Explaining wage differentials by field of study among higher education graduates: Evidence from a large-scale survey of adult skills

Thijs Bol, Jan Paul Heisig

University of Amsterdam, the Netherlands
WZB Berlin Social Science Center, Germany
Freie Universität Berlin, Germany

ABSTRACT

Why do some fields of study in higher education yield higher wage returns in the labor market than others? Human capital perspectives suggest that differences in skills are a major source of between-fields wage differentials. We assess this explanation using data from the Programme for the International Assessment of Adult Competencies (PIAAC). Our pooled analysis of 17,590 graduates from 29 countries indicates that differences in general cognitive (literacy and numeracy) skills matter relatively little, although numeracy skills do play a meaningful role in accounting for the high wages of STEM (science, technology, engineering, and mathematics) graduates. Specific skills, proxied by skill use on the job, explain a substantial portion of between-field wage differentials. Remarkably, we find that the sex composition of the field of study remains important after taking skills into account, particularly for explaining the wage advantage of STEM graduates. Comparative analyses grouping the 29 countries into four institutional clusters—Social-democratic, Conservative, Liberal, and Post-communist—show that these general patterns are broadly similar across different institutional contexts.

1. Introduction

Field of study is an important correlate of labor market outcomes among higher-education graduates (e.g., Kim et al., 2015; Krikkebøen et al., 2016), but the underlying mechanisms are not well-understood. Human capital explanations argue that differences in the distribution of marketable skills are the key explanation (e.g., Arcidiacono, 2004; Reimer et al., 2008; Van de Werfhorst, 2002). For example, Reimer et al. (2008) find that university graduates with a degree in the humanities tend to have higher unemployment risks than graduates from other fields while graduates from “health and welfare” fields have substantially lower occupational status than other graduates. They suspect that differences in the ability levels of students from different fields are an important driver of these differences in labor market attainment.

However, Reimer et al. (2008) cannot examine the role of skills empirically because their data—like most existing surveys of working-age adults—do not contain direct measures of skills. Only a handful of studies have been able to draw on such information.
Using Dutch data, Van De Werfhorst (2002) constructs measures of specific skills that are based on self-reports of the competencies emphasized in the respondent’s university or vocational program. In addition, the small but growing literature on returns to college major in economics includes some studies that incorporate direct measures of skills or cognitive ability (Altonji et al., 2012), but these studies tend to treat skills as a confounder to be controlled in attempting to estimate (causal) effects of field choice on labor market outcomes. The actual contribution of skill differentials to between-field inequalities in labor market attainment therefore remains largely unclear.

Against this background, the primary goal of this article is to investigate the relative importance of skills for between-field wage differentials. While there is a growing literature on field-of-study effects, previous studies do not tell us why some fields pay more than others and to what extent these explanations differ from field to field. For example, cognitive skills might be an important factor behind the relatively high earnings of STEM (i.e., science, technology, engineering, and mathematics) graduates, whereas the advantageous position of law or health graduates might be due to other processes such as social closure (Weeden, 2002).

While differences in general cognitive ability and specific skills have featured prominently in attempts to explain field-of-study effects, they are not the only factors emphasized in the literature. Sex composition has been highlighted as well. Female-dominated fields might have lower status than male-dominated fields and therefore yield lower returns (Ochsenfeld, 2014). An unanswered question is what the importance of both skills and sex composition is.

We use data from the Programme for the International Assessment of Adult Competencies (PIAAC) and apply decomposition methods on a field-by-field basis to investigate these issues. The unique feature of the PIAAC is that it contains high-quality measures of general cognitive skills—literacy and numeracy—based on administering test items to working-age adults. The PIAAC further collected rich information on skill use at work that we use to proxy job- and occupation-specific skills.

Another important strength of our data is their cross-national nature. We leverage this feature in two ways: In the first step of the analysis, we pool graduates from the 29 advanced economies included in our study, thereby minimizing the risk that that our findings reflect national idiosyncrasies, a limitation of previous studies which have mostly focused on single countries. In a second step, we investigate cross-national differences in an explorative way, addressing the shortage of comparative work on labor market inequalities by field of study. While limited sample sizes do not allow for a detailed country-by-country analysis, we can study whether the magnitude of between-field wage differentials and the relative importance of skills and other explanatory factors vary across country clusters—where the latter are defined on the basis of broad institutional similarities.

2. Wage differentials by field of study

Previous studies have documented large wage gaps between higher education graduates from different fields of study. Altonji et al. (2012) show that in the US the average wage difference between the best-paying and worst-paying college major is of a similar magnitude as the college wage premium (see also Kirkeboen et al., 2016). In a study covering 22 European countries, Reimer et al. (2008) find that especially individuals with a humanities degree are more likely to be unemployed.

Fig. 1 depicts wage differentials among the six broad fields of study that we distinguish in our empirical analysis. For each field, it shows the average deviation of log hourly wages from the country-specific mean for higher education graduates. On average, the hourly earnings of engineering graduates are about 9.9% (9.5 log points) higher than those of graduates from all six fields combined. Graduates from the humanities earn about 13.2% (14.2 log points) less than the average graduate. The total wage gap between the two extremes is 23.6 points on the log scale, corresponding to a differential of 26.7%.

2.1. The role of skills

It is often suggested that differences in the skills of graduates are a major source of the wage differentials documented in Fig. 1 (e.g., Altonji et al., 2012; Arcidiacono, 2004; Reimer et al., 2008; Van de Werfhorst, 2002). This perspective is rooted in human capital theory which views education as an investment in human capital that will increase the marginal productivity of workers, and thereby also their wages (Becker, 1962).

Fields of study might be linked to skills through (1) acquisition and (2) selection. Acquisition refers to the possibility that some fields might more effectively enhance the skills of students than others (Gerber and Cheung, 2008; Van de Werfhorst, 2002). Even if all students had identical skill endowments upon entering tertiary education, between-field differences in skills acquisition would lead to between-field differences in skills after graduation. Previous research, however, shows that graduates from different fields differ in their academic abilities already before higher education; that is, choice of field is selective (Arcidiacono 2004; Altonji et al., 2012).

A partly related issue is that skills are multidimensional. Many scholars argue that educational programs at the tertiary level tend to focus on the development of job- and occupation-specific skills (e.g., Van der Werfhorst, 2002). Engineering programs emphasize advanced technical and mathematical skills, whereas business programs might focus on accounting methods and relevant legislation.

1 In a subset of countries, the PIAAC additionally tested respondents in a third domain, “problem-solving in technology-rich environments”. We do not use this information because it is not available for all countries and only for respondents who completed the assessment on a computer (rather than on paper).

2 The PIAAC differentiates between nine broad fields of study. In this paper, we focus on the six largest fields. See Section “Data and methods” for further details.

3 The demeaning purges graduate wages of country differences in the overall level of remuneration.
General schooling at the primary and secondary level, by contrast, puts greater weight on the development of general skills, such as basic literacy and numeracy skills, which are useful in a wide variety of jobs. What role, then, do differences in general and specific skills actually play for the well-documented labor market inequalities between graduates from different fields?

