Skills, earnings, and employment: exploring causality in the estimation of returns to skills

Hampf, Franziska; Wiederhold, Simon; Woessmann, Ludger

Hampf *et al. Large-scale Assess Educ (2017) 5:12*
DOI 10.1186/s40536-017-0045-7

Large-scale Assessments in Education

RESEARCH

Open Access

Skills, earnings, and employment: exploring causality in the estimation of returns to skills

Franziska Hampf[1], Simon Wiederhold[2*] and Ludger Woessmann[3]

*Correspondence: simon.wiederhold@ku.de
[2] KU Eichstaett-Ingolstadt, ifo Institute, and CESifo, Auf der Schanz 49, 85049 Ingolstadt, Germany
Full list of author information is available at the end of the article

Abstract

Ample evidence indicates that a person's human capital is important for success on the labor market in terms of both wages and employment prospects. However, unlike the efforts to identify the impact of school attainment on labor-market outcomes, the literature on returns to cognitive skills has not yet provided convincing evidence that the estimated returns can be causally interpreted. Using the PIAAC Survey of Adult Skills, this paper explores several approaches that aim to address potential threats to causal identification of returns to skills, in terms of both higher wages and better employment chances. We address measurement error by exploiting the fact that PIAAC measures skills in several domains. Furthermore, we estimate instrumental-variable models that use skill variation stemming from school attainment and parental education to circumvent reverse causation. Results show a strikingly similar pattern across the diverse set of countries in our sample. In fact, the instrumental-variable estimates are consistently larger than those found in standard least-squares estimations. The same is true in two "natural experiments," one of which exploits variation in skills from changes in compulsory-schooling laws across U.S. states. The other one identifies technologically induced variation in broadband Internet availability that gives rise to variation in ICT skills across German municipalities. Together, the results suggest that least-squares estimates may provide a lower bound of the true returns to skills in the labor market.

Keywords: PIAAC, Cognitive skills, Education, Labor market, Earnings, Employment, International comparisons

Background

Human capital analysis starts with the assumption that human capital can be acquired through schooling and lifelong learning. While these activities are costly, they are generally expected to entail future benefits, for example in the form of returns in terms of higher wages and increased employability. Following the seminal contributions of Schultz (1961), Becker (1962), and Mincer (1974), thousands of studies have investigated individuals' returns to human capital in the labor market. Human capital can be regarded as skills that make workers more productive in performing their work tasks and as the knowledge and competencies that enable people to generate and adopt new ideas that spur innovation and technological progress. This productivity-enhancing effect of human capital increases a person's wage or allows her to escape unemployment and find a job in the first place.

A key challenge for the work on the role of human capital in modern economies concerns its measurement. Previous empirical literature relies almost exclusively on available quantity-based measures of human capital investment such as educational attainment, which is typically measured by years of schooling. While such measures are certainly related to human capital and, in fact, have been shown to be economically relevant, they nevertheless might be less than perfect approximations of effective human capital. For example, the quality of schooling might change over time and might vary across countries. Approximating an individual's stock of human capital with years of schooling is especially problematic in cross-country comparisons, which implicitly assume that the contribution of each school year to human capital accumulation is independent of the quality of the education system—i.e., that a year of schooling, e.g., in Papua New Guinea creates the same increase in productive human capital as a year of schooling in Japan (Hanushek and Woessmann, 2008, 2015). This can certainly be questioned. Moreover, measures of educational attainment just reflect an individual's human capital at the end of formal schooling, which may not be good indicators of effective human capital when individuals need to constantly adapt their skills to structural and technological change throughout their entire working life.

An alternative approach to human capital measurement is to measure the skills of adults directly. Until fairly recently, almost all of the international evidence on cognitive skills of the adult population came from the International Adult Literacy Survey (IALS) of the mid-1990s (see Hanushek and Woessmann, 2011, for a review). However, skill measures from two decades ago may not accurately capture the situation in economies that have undergone substantial technological change (Autor et al., 2003; Goldin and Katz, 2008; Acemoglu and Autor, 2011). Recently, a new large-scale assessment of the skills of the adult population was conducted—the Programme for the International Assessment of Adult Competencies (PIAAC). Compared to IALS, PIAAC has greater country coverage, considerably larger sample sizes, and tests that cover a wider variety of skills.

