An International Comparison of Earnings Inequality and Skill Distributions

Sabina Albrecht[1]

University of Amsterdam,
PO Box 15867, 1001 NJ Amsterdam, Netherlands

Abstract

This paper uses data about numeracy, literacy and computer skills from over 50,000 adults in 15 countries to analyze international differences in earnings inequality between skill groups. Estimating the canonical supply and demand model for skill, I find that a 10 percent decrease in the relative (net) supply of low-skilled workers increases their relative earnings by 1.5 percent. The model explains almost half of the international differences in earnings dispersion between low-skilled workers and others. This paper is the first to look at computer skills in that framework and to consider a complete labor market that includes women and unemployed people in the sample.

Keywords: Earnings inequality, skill-biased technological change, skill supply, skill demand, skill wage differential.
JEL: J23, J24, J31, O33.

Email address: `s.albrecht@uva.nl` (Sabina Albrecht)

[1]The author is grateful to Hessel Oosterbeek for his guidance and support throughout, to Erik Plug for valuable advice and suggestions, and to seminar participants at the University of Amsterdam, the Tinbergen Institute and the IWAEE 2015 for many helpful and encouraging comments. The author thankfully acknowledges financial support from the Dutch Organization for Scientific Research (*NWO*) for this project.

1. Introduction

Countries differ in their degree of earnings inequality between skill groups. In some countries, such as the United States and the United Kingdom, the wage gap between higher-skilled and lower-skilled workers is much larger than in other countries. This paper tests whether the observed variation stems from differences in supply of and demand for skill across countries. Using newly collected skill data from the OECD Program for the International Assessment of Adult Competencies (PIAAC) on literacy, numeracy and computer skill, I apply the canonical supply and demand model for skill to a cross-section of 15 OECD countries.

Why do we care about skill earnings inequality, and why are supply and demand plausible candidates for an explanation? Skill earnings inequality has a direct relation to overall earnings inequality, a topic that has recently received much attention in the public debate as well as in the academic literature (Autor, 2014). The commonly cited causes for increasing levels of earnings inequality have disparate effects on workers of lower or higher skill. The revolution in information and communication technologies (ICT), for example, has shifted labor demand in favor of those who possess complementary skills. Similarly, off-shoring is assumed to bear down on lower-skilled jobs. Both of these factors increase the pay gap between high- and low-skilled workers through a changed demand for skill. On the other hand, a generally higher-educated labor force affects the supply side of the market and works against increased wage dispersion (OECD, 2011). Factors like these vary in the extent to which they are at work in different countries –, so taken together, the supply of and demand for skill may constitute a useful tool for explaining differences in earnings inequality across countries.

The theory behind the canonical supply and demand model goes back to Tinbergen (1974, 1975) who coined the phrase 'race between technology and education', referring precisely to forces similar to the ICT revolution and an up-skilling of the labor force. While Tinbergen and much of the literature that followed suit were talking about the development of earnings inequality over time, Blau & Kahn (1996) and Leuven et al. (2004) applied his concept to differences in earnings inequality across countries. The two studies come to – at first glance – contradictory but plausible findings. Blau & Kahn (1996) reject the validity of a supply and demand model for skill in the international context and conclude that differences in institutions outweigh any influence of market forces on skill premia. Leuven et al. (2004), however, refute their argument by applying a measure of skill that is more suitable for comparisons across countries, obtained from one of PIAAC's predecessor studies. With data from the 1990s, they find that international differences in supply and demand for skill explain about a third of the differences in earnings inequality between countries. This provides a benchmark estimate and an important lesson for the current paper: an accurate, internationally comparable measurement of skill is crucial for an analysis of skill earnings inequality across countries. Almost twenty years later, such new data including even further developed measures of cognitive skill that encompass the ability to successfully use a computer are now available and the subject of earnings inequality is as topical as ever.

The first studies to make use of the PIAAC data for analyzing earnings and skill distributions are Hanushek et al. (2013), Paccagnella (2015) and Pena (2015). The data show that cognitive skills, as well as formal education, are rewarded differentially across OECD countries (Hanushek et al., 2013; Paccagnella, 2015).[2] Given these findings, a question that follows naturally is to what extent heterogeneous skill prices contribute to differences in earnings *inequality* across countries. To find an answer, Paccagnella (2015) and Pena (2015) decompose the gap in the 90/10-earnings differential between countries into an effect due to the different skill distributions in the countries' populations, and an effect due to different returns to skill.[3] Both studies conclude that purely compositional differences in skill are far less important than disparities in the wage structure or unobservable factors, which is consistent with previous work.[4] Based on these results both studies tentatively conjecture that unobserved institutional factors play a bigger role for international differences in earnings inequality than (supply of and demand for) skill. However, despite the congruence of results across multiple studies, the interpretation offered is not straightforward and often misguided. Skill supply and demand, as captured in the skill composition, and

[2] Hanushek et al. (2013) replace the traditional human capital measure of schooling by numeracy skill in a series of Mincer regressions. The estimated returns range from 12 to 28 percent for a one standard deviation increase in numeracy skill. Paccagnella (2015) complements the estimate of the average return with estimates at different quantiles of the earnings distribution and concludes that returns to human capital favor individuals at the upper end as compared to the lower end of the earnings distribution (i.e., the 90th percentile versus the 10th percentile).

[3] With a slightly different method, Pena (2015) additionally quantifies the contribution of unobservable factors, which in Paccagnella's method is included in the returns component.

[4] See Blau & Kahn (2005); Devroye & Freeman (2001); Fournier & Koske (2012) for earlier applications of the econometric decomposition techniques to similar contexts.

skill prices depend on each other (see Leuven et al., 2004, and even Paccagnella (2015) and Pena (2015) admit to this limitation of their approach). Because of this endogeneity, the decomposition techniques do not allow for a clean interpretation (or an unbiased estimate, for that matter) of the impact of market forces.

The analysis in this paper avoids the pitfall by taking the data to an economic model that *explicitly* relates the relative supply and demand for skill to earnings inequality between groups of different skill. More specifically, each country's population is split into three groups – low-, medium- and high-skilled – based on cutoff values from the skill distribution in a baseline country. Supply of skill is then determined by the number of people in a certain skill group, including unemployed and employed workers. Demand for skill is constructed as an index commonly applied in the literature that takes employment numbers in specific industries and occupations into account. According to the canonical model, net supply (i.e. supply minus demand) of a certain skill group relative to another group correlates negatively with the earnings dispersion between those groups. I estimate this intuitive and simple model on pairwise combinations of countries in a series of regressions for various skill measures.

This provides two contributions to the literature. First, in contrast to earlier studies that use the PIAAC data, the present economic model accounts for the interdependence of skill prices and skill supply and demand, achieving a genuine estimate of the relevance of net supply differences for skill earnings inequality across countries. Second, this paper updates and extends the work by Leuven et al. (2004). I revisit the previous findings relating to literacy and numeracy skill or to formal education with an expanded sample that now also includes women and unemployed people in the labor market. Additionally, I generate new results concerning the novel dimension of computer skill.

The main finding of this paper is that supply and demand for skill remain important for understanding earnings inequality between people of different skill in the 21st century as well. Overall, the supply and demand framework explains around 30 percent of the international variation between skill groups' earnings, in spite of the fact that country-specific labor market institutions may work against this mechanism. The results are particularly pronounced for workers at the bottom of the skill distribution: for them, supply and demand account for almost 50 percent of the international differences in their relative earnings. Based on a measure of broad cognitive skill, the canonical model estimates that a 10 percent decrease in the relative net supply of low-skilled people (i.e. from a 0.33- to a 0.30-share of the population) increases their relative earnings by 1.5 percent. These results persist qualitatively under all measures of cognitive skill, such as literacy and numeracy or computer skill alone. The supply and demand model, however, shows no such correlation based on years of schooling and experience as skill measure as in Blau & Kahn (1996). This confirms the results of Leuven et al. (2004) and reinforces the argument that the indirect variables associated with skill are imperfect at best in describing the true skill level of an individual when it comes to an international comparison.

The remainder of this paper is structured as follows. Section 2 presents the PIAAC data and points out their unique features. Section 3 discusses the supply and demand model of skill and translates the model into an empirical relationship, and Section 4 analyzes the results. Section 5 offers concluding remarks.

2. Data

Description of the PIAAC Survey of Adult Skills

The data for the skill supply and demand analysis come from the recently conducted Survey of Adult Skills as part of the OECD Program for the International Assessment of Adult Competencies (PIAAC). The PIAAC Survey makes internationally comparable skill data, earnings and a variety of background variables accessible for 24 countries, which makes the data set highly suitable for comparative labor market research. It succeeds and extends the International Adult Literacy Survey (IALS, 1994 – 1998) and the Adult Literacy and Lifeskills Survey (ALL, 2003 – 2006), the former having been an important data source for the cited literature (e.g. Blau & Kahn, 2005; Leuven et al., 2004).