An empirical assessment of this question has been hampered by the fact that skills are difficult to measure. A reliable measurement of general cognitive skills requires not only a clear theoretical and conceptual framework, but also a rather time-consuming and therefore costly assessment procedure that is not easily implemented in social surveys. For example, the assessment component of the PIAAC survey used in this article took respondents about 60 min to complete on average (OECD 2013b). Assessment of specifics skills is no less challenging, as the very specificity of these competencies makes measuring them a daunting task.

Some prior studies have been able to draw on high-quality measures of general cognitive abilities. Most prominently, studies in the United States have used high school GPA and SAT scores from the National Longitudinal Study of 1972 (NLS-72) to control for selective enrollment. These studies suggest that selective enrollment makes a noticeable contribution to between-field earnings differentials, even though its precise magnitude can usually not be inferred from the published results (e.g., Hamermesh and Donald, 2008). This is because this mostly economic literature is interested in identifying the causal effect of field of study on earnings and largely treats inequalities in skill before college enrollment as a nuisance to be controlled in a “conditioning on observables” strategy (Altonji et al., 2012).

These adjusted field-of-study effects reveal very little about the explanatory mechanisms: what explains the large wage differentials portrayed in Fig. 1? How big is the contribution of (general and specific) skills and does their importance vary from field to field? These are questions we seek to answer in this article.

2.2.1. General cognitive skills

Literacy and numeracy skills can be understood as general (rather than specific) cognitive skills. In the PIAAC, literacy and numeracy are measured as current skills levels (i.e., after completing tertiary education) and thus capture the combined effect of selective enrollment and between-field differences in skill acquisition during higher education.

Fig. 2 shows the average z-standardized literacy and numeracy scores for the six fields, again after removing country differences in...
the overall level of skills through demeaning. Fields are ordered by their average country-demeaned hourly wages. Graduates from the highest-paying field (“Engineering, manufacturing, and construction”) score high on numeracy but only slightly above average on literacy. Graduates from “Science, mathematics, and computing”, the second-highest-paying field, score highest in both domains. Humanities graduates score low on numeracy (although not lowest), but are among the higher-scoring graduates on literacy. Fig. 2 thus shows a clear, albeit not perfect, field-level relationship between average hourly wages and average cognitive skills. Consistent with previous research, the pattern is more pronounced for numeracy than for literacy skills (Hanushek et al., 2015; Heisig et al., 2019).

2.2.2. Job-specific skills

In addition to this between-field variation in general cognitive skills, there may be systematic differences in job-specific skills (e.g., Autor, 2013). Job-specific skills that are low in supply (and/or in high demand) will yield higher returns. Insofar as fields of study equip graduates with job-specific skills, this is a second pathway through which field choice might affect earnings.

With our cross-sectional data, it is difficult to flesh out when and where workers have acquired job-specific skills. It seems likely that graduates will partly have obtained their job-specific skills during their studies. Medical doctors know how to diagnose patients because they are trained in doing so, sociologists might work with statistical software that they were trained in during their studies. At the same time, it is well-known that many job-specific skills are obtained at the workplace (Thurow, 1975), and it is sometimes suggested that (higher) education may primarily serve as a signal of trainability rather than job-relevant skills (Bills, 2003; Spence, 1973; Thurow, 1975). Similar to most existing studies, we will not be able to disentangle where graduates acquired their job-specific skills, whether in education, on the job, or elsewhere. Nevertheless, the inclusion of job-specific skills in explaining wage differentials is important: either fields of study directly provide these skills, or they provide pathways (e.g., via credentialing) through which students can obtain these skills in the labor market (Bol et al., 2019).

2.2. Sex composition

Previous studies have found that female-dominated fields of study tend to be associated with lower wages than male-dominated fields (Kalmijn and Van der Lippe, 1997; Leuze and Strauß, 2014; Ochsenfeld, 2014; Smyth and Steinmetz, 2008). This points to sex composition as another potentially crucial explanation for wage differences between fields of study, above and beyond the role of general and job-specific skills.

There are several reasons why sex composition might contribute to wage differentials between fields of study. Wage discrimination against women is an obvious one: if women are paid lower wages than otherwise identical men, this will depress the average wage of graduates from more feminized fields. Drawing on the literature on occupational wage differentials, Ochsenfeld (2014) discusses three additional explanations: First, devaluation theory argues that the status associated with tasks and skills is lower when more women take part in them (England 2011). Second, human capital perspectives suggest that women anticipate childbearing and sort themselves into less risky and less-paying fields that emphasize general as opposed to non-portable firm-specific human capital (Alon and DiPrete, 2019). Third, gender role theory argues that social values and norms (e.g., of men as “breadwinners” and women as “carers”) affect women’s preferences and make them select fields that emphasize social and immaterial skills (e.g., creative and artistic skills) with limited labor market value, while eschewing fields that emphasize skills (e.g., mathematical and business skills) with substantial labor market value.

While it is not our primary goal to assess the relative importance of these explanations, two things are worth noting. First, some of the above claims may be impossible to test with labor market data alone. Whether stereotypically female skills are priced fairly relative to stereotypically male skills is a question that involves comparisons with standards of fair pay that cannot come from labor market data themselves. Second, the PIAAC allows us to account for an unusually rich set of specific skills measures, including job-specific skills that capture key “male” as well as “female” tasks. A sizable remaining effect of gender composition would make it difficult to maintain that wage inequalities between more and less “feminized” fields of study are primarily driven by women’s self-selection into fields emphasizing skills with low labor market value because our analysis controls for many of these skills, in contrast to most previous studies.

Table B1 in the online supplement shows that men are substantially overrepresented in STEM fields (i.e., “science, mathematics, and computing” as well as “engineering, manufacturing, and construction”). This overrepresentation of men in the two highest-paying fields is a first indication that sex composition might play an important role for understanding between-field wage differentials. At the same time it remains unclear whether composition matters once the high numeracy skills of STEM graduates and differences in on-the-job skill use are taken into account. Lacking such skill measures, previous studies on wage differences between male- and female-dominated fields of study have generally not been able to investigate this issue (e.g., Leuze and Strauss, 2014).

2.3. Cross-national variation

Our primary goal is to better understand the sources of between-field wage differences by looking at general patterns across a large

---

6 This suggests that, beyond mere sex composition, wage differentials might also be attributable to more complex interactions between sex, family status, and associated differences in care demands or career orientations. In supplementary analyses, we therefore explored the robustness of our main results to including such interactions (see Section 4.1.2 and Appendix C4 in the online supplement).
set of diverse countries. At the same time, it seems very likely that there are cross-national differences in the overall magnitude of between-field wage differences and in how important skills and other factors are for explaining them. Labor markets are organized very differently in different countries (Blau and Kahn, 1999). Higher education and “skill formation systems” more broadly are similarly variable (Arum et al., 2007; Busemeyer, 2009). It is also well-known that the organization of the welfare state is related to the magnitude of gender segregation (Charles and Grusky, 2005). Finally, wage inequality itself tends to differ across countries, and tends to be particularly high in countries with a liberal welfare state and a liberal market economy (Rueda and Pontusson, 2000).

All of these institutional-contextual differences might be important, and studying this in detail will advance our understanding of how the institutional context affects wage inequality between fields of study. At the same time, unraveling the precise contributions of specific institutional factors is beyond the scope of the current article— if only for the technical reason that country-specific sample sizes are not large enough for a country-by-country analysis.

