and across educational groups (Fig. 3), I explore the possibility of constructing a simple summary measure of internal homogeneity. To this end, I run a principal factor analysis of the residual within-group standard deviations of literacy and numeracy skills (i.e., the factor analysis uses six items, the two residual standard deviations for each of the three educational groups). Table 5 in the [Appendix](#) displays detailed results, including the factor loadings (averaged across the ten plausible values). The factor analysis yields a well-defined first factor that loads positively on all six measures of internal heterogeneity. Loadings for the within-group standard deviations of adults with intermediate and high levels of qualification fall between 0.8 and 0.9. Loadings are somewhat lower for the less-educated group, albeit still quite high at 0.539 and 0.617 for literacy and numeracy skills, respectively. The first factor's eigenvalue is 3.646 and Cronbach's alpha is 0.89, indicating strong interrelatedness. As throughout the empirical analysis, all of these values are averaged over the

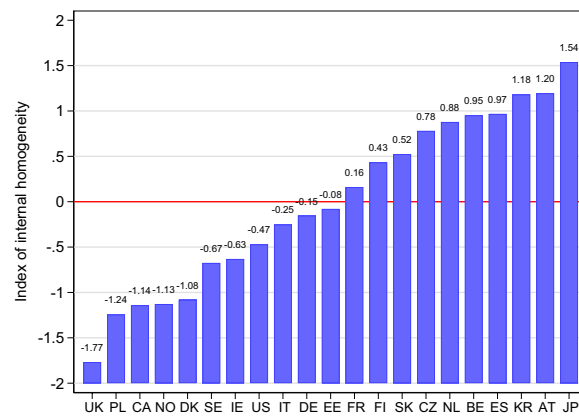


Fig. 4 Index of internal homogeneity. See Fig. 3 for country codes

Table 3 Country-level regressions of internal homogeneity on education system characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Between-school resource inequality	− 0.441* (0.156)							− 0.295 (0.193)
External differentiation (tracking)		0.527* (0.211)				0.692+ (0.356)		0.631+ (0.337)
Vocational orientation			0.229 (0.230)			− 0.242 (0.341)		− 0.234 (0.288)
Standardization of input				− 0.382 (0.264)			− 0.392 (0.307)	− 0.476* (0.205)
Standardization of output					− 0.293 (0.589)		− 0.353 (0.642)	− 0.181 (0.498)
Intercept	0.000 (0.229)	0.000 (0.222)	0.000 (0.253)	0.000 (0.245)	0.196 (0.442)	0.000 (0.235)	0.236 (0.501)	0.121 (0.390)
N	18	18	18	18	18	18	18	18
R ²	0.196	0.280	0.053	0.147	0.016	0.314	0.173	0.614
R ² (adjusted)	0.147	0.236	− 0.003	0.096	− 0.043	0.223	0.064	0.451

Heteroskedasticity-consistent (HC3) standard errors in parentheses.⁺ $p < 0.1$; * $p < 0.05$ (two-tailed tests)

ten plausible values. The second factor loads strongly positively on the within-group standard deviations for the less-educated group (loadings are 0.641 and 0.607 for literacy and numeracy, respectively) and weakly negatively on the within-group standard deviations of the two other groups (with loadings falling between -0.162 and -0.252). The eigenvalue of the second factor is 1.016.

The interpretation of the first factor is straightforward. Loading positively on all within-group standard deviations and displaying high internal consistency, it captures the empirical pattern displayed in Figs. 2 and 3: that some countries are characterized by much higher levels of internal heterogeneity for all educational groups than others. This factor thus corresponds very closely to the construct of internal homogeneity emphasized in the above discussion of skill transparency (internal homogeneity is simply the opposite of internal heterogeneity). The second factor is less well-defined. It essentially

seems to capture the fact that the less-educated group is not always perfectly aligned with the other two groups in terms of internal homogeneity. With the factor loadings for the less-educated and the higher-educated groups going in opposite directions, it can be thought of as capturing the internal heterogeneity of the former relative to the latter.

Given its close correspondence to the theoretical discussion and much better fit statistics, I will concentrate on the first factor in the remaining analysis. I obtain factor scores using regression scoring and multiply the resulting scores by -1 to construct the *index of internal homogeneity*, with higher values indicating that educational degrees send a stronger signal about actual skills. Figure 4 displays the values of the index of internal homogeneity for the 21 countries. Internal homogeneity is lowest in the United Kingdom, Poland, and Canada and highest in Korea, Austria, and Japan.

Internal homogeneity and secondary education systems

In Table 3, I turn to the relationship between internal homogeneity and education systems. The table reports the results of country-level regressions estimated by ordinary least squares. Statistical inference is based on conservative heteroskedasticity-consistent standard errors of the HC3 type (Lewis and Linzer 2005; Long and Ervin 2000). Table 3 is based on the 18 countries for which all five country-level predictors are available. Estonia, France, and Poland are excluded from the analysis because one or more of the country-level predictors are unavailable for them. Table 6 in the Appendix shows the same sequence of models using the maximally available country sample for each specification (i.e., using all countries for which the respective predictors are available). Findings are very similar to those in Table 3.

All explanatory variables except standardization of output, which ranges between zero (completely decentralized examinations) and one (completely centralized examinations; see Table 2 above) are transformed to have a mean of zero and a standard deviation of one in the 18-country sample (z -standardization).⁷ For these predictors, the coefficient estimates can be interpreted as the predicted change in the index of internal homogeneity that is associated with a standard deviation increase in the respective variable. For the standardization of output measure, the estimate is the predicted difference between a country with completely centralized and a country with completely decentralized examinations. I also transformed the index of internal homogeneity to have a mean of zero and a standard deviation of one in the 18 country sample,⁸ so the coefficient estimates for all predictors except standardization of output are in fact fully standardized effects.

Models 1 to 5 enter the five country-level predictors one at a time to explore the bivariate country-level relationships. Coefficient estimates are (marginally) statistically significant for two of the five predictors. In line with expectations, greater resource inequalities among schools appear to reduce the internal homogeneity of educational groups, that is, to render them internally more heterogeneous. According to Model 1, internal homogeneity decreases by more than two fifths of a standard deviation ($b = -0.441$; $p < 0.05$) for every standard deviation increase in resource inequality. Model 2 shows an even

⁷ I did not re-standardize the predictors for the regressions on larger country samples in Table 6, so coefficient estimates are directly comparable.