In addition to measurement, another key challenge in the estimation of returns to human capital on the labor market is causality. Unlike the efforts to identify the causal impact of school attainment on wages or employment using a variety of sources of exogenous variation (for reviews see Card, 1999; Heckman et al., 2006), the literature on returns to skills stops short of providing convincing evidence that the estimated returns can be causally interpreted. There are three main potential threats to causal identification of the relationship between skills and labor-market outcomes. First, measurement error in the skills variable could give rise to classical attenuation bias, implying that least squares estimates of the returns to skills are underestimates of the true impact of skills on wages or employment. Second, different employment patterns could directly affect test scores over the lifecycle, implying problems of reverse causation. For example, better jobs might use and reinforce skills whereas worse jobs or employment breaks might lead to skill depreciation. Third, various omitted variables could bias the estimates. Among others, family background, health, or personality traits could directly influence labor-market outcomes; if also related to skills, these could lead to an omitted variable bias in the analysis of skills.

Using the PIAAC survey, this paper takes a deeper look into the main issues of identification in the estimation of returns to skills for a large set of countries. We explore several approaches designed to deal with the possible sources of bias. First, we make use of the fact that PIAAC tests skills in several domains and provides several plausible values for each skill measure, allowing us to deal with the issue of measurement error. Second, we address reverse causality (e.g., better jobs might reinforce skills) in instrumental-variable (IV) models. These IV models use only the part of the variation in skills that is determined before labor-market entry and is therefore unaffected by job-specific patterns of skill appreciation or skill decay. Third, PIAAC's rich background questionnaire contains several variables that are likely correlated with both skills and labor-market outcomes, but typically remain unobserved in administrative labor-market records (e.g., a person's health status or her parents' education level). We investigate whether controlling for these variables changes the estimated returns to skills, which would indicate that part of the relationship between skills and labor-market outcomes is attributable to these other (typically unobserved) variables. In this analysis, we consider both wages and employment to shed light on the effect of skills at the intensive and extensive margin. While much of the literature focuses on wages only, increased employment may constitute another important dimension of potential returns in terms of participation in the labor market, which may have important repercussions for societal participation more generally.

Our baseline least squares estimations of the returns to skills suggest that going up one (out of five) PIAAC proficiency levels in numeracy skills is associated with an average increase in hourly wages of about 20 percent and an increase in the likelihood of being employed of about 8% points on average across the participating countries. There is a wide variation in the returns across the 32 countries in our sample, though, ranging from wage increases of 10% in Greece to 47% in Singapore and from employment increases of 2.4% in Indonesia (Jakarta) to 14.2% in Spain. Estimated returns to skills in the different IV models are consistently larger than estimates derived from least squares models. When addressing measurement error in the skill variable, estimated wage returns in the pooled sample increase by approximately 10%, while they more than double in size in specifications dealing with reverse causality. Moreover, estimated returns only slightly decrease when we control for parental education or health as potential omitted variables, suggesting that the empirical relevance of concerns from omitting family-background and health measures may be limited.

One of our most striking findings is that the described pattern of results across the different specifications is remarkably consistent across the diverse set of countries in our sample.[1] In fact, most of the IV models lead to larger estimates of the returns to skills in every single participating country. Results are also robust to a number of alternative specifications, including different samples and additional controls.

[1] Real GDP per capita (at constant national prices) ranges from $18,609 in Turkey to $82,297 in Norway, a difference by a factor of 4.

The second part of our analysis focuses on the United States and Germany, as these countries provide two "natural experiments" that induce quasi-exogenous variation in skills. These approaches more credibly identify causal effects than the above approaches that separately address the main types of possible bias. The first approach, suggested by Hanushek et al. (2015), exploits statewide compulsory schooling requirements that led to changes in educational attainment and, therefore, skills in the United States. These state-level changes in schooling requirements can be used as instrumental variables to examine the impact of skills on wages.[2] Identification of these effects is achieved by exploiting variation in the timing of the law changes across states over time such that different birth cohorts within each state have different compulsory schooling requirements.