In the first part of our empirical analysis, we will therefore pool graduates from all countries and look for general patterns, with a particular focus on the role of skills in understanding between-field wage differences. In a second step, we will then exploratively address the issue of cross-national variation: Specifically, we will investigate whether the ordering of fields with respect to average wages, the size of between-field wage differentials, and the relative importance of the different explanatory factors vary across clusters of countries characterized by similar labor market and welfare state institutions.

3. Data and methods

3.1. Data and sample

We use data from the first and second round of the PIAAC. Data collection took place in 2011/2012 for 24 first-round countries and in 2014/2015 for 9 second-round countries. We use 29 of the 33 countries, dropping Cyprus and Russia because of concerns about data quality and Australia and Indonesia because they do not provide public use files. Our sample includes the following countries: Austria, Sweden, Denmark, Norway, Finland, Iceland, Estonia, Latvia, Lithuania, and Luxembourg.

We use the June 28, 2016, versions of the public use files available at http://www.oecd.org/skills/piaac/publicdataandanalysis/. For the United States, we use the Combined 2012/2014 U.S. International PUF available under the same address. This data set includes additional cases from a second round of data collection. For Germany, we use version 1.1.0 of the so-called “Prime Age” data (Solga and Heisig, 2015) which include additional cases from an oversample of East German respondents. The sampling weights used in all analyses correct for this oversampling.
Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Ireland, Israel, Italy, Japan, Lithuania, Netherlands, New Zealand, Norway, Poland, Singapore, Slovak Republic, Slovenia, South Korea, Spain, Sweden, Turkey, United Kingdom, and the United States.

The target population of the PIAAC is the noninstitutionalized, working-age population (ages 16 to 65) with primary residence in the country at the time of data collection. We confine the analysis to respondents with a tertiary degree (levels five and six of the 1997 version of the International Standard Classification of Education; ISCED), dropping graduates from three minor fields (“general programmes”, “agriculture and veterinary medicine”, and “services”) that are much smaller than the others.\(^8\) In the main analysis, we do not include respondents with a degree in ISCED 5B, which mostly consists of short advanced vocational and professional programs with an applied focus (Schneider 2008).\(^9\) Robustness checks show that our main conclusions do not change when these cases are included (see Appendix C2 in the online supplement). We further restrict the analysis to “prime-age” wage and salary workers who were 25–54 years old at the time of interview (hourly wages are unavailable for self-employed respondents). We exclude respondents who obtained their highest degree in another country than where they were surveyed because the labor market value of foreign degrees can be very different from that of domestic ones. We also drop 290 respondents with missing information on at least one of the characteristics defining the sample.

Our primary approach to handling missing values is to use multiple imputation via chained equations (10 imputations), but we drop 105 cases with missing information on field of study, foreign-birth/foreign-language status, and computer use at work because doing so substantially simplifies the imputation procedure. The final sample comprises 17,590 higher education graduates. Appendix A in the online supplement provides further details on the treatment of missing data and the imputation procedure.

Our analyses account for unequal selection probabilities and response rates using the survey weights provided in the PIAAC, rescaled to sum to the unweighted sample size for each country. We ran sensitivity analyses to explore the impact of alternative weighting choices (see Appendix C1 in the online supplement).

### 3.2. Measures

Our goal is to better understand the marked between-field differences in log hourly wages shown in Fig. 1 above. In the main analysis, we use a relatively noisy measure of log hourly wages (the log of the median wage for a respondent’s wage decile) because the exact wage is not available for all countries. Appendix C3 in the online supplement shows that results based on exact wages are similar to the main analysis, despite the fact that we have to drop six countries where these are unavailable.

Higher education graduates are categorized into six broad fields of study: 1) teacher training and education science; 2) humanities, languages, and arts; 3) social science, business, and law; 4) science, mathematics, and computing; 5) engineering, manufacturing, and construction; 6) health and welfare. This is the level of detail provided in the PIAAC data, so more fine-grained differentiation is not possible. For readability, we will sometimes refer to the six broad fields using the first or first two subfields (e.g., use “science” or “science, mathematics” to refer to “science, mathematics, and computing”).

The PIAAC data are uniquely suited for our purposes because they provide cross-nationally comparable, high-quality measures of cognitive skills for working-age adults. All participating countries administered test items to assess the reading and text comprehension skills (literacy) and practical mathematical skills (numeracy) of participants (for details on the skill domains and overall assessment framework, see OECD, 2013a; 2013b). Each participant received only a relatively small number of test items to limit respondent burden, rendering individual competence estimates quite unreliable. The PIAAC provides ten plausible values rather than a single competence score to reflect this uncertainty. All results reported in the paper and online supplement were obtained by running the respective analysis ten times, once for each plausible value, and applying the appropriate rules to obtain final point estimates and statistical inference (Little and Rubin, 2002).

The PIAAC also provides detailed information on respondents’ skill use at work, which we use to proxy job-specific skills. This information was collected using a large number of questions about the frequency with which respondents perform various tasks at work, measured on a five-point scale ranging from “Never” or “None of the time” to “Every day” or “All of the time”. We use all 46 skill use measures available in the PIAAC, but following OECD (2013a: Chapter 4) we group them into 12 different skill use domains for a more concise presentation. These domains are (with the number of underlying items in parentheses): cooperative skills (1); problem-solving skills (1); physical strength (1); dexterity (1); influencing (7); planning/self-organizing (2); reading (8); writing (4); task discretion (4); ICT (8)\(^10\); learning (3); numeracy (6). The individual items grouped under these headings as well as their original PIAAC variable names can be found in Table B3 in the online supplement.

In addition to the skill measures, our decompositions include the sex of the respondent (dummy variable for being male) to assess the role of field differences in sex composition. We also include several potential confounders: parental education (three categories: no

---

\(^8\) The largest of the three fields is “agriculture and veterinary medicine” with 395 graduates (after application of all sample restrictions). This amounts to only 2.1% of graduates with valid information on field of study (unweighted). The smallest field included in our analysis, “humanities, languages, and arts”, comprises 2066 respondents (11.1%).

\(^9\) Schneider (2008) therefore suggests that many 5B programs would better fit into ISCED level 4 (non-tertiary post-secondary education).

\(^10\) The eight measures of ICT skill use include a dummy variable indicating whether the respondent reported any computer use at work. In the survey this variable was used as a filtering question, and only respondents who reported any computer use were then questioned concerning the frequency with which they perform seven ICT-related tasks (e.g., writing emails or using a word processor). We coded respondents who reported no computer use at work to the lowest category (“never”) on these measures of specific ICT tasks.
parent has completed upper secondary education; at least one parent has completed upper secondary education; at least one parent has completed tertiary education); age (five-year groups); foreign-birth/foreign-language status (four categories: born in survey country and test language is first language; born in foreign country and test language is not first language; born in foreign country and test language is first language; born in foreign country and test language is not first language); years of education (i.e., the number of years typically required to obtain the respondent’s highest degree); actual work experience (seven categories)\(^{14}\). We also include an indicator for part-time work (\(<=30\) h per week) to account for the possibility that individuals (and especially women) with greater non-work (care and family) demands may be overrepresented in certain fields. In supplementary analyses (see Appendix C4 in the online supplement), we also included partnership status and number of children as well as their interactions with sex. These predictors are not part of the main specifications because they complicate the interpretation of the role of sex composition (see next section).