⁸ By construction, it has a mean of zero and a standard deviation of one in the full sample of 21 countries.

stronger effect for the extent of tracking. Again, the direction is consistent with expectations. A standard deviation increase in the index of external differentiation is associated with an increase in internal homogeneity by more than half a standard deviation ($b = 0.527$; $p < 0.05$). None of the other predictors show a clear bivariate relationship with internal homogeneity.

Model 6 simultaneously includes the indices of external differentiation and vocational orientation, two aspects of secondary education that are often studied in conjunction. In this specification, the coefficient of external differentiation is now only (marginally) significant at the 10% level ($b = 0.692$). It is worth emphasizing, however, that the loss of statistical significance compared to Model 2 is solely due to the lower precision of the coefficient estimate (the standard error increases from 0.211 to 0.356) and not to an attenuation of the effect size (which even increases noticeably). The loss of precision relative to the bivariate specification is due to the high correlation between the indices of tracking and vocational orientation noted above ($r = 0.681$). The coefficient of the vocational orientation index changes quite substantially from the bivariate specification (Model 3) to Model 6. In Model 3, it is positive but statistically insignificant ($b = 0.229$; $p > 0.1$). When the extent of tracking is controlled in Model 6, it changes sign and becomes negative ($b = -0.242$; $p > .1$). Taken together, these results provide relatively strong evidence for the expected positive relationship between the extent of tracking in secondary education and the internal homogeneity of educational groups. The vocational orientation of the education system shows no clear relationship with internal homogeneity.

Model 7 includes the two measures of standardization simultaneously. Coefficient estimates are relatively similar to the bivariate results in Models 4 and 5. For standardization of output in the form of centralized examinations, there is no evidence that it is related to the internal homogeneity of educational groups. Not only is the coefficient estimate quite imprecise and statistically insignificant, at somewhat more than a third of a standard deviation ($b = -0.353$) it is also very moderately sized—recall that the unit change represents the maximum effect (i.e., the difference between fully centralized and fully decentralized examinations) rather than the effect of a standard deviation change in this case. While not attaining statistical significance, the estimated effect of standardization of input is negative and more substantially sized at -0.382 and -0.392 in Models 4 and 7, respectively. The direction of the effect is contrary to expectations, however, as I speculated that a greater standardization of input (textbooks, school supplies, course content, types of courses) should increase the internal homogeneity of educational groups.

The final specification in Table 3, Model 8, includes all five predictors simultaneously. Given the limited degrees of freedom, this model must be viewed with caution. That said, the effect of external differentiation appears very robust, being of similar magnitude as in Models 2 and 6 and staying (marginally) statistically significant at the 10% level. The coefficient estimate of between-school resource inequality changes more substantially compared with the bivariate specification, declining from -0.441 in Model 1 to -0.295 in Model 8. It is also far from reaching statistical significance in Model 8. Overall, the findings on the role of between-school resource inequality are therefore ambiguous. To a considerable extent, the measure seems to pick up the effect of external differentiation/tracking in the bivariate specification (as noted above, there is a relatively

strong negative correlation between the two measures; see Table 4).⁹ At the same time, the expected negative coefficient remains of non-negligible size in Model 8. It becomes even stronger and statistically significant in sensitivity analyses that use a richer set of lower-level predictors (see the “[Robustness checks](#)” section below). Thus, there is at least some suggestive evidence for the expected negative relationship between school-level resource inequality and the homogeneity of educational groups. Finally, the unexpected negative coefficient on the standardization of input measure increases in absolute size compared to Models 4 and 7 and even becomes statistically significant at the 5% level ($b = 0 - .476$). This counterintuitive result proves robust in the supplementary analyses considered in the next section. It should be further investigated in future research.

I also estimated the same sequence of models as in Table 3 with the second factor from the factor analysis (i.e., the one capturing the homogeneity of the less-educated relative to the higher educated groups). The results are displayed in Table 7 in the [Appendix](#) and provide essentially no evidence for systematic relationships between the relative internal homogeneity of the less educated and the education system characteristics.

Robustness checks

I conducted a series of supplementary analyses to assess the robustness of the findings concerning country differences in the internal homogeneity of educational groups and their relationships with the education system characteristics.

In a first check I used a more fine-grained measure of highest educational degree. In particular, I used a five-category measure that is based on a six-category variable provided as part of the PIAAC public use files. The original variable differentiates among the following levels: below upper secondary (ISCED 0–2, 3C short), upper secondary (ISCED levels 3A, 3B, 3C long), post-secondary, but non-tertiary (ISCED 4A, 4B, 4C), professional tertiary (ISCED 5B), bachelor’s (ISCED 5A short), and research degree at the master’s level and above (e.g., PhD; ISCED 5A long/6). It is not possible to implement this level of disaggregation for all countries because some of the categories are very small in some countries. This applies to the ‘post-secondary, but non-tertiary’ category in particular.¹⁰ As in the main analysis, I therefore collapsed it with the upper secondary group to obtain a five-category measure. I restricted this supplementary analysis to the 12 countries with at least 40 observations for each of the five educational categories because estimating within-group standard deviations based on fewer cases would be dubious.¹¹ I ran a factor analysis similar to the one from the main analysis, but this time with 10 rather than 6 residual within-group standard deviations (one for each combination of the 5 educational groups and 2 skill domains). The average eigenvalue of the first factor across the 10 plausible values was 4.85 and the average value of Cronbach’s alpha was 0.87.¹² As in the main analysis, I constructed an index of internal homogeneity by

⁹ Indeed, the attenuation occurs already when the index of external differentiation is the only additional predictor in the model (results available upon request).

¹⁰ In addition, the United Kingdom has to be excluded from this supplementary analysis because its public use file groups all higher education graduates together.

¹¹ These countries are Canada, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Japan, Spain, Sweden, United States.

¹² Given the small number of only 12 cases (and the fact that the number of within-group standard deviations is only slightly lower at 10), this factor clearly has to be viewed with caution, but the consistency with the results in the main analysis is reassuring.

reverse coding the scores for the first factor. The country-level correlation between the homogeneity measure used in the main analysis and the one based on the five-category education measure is 0.98, suggesting that results for the main analysis would look similar if it were possible to use a more fine-grained education measure for all countries.