Due to some data limitations the analysis concentrates on samples from 15 countries. These countries are Belgium (Flanders), the Czech Republic, Denmark, Estonia, Finland, Germany, Ireland, Japan, Korea, the Netherlands, Norway, Poland, the Slovak Republic, the United Kingdom (England and Northern Ireland) and the United States.[5] The data collection extended over the period from August 2011 to April 2012 for these countries and took 230 days on aver-

[5] Precluded from the analysis are Australia and Cyprus because public-use files were not available at the time of writing; Austria, Canada and Sweden because of missing data on earnings; France, Italy and Spain because these countries have chosen not to participate in the computer skill assessment; and the Russian Federation because of a considerably smaller sample size than other countries.

age. The background information was collected in a computer-aided interview in the country's official language(s) that took about 30 to 45 minutes. In a second part of the interview the cognitive test was taken with no restriction on time.[6] The response rate was sufficiently high as judged by PIAAC's Technical Standards and Guidelines. All countries offered a modest (monetary or non-monetary) incentive to respondents in order to help reduce non-response bias.[7] The target population of the study comprised non-institutionalized adult residents between the age of 16 and 65, regardless of their citizenship, nationality or language. While some countries deviated from the PIAAC standard sampling design for the purpose of further national use of the data, strict quality controls by the PIAAC Consortium assured that the final probability-based samples were representative of the target population (OECD, 2013d, ch.10, 14).

This paper restricts the sample to those active in the labor market. The sample includes 18 to 65 year old women and men who are either employed as wage and salary workers or are unemployed. Since the analysis is based on available background information about education, work experience, earnings, occupation and industry affiliation, respondents with missing observations in any of these variables are dropped from the sample. Overall, the per-country sample sizes range from 2597 (United States) to 4753 (Denmark) observations.

Definition of skill measures

The uniqueness of the PIAAC survey lies in assessing cognitive skills of the participants in three dimensions – literacy, numeracy and computer skill –, which contain the exact same information for every country. To give an impression of what these skill dimensions measure, I quote the definitions from the PIAAC Technical Report (OECD, 2013d):

Literacy (including reading components): understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential;

[6]The average time taken to complete the cognitive test was 50 minutes (see key fact sheet made available by the OECD (2013a)).

[7] The response rate was 50% or higher and without significant non-response bias for all countries. This was assessed in several Non-Response Bias Analyses (basic, extended or item-related), in which all countries were (at least) required to compare response rates for different subgroups and to compare the distribution of auxiliary variables (correlated with proficiency) for respondents and nonrespondents. For more detailed information see OECD (2013d), ch.16.3.

Numeracy: 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;

Computer skill: [8] using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks; more specifically, the ability to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks.

The domains literacy and numeracy are to a large extent based on the measurements in the previous OECD skill studies, whereas the computer skill domain introduces a new, modern dimension to skill measures.[9] The potential scale for all three measures ranges from 0 to 500 points. For the purpose of the supply and demand analysis I rely in turn on one skill measure comprising both literacy and numeracy (due to high correlations between the two, as outlined below), which is called S_{litnum}; on the computer skill measure alone, $S_{computer}$; and on an all-encompassing measure of cognitive skill, S_{PIAAC}.[10]

Internationally less comparable but more often available and used in the relevant literature are the measures of formal education, which translates acquired levels of education in ISCED classification into years of schooling, and work experience. Blau & Kahn (1996) and Leuven et al. (2004) combine the two in a weighted average, with weights stemming from a regression of wages on years of schooling, a second-order polynomial in work experience and country dummies. In order to assure comparability of my results with the previous literature and to assess the information content of this classical measure of human capital,

[8] Formally, this dimension is called ‘Problem solving in technology-rich environments’ (PSTRE). For reasons of clarity and brevity this paper uses ‘computer skill’ in lieu thereof.

[9] Computer skill is measured only for those survey participants with at least some experience in using a computer. Missing values are imputed to retain a representative sample; Appendix A provides further details and a critical discussion.

[10] All PIAAC-based measures are averages of the plausible values of the respective skill measure(s) in the data set, scaled by the factor $\frac{1}{100}$. To give you an example, S_{litnum} is constructed as the average of PVLIT1 to PVLIT10 and PVNUM1 to PVNUM10, divided by 100.

results based on the Blau-Kahn measure, S_{BK}, are reported as well.[11] The Blau-Kahn measure is an indirect measure of skill as opposed to cognitive skill. Since my analysis compares skill across countries, the PIAAC-based measures may be superior to an indirect measure, given the large variety of school education and post-school training systems across the world. On the other hand, cognitive skills are not as easily observed by employers as diplomas and training certificates, with the consequence that earnings might not accurately reflect differences in those skills.

2.1. Some descriptive statistics

The supply and demand analysis of skill depends on observed differences in skill supply, relative earnings and the employment sector across countries. Looking at the descriptive statistics of certain key variables, some dissimilarities between countries become apparent. Table B.1 in the appendix shows averages and standard deviations by country.

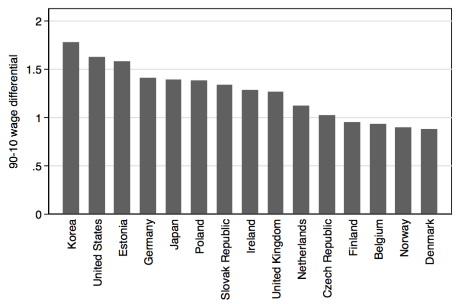
Differences are most notable in the earnings measure. Earnings denote gross hourly earnings of wage and salary workers and are PPP-corrected for \$US.[12] They range from a low of \$US 9 to 10 in the Eastern European countries and Estonia to a high of \$US 23 to 24 in Norway, Denmark and the United States. Similarly, their spread varies considerably across countries. The divergence in the second moment carries over to other aspects of the earnings distributions. Figure 1 shows common measures of overall earnings inequality. The log wage differential between the 90th and 10th percentile as well as the split into the two

[11] The obtained weights from the worldwide regression are 0.092 for education, 0.372 for experience, -0.058 for the square of experience and 1.277 for the constant. Note that as in Blau & Kahn (1996) experience is scaled by a factor of $\frac{1}{10}$.

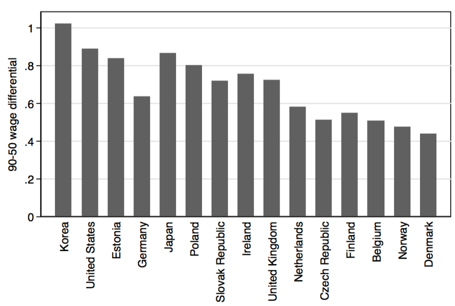
[12] The data was collected from a set of questions that allowed respondents to choose the time interval for which they report their earnings. The PIAAC team combined all pieces of information into an hours-corrected earnings measure and performed quality checks by looking at the individual earnings distributions in the countries (OECD, 2013d).

halves of the distribution are comparable to statistics reported by the OECD (2013b) and display Korea as the country with highest earnings inequality, regardless of the measure. Likewise, the Nordic countries and Belgium always form the group with lowest earnings inequality. Germany stands out as having a higher degree of inequality in the lower half of the distribution, whereas inequality is centered in the top half of the distribution for Japan.[13]

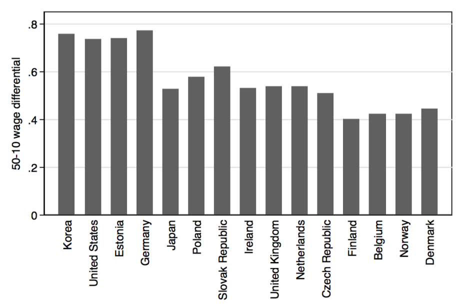
Figure 1: Pattern of log earnings inequality



(a) 90-10 log wage differential



(b) 90-50 log wage differential



(c) 50-10 log wage differential

Countries differ also in their mean achievements in terms of cognitive skills and years of education and work experience. On average the survey respondents spent around 13 to 14 years in school and have acquired 18 to 19 years of work experience, with some exceptions. When translated into the Blau-Kahn measure

[13] For a detailed descriptive analysis of both earnings distributions and skill distributions see Paccagnella (2015).

of human capital, disparities with respect to the cognitive skill measures become obvious. Countries that score high on average in the classical measure based on schooling and experience are not necessarily those that do well on the cognitive scores.[14]

Within countries, the associations between the dimensions of cognitive skill are strong, as Table 1 reveals. For all countries literacy and numeracy are correlated with a coefficient around 0.9, which is the reason for combining the two dimensions into one measure. Computer skill is clearly positively correlated with literacy and numeracy, but to a lesser extent. This justifies the use of the pooled PIAAC measure S_{PIAAC} and at the same time leaves room for potential differences when looking at computer skill individually.