All variables, that is, the hourly wage and all predictors, are country-demeaned to remove the effects of country differences in the overall levels of the variables.

3.3. Empirical strategy

We use the twofold version of the so-called Oaxaca-Blinder decomposition approach, originally proposed by Kitagawa (1955). The decompositions allow us to assess whether differences in explanatory variables (or “endowments”) can account for between-field differences in log hourly wages.\(^{12}\) The approach decomposes \(\Delta \hat{w} = Y_B - Y_A\), the observed difference in average log wages between two groups A (e.g., graduates from “humanities, languages, and arts”) and B (e.g., graduates from “science, mathematics, and computing”), as follows (for details, see Fortin et al., 2011, p. 36ff.):

\[
\Delta \hat{w} = \Delta \hat{w}^c + \Delta \hat{w}^e \tag{1}
\]

The term \((\bar{X}_B - \bar{X}_A)\hat{\beta}_A\) is the composition effect, often also referred to as the “endowment effect” or the “explained part”. The term \(\bar{X}_B (\hat{\beta}_B - \hat{\beta}_A) = \Delta \hat{w}^c\) is what Fortin et al. (2011) characterize as the “wage structure effect” and what others have labeled the “unexplained part” (e.g., Jann, 2008). Intuitively, the composition effect \(\Delta \hat{w}^c\)—our primary focus in this paper—is the portion of the overall difference attributable to group differences in the average levels of the explanatory variables (e.g., skills).

In Equation (1), the coefficient vector used in computing the explained part \(\hat{\beta}_A\), comes from a wage regression for one of the two groups being compared, arbitrarily labeled A. Other approaches to estimating the coefficient vector are common as well (see Fortin et al., 2011). When using a vector of reference coefficients that is not based on one of the two groups (call it \(\hat{\beta}^*\)) the unexplained part of the decomposition becomes slightly more complicated (see Fortin et al., 2011, p. 48):

\[
\Delta \hat{w} = (\bar{X}_B - \bar{X}_A)\hat{\beta}^* \hat{s}^c + \bar{X}_B (\hat{\beta}_B - \hat{\beta}^*) + \bar{X}_A (\hat{\beta}_A - \hat{\beta}^*) \hat{s}^e \tag{2}
\]

In addition to the total contribution of compositional differences to the observed wage difference, one can also quantify the contribution of individual variables or sets of variables, often referred to as a “detailed” decomposition. These detailed decompositions allow us to assess the relative importance of the various explanations for wage differentials among fields. The contribution of a particular variable is the difference in group means, weighted by the corresponding regression coefficient. For categorical variables with more than two levels (e.g., parental education), the contribution of an individual category (e.g., of having a parent with a tertiary-level degree) is sensitive to the (arbitrary) choice of the reference category (Fortin et al., 2011). We will therefore generally report the combined contribution of all categories, which does not depend on the reference category. These considerations also make clear that there is no straightforward way of separating the unique contributions of variables that are part of an interaction (e.g., the interactions between the graduate’s sex and family-related variables such as number of children mentioned in the previous section).\(^{14}\)

Detailed decompositions of the wage structure component face more serious complications. In contrast to the explained part, even the overall contribution of the J-I dummy variables representing a categorical predictor with J categories will be ambiguous and

\(^{11}\) The categories (in years of work experience) are: less than 5; 5–10; 10–15; 15–20; 20–30; 30–40; 40 or more.

\(^{12}\) We implemented the decompositions in Stata 15 using the oaxaca program by Jann (2008). Figures were constructed using the plotplain package by Bischof (2017).

\(^{13}\) We include demeaned indicators for field of study in addition to the demeaned values of the explanatory variables in these regressions, as it is generally recommended to include group dummies when estimating the reference coefficients using a pooled sample (Jann, 2008; Fortin et al., 2011).

\(^{14}\) Including the interactions of two categorical variables with \(k_1\) and \(k_2\) categories amounts to including dummy variables for the \(k_1k_2\) groups formed by the intersection of the two variables (leaving one group out as the reference category).
depend on the omitted category (Fortin et al., 2011; Jann, 2008). Some solutions to this problem have been proposed (e.g., Gardeazabal and Ugidos, 2004). Unfortunately, as emphasized by Fortin et al. (2011), even these approaches generally involve some arbitrary decisions that will affect the estimated contributions. We therefore avoid detailed decompositions of the unexplained part.

We focus on pairwise comparisons between graduates from “humanities, languages, and arts”—the lowest-earning group—and the other five fields of study. This focus on the pairwise comparisons involving humanities graduates is less restrictive than it may seem: If we know the explained part for the comparison between two groups A (e.g., humanities) and B (e.g., science), and we also know the explained part for the comparison between A and C (e.g., teaching), and we also know the explained part for the comparison between B and C, we can calculate the explained part for the comparison between A and C, \( \Delta \mu_{AC} = (X_C - X_B)\beta^* \), by simple subtraction because \( \Delta \mu_{(AC)} = \Delta \mu_{(AB)} = (X_C - X_A)\beta^* - (X_B - X_A)\beta^* = (X_C - X_B)\beta^* \).

In the first step of the analysis, we pool cases from all 29 countries to maximize power, minimize the impact of national idiosyncrasies, and detect general patterns. As noted above, the coefficient vector used in these decompositions comes from a regression pooling graduates from all six fields and all 29 countries. Standard errors are corrected for clustering at the country level. With 29 countries, the number of clusters in the pooled analysis somewhat smaller than is typically recommended in the literature (Cameron and Miller, 2015). This might result in slightly anti-conservative statistical inference (i.e., downward-biased standard errors), but simulation work fortunately suggests that inaccuracies should be limited (Cameron and Miller, 2015; Heisig et al., 2017).

In the second step of the analysis, we repeat the decomposition exercise separately for four groups of countries characterized by broadly similar institutional environments: a social-democratic, a conservative, a liberal, and a post-communist group (the countries included in these groups are listed below). This second part of the analysis is more explorative. We mainly want to explore how consistent the results from the pooled analysis play out in the different institutional setting, and whether in what ways specific groups of countries deviate from the broader patterns. With some country clusters comprising as few as five countries, concerns about the few-cluster performance of cluster-robust variance estimation loom large for the comparative part of the analysis, which is why we will mainly focus on point estimates and would like to emphasize that the few inferential quantities we report should be taken with a grain of salt. We also note that the vector of reference coefficients is estimated separately by country group for this part of the analysis.

Table B1 in the online appendix shows the field-specific averages of the country-demeaned explanatory variables. Table B2 reports the wage regression for estimating the reference coefficients for the pooled analysis.