In a second check, I explored how results change when using a broader age range of 16–54 rather than 20–34 (excluding, as before, anyone enrolled in full-time education or with a foreign degree). In this analysis, I reverted to the three-category measure of educational attainment again to ensure full consistency with the main analysis (except with respect to the age restriction). Again, I repeated the factor analysis to construct an index of internal homogeneity for the larger sample. The average eigenvalue of the first factor across the ten plausible values was 4.06 and the average value of Cronbach's alpha 0.91. The correlation with the index used in the main analysis (based on respondents aged 20–34) was a reassuring 0.91. I also reestimated the country-level regressions of internal homogeneity on education system characteristics displayed above. The results, reported in Table 8 in the [Appendix](#), are very similar to the main analysis (cf. Table 3). Evidence for a positive relationship between external differentiation and internal homogeneity is somewhat weaker than in the main analysis, with the corresponding coefficient estimates no longer being significant at the 10% level in Models 6 and 8. The same holds for the unexpected negative relationship between internal homogeneity and standardization of input where the coefficient estimate in Model 8 is only significant at the ten (rather than the five) percent level when the broader age group is used.

In a third analysis, I explored the impact of adjusting for a richer set of individual-level characteristics before calculating the residual within-group standard deviations of literacy and numeracy scores that form the basis of the homogeneity measure. In addition to sex, age, and foreign-birth/foreign-language status, I included the following predictors: parental education (three categories: both parents below upper secondary degree; at least one parent attained upper secondary degree; at least one parent attained tertiary degree), participation in adult education and training during the 12 months before the interview (four indicators for participation in formal education for job-related reasons, in formal education for non-job-related reasons, in non-formal education for job-related reasons, and in non-formal education for non-job-related reasons), employment status (three categories: employed; unemployed; out of labor force), work experience (four categories: currently working; worked within last 12 months before interview; left paid work more than 12 months before the interview; no work experience), occupation in current or last job (if respondent worked within the last 5 years before the interview; ten groups based on the first digit of the International Standard Classification of Occupations 2008), field of study (nine groups; only for respondents with a tertiary degree or an upper secondary degree from a vocational program, interacted with whether highest degree is at the upper secondary or tertiary level), living with a partner (indicator variable), number of children (five categories: 0, 1, 2, 3, 4 or more).¹³ Residual standard deviations for constructing the homogeneity measure were computed within the same three educational groups used in the main analysis. However, I included all available

¹³ The sample size for this analysis is somewhat smaller than for the main analysis (28,412 rather than 30,646 cases) due to missing values on the additional predictors (primarily on parental education).

categories of the six-category measure used in the first robustness check described above in the country-specific regressions and additionally added a dummy indicating whether respondents with a degree at the upper secondary level obtained their degree in a vocationally oriented program.¹⁴

The principal factor analysis of the residual standard deviations within educational groups yielded an average eigenvalue of 3.81 in this sensitivity analysis. The average value of Cronbach's alpha was 0.89. The correlation of the resulting index of internal homogeneity with the one used in the main analysis (i.e., the one constructed after adjusting for a much smaller set of covariates) was 0.94. I also reran the country-level regressions on education system characteristics for the resulting index of internal homogeneity (see Table 9 in the Appendix). Results are generally similar to those from the main analysis. Evidence for a positive relationship between external differentiation and internal homogeneity is somewhat stronger, with the corresponding coefficient estimate now being statistically significant at the 5% level in all specifications (Models 2, 6, and 8). The negative relationship between internal homogeneity and standardization of input similarly persists and remains statistically significant at the five per cent level in Model 8. In addition, evidence for the expected negative relationship between internal homogeneity and between-school inequality of educational resources is clearer than in the main analysis: Effect sizes become somewhat stronger and the coefficient attains (marginal) significance at the 10% level in Model 8 (i.e., the model including all country-level predictors simultaneously).

In a final sensitivity analysis, I investigated the influence of individual country cases on the regression results by calculating *DFBETA* influence statistics for Model 8 in Table 3 (i.e., the model including all country-level predictors simultaneously). A widely used influence statistic, *DFBETA* measures the impact of a given case on a coefficient estimate as the difference in the full-sample estimate and the estimate when the case is excluded from the sample, expressed in terms of standard errors in the reduced sample (i.e., excluding the case in question).¹⁵ A value of -1 thus means that inclusion of the country shifts the coefficient estimate one standard error in the negative direction. Commonly used cutoff values for *DFBETA* are ± 1 and $\pm 2/\sqrt{n}$. Observations whose *DFBETA* value exceeds ± 1 shift the coefficient estimate in question by more than one standard error and must be considered highly influential by almost any standard. The $\pm 2/\sqrt{n}$ is more conservative and depends on the sample size. In the present case where $n = 18$ it is approximately ± 0.47 .

Reassuringly, Fig. 7 in the Appendix shows that none of the countries even comes close to reaching the ± 1 threshold for any of the coefficients. In a few cases, *DFBETA* exceeds the more conservative cutoff of ± 0.47 . Concerning external differentiation, which showed the most consistent relationship with internal homogeneity in Table 3,

¹⁴ As noted above (see note 10), not all six categories are available for all countries.

¹⁵ Formally, $DFBETA_{ij}$, the value of the statistic for the i th coefficient and the j th case, is defined as

$$\frac{\hat{\beta}_i - \hat{\beta}_{i(-j)}}{se(\hat{\beta}_{i(-j)})},$$

where $\hat{\beta}_i$ is the full sample estimate, $\hat{\beta}_{i(-j)}$ is the estimate with the j th case dropped, and $se(\hat{\beta}_{i(-j)})$ is the standard error of that estimate.