Table 1: Correlations between skill measures

	literacy -numeracy	literacy - computer	numeracy - computer	literacy - S_{BK}	numeracy - S_{BK}	computer - S_{BK}	N
Belgium	0.910	0.622	0.642	0.368	0.395	0.216	2761
Czech Republic	0.837	0.492	0.546	0.249	0.334	0.099	2893
Denmark	0.912	0.643	0.621	0.304	0.375	0.181	4753
Estonia	0.881	0.505	0.552	0.213	0.289	0.102	4306
Finland	0.880	0.617	0.607	0.221	0.278	0.104	3411
Germany	0.906	0.556	0.629	0.280	0.342	0.098	3431
Ireland	0.896	0.547	0.552	0.396	0.400	0.308	3194
Japan	0.892	0.407	0.437	0.279	0.392	0.133	3301
Korea	0.913	0.521	0.565	0.342	0.399	0.282	3239
Netherlands	0.914	0.660	0.642	0.310	0.348	0.204	3227
Norway	0.922	0.648	0.655	0.271	0.338	0.122	3078
Poland	0.878	0.525	0.538	0.321	0.334	0.192	4464
Slovak Republic	0.879	0.453	0.465	0.233	0.326	0.154	2769
United Kingdom	0.891	0.544	0.607	0.295	0.308	0.137	4514
United States	0.922	0.704	0.700	0.507	0.530	0.384	2597

Notes: Within-country correlations calculated using sampling weights.

[14] According to S_{PIAAC} Finland, the Netherlands and Norway come out on top of the ranking. These countries would take ranks 12, 5 and 2 respectively under a ranking based on years of schooling and experience. Similarly, the first three countries under S_{BK} are ranked 13th, 3rd and 12th under S_{PIAAC}. Spearman's rho gives a rank correlation of 0.129.

2.2. Mincer earnings regressions

All human capital variables are related to hourly earnings in a meaningful way. Mincer (1974)-type of regressions show a coherent direction of association for all countries, but great variation in the magnitude of estimates. In order to assess the relative and individual importance of schooling versus the comprehensive measure of cognitive skill, I estimate three specifications of earnings regressions.

Table 2: Earnings regressions on variables of human capital

Dependent variable: *log earnings*	educ.	(s.e.)	S_{PIAAC}	(s.e.)	educ.	(s.e.)	S_{PIAAC}	(s.e.)	N
Belgium	0.069	(0.003)	0.158	(0.007)	0.051	(0.003)	0.087	(0.008)	2679
Czech Republic	0.071	(0.005)	0.160	(0.014)	0.052	(0.006)	0.097	(0.015)	2607
Denmark	0.062	(0.003)	0.118	(0.011)	0.053	(0.003)	0.059	(0.011)	4467
Estonia	0.075	(0.004)	0.188	(0.009)	0.054	(0.004)	0.122	(0.010)	3960
Finland	0.062	(0.002)	0.122	(0.009)	0.054	(0.002)	0.058	(0.009)	3224
Germany	0.101	(0.005)	0.216	(0.011)	0.078	(0.005)	0.120	(0.012)	3286
Ireland	0.070	(0.005)	0.169	(0.016)	0.052	(0.005)	0.097	(0.017)	2774
Japan	0.078	(0.005)	0.179	(0.013)	0.058	(0.006)	0.122	(0.015)	3248
Korea	0.085	(0.004)	0.209	(0.013)	0.069	(0.005)	0.085	(0.016)	3103
Netherlands	0.080	(0.004)	0.155	(0.011)	0.065	(0.004)	0.081	(0.012)	3088
Norway	0.061	(0.003)	0.130	(0.009)	0.047	(0.004)	0.089	(0.009)	2984
Poland	0.093	(0.005)	0.195	(0.013)	0.076	(0.006)	0.090	(0.014)	3866
Slovak Republic	0.089	(0.005)	0.210	(0.012)	0.068	(0.006)	0.127	(0.013)	2505
United Kingdom	0.086	(0.005)	0.202	(0.010)	0.063	(0.005)	0.153	(0.012)	4209
United States	0.098	(0.005)	0.224	(0.018)	0.077	(0.006)	0.093	(0.021)	2350

Notes: Least squares regressions weighted by sampling weights. Robust standard errors in parentheses. All coefficients are statistically significant at 1%. Regressions control for gender, experience, experience^2. S_{PIAAC} is standardized within each country. Each separated section corresponds to one specification.

Table 2 reports the results in three separated sections. Earnings are generally concavely associated with work experience and positively with education or skill. In a classical Mincer regression of log earnings on education and a quadratic in work experience, one additional year of schooling is associated with an increase in earnings between 6.1 (Norway) and 10.1 (Germany) percent. The estimates are comparable in what they convey about relative effects across countries, but not identical to Hanushek et al. (2013)'s estimates due to a less restricted sample. Replacing education by S_{PIAAC} shows that earnings are also strongly related to cognitive skills. A one standard deviation increase in cognitive skill is associated with a 12.2 (Finland) to 22.4 (United States) percent increase in earnings.[15] Including both variables of human capital side by side proves that they each pick up some of the variation in earnings and remain statistically significant, but reduce the importance of the other. At the same time, the overall fit increases. This points to the fact that the PIAAC scores and years of schooling at least partly impact earnings through different channels. When the comprehensive PIAAC score is split up into its three components (estimates are not reported here), the coefficients on education remain unchanged and the association of S_{PIAAC} is distributed across literacy, numeracy and computer skill. Due to the strong correlation between the dimensions not all coefficients are significant, but which type of skill dominates differs between the countries.[16]

The insight that emerges from all three specifications of Mincer regressions is that even though the estimated coefficients on education or cognitive ability do not reflect a causal relationship, they nevertheless are consistently positive and economically significant across countries and therefore predictive for average earnings. This provides the premise for examining both the indirect measure of skill based on education and experience and the PIAAC-based measures with respect to earnings inequality.

3. The supply and demand model of skill

The canonical model for the supply and demand of skill represents an appealing route to do so because of its simplicity and intuitive groundings. The supply and

[15] Note that these effects appear to be somewhat smaller than the effect of education since the within-country standard deviation of education lies around 2.5 years for all countries. However, as Paccagnella (2015) already notes, such direct comparisons have to be taken very cautiously because of the different metrics of the two variables.

[16] For further analysis of earnings regressions including heterogeneous effects by age and other explanatory variables, see Hanushek et al. (2013).

demand model formalizes Tinbergen's race of education and technology. Tinbergen looked at the longterm movements in income inequality between graduate and other labor in developed countries as the "net balance of [...] conflicting effects" (Tinbergen, 1975, p.79) arising from demand and supply factors. The individual's supply behavior originates from her utility maximization and ultimately results in a certain share of a population obtaining higher education or skill. Demand for highly educated or skilled workers follows from the production function and increases with augmented capital and technological development, thus coining the phrase 'skill-biased technological change'. The interplay of demand and supply factors determines the income ratio between higher and lower skilled labor, and how income inequality between skill groups evolves is therefore decided in a race between technology and (the supply of) education. Intuitively, because workers of different skill groups are considered to be imperfect substitutes, higher relative net demand for one skill group versus the other results in relatively higher earnings for this skill group.

3.1. The empirical model

The competitive framework of the origination of skill group inequality was first applied by Blau & Kahn (1996) to explain the different levels of earnings inequality across countries.[17] Translating the mechanism into an international context requires the additional assumption that there are international barriers to the mobility of capital, labor or goods so that skill prices are not equalized. Then, a specific supply and demand structure for skill in one country implies distinct returns to skill levels as compared to another country. In order to empirically test this relationship, Blau & Kahn (1996) employ demand and supply indices for skill groups that are relative across two dimensions. First, supply

[17] The cross-country methodology is an adaptation of the partial equilibrium framework developed by Katz & Murphy (1992) to study changes in skill earnings inequality over time.

and demand for skill in one country is always measured in reference to a *baseline country*. Second, the difference in supply and demand *between two skill groups* is used to explain their relative wages.

Leuven et al. (2004) develop Blau and Kahn's indices slightly further and I report their formulas here. Low, middle and high skill groups are defined by cutoff values from the skill distribution in the baseline country that split the population in the baseline country into three equally sized groups. This absolute perspective on skill creates variation in skill supply as the distributions in the countries differ. Appendix C illustrates the skill group classification, as well as supply and demand indices, for the United States as the (arbitrarily chosen) baseline country.