4. Results

4.1. Pooled analysis of graduates from 29 countries

Table 1 presents the results of the decomposition for all countries combined. The table has five columns corresponding to the five pairwise comparisons between graduates from “humanities, languages, and arts” and each of the other fields. For each pairwise comparison, the three rows at the top show the observed log wage differential as well as its (aggregate) decomposition into an unexplained part and a part that is explained by compositional differences in the covariates. The remaining rows report the contributions of the different covariates to the explained part. The observed wage difference is always positive because “humanities, languages, and arts” is the lowest-paying field. The explained part and its individual components can be negative, however. Negative contributions of compositional differences indicate that the wage differential would be even larger in the absence of these differences.

As shown already in Fig. 1 above, the observed wage differential to graduates from the humanities is highest for graduates from “engineering, manufacturing, and construction” (23.6 log points), with “science, mathematics, and computing” (20.2 log points) a close second. The differences between humanities and the other fields are smaller: 16.5 log points for “social science, business, and law”, 12.9 log points for “health and welfare”, and only 5.0 log points for “teaching and education.” All differences are highly statistically significant, except for the difference to “teaching and education”.

The extent to which the covariates explain the wage gaps differs markedly across the five comparisons. For the two STEM fields, we are able to explain half of the gap: 11.8 log points (49.8%\(^{15}\) of the observed gap) for engineering and 10.3 log points (50.8%) for science and mathematics. For “teaching and education” we can account for 2.2 log point (44.2%). For the remaining two fields, the explanatory power is lower. Taken together, the covariates account for 5.3 log points (31.9%) of the observed wage gap between the humanities and “social science, business, and law”. For “health and welfare” we even find a substantial negative explained part (–5.0 log points).

We next turn to the individual contributions of the various explanatory factors. While the literature pays a lot of attention to general

---

15 Thus, if we find that compositional differences account for 5 log points of the wage gap between A and B and for 10 log points of the gap between A and C, we know that they explain 5 log points of the gap between B and C. This argument holds for the total composition effect as well as for individual components in the detailed decomposition.

16 Alternative approaches provide no easy solutions to these challenges. In particular, mixed-effects (multilevel) models with random intercepts and slopes appear much less attractive. Recent work by Heisig and Schaeffer (2019) indicates that accurate statistical inference for lower-level predictors requires these models to include the corresponding random slope term. Given the numerous lower-level predictors of interest in our study, this would require estimation of intractably complex models with large numbers of random slopes.

17 All percentages reported in the following are based on unrounded estimates and may therefore differ slightly from the percentages implied by the rounded estimates reported in the text and tables.
Table 1
Kitagawa-Oaxaca-Blinder decompositions of hourly wage gap between graduates from humanities, languages, and arts and five other fields.

<table>
<thead>
<tr>
<th></th>
<th>Engineering, manufacturing, construction</th>
<th>Health, welfare</th>
<th>Science, mathematics, computing</th>
<th>Social science, business, law</th>
<th>Teaching, education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed difference</td>
<td>0.236***</td>
<td>0.129***</td>
<td>0.202***</td>
<td>0.165***</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.029)</td>
<td>(0.021)</td>
<td>(0.018)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Unexplained</td>
<td>0.119***</td>
<td>0.179***</td>
<td>0.099***</td>
<td>0.112***</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.024)</td>
<td>(0.018)</td>
<td>(0.015)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Explained</td>
<td>0.118***</td>
<td>−0.050***</td>
<td>0.103***</td>
<td>0.053***</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.012)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual contributions to explained part (endowment/composition effects)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General skills</td>
</tr>
<tr>
<td>Numeracy skills</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Literacy skills</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Specific skills</td>
</tr>
<tr>
<td>Cooperative skills</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Problem-solving skills</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Physical strength</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Dexterity</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Influencing</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Planning/self-organizing</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Reading</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Task discretion</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>ICT use</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Learning</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Mathematics/numeracy use</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Writing</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total specific skills</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

| Demographic factors                                                     |
| Sex                                                                      | 0.044***       | −0.009***        | 0.028***                         | 0.015**                        | −0.008***         |
|                            | (0.004)                                  | (0.002)         | (0.003)                          | (0.003)                        | (0.002)           |
| Age                                                                     | 0.003          | −0.002           | −0.001                           | −0.002                         | 0.006*            |
|                            | (0.002)                                  | (0.002)         | (0.002)                          | (0.001)                        | (0.003)           |
| Parental education                                                     | −0.001*        | −0.002**         | 0.000                            | −0.001**                       | −0.003***         |
|                            | (0.000)                                  | (0.000)         | (0.000)                          | (0.000)                        | (0.001)           |
| Foreign-birth/foreign-language status                                   | 0.000          | 0.000            | 0.000                            | 0.000                          | 0.001             |
|                            | (0.000)                                  | (0.000)         | (0.000)                          | (0.000)                        | (0.000)           |
| Years of education                                                     | 0.000          | 0.000            | 0.007***                         | −0.001                        | −0.004*           |
|                            | (0.002)                                  | (0.002)         | (0.002)                          | (0.001)                        | (0.002)           |
| Work experience                                                         | 0.006          | 0.004            | −0.002                           | 0.003                          | 0.018***          |
|                            | (0.003)                                  | (0.004)         | (0.004)                          | (0.002)                        | (0.004)           |
| Part-time work (< 30 h/week)                                            | −0.026***      | −0.006           | −0.019***                        | −0.021**                       | 0.007*            |
|                            | (0.003)                                  | (0.003)         | (0.003)                          | (0.003)                        | (0.003)           |

Note: Results are based on 17,590 higher education graduates from 29 countries. See Section 3.1 for a list of the countries. Standard errors in parentheses. Standard errors are corrected for clustering at the country level. *p < 0.05, **p < 0.01, ***p < 0.001 (two-tailed tests). Sources: PIAAC, rounds 1 and 2, authors’ calculations.

Skills, our results do not show them to be particularly important. Only numeracy skills make a meaningful contribution, and only in the case of the STEM fields. In the case of engineering, numeracy skills account for 1.6 log points (or 6.8%) of the overall gap of 23.6 log points. In the case of science and mathematics, a slightly larger portion of the observed gap, 8.7% (or 1.8 out of 20.2 log points), is attributable to differences in numeracy skills. Literacy skills do not explain the wage gaps between the different fields of study. This is partly because between-field compositional differences are smaller for literacy than for numeracy skills (see Fig. 2 above). Much more
important, however, is the fact that the relationship between literacy skills and wages is practically zero when numeracy skills and on-the-job skill use are controlled (see Table B2 in the online supplement).

Turning to the specific skills used on the job, the comparatively high earnings of graduates from STEM fields are partly explained by workers from those fields doing more mathematical tasks, accounting for 2.4 and 2.3 log points of the wage advantage of engineering and science and mathematics graduates, respectively. We also find positive and statistically, albeit smaller, positive contributions of mathematics and numeracy use to the wage gaps between humanities graduates and graduates from the three remaining fields. Reading at work is a factor as well, explaining 1.6 points of the humanities-engineering and 1.2 points of the humanities-science gap (Table B3 in the online supplement shows that out of the eight different types of reading activities covered in the PIAAC, reading professional journals and especially reading diagrams, maps, and similar items are the most important factors here).