Fig. 7 indicates that both the UK and the German case exert a relatively strong influence on the results. Yet with *DFBETA* statistics of, respectively, 0.71 and -0.66 their effects work in opposite directions and should largely neutralize each other. For standardization of input, all *DFBETA* values fall within the ± 0.47 threshold, so the unexpected negative coefficient of this variable cannot be attributed to a single influential observation. With respect to between-school resource inequality, the Italian case pulls the coefficient estimate upwards, away from the hypothesized direction (*DFBETA* = 0.59). This underscores the overall interpretation that there is some, albeit no fully conclusive evidence, for the expected negative effect of between-school inequalities on internal homogeneity. For the remaining two predictors, vocational orientation and standardization of output, Fig. 7 indicates that the null results in the main analysis are not due to individual influential cases suppressing an association.¹⁶ Finally, it is worth noting that Korea—where the group of less-educated adults is very small with a population share of only about 3% (see Table 1 above)—generally has very little influence on the regression results.

Taken together, these robustness checks instill additional confidence in the conclusion that there are robust and substantial country differences in the internal homogeneity of educational groups. They further indicate that the country-level relationships between internal homogeneity and external differentiation as well as standardization of input are robust.

Skill transparency and labor market inequalities

The preceding analysis has shown that the internal homogeneity of educational groups differs considerably across countries and found some evidence that this might be related to structural characteristics of education systems. As discussed in the initial sections of the paper, these cross-national differences in internal homogeneity (and skills gaps) might help to account for country variation in the labor market attainment of different educational groups. This possibility is investigated in a closely related study by Heisig et al. (2016) that focuses on differences in occupational status between less- and intermediate-educated adults. Heisig et al. (2016) use a broader age restriction than the main analysis in the present paper, including all respondents between ages 16 and 54 in their analysis (as in the present study, respondents are excluded if they are enrolled in full-time education or obtained their highest degree in a foreign country). Occupational status is measured using the International Socio-Economic Index of Occupational Status (Ganzeboom et al. 1992), a standard measure of labor market attainment in the (comparative) sociological literature. Because Heisig et al.'s (2016) analysis focuses on differences between less- and intermediate-educated adults, the index of internal homogeneity used in the analysis is based only on standard deviations of literacy and numeracy skills within these two groups (ignoring those with high levels of qualification). However, the measure is highly correlated with the one based on all three educational groups (cf. Fig. 4 above).

¹⁶ There are cases with *DFBETA* values exceeding ± 0.47 , but omitting these from the analysis does not lead to the emergence of a clear association with internal homogeneity for either variable (results available upon request).

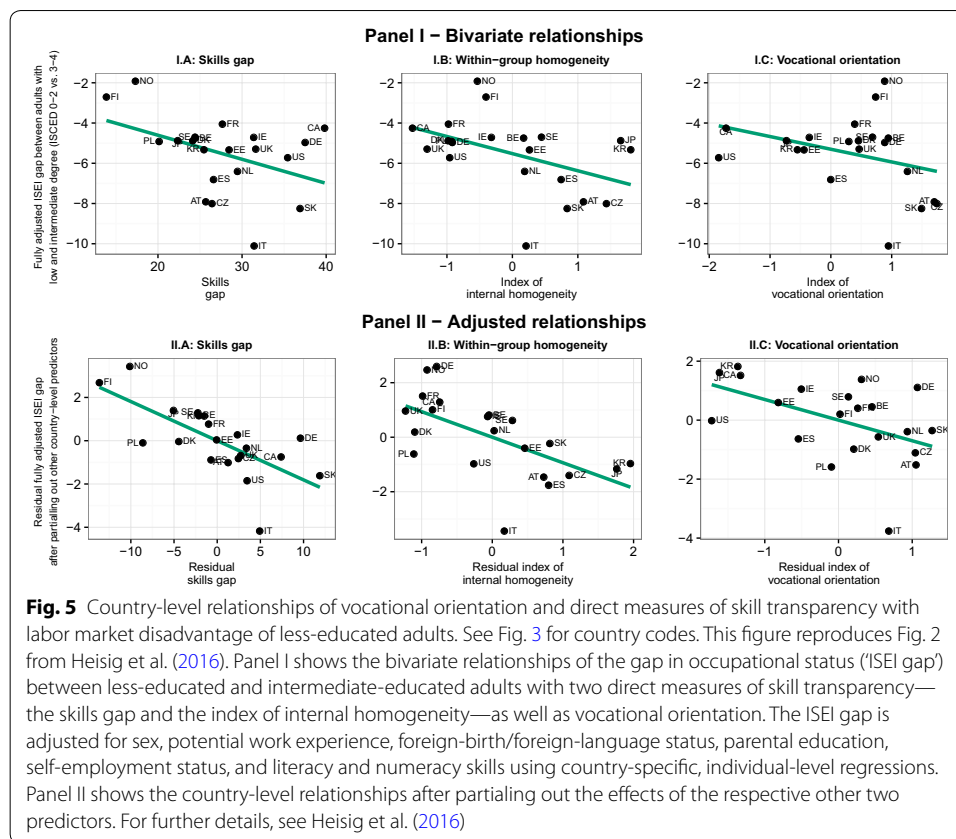


Figure 5 reproduces Fig. 2 from Heisig et al. (2016). Panel I shows the bivariate relationships. More specifically, Subgraphs I.A and I.B show the relationships for the two direct measures of skill transparency, the skills gap and the index of internal homogeneity. Subgraph I.C shows the relationship for vocational orientation, which previous research has found to be an important predictor of labor market inequalities by educational attainment (e.g., Bol and Van de Werfhorst 2011). Panel II shows the relationships between the ISEI gap and the three predictors after partialling out the effects of the respective other two country-level predictors (as in a conventional multiple regression). Note that the ISEI gap is adjusted using country-specific, individual-level regressions. Most importantly, these regressions include literacy and numeracy skill to account for direct effects of cognitive skills. In addition, they also control for sex, potential work experience, foreign-birth/foreign-language status, parental education, and self-employment status as additional predictors. For further details, see Heisig et al. (2016).

The most important result in Fig. 5 is that the two direct measures of the signaling value of educational degrees, the skills gap between less- and intermediate-educated adults and the index of internal homogeneity, are associated with the labor market disadvantage of less-educated adults. Again, it is important to note that this association holds after adjusting for the direct effects literacy and numeracy skills at the individual level. Effect sizes for the two direct measures of skill transparency are substantial, being larger than for vocational orientation, a well-established predictor of labor market inequalities by educational attainment. Formal country-level regressions reported in Heisig et al. (2016) confirm

these findings and show that the coefficient estimates on both skill transparency measures are statistically significant.