The *skill supply index* in reference to the baseline country is a count of the representation of skill group k in the workforce (including currently employed and unemployed persons) of country j relative to the baseline country b on a log scale.[18]

$$s_{k,j} = \ln\left(\frac{S_{k,j}}{S_{k,b}}\right)$$

with $S_{k,j}$ the share of skill group k in country j
$S_{k,b}$ the share of skill group k in baseline country b ($= \frac{1}{3}$ by construction).

The *skill demand index* in reference to the baseline country focuses on employed people and measures the degree to which the occupation-industry structure in one country j favors the skill group k relative to the baseline country (Blau & Kahn, 1996; Leuven et al., 2004). It sums over the weighted differences in employment in occupation-industry cells between two countries.[19] The differ-

[18] Leuven et al. (2004) and Blau & Kahn (1996) focus on employed workers also in their supply indices. I deviate from this because I assume unemployed persons to be 'ready to be hired' and therefore being part of the supply of skill on the labor market. A robustness check shows that when the analysis is conducted on employed persons only, the results are very similar.

[19] Occupation-industry cells are determined by a 3x6-grid of three major groups of occupations (managers and professionals; clerical and sales workers; craft, trade, operators,

ences in employment are measured regardless of the skill group affiliation of the employees. The relation to skill groups is introduced through the weights, which are constructed as the share of the skill group employed in the individual occupation-industry cells in the baseline country, scaled by the skill group's total share of employed in the baseline country. The final number is also transformed into its logarithm and centered around zero.

$$d_{k,j} = \ln\left(1 + \sum_o c_{ok}\frac{\Delta E_o}{E_{k,b}}\right)$$

with	c_{ok}	the share of skill group k of employed in occupation-industry cell o in the baseline country b
	ΔE_o	the difference in shares of total labour input employed in cell o between country j and b
	$E_{k,b}$	the share of total labour input accounted for by skill group k in baseline country b.

Subtracting demand from supply gives the net supply of skill group k in country j.

$$NS_{k,j} = s_{k,j} - d_{k,j}$$

Finally, the difference between the net supply indices of two skill groups k and l gives their *relative net supply.*

$$NS_{k,j} - NS_{l,j} \tag{1}$$

For these two skill groups earnings differentials are calculated as the log of the ratio of average earnings in skill groups k and l in country j.

$$W_{k/l,j} = \ln\left(\frac{W_{k,j}}{W_{l,j}}\right)$$

In comparison to the baseline country b the *relative skill earnings differential* is

assemblers, elementary (laborers), service workers) and six industries (agriculture; mining, manufacturing and construction; transportation, communication and public utilities; trade; finance, insurance, real estate and services; government).

then

$$W_{k/l,j} - W_{k/l,b} \tag{2}$$

The supply and demand model predicts that if one country has a larger relative net supply of one skill group as compared to the baseline country, the skill group should fare worse in terms of relative earnings in that country compared to the baseline country, i.e., the relative skill earnings differential should be negatively correlated to the relative skill net supply. To give you an example, the larger the relative net supply of high skilled workers in Finland compared to the United States, the worse off are high skilled workers in relative terms in Finland than they are in the United States (possibly interpreted as bargaining power). The following regression equation, which combines equations (1) and (2), expresses this relationship:

$$W_{k/l,j} - W_{k/l,b} = \alpha + \beta(NS_{k,j} - NS_{l,j}) + \varepsilon_j \tag{3}$$

with β having a negative sign. In terms of Tinbergen's original idea, β can be thought of as the inverse of the elasticity of substitution between the two skill groups (Tinbergen, 1975, p. 85).

4. Results

I assess the validity of the theoretical relationship expressed in equation (3) in separate regressions for each pairwise comparison between two skill groups, i.e. high versus low, high versus middle and middle versus low skill, and in pooled regressions with all skill groups. The former estimations are informative on whether market forces influence earnings differentials only for certain parts of the skill distribution, whereas pooled regressions of all pairs of relative net supply and relative earnings differentials give an indication of how much variation the model explains overall. Importantly, I repeat this sequence of regressions for each measure of skill as defined in Section 2. On the one hand, this allows verifying to some extent the inference drawn from a single skill measure. On the other hand, it provides insight into whether individual dimensions of skill add to the explanatory power of the model in the international context.

Paccagnella (2015) and Pena (2015) both choose to look at earnings inequality in reference to one specific baseline country, which is in Paccagnella's case the United States, in Pena's case the United Kingdom. This is commonly done in the literature but remains an arbitrary choice that may influence the outcome. To avoid this potential bias the supply and demand indices are calculated with reference to each of the 15 countries in turn. The estimations of equation (3) therefore cluster standard errors at the country level in order to account for the underlying dependencies, are heteroskedasticity-robust and apply a small sample-correction factor as proposed in Cameron et al. (2008).[20]

Figure D.1 in the Appendix complements the regression results with a graphical representation of the data. The supply and demand model deduces a negative correlation of earnings differentials and net supply differences between two countries. Any data point in the scatter plot is thus predicted to lie either in the second or fourth quadrant, and a fitted regression line has a negative slope. While the regression models yield estimates of the slope, the graphs show which quadrants the data points are scattered in.

4.1. Cognitive skill measures

Table 3 displays the regression results under the three skill measures based on the cognitive test scores from PIAAC. Given the strong correlation between

[20] A world average could also serve as the baseline; the findings are robust to this exercise, but for reasons of comparability with Leuven et al. (2004) and generality of the model data points are constructed for every country with respect to every other country.

Table 3: Regressions on net supply

	S_{litnum}			$S_{computer}$			S_{PIAAC}		
Dependent variable: *Earnings differentials*									
	Net supply	(s.e.)	R^2	Net supply	(s.e.)	R^2	Net supply	(s.e.)	R^2
high – low	-0.124	(0.028)***	0.405	-0.148	(0.021)***	0.322	-0.160	(0.021)***	0.466
med – low	-0.119	(0.023)***	0.477	-0.122	(0.032)***	0.239	-0.151	(0.026)***	0.317
high – med	-0.039	(0.025)	0.030	-0.062	(0.035)*	0.041	-0.052	(0.031)	0.044
pooled	-0.107	(0.023)***	0.299	-0.123	(0.022)***	0.206	-0.131	(0.023)***	0.292

Notes: Robust standard errors that take clustering at the country level into account in parentheses. *,**,*** denote significance at 10%, 5% and 1%.

performance in the three dimensions, it is not surprising that they are very similar in nature. Under all PIAAC-based measures the estimated coefficient on relative net supply of any two skill groups is negative, *confirming* the prediction of the supply and demand model. In general, the supply and demand model holds consistently and is able to explain 20 to 30 percent of the overall variation in between-group inequality between any two skill groups (referring to the pooled regressions).

The results are clearest, however, for low-skilled workers: a 10 percent decrease in relative net supply of low-skilled workers (i.e. from a 0.33- to a 0.30-share of the population) increases their relative earnings between 1.2 and 1.6 percent, and the explanatory power of the canonical model reaches up to 47 percent (depending on the skill measure and the comparison group). The – in absolute value – larger coefficients on net supply in the regressions with high- and low-skilled workers as compared to middle- and low-skilled workers are also in line with the model as they represent the lower degree of substitutability between the two skill groups that are farther apart.

One drawback is the non-significant estimate on the net supply of high- versus middle-skilled workers. While the sign goes in the right direction, it is clear by the R-squared that the canonical model fails in explaining earnings differences in this category. This could have to do with the fact that the type of skills elicited in the PIAAC test are not as important for what determines earnings at the higher end of the skill distribution. Instead, non-cognitive components such as managerial skills may play a bigger role.

Up to this point, the outcomes discussed applied to all three measures of cognitive skill. Looking at them separately reveals that both a measure only based on literacy and numeracy skills and a measure based on computer skill have explanatory power on their own. Especially S_{litnum} is suitable for explaining earnings differences for the middle-skilled versus the low-skilled, a finding consistent with Leuven et al. (2004). Computer skill seems to add information that is reflected in relative earnings particularly when contrasting high- to low-skilled workers. Here, a skill measure based on computer skill only is able to explain 32 percent of the international variation in earnings differentials (or, expressed differently, combining the literacy and numeracy measure with computer skill raises the explanatory power of the model roughly from 40 to 47 percent), which gives support to the hypothesis of an increased role of computer skill in the labor market (OECD, 2013c, p.3). In general, the R-squareds are lower as compared to S_{litnum} or S_{PIAAC}. This has perhaps to do with the fact that measuring skill solely through the ability to use computer technology is too unidimensional and not as relevant for certain types of jobs in lower occupations and certain industries; nevertheless, given the broad concordance of estimates, one cannot negate the importance of possessing computer skill in the labor market and their relative valuation in earnings.