Some further noteworthy findings are that cooperative skills matter very little, while the use of problem-solving skills partly explains the disadvantage of humanities relative to “engineering, manufacturing, and construction” (0.6 log points), “science, mathematics, and computing” (0.4 log points), and “social science, business, and law” (0.5 log points). Interestingly, writing is also important when it comes to explaining the wage advantages of STEM graduates, accounting for 0.7 log points of the engineering-humanities and for 0.6 log points of the science-humanities gap. Influencing skills are particularly important for understanding why the wages of “teaching and education” graduates are higher than those of humanities graduates (2.4 log points). Planning, task discretion and learning have only very little importance—except for the humanities vs. “health and welfare” contrast where task discretion and learning make statistically significant contributions in the “wrong direction”.

A potential concern is that the results in Table 1 understate the importance of general literacy and numeracy skills because their contributions to between-field wage differentials are partly mediated by on-the-job skill use, our proxy for specific skills. To assess this issue, we repeated the decompositions without the skill-use measures (i.e., with only literacy, numeracy, and the demographic factors). The results, reported in Table B4 in the online supplement, are qualitatively similar to the main analysis: numeracy skills continue to play a far bigger role than literacy skills whose estimated contributions generally remain very small. Table B4 shows that numeracy skills continue to matter primarily when it comes to explaining the high wages of STEM graduates, now explaining 2.3 log points of the engineering-humanities and 2.5 log points of the science-humanities gap (rather than 1.6 and 1.8 log points in Table 1). This means that they account for somewhat less (engineering) or somewhat more (science) than a tenth of the observed wage advantage, so their explanatory power remains relatively modest even when we exclude the measures of specific skills.

The final rows of Table 1 show the contributions of the remaining variables included in the decomposition. We focus on the role of sex composition, which stands out as a major source of the wage differentials between graduates from the humanities and the STEM fields. It accounts for almost a fifth of the overall gap in the case of “engineering, manufacturing, and construction” (4.4 out of 23.6 log points, or 18.5%). For “science, mathematics, and computing”, the contribution is smaller, but still very substantial at 2.8 out of 20.2 log points (14.0%).

Sex thus remains a major factor even when cognitive skills and specific skills are accounted for—something that previous studies could not do due to data limitations (e.g., Leuze andStrauß, 2014). The persistence of substantial gender composition effects in the presence of detailed and high-quality measures of general and specific skills suggests that gender role theory alone cannot explain why more feminized fields of study tend to be associated with lower wages. If women’s self-section into fields that focus on skills with low labor market value were the primary driver of between-field wage differentials, sex composition should no longer play an important role once these skills are accounted for.19

4.1.1. Overall importance of the different factors

The detailed decompositions in Table 1 are informative, but it is difficult to infer the overall contributions of these factors to cross-field wage variation from these results. In Table 2 we therefore investigate to what extent adjusting for compositional differences with respect to (different subsets of) the skill measures, and sex reduces the overall between-field variation in (log) wages. We first computed the (between-field) variance of the unadjusted mean wage for each field, that is:

\[ \Sigma_{i=1}^{f} \left( \mu_i - \bar{\mu} \right)^2 / 5 \]

where \( f \) indexes each of the six fields of study that we consider, \( \mu_i \) denotes the mean (country-demeaned) log wage of graduates from field \( f \) and \( \bar{\mu} \) is the mean of the field-specific means (i.e., \( \Sigma \mu_i / 6 \)). The between-field variance thus calculated is .00822, corresponding to a standard deviation of .091 or 9.1 log points.

We then compare the between-field variance of the unadjusted mean wages to the between-field variance of field-specific means that are adjusted for differences in the focal explanatory variables. For example, we calculated, for each field, the counterfactual wage that graduates from that field would earn if their numeracy skill composition matched that of graduates from “humanities, languages,

---

18 According to Tables B1 and B3 in the online supplement, a major reason for this is that STEM graduates more often write reports.
19 An interpretation of our findings in relation to “devaluation” explanations is ambiguous. When devaluation occurs at the level of the six broad fields of study, it would be subsumed in the unexplained portions of the wage differentials. Under that assumption, devaluation would equally affect men and women in more feminized fields of study. However, devaluation most likely also occurs at the level of specific fields within the broader fields of study, or at the level of occupations. In that scenario it would be partially picked up by the coefficient on being male and therefore be part of the explained part.
and arts.” Using the skill composition of humanities graduates is arbitrary, but this choice does not affect the results because any change in the reference composition affects the counterfactual means for all fields in the same way, thereby leaving the between-field variance unaltered.20

Table 2 shows that adjusting for all skill measures (i.e., numeracy skills, literacy skills, and the “specific skills” measures of on-the-job skill use) reduces the variance of the field-specific mean wages from .00822 to .00384, that is, by 53.2% (see column 4). Columns 1 to 3 in Table 2 show the reductions achieved when adjusting the field-specific means for one skill dimension at a time. Note that in contrast to the decompositions of the gap reported in Table 1, the reductions achieved by the individual skill dimensions do not add up to the reduction achieved by all three dimensions combined because the calculations reported here do not have the additive linear structure of the above decompositions.

Consistent with the pairwise comparisons in Table 1 above, the results in Table 2 suggest that differences in literacy skills play no role for wage differences according to field of study. By contrast, differences in both numeracy skills and on-the-job skill use make a meaningful contribution. Adjusting the composition of graduates with respect to numeracy reduces the between-field variance of average wages by 16.3 percent. Table 2 underlines the importance of specific skills: adjusting for our set of on-the-job skill measures reduces the between-field variance by 46.6 percent. The final column of Table 2 again highlights the importance of gender composition. When we only take sex into account, we are able to explain about 32.6% of the between-field wage variance.

4.1.2. Supplementary analyses and robustness checks

We conducted several supplementary analyses to extend the pairwise decomposition analysis in Table 1 and to assess the robustness of these results. The full results and additional details can be found in the online supplement.

In a first set of analyses (Appendix B5), we explored the impact of using alternative reference coefficients. Our main results are based on the coefficients from a wage regression for the pooled sample comprising graduates from all fields. A common alternative would be to use the coefficients from a regression based on respondents from one field only (e.g., only humanities graduates or only graduates from the respective comparison group). Using a different set of reference coefficients will lead to different decomposition results, an issue that is also known as the “index number problem” (Oaxaca 1973).

Using alternative reference coefficients does not qualitatively change our results. The overall size of the explained and unexplained portions of the wage gaps between humanities graduates and graduates from the other fields remain similar and so do the contributions of the different sets of explanatory variables. That being said, there are a few noteworthy specific findings. We find that the male-female wage differential is considerably larger in engineering than among graduates as a whole (and also than among humanities graduates), a result worthy of further investigation. It also implies that the contribution of sex composition to the humanities-engineering differential is even larger than in the main analysis when the reference coefficients are based on engineering graduates only.21 As for the contributions of general and specific skills, results are similar to the main analysis. Irrespective of the sample used to estimate the reference coefficients, literacy skills do not seem to matter at all, numeracy skills account for a modest but meaningful portion of the wage advantage of STEM graduates, and specific skills generally make substantial contributions to the wage differentials between humanities graduates and the comparison fields (with the exception of the gap between humanities and “health and welfare” graduates). The exact magnitude of the specific skills contribution depends on the choice of reference coefficients to a noticeable extent, but findings remain qualitatively similar to the main analysis. The most noteworthy difference emerges for the humanities-health gap where the combined contribution of all specific skills starts to go in the “right” direction and becomes modestly positive when the reference coefficients are based on the “health and welfare” sample rather than all graduates combined.22

In a second extension of the main analysis, we conducted the decomposition separately for men and women. The gender-specific results (Appendix B6) are qualitatively consistent with the key findings from the main analysis. The ranking of fields in terms of average wages is similar, although the advantages of graduates from the high-earning fields appear to be larger among men. Crucially, our key insights concerning the importance of differences in general and specific skills also carry over to the gender-specific analyses, with some nuanced findings concerning gender differences in the role of individual skill use domains.