Conclusions

The relationship between formal qualifications and skills plays a crucial role in many prominent explanations of labor market stratification and of social inequality more broadly. By providing high-quality, internationally comparable data on the cognitive skills of working-age adults, PIAAC offers great potential for illuminating the complex interplay of formal qualifications and skills in shaping social inequalities.

The main goal of the present study was to extend the work of Heisig and Solga (2015) and Park and Kyei (2011) on cross-national variation in skills gaps by examining another crucial dimension of the qualification-skill nexus: the internal homogeneity of educational groups in terms of cognitive skills. I found that the internal homogeneity of educational groups (after accounting for sex, age, and foreign-birth/foreign-language status) varies considerably across the 21 advanced economies included in the analysis. Moreover, the extent of internal homogeneity is similar across skill domains (literacy and numeracy) and across educational groups (less-, intermediate-, and high-educated adults). Using factor analysis, I therefore constructed a one-dimensional summary measure, the index of internal homogeneity.

Country-level regressions relating the index of internal homogeneity to education system characteristics provide relatively strong evidence that the internal homogeneity of educational groups is higher in countries with stronger tracking in secondary education. This finding bolsters a crucial (but so far largely untested) assumption of previous cross-national studies on labor market returns to formal qualifications (Andersen and Van de Werfhorst 2010; Bol and Van de Werfhorst 2011): that stronger tracking increases the skill transparency (i.e., the signaling value) of educational degrees.

A likely explanation for the relationship between tracking and internal homogeneity is that the importance of teachers and other gatekeepers in allocating students to educational programs raises the salience of academic considerations. By contrast, other factors and in particular educational aspirations (which in turn are related to social background) should play a relatively larger role when systems are comprehensive or 'choice-driven' (Jackson et al. 2012). This interpretation squares well with other research on the consequences of teacher 'gatekeeping'. Intra-German research on the impact of binding teacher recommendations is a good example. While Germany is commonly considered a country with strong tracking, teacher recommendations for secondary school tracks carry greater force in some federal states than in others and this in turn moderates the strength of social background effects on educational transitions: Background effects tend to be weaker when teacher recommendations are binding than when they are non-binding (e.g., Neugebauer 2010).

There also was some evidence that internal homogeneity declines with the extent of standardization of input (e.g., course content, textbooks). This is counterintuitive, as one would expect such standardization to result in more homogeneous learning environments and eventually more homogeneous learning outcomes (i.e., skills distributions). This finding thus warrants further investigation. Some of the results further suggest that between-school inequality of educational resources (Park and Kyei 2011)

reduces the internal homogeneity of educational groups. Such a relationship is very plausible on a theoretical level, but the evidence is not as clear as it is for tracking and standardization of input (a substantial portion of the bivariate association seems to be due to differences in tracking, which is quite highly correlated).

No clear relationships with internal homogeneity were found for two other education system characteristics that have featured prominently in previous research: vocational orientation of upper secondary education (Bol and Van de Werfhorst 2016; Heisig and Solga 2015) and standardization of output through centralized exit examinations (Allmendinger 1989; Bol and Van de Werfhorst 2016).

To illustrate the potential of measuring the signaling value of educational degrees more directly, I have presented findings from a related study by Heisig et al. (2016). The study investigates cross-national differences in the labor market disadvantage of less-educated relative to intermediate-educated adults (in terms of occupational status). Consistent with the signaling explanation, it finds that the disadvantage of less-educated adults increases with the size of the skills gap and with the internal homogeneity of the two groups. This holds even after accounting for the direct, individual-level effects of literacy and numeracy skills as well as other key observables. These findings suggest that signaling (and human capital) theory can partly account for the existence of labor market inequalities by level of education as well as for cross-country variation in their magnitude. They do not, however, imply that the two approaches are sufficient for explaining labor market inequalities entirely. In particular, processes of rent seeking and occupational closure, which are emphasized by some versions of credentialism (see, e.g., Bol and Weeden 2014; Di Stasio 2017), may still play an important role in the generation of advantages for better-educated workers.

The present study inevitably has some limitations. A first one is that it was not possible to adjust literacy and numeracy skills for all characteristics that might be readily visible to employers (e.g., GPA, degree-conferring institution, or detailed immigration history) and that, due to limited sample sizes, internal homogeneity had to be defined at the level of aggregate educational groups rather than detailed categories. Robustness checks using additional individual-level covariates and more fine-grained educational categories produced reassuring results. Still, richer information on respondents' educational and employment biographies as well as larger sample sizes would clearly allow for a more rigorous analysis.

A second limitation is that the analysis did not account for country differences in skill development after formal education. For example, countries might differ in the opportunities for, and social inequalities in, further training participation and this might in turn influence the internal homogeneity of educational groups. PIAAC provides several measures of training participation in the last twelve months before the interview, which were used in one of the sensitivity analyses reported above. However, these variables inevitably provide a very incomplete picture of inequalities in adult training experiences.

The present analysis also suggests interesting questions for future research. One promising direction would be to more directly examine employer perceptions and employer decision-making, as some recent studies have done using field and survey experiments (e.g., Di Stasio 2017; Protsch and Solga 2015). The present analysis could

not show directly that skills gaps and the extent of internal homogeneity actually play a role in employer decision-making. The fact that these measures help to account for labor market inequalities similarly only indirectly supports this notion. It would be bolstered enormously if it could be shown that employers care about skills gaps and internal homogeneity and that they have reasonably accurate perceptions of the relationship between formal qualifications and skills.

Another fruitful direction might be to extend the approach proposed in this article—using large-scale assessment data to measure signal strength—to other readily observable worker characteristics. For example, it would be worthwhile to explore if direct measures of signal strength can also account for (cross-national differences in) labor market inequalities associated with immigration background or age.¹⁷ In any case, the analysis has shown that PIAAC, and large-scale assessments of adults more generally, have the potential to greatly advance our understanding of labor market stratification by making elusive concepts such as skills or the signaling value of educational degrees amenable to empirical analysis.

Authors' contributions

JPH developed the research question, conducted the analysis, and wrote the manuscript. The author read and approved the final manuscript.