The supply of and demand for cognitive skills are predictive for earnings inequality when it comes to international variation; but how useful are years of schooling and experience in that context?

4.2. Education and experience

When the less direct measure of skill, the composite of years of schooling and experience, is used to classify multiple countries into different skill groups, the picture changes drastically. Table 4 shows the results.

Table 4: Blau and Kahn's measure

	S_{BK}		
Dependent variable: *Earnings differentials*			
	Net supply	(s.e.)	R^2
high – low	0.059	(0.038)	0.038
med – low	0.031	(0.044)	0.008
high – med	0.003	(0.025)	0.000
pooled	0.035	(0.027)	0.015

Notes: Robust standard errors that take clustering at the country level into account in parentheses. *,**,*** denote significance at 10%, 5% and 1%.

Under this measure the supply and demand model does not fit the data, having literally zero explanatory power. The regression model estimates no significant relationship between the net supply structure and earnings differentials between any two skill groups. If anything, the elasticity estimates carry a *positive* sign. This pattern is not reconcilable with the supply and demand model for skill but coincides with the finding in Blau & Kahn (1996), where a relative abundance of high- versus low-skilled workers in the United States positively correlates with higher relative wages than in other countries (p. 822).[21]

Years of education and experience appear to not be suitable for characterizing international differences in earnings inequality. This comes as no surprise given the international dissimilitude of (post-)educational systems. In contrast, the PIAAC measures are designed to assess basic competencies of the working population in every country "that are relevant to adults in many social contexts and work situations, and necessary for fully integrating and participating in the

[21] Appendix C elaborates on the direct comparison of the two data sets.

labor market, education and training, and social and civic life", as stated in a summarizing key fact sheet about the Survey of Adult Skills (OECD, 2013a). Given this objective and the findings in Tables 3 and 4, a plausible reading of the results is that the PIAAC cognitive scores succeed in making skills internationally comparable, both overall and especially so when describing the lower end of the skill distribution. For comparing only the top two skill groups, most likely other, more sophisticated skills are reflected in relative earnings as well, so that the prediction of the elasticity of substitution between the highest skill groups is obscured by the simplicity of the assessed tasks.

4.3. The roles of supply and demand

The data prove that the net supply situation of skill in a country (compared to a baseline country) is predictive for earnings inequality when skill is measured on an internationally comparable basis. For those skill measures it is instructive to look at the influence of supply and demand indices separately.

Table 5: Regressions on supply and demand

	S_{PIAAC}					
Dependent variable: *Earnings differentials*						
	Supply	(s.e.)	Demand	(s.e.)	R^2	H0: $\beta_S = -\beta_D$
high – low	-0.162	(0.025)***	0.181	(0.090)*	0.466	0.7948
med – low	-0.141	(0.031)***	0.003	(0.083)	0.348	0.0427
high – med	-0.051	(0.032)	0.038	(0.123)	0.044	0.9022
pooled	-0.129	(0.027)***	0.100	(0.097)	0.293	0.7047

Notes: Robust standard errors that take clustering at the country level into account in parentheses. *,**,*** denote significance at 10%, 5% and 1%.

Table 5 shows coefficients for the split indices and adds p-values of two-sided hypothesis tests for the equality of coefficients in absolute terms.[22] As theory predicts, relative skill supply is negatively associated with relative earnings. All coefficients (except on the high- versus middle-skilled comparison) are significant

[22] Results are only shown for the measure S_{PIAAC}; S_{litnum} and $S_{computer}$ yield similar estimates and all qualitative statements hold for these measures as well.

at 1 percent and are quite similar to the estimates on net supply in Table 3. The coefficients on demand are positive and largely of a similar absolute size as the supply estimates, but statistically indistinguishable from zero. This is due to the fact that there is less variation in demand indices as compared to supply indices; the overall standard deviation of relative demand is 0.131 (mean: 0.007), whereas the overall standard deviation of supply is 0.477 (mean: -0.014).[23] This results in large standard errors for the demand estimates and non-significant results, but the hypothesis of equality of demand and supply estimates in absolute terms cannot be rejected in any of the cases except the middle- versus low-skilled comparison.[24] Significance issues aside, according to the data a 10 percent increase in the supply of middle- or medium-skilled workers decreases their earnings relative to low-skilled workers by 1.4 to 1.6 percent, inequality thus declines. A 10 percent higher relative demand for high-skilled workers, however, increases inequality between high- and low-skilled workers by 1.8 percent. Stronger demand for middle-skilled workers, in contrast, does not seem to affect the relative earnings of low-skilled competitors.

This tentative exercise in distinguishing the impacts of supply and demand completes the picture that the canonical model draws within the scope of this work.

5. Concluding remarks

Applying the canonical supply and demand model of skill to the PIAAC data shows that a substantial amount of earnings inequality between skill groups can be accommodated by a simple partial equilibrium framework. Depending on

[23] For an example of supply and demand indices for one baseline country, see Table C.1 in the appendix.

[24] Note that under S_{litnum} and $S_{computer}$ this is not the case; there is still a probability of 0.2639 and 0.1995 of obtaining the observed difference in size of estimates, assuming the null hypothesis is true.

the measure of skill used to categorize the international population into groups, the conclusions about the validity as well as the power of the model differ. All skill measures based on the newly collected data from the OECD Program of the International Assessment of Adult Competencies confirm the predictive power of supply of and demand for skill for relative earnings differentials. Up to 47 percent when looking at the relative earnings of the lowest skill group, or 30 percent of overall between-group inequality are explained. Persistent results based on the ability to make use of computer technology prove the importance of this 'novel' skill dimension for the labor market. As in Leuven et al. (2004) and Blau & Kahn (1996), the supply and demand model has no explanatory power under a skill measure based on years of schooling and experience, strengthening the argument that years of education and work experience cannot be easily compared across countries. The results are robust to various checks of the data and are not unequivocally driven by either supply or demand.

An interesting discovery is that the results are very much in line with what Leuven et al. (2004) estimate from data of the 1990s. This level of congruence is remarkable for several reasons. The first is that the countries included in the two studies were not the same. Only 8 countries overlap between the two samples, whereas 14 countries appear in only one of the two data sets. Amongst those 14 are countries from very different world regions such as South America (Chile in Leuven et al. (2004)) or Asia (Japan and Korea in this study). The accordance of estimates and R-squareds seems to suggest that regardless of which OECD countries are compared to each other, differences in skill net supply explain about a third of the differences in earnings inequality between skill groups.

The second reason why this finding is remarkable is because the canonical supply and demand model for skill has lately been criticized as being 'too simplistic' for modern economies. Starting with a widely-cited contribution by

Autor, Levy & Murnane (2003), a literature has emerged that focuses on occupational tasks rather than skills in order to explain changes in earnings inequality (Acemoglu & Autor, 2011; Firpo et al., 2012; Goos et al., 2014, among others). The idea behind a task-based model is that skills are portable across tasks and that tasks are the unit that produces output. Changes in labor market conditions and technology therefore primarily influence the allocation of skills to tasks (Acemoglu & Autor, 2011).[25] This distinction allows technological change to be routine-biased (i.e. biased against routine tasks) rather than skill-biased (i.e. biased in favor of higher skill), and explains the job polarization phenomenon that is observed in employment data of many countries (Autor, 2015; Michaels et al., 2010; Goos & Manning, 2007; Goos et al., 2009). Job polarization describes changes in the employment structure over time and is therefore also related to inequality changes over time within one country. With its cross-sectional cross-country analysis, this paper takes on a different perspective and it is not clear *a priori* how job polarization (in some of the countries, or to differing degrees) should affect the outcome. However, in light of a generally more complex relationship between supply and demand for skill and skill wage inequality, achieving the same explanatory power with the canonical model as 20 years earlier is a noteworthy result.

While a more complex model could perhaps do an even better job at explaining the prevailing empirical patterns, the results speak for the fact that the canonical supply and demand model for skill provides a good benchmark for looking at skill earnings inequality, especially in a cross-country context. Consequently, policy makers should consider the insights that we gain from skill supply and demand as a tool for shaping (earnings) inequality (Autor,

[25] See Gathmann & Schoenberg (2010) for an empirical measurement of the portability of skills across occupations.