While changes in compulsory schooling laws across states over time are likely to affect skills in general, Germany provides a unique setting to investigate the wage effect of domain-specific skills, namely, the capacity to master information and communication technologies (i.e., ICT skills). It has recently been argued that ICT skills are central in modern labor markets and, according to the former Vice President of the European Commission, Neelie Kroes, can be regarded as "the new literacy".[3] However, existing evidence on the returns to ICT skills is scarce and purely descriptive because of the difficulty to find a source of variation in ICT skills that is independent of a person's overall ability. Falck et al. (2016) use technological peculiarities that led to variation in broadband availability at a very fine regional level within Germany, inducing differences in ICT skills developed by performing ICT-related tasks. Specifically, in traditional telephone networks, the distance between a household and the main network node ("last mile") was irrelevant for the quality of voice-telephony services; however, when these networks became the basis for broadband Internet, the last-mile distance turned out to play a crucial role for broadband availability. Beyond a certain distance threshold, high-speed Internet access was not feasible without major infrastructure investment, a situation that excluded a considerable share of German municipalities from early broadband Internet access. The variation in ICT skills induced by differences in early broadband access is independent of a person's overall ability and can therefore be used to estimate a plausibly causal effect of ICT skills on wages.

The evidence on the returns to skills from the natural experiments in the United States and Germany corroborates the findings from the international analysis; estimated returns to skills in the IV models are again considerably larger than the least squares estimates. This suggests that the least squares results may provide a lower bound of the true returns to skills in the labor market.

This striking pattern of results, holding for various sources of exogenous variation in skills and across different contexts, yields important implications for policy. It suggests that policies of skill development—even if based on standard (least squares) results on

[2] Previous literature has also used compulsory schooling laws to investigate the effect of increased schooling on mortality, incarceration, and the social returns to schooling, among others (Acemoglu and Angrist, 2001; Lochner and Moretti, 2004; Lleras-Muney, 2005).

[3] http://getonlineweek.eu/vice-president-neelie-kroes-says-digital-literacy-and-e-skills-are-the-new-literacy/ . Accessed September 19, 2016).

skill returns—are not pursuing overly optimistic outcomes. However, this paper is only the starting point toward gaining a better understanding of causality in the estimation of returns to skills, and substantially more work needs to be done to show the robustness and generalizability of our results.

The paper proceeds as follows: "Previous literature on labor-market returns to human capital" summarizes the previous literature on the returns to human capital in the labor market. "The PIAAC data" briefly describes the PIAAC data. "Empirical strategy" outlines the empirical strategy for the returns-to-skills estimations and discusses the main potential threats to causal identification of the relationship between skills and labor-market outcomes. "Returns to general skills: explorations into causality" presents results on the returns to general skills from empirical models that separately deal with the main types of possible bias, including evidence from a natural experiment that exploits changes in compulsory schooling laws in the United States. "Returns to ICT Skills: Evidence from Peculiarities in Broadband Technology in Germany" presents results on the returns to a domain-specific skill, namely ICT skills, from a natural experiment that exploits peculiarities in broadband technology across German municipalities. "Conclusion" concludes.

Previous literature on labor-market returns to human capital

Starting with the seminal work of Becker (1962) and Mincer (1974), a substantial body of research has shown that human capital has positive effects on an individual's labor-market success. Most analyses rely on the Mincerian wage regression, derived from a theoretical framework of optimal human capital investment, which allows estimating the rate of return to schooling (Mincer, 1970, 1974).[4] A large amount of evidence exists on the returns to schooling, and the overwhelming majority of studies find a positive relationship between schooling and individual earnings: on average, an additional year of schooling is associated with roughly a 10% increase in earnings (Psacharaopoulos and Patrinos, 2004). However, the authors also show that estimated returns to schooling vary significantly between studies and contexts; for instance, returns appear to be higher in low-income countries, for women, and for lower levels of schooling.