In addition to these extensions of the main analysis, we conducted four sets of robustness tests. Specifically, we examined the sensitivity of the main decomposition results in Table 1 to alternative ways of weighting the individual cases (Appendix C1), to including graduates from ISCED category 5B (short and applied programs) in the analysis (Appendix C2), to using the exact hourly wage instead of the medians of the wage decile (exact wages are not available for six countries in our sample; Appendix C3), and to including partnership status and number of children as well as their interactions with sex among the control variables (Appendix C4). None of these additional analyses challenged our main findings.

20 In calculating these adjustments, we used the same coefficients (i.e., the same “skill prices”) as in Table 1, namely those from a wage regression based on the pooled sample comprising the graduates from all six fields and including all explanatory variables considered in this paper (so the skill prices assumed are “net” prices controlled for the other covariates).
21 In the main results sex composition accounts for 4.4 log points (see Table 1). When the reference coefficients are based on engineering graduates the point estimate increases to 7.8 log points.
22 More specifically, the estimated combined contribution of all specific skills domains increases from −3.0 to +3.1 log points. Even so, the humanities-health comparison remains the one where we are least successful at explaining the wage gap, with all covariates explaining 1.2 log points out of an observed wage differential of 12.9 log points when the reference coefficients are based on “health and welfare” graduates rather than all graduates combined (in the latter case, the “explained” part is −5.0 log points, see Table 1 above).
Table 2
Proportion of between-field variance in log hourly wages explained by compositional differences in key covariates.

<table>
<thead>
<tr>
<th></th>
<th>(1) Unadjusted</th>
<th>(2) Numeracy only</th>
<th>(3) Literacy only</th>
<th>(4) Specific skills only</th>
<th>(5) All skill measures</th>
<th>(6) Sex only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>.00822</td>
<td>.00688</td>
<td>.00821</td>
<td>.00439</td>
<td>.00384</td>
<td>.00554</td>
</tr>
<tr>
<td>Proportion explained</td>
<td>–</td>
<td>.1629</td>
<td>.0007</td>
<td>.4656</td>
<td>.5324</td>
<td>.3259</td>
</tr>
</tbody>
</table>

Note: Results are based on 17,590 higher education graduates from 29 countries. See Section 3.1 for a list of the countries.
Sources: PIAAC, rounds 1 and 2, authors’ calculations.

4.2. Comparative results

In the analysis so far, all countries have been grouped together. In this section, we explore whether the answers to our research questions depend on the institutional context. Since the relatively small number of observations per country does not allow for country-by-country analyses, we will look at variation across clusters of countries that share key institutional similarities. We differentiate the following clusters: social democratic (Denmark, Finland, Netherlands, Norway, Sweden); conservative (Austria, Belgium, France, Germany, Greece, Italy, Spain, Turkey); liberal (Canada, Chile, Ireland, Israel, Japan, New Zealand, Singapore, South Korea, United Kingdom, United States); post-communist (Czech Republic, Estonia, Lithuania, Poland, Slovak Republic, Slovenia). While some of our classification decisions may be debatable, we believe that they accord relatively well with the distinctions drawn in the literatures on skill formation, labor market, and welfare state systems.

In Fig. 3 we show the results for the four country clusters separately. The figure consists of two panels. Panel A shows the observed wage gap between humanities graduates and graduates from the other five fields (markers). The bars show the portion of the wage gap that remains unexplained after accounting for all of the explanatory variables (see Table 1 above). The arrows accordingly visualize the explained part of the gap. Panel B then zeroes in on the contributions of the three focal sets of covariates to the explained part: general skills (solid markers), specific skills (light/grey markers), and sex (hollow markers). The full decomposition results can be found in Tables B7 to B10 in the online supplement.

Panel A shows that the observed between-field wage differentials vary quite substantially across countries. For all five comparison fields, the wage advantage over humanities graduates is largest in countries in the liberal cluster, in line with the comparatively high wage inequality in these countries (Rueda and Pontusson, 2000). A more surprising finding is that some of the between-field wage differences are also large in social democratic countries. The wage advantages of graduates from the two STEM fields and “social sciences, business, law” are second-largest in the social-democratic cluster. Given the explorative nature of this comparative analysis and the huge number of possible contrasts, we do not report inferential statistics in Fig. 3. We note, however, that the more substantial regime differences in the observed wage gap are generally statistically significant. For example, two-sided tests show the large humanities-engineering differential in the liberal cluster to be significantly different from the differential in the conservative (p < .01) clusters, albeit not from the differential in the social-democratic group (p = .21).

While there is stark variation in the overall magnitude of between-field wage inequality across institutional contexts, the extent to which our decompositions explain this inequality is less variable. Panel A shows that the combined contributions of all variables in explaining the difference between humanities and engineering or humanities and science are qualitatively similar. In particular, the findings resemble the pooled analysis in Table 1 in that we are most successful in explaining the differentials between humanities graduates on the one and STEM graduates on the other hand—and least successful (in fact generally not successful at all) in explaining the humanities-health differential. Despite this pattern of broad similarity, some cluster-specific results do stand out. The most striking one in Panel A comes from the post-communist cluster where our variables explain a much higher proportion of the observed gaps between humanities and STEM graduates and where we are particularly unsuccessful when it comes to explaining the humanities health gap.

In Panel B we start to unravel these results by showing the relative contributions of the three factors of primary interest: general skills, specific skills, and sex. Many results are both similar across country groups and to the pooled decomposition results in Table 1. This applies to the contributions of general (literacy and numeracy) and specific skills. In line with the pooled decomposition, general skills show a meaningful contribution primarily for the two humanities-STEM gaps, and even here their contribution remains modest, generally below 2.5 log points. Again resembling the pooled results, specific skills are far more important for explaining the wage advantages of STEM graduates and also of those from “social sciences, business, and law”. The findings on the role of sex composition are more complex and here it is again the post-communist cluster that stands out. Sex composition makes a meaningful contribution to the wage advantages of STEM vs. humanities graduates in all country groups, including the social-democratic, conservative, and liberal clusters where its role is comparable to that of general skills for the humanities-science gap and slightly larger than that of general skills for the humanities-engineering gap. In the post-communist

---

23 We note again (see also Section 3.3 above) that these p-values may be downward-biased because of the small number of countries per group and the associated problems for the standard cluster-robust variance estimation employed here. However, the illustrative p-values reported in the text indicate that this bias would have to be very substantial in order to completely overthrow the conclusion that at least the more sizable cross-regime differences displayed in Fig. 3 are statistically significant.
Fig. 3. Cross-national differences.