2014). For example, by promoting educational programs and providing broader access to postsecondary education, policy can steer the supply of skill and work towards a moderation of skill earnings inequality. Alternatively, skill demand can be influenced through taxation and investment in such a way that it benefits skill groups that are currently oversupplied. Comparing one country's situation to, say, a neighboring country should always take differences in supply and demand for skill between the countries into account. Only after these differences are removed, other factors such as labor market institutions and regulations may give insight into additional drivers of earnings inequality.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. chapter 12. (pp. 1043–1171). Elsevier volume 4 of *Handbook of Labor Economics*.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, *344*, 843–851.

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, *29*, 3–30.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*, 1279–1333.

Blau, F. D., & Kahn, L. M. (1996). International differences in male wage inequality: Institutions versus market forces. *Journal of Political Economy*, *104*, 791–836.

Blau, F. D., & Kahn, L. M. (2005). Do cognitive test scores explain higher US wage inequality? *The Review of Economics and Statistics*, *87*, 184–193.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, *90*, 414–427.

Devroye, D., & Freeman, R. B. (2001). *Does inequality in skills explain inequality in earnings across advanced countries?*. NBER Working Papers, National Bureau of Economic Research, Inc.

Firpo, S., Fortin, N. M., & Lemieux, T. (2012). *Occupational tasks and changes in the wage structure*. Textos para discusso, Escola de Economia de Sao Paulo, Getulio Vargas Foundation (Brazil).

Fournier, J.-M., & Koske, I. (2012). *Less income inequality and more growth Are they compatible? Part 7. The drivers of labour earnings inequality An analysis based on conditional and Unconditional Quantile Regressions*. OECD Economics Department Working Papers, OECD Publishing.

Gathmann, C., & Schoenberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, *28*, 1–49.

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *The Review of Economics and Statistics*, *89*, 118–133.

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. *American Economic Review*, *99*, 58–63.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2013). *Returns to skills around the world: Evidence from PIAAC*. OECD Education Working Papers OECD Publishing.

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, *107*, 35–78.

Leuven, E., Oosterbeek, H., & van Ophem, H. (2004). Explaining international differences in male skill wage differentials by differences in demand and supply of skill. *Economic Journal*, *114*, 466–486.

Michaels, G., Natraj, A., & Reenen, J. V. (2010). *Has ICT polarized skill demand? Evidence from eleven countries over 25 Years*. Technical Report.

Mincer, J. A. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research, Inc.

OECD (2011). *Divided we stand: Why inequality keeps rising*. Technical Report OECD Publishing.

OECD (2013a). Key facts about the Survey of Adult Skills (PIAAC). Retrieved November 10 2013, from http://www.oecd.org/site/piaac/publicdataandanalysis.htm.

OECD (2013b). *OECD Employment Outlook 2013*. Technical Report OECD Publishing.

OECD (2013c). *OECD Skills Outlook 2013: First results from the Survey of Adult Skills*. Technical Report OECD Publishing.

OECD (2013d). Technical Report of the Survey of Adult Skills (PIAAC). Retrieved August 8 2014, from http://www.oecd.org/site/piaac/publicdataandanalysis.htm.

Paccagnella, M. (2015). *Skills and wage inequality: Evidence from PIAAC*. OECD Education Working Papers 114, OECD Publishing.

Pena, A. A. (2015). Revisiting the effects of skills on economic inequality: Within- and cross-country comparison using PIAAC. Retrieved May 21 2015, from http://economics.colostate.edu/author/aalves/.

Tinbergen, J. (1974). Substitution of graduate by other labor. *Kyklos*, *27*, 217–226.

Tinbergen, J. (1975). *Income Distribution. Analysis and Policies*. North-Holland Publishing Company.

Appendix A. Imputation of missing computer skill scores

In light of the technological revolution that has reached both private life and the workplace it is especially interesting to study the influence of supply and demand of computer skill on inequality. However, part of the population does not possess the basic knowledge necessary to participate in a direct assessment of computer skills. This leads to missing computer skill scores for a substantial fraction of each country's population, which are displayed in Table A.1. The Technical Report (OECD, 2013d) names three reasons for a missing computer skill score:

1. The individual reports to not have any prior experience in using a computer.
2. The individual reports to have used a computer before but fails a basic ICT core test.[26]
3. The individual opts out of the computer-based assessment.[27]

Logit regressions in Table A.2 show that the probability of not having participated in the computer skill assessment is positively related to a higher age group and working in a lower type of occupation, and negatively associated with earnings, experience, education and higher cognitive scores in the other dimensions. These relations are stronger for the first category of missing than for the second, and partly disappear for the third category. The data are thus not missing at random. Leaving these observations out of a supply and demand analysis would distort the indices by imposing the assumption that these individuals play no role in the labor market.

To avoid this problem computer skills are imputed for those missing in accordance with the stated reason for why they are missing. Since reasons 1. and 2. indicate with certainty that the subject has no computer skills (as defined by the scale of the assessment), the imputed score is 0. Category 3. is more ambiguous as the reason for opting out is unspecified. The score for computer skill is imputed as the prediction based on the variables education, experience, age, male, literacy, numeracy and an indicator for the lowest type of occupation, with weights being the estimated coefficients of a regression of observed computer skill on these variables.

The imputation allows to make use of the novel skill dimension to examine the importance of computer skills, as well as to construct a comprehensive measure of cognitive skill covering all three dimensions. It is important to note that while imputing computer skill scores introduces some uncertainty about the obtained results

[26] The ICT core test assesses whether the subject commands the skills necessary to follow the actual test, i.e. scrolling, clicking and the like (OECD, 2013c, p. 88).

[27] Since the evaluation of computer skills was a voluntary component, 'opting out' need not be related to any reason in particular. It is likely, however, that individuals opt out because they do not feel comfortable enough taking the assessment.

Table A.1: Missing computer skill scores

	missing	no computer experience	failed ICT core test	opted out
Belgium	0.11	0.04	0.03	0.04
Czech Republic	0.20	0.06	0.02	0.12
Denmark	0.12	0.01	0.06	0.05
Estonia	0.26	0.06	0.03	0.16
Finland	0.11	0.01	0.03	0.07
Germany	0.13	0.05	0.03	0.05
Ireland	0.23	0.05	0.04	0.14
Japan	0.31	0.07	0.10	0.14
Korea	0.26	0.12	0.09	0.05
Netherlands	0.07	0.01	0.03	0.03
Norway	0.09	0.00	0.04	0.04
Poland	0.32	0.06	0.07	0.19
Slovak Republic	0.33	0.18	0.02	0.12
United Kingdom	0.08	0.02	0.04	0.03
United States	0.13	0.04	0.04	0.06

Note: All values are fractions of the total populations.

Table A.2: Logit regressions of the probability of missing computer skill score

	(1) missing	(2) no exp	(3) fail ICT	(4) opt out	(5) missing	(6) no exp	(7) fail ICT	(8) opt out
earnings	-0.722***	-0.860***	-0.0854	-0.547***				
	(0.116)	(0.220)	(0.0871)	(0.122)				
age group	0.376***	0.588***	0.125***	0.168***	0.371***	0.578***	0.123***	0.169***
	(0.0303)	(0.0382)	(0.0277)	(0.0382)	(0.0313)	(0.0378)	(0.0278)	(0.0404)
male	0.186**	0.265***	0.339***	-0.139*	0.0201	0.0905	0.318***	-0.269***
	(0.0824)	(0.0826)	(0.0807)	(0.0754)	(0.0784)	(0.0620)	(0.0810)	(0.0682)
education	-0.0911***	-0.228***	0.0124	-0.0387	-0.123***	-0.234***	0.00278	-0.0621***
	(0.0277)	(0.0303)	(0.0162)	(0.0245)	(0.0253)	(0.0308)	(0.0176)	(0.0234)
experience	-0.224***	-0.286***	-0.184***	0.0383	-0.302***	-0.343***	-0.194***	-0.0215
	(0.0557)	(0.0766)	(0.0451)	(0.0695)	(0.0633)	(0.0661)	(0.0460)	(0.0784)
numeracy	-1.120***	-1.145***	-1.399***	-0.310	-1.167***	-1.116***	-1.420***	-0.371**
	(0.218)	(0.196)	(0.299)	(0.198)	(0.208)	(0.200)	(0.283)	(0.181)
occupation	0.497***	0.957***	0.239***	0.367***	0.612***	1.089***	0.231***	0.469***
	(0.0431)	(0.0456)	(0.0606)	(0.0626)	(0.0469)	(0.0801)	(0.0573)	(0.0631)
unemployed					0.0234	0.105	-0.210***	0.157**
					(0.0955)	(0.175)	(0.0548)	(0.0729)
Constant	1.585***	-0.855*	-0.292	-1.399***	0.335	-2.976***	-0.0577	-2.476***
	(0.526)	(0.494)	(0.512)	(0.473)	(0.520)	(0.734)	(0.542)	(0.486)
Observations	48,350	48,350	48,350	48,350	51,938	51,938	51,938	51,938

Note: Robust standard errors account for clustering by country. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

in the supply and demand analysis, the imputation does not affect the analysis based solely on literacy and numeracy skills or the formal measures of years of schooling and experience. Eyeballing the results in Table 3 proves that the obtained estimates for the computer skill measure are not unsound.