Acknowledgements

This research was supported by a Grant from the German Federal Ministry of Education and Research (Grant Number PLI3061).

Competing interests

The author declares that he has no competing interests.

Appendix

See Tables 4, 5, 6, 7, 8, 9 and Figs. 6, 7

Table 4 Pairwise Correlations between country-level predictors

	Between-school inequality of instructional resources	External differentiation index (tracking)	Vocational orientation index	Standardization of input	Standardization of output
Between-school inequality of instructional resources	1				
External differ- entiation index (tracking)	− 0.477*	1			
Vocational orien- tation index	− 0.521*	0.681*	1		
Standardization of input	− 0.0923	0.183	0.233	1	
Standardization of output	0.140	− 0.0163	0.126	− 0.0671	1
Observations	18				

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

¹⁷ In a recent study using data on 15-year-olds from PISA, Spörlein and Schlueter (2018) compare the mean competence levels and the internal homogeneity of native-born adolescents with first- and second-generation immigrants. They do not investigate possible links to labor market inequalities, however.

Table 5 Results of factor analysis

	Factor loadings			Uniqueness
	Factor 1	Factor 2	Factor 3	
SD of literacy, low education	0.539	0.641	0.054	0.241
SD of literacy, interm. education	0.876	− 0.252	0.020	0.112
SD of literacy, high education	0.844	− 0.185	− 0.078	0.124
SD of numeracy, low education	0.617	0.607	0.075	0.202
SD of numeracy, interm. education	0.824	− 0.216	0.048	0.118
SD of numeracy, high education	0.882	− 0.162	− 0.069	0.124
Eigenvalue	3.646	1.016	0.368	

Factor analysis method is principal factor analysis. All values are averages across ten plausible values. SD = Standard deviation (as discussed in the text, this is the residual within-group standard deviation of literacy/numeracy skills after accounting for age, sex, and foreign-birth/foreign-language status)

interm. intermediate

Table 6 Country-level regressions of internal homogeneity on education system characteristics using all available country cases for each specification

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Between-school resource inequality	− 0.436* (0.155)							− 0.295 (0.193)
External differentiation (tracking)		0.523* (0.203)				0.675+ (0.342)		0.631+ (0.337)
Vocational orientation			0.231 (0.217)			− 0.224 (0.332)		− 0.234 (0.288)
Standardization of input				− 0.341 (0.258)			− 0.386 (0.304)	− 0.476* (0.205)
Standardization of output					− 0.297 (0.567)		− 0.374 (0.625)	− 0.181 (0.498)
Intercept	− 0.015 (0.216)	− 0.049 (0.209)	− 0.056 (0.222)	− 0.071 (0.221)	0.196 (0.441)	− 0.046 (0.220)	0.238 (0.499)	0.121 (0.390)
N	19	20	21	21	19	20	19	18
R ²	0.192	0.254	0.051	0.111	0.017	0.281	0.170	0.614
R ² (adjusted)	0.145	0.214	0.004	0.066	− 0.039	0.197	0.068	0.451

Heteroskedasticity-consistent (HC3) standard errors in parentheses

+ $p < 0.1$; * $p < 0.05$ (two-tailed tests)

Table 7 Country-level regressions of second factor (internal homogeneity of less-educated relative to higher-educated groups) on education system characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Between-school resource inequality	− 0.192 (0.354)							− 0.446 (0.543)
External differentiation (tracking)		− 0.214 (0.244)				− 0.016 (0.393)		− 0.160 (0.377)
Vocational orientation			− 0.302 (0.247)			− 0.291 (0.409)		− 0.402 (0.548)
Standardization of input				− 0.052 (0.319)			− 0.075 (0.309)	0.015 (0.301)
Standardization of output					− 0.750 (0.595)		− 0.762 (0.602)	− 0.495 (0.602)
Intercept	0.000 (0.265)	0.000 (0.255)	− 0.000 (0.246)	0.000 (0.267)	0.502 (0.425)	− 0.000 (0.260)	0.510 (0.442)	0.331 (0.434)
N	18	18	18	18	18	18	18	18
R ²	0.044	0.049	0.099	0.012	0.109	0.116	0.130	0.382
R ² (adjusted)	− 0.003	0.000	0.060	− 0.045	0.059	0.011	0.017	0.123

Heteroskedasticity-consistent (HC3) standard errors in parentheses

Table 8 Country-level regressions of internal homogeneity on education system characteristics; homogeneity measure based on respondents aged 16–54 rather than 20–34

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Between-school resource inequality	− 0.341 ⁺ (0.175)							− 0.137 (0.186)
External differentiation (tracking)		0.534* (0.227)				0.584 (0.373)		0.587 (0.387)
Vocational orientation			0.324 (0.248)			− 0.074 (0.372)		− 0.046 (0.309)
Standardization of input				− 0.365 (0.315)			− 0.362 (0.330)	− 0.470 ⁺ (0.221)
Standardization of output					0.145 (0.551)		0.089 (0.624)	0.152 (0.491)
Intercept	0.088 (0.239)	0.088 (0.222)	0.088 (0.248)	0.088 (0.252)	− 0.008 (0.415)	0.088 (0.240)	0.029 (0.485)	− 0.013 (0.371)
N	18	18	18	18	18	18	18	18
R ²	0.118	0.289	0.107	0.135	0.004	0.293	0.137	0.532
R ² (adjusted)	0.063	0.245	0.052	0.082	− 0.058	0.199	0.023	0.336

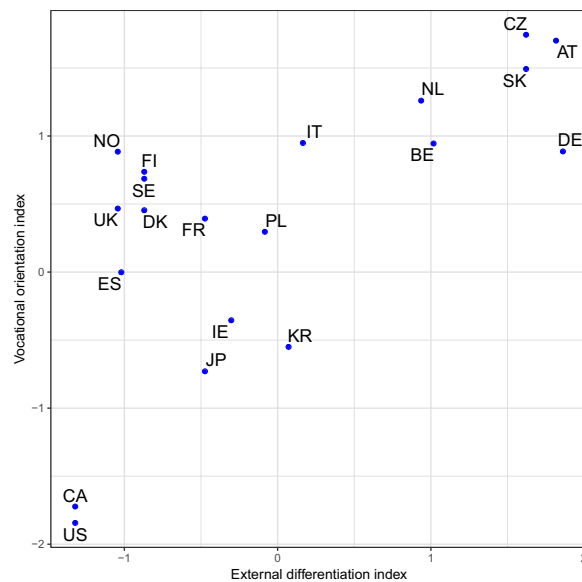
Heteroskedasticity-consistent (HC3) standard errors in parentheses