While most of the early evidence on the returns to education has been purely descriptive, more recent studies try to tackle potential endogeneity issues in the returns estimation.[5] These studies aim to give a causal interpretation to estimated returns to education by exploiting variation in education stemming from changes in compulsory schooling laws and in restrictions on child labor, variation in education stemming from differences

[4] In particular, the Mincer equation models the logarithm of individual earnings as a function of years of formal education, a quadratic polynomial in years of (potential) experience and potentially other covariates. The Mincer earnings function is widely used in empirical economics to estimate the returns to formal education. See Heckman et al. (2006) for a discussion under which conditions the coefficient on schooling in a Mincer equation estimates the rate of return to schooling.

[5] The term "endogeneity" stems from the idea that a certain variable (e.g., education) cannot be viewed as exogenous to the model of interest, as it should be, but that it is rather endogenously determined within the model—depending on the outcome (i.e., reverse causality) or being jointly determined with the outcome by a third factor (i.e., omitted variable bias). Because of the problem of endogeneity, estimates of the association between the variable and outcome based on correlations will be biased estimates of the causal effect of the variable on the outcome. We describe potential endogeneity problems in the returns to skills estimation and approaches to tackle these issues in "Empirical strategy".

in the distance to the nearest educational institution, and variation in education occurring between siblings and twins.[6] Frequently, pursuing these more demanding identification strategies even leads to larger estimates of the returns to education (e.g., Oreopoulos, 2006).[7] The overall conclusion of earlier studies estimating simple Mincerian wage equations, however, is confirmed: education has a strong causal impact on earnings.[8]

While the empirical literature on returns to education relies almost exclusively on school attainment as a measure of human capital, such measure may in fact be a poor approximation of an individual's effective human capital. In recent work on the macroeconomic effect of human capital on a country's economic growth, it has been shown that educational outcomes (the cognitive skills people have actually learned), not just attainment (how long people stayed in school), are more reliable proxies of human capital. Hanushek and Woessmann (2012, 2015) measure a country's stock of human capital as the average test score on all international student achievement tests in math and science between 1964 and 2003. Estimating cross-country growth regressions, they find strong support for a positive association between human capital and long-run growth. When the stock of human capital is instead measured by the average years of schooling of the population, the association with economic growth is much weaker, and the model accounts for only one quarter of the cross-country variation in long-run growth (rather than three quarters with achievement). In fact, once differences in achievement are taken into account, there is no separate relationship whatsoever between years of schooling and economic growth. Several rigorous analyses, detailed in Hanushek and Woessmann (2012, 2015), indicate that the achievement-growth picture indeed depicts a causal effect of better educational achievement on economic growth. The results suggest that the quantity of education matters for growth only insofar as it in fact leads to better knowledge and skills of the population. It is what people know and can do that matters for economic growth, not how long it took them to reach that achievement. This evidence strongly calls for a focus on educational outcomes, not just attainment.

However, unlike the case of the returns to school attainment, analysis of the returns to cognitive skills on the labor market has had to rely on a small number of specialized data sets. While assessments of the achievement of students are common, tested students are seldom followed from school into the labor market where the impact of differential skills can be observed. In fact, evidence incorporating direct measures of cognitive skills is mostly restricted to early-career workers in the United States.[9] A notable exception is the work based on the international IALS data of adult skills in the mid-1990s.[10]

6 See, for instance, Angrist and Krueger (1991), Harmon and Walker (1995), Ashenfelter and Rouse (1998) and Oreopoulos (2006). Card (1999) and Heckman et al. (2006) provide comprehensive overviews and Woessmann (2016) provides a less technical summary.

7 Recent evidence suggests, however, that estimates based on changes in compulsory schooling laws may in fact be smaller, after all; see Pischke and von Wachter (2008), Devereux and Hart (2010), Grenet (2013), and Stephens and Yang (2014).

8 Education may also exhibit non-monetary returns at the individual as well as the societal level, including higher fringe benefits, higher job satisfaction, reduced crime, improved health, and good citizenship (for reviews, see Lochner, 2011 and Oreopoulos and Salvanes, 2011).

9 See, for example, Bishop (1989), Murnane et al. (1995), Neal and Johnson (1996), Mulligan (1999), Murnane et al. (2000), Lazear (2003), and Chetty et al. (2011). Bowles et al. (2001) provide an early survey of studies of achievement effects, and Hanushek and Woessmann (2008) and Hanushek and Rivkin (2012) survey more recent evidence.