As an additional robustness check to the applied imputation mechanism the supply and demand model is replicated under two alternatives. These mechanisms are chosen in such a way that they provide one rather 'optimistic' and one rather 'pessimistic' estimation of missing computer skill scores.

Pessimistic imputation Individuals reporting no prior computer experience, as well as individuals who fail the basic ICT core test are assigned a zero score for their average computer skill value. Individuals who opt out of the computer-based assessment receive an extremely low score of 0.85[28].

This mechanism can be thought of as assuming that individuals who opt out of the computer-based assessment feel uncomfortable working with a computer because they have very poor computer skills.

Optimistic imputation All missing computer skill scores are imputed as out-of-sample predictions based on available information on age, gender, education, experience, average literacy and numeracy score and occupation. Coefficients come from a regression of observed average computer skill scores on the aforementioned characteristics.

This mechanism can be thought of as a forward looking scenario, assuming that in the near future it will be less likely that individuals have no computer skills at all.

Table A.3 shows the estimation results for regressions of skill group earnings differentials on relative net supply applying the two alternative imputation mechanisms for missing computer skill scores. In the direct comparison with Table 3 a few observations are salient. First, under the 'pessimistic imputation' standard errors are slightly smaller, leading to significant estimates even for contrasting high- to middle-skilled workers. The explanatory power of the model for this category is greater but remains limited at 8 to 11 percent. Second, under the 'pessimistic imputation' the estimated net supply coefficients according to the measure $S_{computer}$ are smaller in absolute value and the explanatory power is generally lower, but qualitative nuances are the same. Changed imputed scores are of little consequence for the combined PIAAC measure S_{PIAAC}. Third, the 'optimistic imputation' results in similar estimates as the applied imputation method, with only slight differences in explanatory power shifted from the high- versus low-skilled contrast to the middle- versus low-skilled comparison. In

[28] The value stems from OECD (2013b) where individuals with missing literacy or numeracy scores are treated as scoring 85 points on average.

summary, elasticity estimates from the supply and demand model are reliable across a range of methods of imputation of missing values; regardless of the method, around 30 percent of overall earnings inequality between skill groups can be explained by the model.

Table A.3: Net supply regressions under alternative imputation mechanisms

Optimistic imputation						
	$S_{computer}$			S_{PIAAC}		
Dependent variable: *Earnings differentials*						
	Net supply	(s.e.)	R^2	Net supply	(s.e.)	R^2
high – low	-0.132	(0.025)***	0.260	-0.131	(0.028)***	0.376
med – low	-0.136	(0.028)***	0.350	-0.124	(0.025)***	0.450
high – med	-0.077	(0.035)**	0.064	-0.050	(0.028)*	0.043
pooled	-0.119	(0.023)***	0.203	-0.113	(0.024)***	0.277
Pessimistic imputation						
	$S_{computer}$			S_{PIAAC}		
Dependent variable: *Earnings differentials*						
	Net supply	(s.e.)	R^2	Net supply	(s.e.)	R^2
high – low	-0.085	(0.015)***	0.211	-0.123	(0.016)***	0.390
med – low	-0.051	(0.017)***	0.165	-0.101	(0.023)***	0.327
high – med	-0.073	(0.024)***	0.076	-0.071	(0.022)***	0.106
pooled	-0.071	(0.012)***	0.158	-0.108	(0.015)***	0.299

Notes: Robust standard errors that take clustering at the country level into account in parentheses. *,**,*** denote significance at 10%, 5% and 1%.

Appendix B. Descriptive statistics

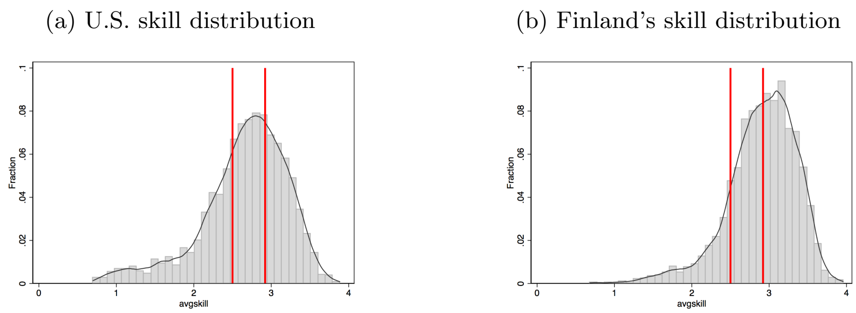
Table B.1: Mean (standard deviation) of key variables

	N	empl.	male	age		education		experience		earnings	
Belgium	2761	2679	0.52	41.09	(11.18)	12.99	(2.59)	19.60	(11.55)	20.24	(10.85)
Czech Republic	2893	2607	0.53	40.15	(11.33)	13.27	(2.52)	18.51	(11.68)	9.11	(7.84)
Denmark	4753	4467	0.50	41.09	(12.32)	13.01	(2.59)	21.29	(12.42)	24.06	(11.49)
Estonia	4306	3960	0.45	40.30	(12.43)	12.50	(2.60)	18.62	(12.49)	9.97	(9.51)
Finland	3411	3224	0.49	41.12	(12.40)	13.00	(2.88)	18.34	(12.41)	18.65	(7.98)
Germany	3431	3286	0.52	41.34	(11.98)	13.73	(2.54)	19.32	(12.48)	18.74	(13.11)
Ireland	3194	2774	0.49	37.93	(11.60)	15.36	(2.89)	16.40	(11.09)	22.19	(16.30)
Japan	3301	3248	0.56	41.51	(12.50)	13.32	(2.35)	18.68	(12.03)	16.60	(17.31)
Korea	3239	3103	0.57	39.27	(11.33)	13.38	(2.98)	12.66	(9.95)	17.58	(18.35)
Netherlands	3227	3088	0.52	39.68	(12.46)	13.61	(2.48)	18.60	(11.60)	20.87	(14.61)
Norway	3078	2984	0.49	39.74	(12.72)	14.44	(2.41)	18.26	(12.03)	24.43	(12.31)
Poland	4464	3866	0.53	38.60	(11.75)	13.37	(2.93)	15.63	(11.78)	9.71	(10.99)
Slovak Republic	2769	2505	0.52	40.22	(11.31)	13.59	(2.58)	18.14	(11.68)	10.00	(18.19)
United Kingdom	4514	4209	0.51	38.82	(12.23)	13.19	(2.29)	18.74	(12.12)	18.86	(14.55)
United States	2597	2350	0.50	median: 40-44		13.85	(2.92)	20.32	(12.33)	23.44	(21.01)
		S_{BK}		S_{PIAAC}		literacy		numeracy		computer skill	
Belgium		2.916	(0.281)	2.765	(0.509)	2.817	(0.424)	2.870	(0.451)	2.608	(0.828)
Czech Republic		2.923	(0.288)	2.686	(0.490)	2.758	(0.380)	2.772	(0.407)	2.527	(0.918)
Denmark		2.931	(0.304)	2.770	(0.484)	2.770	(0.414)	2.852	(0.450)	2.689	(0.756)
Estonia		2.843	(0.309)	2.680	(0.491)	2.787	(0.402)	2.761	(0.407)	2.493	(0.890)
Finland		2.886	(0.325)	2.888	(0.465)	2.971	(0.419)	2.914	(0.438)	2.780	(0.711)
Germany		2.968	(0.323)	2.686	(0.535)	2.733	(0.430)	2.764	(0.473)	2.561	(0.907)
Ireland		3.085	(0.304)	2.612	(0.522)	2.727	(0.420)	2.631	(0.465)	2.477	(0.910)
Japan		2.925	(0.274)	2.784	(0.547)	3.004	(0.352)	2.927	(0.398)	2.421	(1.186)
Korea		2.837	(0.315)	2.565	(0.575)	2.756	(0.378)	2.677	(0.410)	2.261	(1.174)
Netherlands		2.957	(0.289)	2.851	(0.475)	2.913	(0.430)	2.869	(0.448)	2.771	(0.702)
Norway		3.022	(0.310)	2.801	(0.497)	2.833	(0.427)	2.842	(0.490)	2.729	(0.733)
Poland		2.878	(0.324)	2.527	(0.584)	2.720	(0.427)	2.661	(0.449)	2.199	(1.136)
Slovak Republic		2.946	(0.295)	2.631	(0.536)	2.789	(0.337)	2.835	(0.394)	2.268	(1.145)
United Kingdom		2.914	(0.267)	2.751	(0.483)	2.820	(0.425)	2.728	(0.477)	2.705	(0.745)
United States		2.997	(0.334)	2.624	(0.584)	2.740	(0.479)	2.592	(0.551)	2.541	(0.879)

Note: Male stands for the fraction of males in the sample. Age is continuous and ranges from 18 to 65, with exception of the United States where it is grouped into 5 year-intervals from 20 to 65. Education refers to years of schooling. Experience denotes years worked. Earnings are gross hourly earnings in $US. All statistics are calculated using sampling weights.