Note: Results are based on 17,590 higher education graduates from 29 countries. See text for a list of the countries and their grouping into the four country clusters (social-democratic, conservative, liberal, post-communist). In Panel A, markers indicate the observed (unadjusted) earnings differential between graduates from the respective field and graduates from humanities, languages, and arts. Bars indicate the unexplained part after accounting for differences in literacy/numeracy skills, on-the-job skill use, gender, and further covariates (see text and Table 1 above). Arrows indicate the explained part of the earnings differential. Panel B shows the estimated contributions of the different sets of covariates to the explained part of the gap, specifically the contributions of general skills (light filled markers), specific skills (dark filled markers), and sex (hollow markers). Full results, including the contributions of the predictors not shown here, are provided in Tables B7 to B10 in the online supplement.

Sources: PIAAC, rounds 1 and 2, authors’ calculations.
cluster, sex composition is dramatically more important, however. For the humanities-engineering gap, it is the single most important factor, its contribution exceeding the combined impact of the rich set of skill-use measures used to capture specific skills. The contribution of sex composition to the humanities-science gap is not quite as dramatic, but still substantially larger than in the other country groups. Clearly, this important role of sex composition for explaining the wage advantages of STEM graduates in the post-communist countries is also the main reason why the aggregate explained part of the gap in Panel A is particularly large for this cluster.

In sum, Fig. 3 highlights some important cross-national differences. The magnitude of between-field wage inequality varies starkly across country clusters, and in the post-communist country cluster sex plays a strikingly important role for the higher wages of STEM graduates. Future analyses should investigate the (institutional) origins of these differences in greater detail. Nevertheless, we would argue that the cross-cluster similarities in Fig. 3 are equally striking. Across very different groups of countries, our explanatory variables account for broadly similar portions of the various pairwise wage gaps. More importantly, we find qualitative support for our main findings on the role of skills: general skills (particularly numeracy) matter mostly for the advantages of STEM graduates and even here their contribution is modest. Differences in specific skills, as captured by detailed measures of on-the-job skill use, are more important for explaining wage differences across fields of study.

5. Conclusions

A substantial literature has shown that the labor market attainment of higher education graduates is related to their field of study (Altonji et al., 2012; Reimer et al., 2008). While several studies have speculated about the importance of skills for understanding between-field differentials in labor market outcomes, very few have been able to map the explanatory power of skills empirically. Some scholars have highlighted the role of other factors, particularly the sex composition of fields of study (Leuze and Strauss, 2014; Ochsenfeld, 2014). In this article, we have used data from a large cross-national survey of adult skills to provide novel and more direct evidence on the actual importance of these explanations.

Our first main finding is that the importance of skills depends on what skills one looks at. General cognitive skills seem to matter relatively little for understanding between-field wage gaps. Only for the STEM fields did we find a meaningful contribution of general cognitive skills. STEM graduates’ higher numeracy skills partly explain why their average wages are higher than those of graduates from the humanities, but the effects are not very large. Using only numeracy skills, we were able to explain about 16% of the between-field wage variance. Job-specific skills, measured by the tasks that workers perform on the job, are more important for understanding between-field wage differentials: almost 47% of the between-field wage variance are explained by the unusually rich set of specific task measures included in our analysis. We conclude that skills do matter, but more so the specific skills that workers can put to use in their jobs and to a lesser extent the general cognitive abilities that workers possess. A limitation of our study is that we were not able to disentangle where students acquired their (general and job-specific) skills—whether in school, higher education, at the workplace, or in yet other settings. An important direction for future research is to study these questions about skills acquisition in more detail using longitudinal data.

The influence of general cognitive skills might partly be mediated by on-the-job skill use. For example, a certain level of general numeracy skills is required to successfully complete numerical tasks on the job. However, supplementary analyses showed that our results remain qualitatively similar when the skill-use measures are excluded from the decompositions. The contributions of literacy skill differentials to the between-field wage gaps remain tiny. Numeracy skills, while becoming somewhat more important, continue to play a substantial role only for explaining the high wages of STEM graduates; but even for these fields they do not account for much more than one tenth (engineering) or one eighth (science) of the wage advantage over humanities graduates.

Another key finding from our analysis is that sex composition remains a major factor in explaining between-field wage differentials even after accounting for an unusually rich and high-quality set of skills measures. Sex composition is particularly important in accounting for the high wages of STEM graduates. These gender effects could partly be compositional: direct wage discrimination against women would lower the average wage for more feminized fields.

At the same time, processes might also operate on the field level. Previous research shows that occupational wages tend to drop when occupations feminize (Levanon et al., 2009). A similar argument can be made for fields: female-dominated fields might have lower status than male-dominated fields, and therefore yield lower returns (Ochsenfeld, 2014). In this article, we cannot disentangle these different explanations (gender composition and devaluation), but our findings do more than simply confirm previous results: The fact that sex composition continues to play such an important role even after incorporating rich and high-quality measures of general and specific skills is difficult to reconcile with the notion that these differentials simply reflect differences in (typically unobserved) human capital (Ochsenfeld, 2014; Tam, 1997).

The first part of our analysis pooled respondents from 29 countries to minimize the impact of national peculiarities and uncover the general forces that shape labor market inequalities among higher education graduates in advanced economies. An obvious downside of this approach is to implicitly ignore country differences in education systems, labor market structures, and welfare state arrangements—all factors that might influence the extent of labor market inequalities among graduates from different fields and the role of skills in particular.

In the second part of the analysis, we therefore explored cross-national variation in the magnitude and sources of between-field differentials differs across clusters. Differentials were largest in the liberal cluster. Perhaps surprisingly, they were also large in the social-democratic group, at least when it comes to the advantages of graduates from the STEM fields and from “social sciences, business, and law”. Inequalities by field of study were smaller in the post-communist and conservative clusters. While overall wage differentials thus vary across the clusters, the extent to which skills can account for them between-field wage differences was rather similar. Almost
all comparisons confirmed our main findings on the role of skills: It is primarily specific skills that matter. General (numeracy) skills are relevant for explaining the high wages of STEM graduates but do not appear to be a major factor in shaping wage differences among the remaining fields. One notable deviation from the pattern of similarity across country clusters is that sex composition appears to be exceptionally important for understanding between-field inequalities in post-communist countries. It is clear that these cross-national results are only a first step that needs to be followed up on in future work.

We conclude with two limitations of our study. A first limitation is the coarse grouping of fields of study in our data that forced us, for example, to pool social scientists with law graduates. Future work should extend our analysis by using more detailed fields of study. The second limitation concerns the content of our skill measures and the modeling of their effects. Future research should refine our analysis by incorporating further skill dimensions, including so-called “non-cognitive” skills (e.g., grit, motivation), and by exploring the possibility of “skill complementarity”, that is, the possibility that the effects of different types of skills interact and reinforce one another. Despite these limitations, the findings of our study help to paint a clearer and more nuanced picture of the forces shaping labor market inequalities by field of study in advanced economies.

Acknowledgments

This research has been supported by a Veni grant from the Netherlands’ Organization for Scientific Research (NWO) (451-15-001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://10.1016/j.ssresearch.2021.1025940.

References