10 See, for example, Leuven et al. (2004) and Hanushek and Zhang (2009), among others; see the review in Hanushek and Woessmann (2011).

More recently, using data from the PIAAC survey of adult skills over the full lifecycle in 23 countries in 2011–12, Hanushek et al. (2015) show that the focus on early-career workers in previous studies leads to an underestimation of the actual returns to skills by about one quarter. For prime age workers, going up one (out of five) PIAAC proficiency levels is associated with an 18% increase in hourly wages.[11]

The PIAAC data

PIAAC was developed by the OECD and the data were collected between August 2011 and March 2012 (first round) and between April 2014 and March 2015 (second round). PIAAC provides internationally comparable data about skills of the adult populations in 33 countries.[12] In each country, at least 5000 adults participated in the PIAAC assessment, providing considerably larger samples than in IALS, the predecessor of PIAAC. In each participating country, a representative sample of adults between 16 and 65 years of age was interviewed at home in the language of their country of residence. The standard survey mode was to answer questions on a computer, but for respondents without computer experience or sufficient computer knowledge there was also the option of a pencil-and-paper interview.

PIAAC was designed to measure key cognitive and workplace skills needed for individuals to advance in their jobs and participate in society. The survey included an assessment of cognitive skills in three domains: numeracy, literacy, and ICT (called "problem solving in technology-rich environments" in PIAAC).[13] The tasks respondents had to solve were often framed as real-world problems, such as maintaining a driver's logbook (numeracy domain) or reserving a meeting room on a particular date using a reservation system (ICT domain). The domains , described in more detail in OECD (2013), refer to key information-processing competencies and are defined as:

Literacy: Ability to understand, evaluate, use and engage with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential.

Numeracy: 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;

ICT skills: Ability to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks.

[11] Other research using the PIAAC data investigates—among others—the effect of teacher skills on student achievement (Hanushek et al., 2014), the role of skill mismatch for earnings (Levels et al., 2014; Perry et al., 2014), skill depreciation over the lifecycle (Barrett and Riddell, 2016), and the effect of vocational education on lifecycle employment (Hampf and Woessmann, 2016).

[12] Participating countries in the first round were Australia, Austria, Belgium (Flanders), Canada, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, the Netherlands, Norway, Poland, the Russian Federation, the Slovak Republic, Spain, Sweden, the United Kingdom (specifically England and Northern Ireland), and the United States. In the second round, the following countries participated: Chile, Greece, Indonesia (Jakarta only), Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey. We do not include the Russian Federation in the subsequent analyses because its data are still subject to change and are not representative of the entire population because of the lack of the Moscow municipal area (OECD, 2013).

[13] Participation in the ICT domain was optional; Cyprus, France, Italy, and Spain (first round) as well as Indonesia (second round) did not participate in this domain.

PIAAC measures each of the three skill domains on a 500-point scale.[14] All three scales are intended to measure different dimensions of a respondent's skill set, although a person who performs well in literacy usually tends to have relatively higher numeracy and ICT scores, too. IALS suffered from pairwise correlations of individual skill domains that exceeded 0.9, making it virtually impossible to distinguish between different skills. The skill domains in PIAAC are less strongly correlated with an individual-level correlation between numeracy and literacy (ICT skills) of 0.86 (0.66); the correlation between literacy and ICT skills is at 0.70.[15]

Before the skill assessment, all participants responded to a background questionnaire that gathered information about labor-market status, earnings, education, experience, and demographic characteristics of the respondents. The measure of experience refers to actual work experience and was collected as the number of years where at least six months were spent in paid work.