Appendix C. Example of supply and demand indices

Figure C.1: Skill cutoffs according to U.S. distribution



Using the United States as an example for the baseline country, Figure C.1 shows the categorization of another country's population into absolute skill groups. The U.S. skill distribution (skill measure: S_{PIAAC}) in Panel (a) is split into three equal parts, where the vertical lines mark the cutoff values. When overlayed with the Finnish skill distribution in Panel (b), the low skill group is noticeably smaller and the high skill group larger. That means, when compared to the United States, Finland has a lower supply of low-skilled workers and a larger supply of high-skilled workers.

Table C.1: Supply, demand and net supply under S_{PIAAC}

	Supply			Demand			Netsupply		
	low	med	high	low	med	high	low	med	high
Belgium	-0.327	-0.060	0.290	0.022	0.018	-0.038	-0.349	-0.078	0.328
CzechRep	-0.175	0.094	0.060	0.186	0.022	-0.229	-0.361	0.072	0.290
Denmark	-0.407	0.036	0.261	0.027	-0.004	-0.021	-0.434	0.040	0.282
Estonia	-0.138	0.103	0.021	0.102	0.005	-0.107	-0.241	0.098	0.128
Finland	-0.732	-0.034	0.440	0.084	-0.012	-0.069	-0.815	-0.022	0.509
Germany	-0.105	-0.031	0.122	0.139	0.014	-0.159	-0.244	-0.045	0.281
Ireland	-0.008	0.102	-0.105	0.098	0.003	-0.100	-0.106	0.099	-0.005
Japan	-0.283	-0.276	0.398	0.141	0.008	-0.155	-0.424	-0.285	0.553
Korea	0.042	0.027	-0.073	0.173	0.044	-0.238	-0.131	-0.017	0.165
Netherlands	-0.603	-0.072	0.420	-0.027	0.014	0.010	-0.577	-0.085	0.410
Norway	-0.480	-0.025	0.341	0.025	0.004	-0.027	-0.505	-0.029	0.368
Poland	0.189	-0.026	-0.201	0.133	0.023	-0.162	0.056	-0.049	-0.040
SlovakRep	-0.101	0.061	0.033	0.143	0.007	-0.156	-0.245	0.054	0.190
United Kingdom	-0.256	0.020	0.187	0.070	-0.005	-0.063	-0.326	0.025	0.250
United States					baseline country				

Notes: Calculated using sample weights. Skill measure used is S_{PIAAC}.

Table C.1 shows supply, demand and net supply of skill groups for all countries compared to the United States. It becomes clear that most countries have fewer low-skilled workers in terms of the U.S. skill distribution, but that demand for this skill group tends to be higher in these countries as compared to the United States. This results in negative net supply for all countries except Poland. Roughly the opposite holds for high-skilled workers; for the middle skill group differences to the United States are not as systematic.

The comparison of Table C.1 to Tables 5 – 7 in Blau & Kahn (1996), which also display demand, supply and net supply with reference to the United States, is interesting. There is little overlap in countries considered and the underlying samples are different with respect to gender and the exclusion of self-employed or unemployed, but nevertheless the big picture of demand indices is comparable, suggesting that the demand structure for skill has not changed that much over time. The supply indices, however, are diametrically opposed for the low and high skill groups. Applying the same skill measure as Blau and Kahn used in the current data set shows that this is not a product of a change in supply structure over time, but entirely accounted for by the different measurement of skill. This finding provides a strong argument for the superiority of cognitive skill measures as opposed to a composite of years of education and experience in the international context.

Table C.2: Supply, demand and net supply under S_{BK}

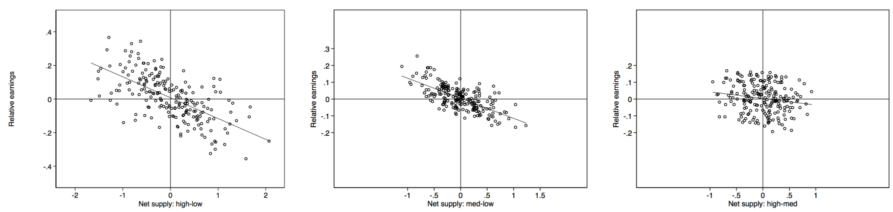
	Supply			Demand			Netsupply		
	low	med	high	low	med	high	low	med	high
Belgium	0.178	0.074	-0.319	-0.008	0.044	-0.037	0.186	0.030	-0.282
CzechRep	0.212	0.273	-0.808	0.135	0.118	-0.293	0.077	0.154	-0.515
Denmark	0.148	-0.007	-0.166	0.023	0.012	-0.033	0.125	-0.020	-0.132
Estonia	0.517	-0.327	-0.509	0.084	0.028	-0.114	0.433	-0.356	-0.395
Finland	0.348	-0.233	-0.235	0.072	0.031	-0.104	0.276	-0.264	-0.132
Germany	-0.013	0.220	-0.269	0.120	0.079	-0.218	-0.133	0.141	-0.051
Ireland	-0.222	-0.398	0.427	0.122	0.023	-0.151	-0.344	-0.421	0.578
Japan	0.203	0.091	-0.389	0.150	0.079	-0.258	0.053	0.012	-0.132
Korea	0.442	-0.088	-0.643	0.192	0.093	-0.338	0.250	-0.181	-0.305
Netherlands	0.076	0.004	-0.088	-0.017	0.012	0.003	0.093	-0.008	-0.091
Norway	-0.104	-0.207	0.253	0.039	0.008	-0.044	-0.142	-0.215	0.297
Poland	0.389	-0.112	-0.463	0.104	0.060	-0.173	0.285	-0.172	-0.290
SlovakRep	0.169	0.227	-0.583	0.107	0.071	-0.192	0.062	0.156	-0.391
United Kingdom	0.313	-0.091	-0.331	0.070	0.050	-0.122	0.243	-0.141	-0.208
United States					baseline country				

Notes: Calculated using sample weights. Skill measure used is S_{BK}.

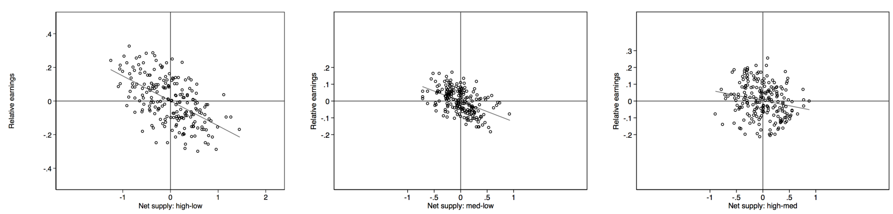
Appendix D. Figures

Figure D.1: Graphical representation of relative net supply and relative earnings differentials

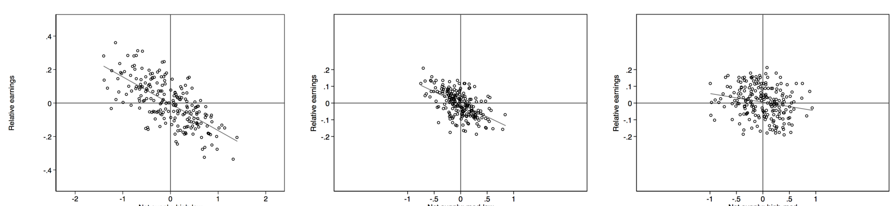
(a) S_{litnum}: high vs. low (b) S_{litnum}: med vs. low (c) S_{litnum}: high vs. med



(d) $S_{computer}$: high vs. low (e) $S_{computer}$: med vs. low (f) $S_{computer}$: high vs. med



(g) S_{PIAAC}: high vs. low (h) S_{PIAAC}: med vs. low (i) S_{PIAAC}: high vs. med



(j) S_{BK}: high vs. low (k) S_{BK}: med vs. low (l) S_{BK}: high vs. med