⁺ p < 0.1; * p < 0.05 (two-tailed tests)

Table 9 Country-level regressions of internal homogeneity on education system characteristics; within-group standard deviations adjusted for a richer set of covariates than in main analysis (see the “[Robustness checks](#)” section for details)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Between-school resource inequality	− 0.537* (0.163)							− 0.339+ (0.150)
External differentiation (tracking)		0.617* (0.172)				0.669* (0.274)		0.607* (0.225)
Vocational orientation			0.379+ (0.207)			− 0.077 (0.290)		− 0.106 (0.201)
Standardization of input				− 0.318 (0.259)			− 0.324 (0.290)	− 0.437* (0.149)
Standardization of output					− 0.157 (0.587)		− 0.206 (0.650)	− 0.062 (0.364)
Intercept	0.090 (0.192)	0.090 (0.178)	0.090 (0.217)	0.090 (0.229)	0.195 (0.478)	0.090 (0.190)	0.228 (0.545)	0.131 (0.280)
N	18	18	18	18	18	18	18	18
R ²	0.345	0.458	0.173	0.122	0.007	0.466	0.136	0.800
R ² (adjusted)	0.304	0.423	0.124	0.071	− 0.052	0.393	0.025	0.715

Heteroskedasticity-consistent (HC3) standard errors in parentheses

+ $p < 0.1$; * $p < 0.05$ (two-tailed tests)**Fig. 6** External differentiation and vocational orientation of upper secondary education systems

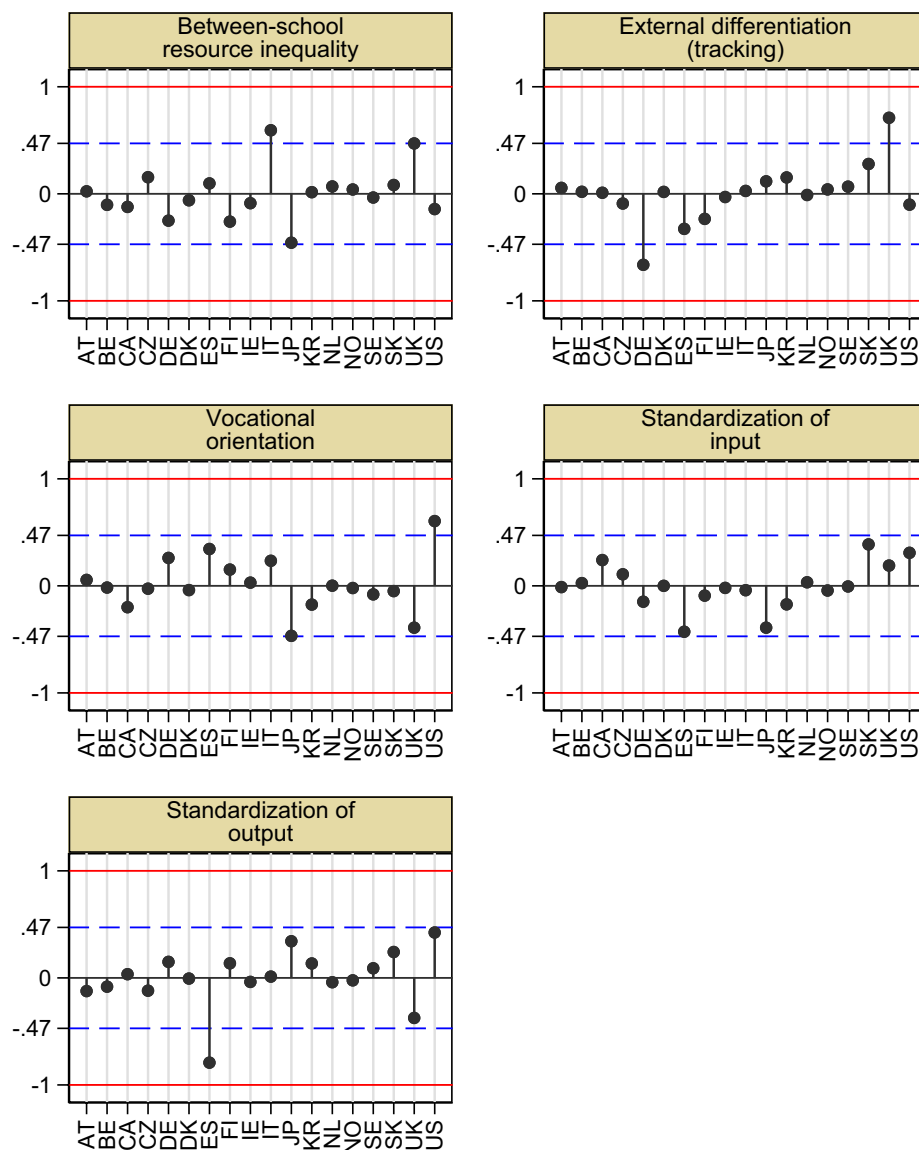


Fig. 7 DFBETA Influence statistics for Model 8 in Table 3

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 22 December 2016 Accepted: 23 August 2018

Published online: 04 September 2018

References

- Abrassart, A. (2013). Cognitive skills matter: The employment disadvantage of low-educated workers in comparative perspective. *European Sociological Review*, 29(4), 707–719. <https://doi.org/10.1093/esr/jcs049>.
- Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30(2), 175–187.
- Allmendinger, J. (1989). Educational systems and labor market outcomes. *European Sociological Review*, 5, 231.
- Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116, 313–350.

- Andersen, R., & Van de Werfhorst, H. G. (2010). Education and occupational status in 14 countries: The role of educational institutions and labour market coordination. *The British Journal of Sociology*, 61(2), 336–355.
- Arrow KJ. (1973, July). Higher education as a filter. *Journal of Public Economics* 23:193–216. Retrieved 1 March 2018, from <http://www.sciencedirect.com/science/article/pii/0047272773900133>. [https://doi.org/10.1016/0047-2727\(73\)90013-3](https://doi.org/10.1016/0047-2727(73)90013-3).
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5), 9–49. <https://doi.org/10.2307/1829103>.
- Berg, I. (1971). *Education and jobs: The great training robbery*. Boston: Beacon Press.
- Bills, D. B. (2003). Credentials, signals, and screens: Explaining the relationship between schooling and job assignment. *Review of Educational Research*, 73(4), 441–449.
- Bills, D. B. (2016). Congested credentials: The material and positional economies of schooling. *Research in Social Stratification and Mobility*, 43, 65–70.
- Bills, D. B., Di Stasio, V., & Gërkhani, K. (2017). The demand side of hiring: Employers in the labor market. *Annual Review of Sociology*, 43(1), 291–310. <https://doi.org/10.1146/annurev-soc-081715-074255>.
- 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. (2016). Measuring educational institutional diversity: tracking, vocational orientation and standardisation. In A. Hadjar & C. Gross (Eds.), *Education systems and inequalities. International comparisons* (pp. 73–94). Bristol: Policy Press.
- Bol, T., & Weeden, K. A. (2014). Occupational closure and wage inequality in Germany and the United Kingdom. *European Sociological Review*, 31(3), 354–369.
- Brown, D. K. (2001). The social sources of educational credentialism: Status cultures, labor markets, and organizations. *Sociology of Education*, 74, 19–34. <https://doi.org/10.2307/2673251>.
- Collins, R. (1979). *The credential society: A historical sociology of education and stratification*. New York: Academic Press.
- Di Stasio, V. (2017). Who is ahead in the labor queue? Institutions' and employers' perspective on overeducation, undereducation, and horizontal mismatches. *Sociology of Education*, 90(2), 109–126.
- Di Stasio, V., Bol, T., & Van de Werfhorst, H. G. (2016). What makes education positional? Institutions, overeducation and the competition for jobs. *Research in Social Stratification and Mobility*, 43, 53–63.
- Gamoran, A. (2000). Is ability grouping equitable? In R. Arum & I. Beattie (Eds.), *The structure of schooling* (pp. 235–240). New York: McGraw Hill.
- Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). Standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Green, A. D., & Pensiero, N. (2016). The effects of upper secondary education and training systems on skills inequality: A quasi-cohort analysis using PISA 2000 and the OECD survey of adult skills. *British Educational Research Journal*, 42, 756–779.
- Heisig, J. P., Gesthuizen, M., & Solga, H. (2016). Human capital or signaling? Differences in skills distributions and the labor market disadvantage of less-educated adults across 21 countries. Preprint available on SocArXiv (Dec 9, 2016). <https://osf.io/preprints/socarxiv/wc4s9/>.
- Heisig, J. P., Schaeffer, M., & Giesecke, J. (2017). The costs of simplicity: Why multilevel models 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. (2015). 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.
- Jackson, M., Jonsson, J. O., & Rudolph, F. (2012). Ethnic inequality in choice-driven education systems a longitudinal study of performance and choice in England and Sweden. *Sociology of Education*, 85(2), 158–178.
- 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.
- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217–224.
- Massing, N., Schneider, S. L. (2017, December). Degrees of competency: the relationship between educational qualifications and adult skills across countries. *Large-scale Assessments in Education*. 561–34.
- Mincer, J. (1970). The distribution of labor incomes: A survey with special reference to the human capital approach. *Journal of Economic Literature*, 8(1), 1–26.
- Neugebauer, M. (2010). Bildungsungleichheit und Grundschulempfehlung beim Übergang auf das Gymnasium: Eine Dekomposition primärer und sekundärer Herkunftseffekte. *Zeitschrift für Soziologie*, 39(3), 202–214.
- OECD. (2006). *Education at a Glance 2006*. Paris: OECD Publishing.
- OECD. (2013a). *Skills outlook 2013: First results from the Survey of Adult Skills*. Paris: OECD.
- OECD. (2013b). *Technical report of the Survey of Adult Skills (PIAAC)*. Paris: OECD Publishing.
- Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual review of sociology*, 34, 181–209.
- Park, H. (2007). South Korea: Educational expansion and inequality of opportunity for higher education. In Y. Shavit, R. Arum, & A. Gamoran (Eds.), *Stratification in higher education: A comparative study of stratification in higher education: A comparative study* (pp. 87–112). CA: Stanford University Press: Stanford.
- Park, H., & Kyei, P. (2011). Literacy gaps by educational attainment: A cross-national analysis. *Social Forces*, 89(3), 879–904.
- 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. (1998). *From school to work. A comparative study of educational qualifications and occupational destinations*. Oxford: Clarendon Press.
- Sørensen, A. B. (2000). Toward a sounder basis for class analysis. *American Journal of Sociology*, 105(6), 1523–1558.
- Soskice, D. (1994). Reconciling markets and institutions: The German apprenticeship system. In M. Lynch (Ed.), *Training and the private sector* (pp. 26–60). Chicago: University Of Chicago Press.
- Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87(3), 355–374.

- Spörlein, C., & Schlueter, E. (2018). How education systems shape cross-national ethnic inequality in math competence scores: Moving beyond mean differences. *PLoS ONE*, 13(3), 1–21.
- Stiglitz, J. E. (1975). The theory of 'Screening,' education, and the distribution of income. *The American Economic Review*, 65, 283–300.
- Thurow, L. C. (1979). A job competition model. In M. J. Piore (Ed.), *Unemployment and inflation: Institutional and structural views* (pp. 17–32). New York: M.E. Sharpe.
- Van de Kerckhove, W., Mohadjer, L., Krenzke, T. (2013). *Treatment of outcome-related non-response in an international literacy survey*. (Paper presented at the American Statistical Association Joint Statistical Meetings 2013 Montreal, Canada): Survey research methods section.
- Van de Werfhorst, H. G. (2011). Skill and education effects on earnings in 18 countries: The role of national educational institutions. *Social Science Research*, 40(4), 1078–1090.
- Weiss, A. (1995). Human capital vs. signalling explanations of wages. *The Journal of Economic Perspectives*, 9(4), 133–154.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