Following Hanushek et al. (2015), most part of our analysis focuses on workers aged 35–54 who are full-time employed,[16] because prime-age earnings best approximate lifetime earnings.[17] In the econometric analysis, we standardize skills to have mean zero and standard deviation (SD) one and always employ the sample weights provided in PIAAC. In the pooled sample of 32 countries, one SD in numeracy skills is 53 PIAAC points, which is roughly equivalent to one out of five proficiency levels in PIAAC.[18] Note that one SD in numeracy skills is about twice the learning progress made by school-attending PIAAC respondents between lower secondary and upper secondary education, which amounts to 24 points across the countries in our sample.[19]

Empirical strategy

Following Hanushek et al. (2015) and Falck et al. (2016), we estimate returns to skills in a general Mincer framework that relates a person's human capital to earnings in the labor market (see "Previous literature on labor-market returns to human capital"). Specifically, we estimate the following individual-level wage regression:

[14] PIAAC provides 10 plausible values for each respondent and each skill domain. We employ all plausible values in the least squares estimations in Tables 1 and 2 (using Stata's repest command). In the instrumental-variable models, however, there is no straightforward way to add the imputation error to the variance estimator allowing the computation of correct standard errors. Therefore, we use the first plausible value of the PIAAC scores in each domain in the instrumental-variable estimations. We carefully checked whether using just the first plausible value affects our results, and found estimated returns to skills to be very similar across the range of plausible values. See Additional file 1: Tables S1 and S2 for the estimation results using each of the 10 plausible values in the pooled sample. Below, we also report results using other plausible values as instruments for the first plausible value. See also Perry et al. (2014) for a discussion of the plausible values in PIAAC.

[15] These numbers refer to the pooled sample of full-time employees aged 35–54 years.

[16] Full-time employees are defined as those working at least 30 hours per week. Since Australia and Austria did not publish information on working hours in the PIAAC Public Use File, the full-time working status is based on a question of whether a respondent works full-time. The Canadian sample includes full-time and part-time workers because the available data do not report working hours or work status.

[17] For obvious reasons, we do not restrict the sample to full-time workers in the employment regressions. In the ICT-skills analysis, which is restricted to West German municipalities ("Returns to ICT Skills: Evidence from Peculiarities in Broadband Technology in Germany"), we also include part-time workers and expand the considered age range to 20–65 years to be able to exploit more variation in ICT skills.

[18] For descriptive statistics on participants' characteristics for each PIAAC country, see Table 1 in Hanushek et al. (2015) and Table A-1 in Hanushek et al. (2017b).

[19] We calculated this "ISCED-level equivalent" by regressing numeracy skills of PIAAC respondents aged 16–18 years in the 32 sample countries on an indicator that takes the value 1 if the respondent is currently in upper secondary education (ISCED 3A-B, C long); 0 if the respondent is currently in lower secondary education (ISCED 2, 3C short). Regressions control for gender, age, number of books at home, a migrant indicator, and country fixed effects. The estimate provides an approximation of how much students learn on average transiting from lower secondary to upper secondary education.

$$\log y_{in} = \beta_0 + \beta_1 C_{in} + \boldsymbol{X_{in}\beta_2} + \varepsilon_{in}. \tag{1}$$

Depending on the specification, y_{in} is either gross hourly wages[20] earned by individual i living in country n or the individual's employment status.[21] C_{in} refers to the individual's cognitive skills measured in PIAAC. $\boldsymbol{X_{in}}$ is a vector of individual-level variables including gender and a quadratic polynomial in actual work experience (in the specifications with wage as outcome) or in age (in the specifications with employment status as outcome).[22] We estimate labor-market returns to skills without accounting for years of schooling, which is one of several inputs into cognitive skills.[23] ε_{ic} is a standard error term. The coefficient of interest is β_1, which shows the wage change in percent or the change in the employment probability in percentage points when skills increase by one SD.[24]

In this basic regression framework, β_1 cannot necessarily be interpreted as the causal effect of cognitive skills on labor-market success. The most obvious reasons for β_1 being a biased estimate of the true returns to skills are measurement error, reverse causality, and omitted variables (for a discussion, see also Hanushek et al., 2015). Measurement error may occur if the skills measured in PIAAC are an error-ridden measure of the human capital relevant in the labor market. Errors in the measurement of cognitive skills can also occur if PIAAC respondents had a bad testing day or solved tasks correctly or incorrectly simply by chance. Such measurement error in the assessment of an individual's skills will bias the coefficient on skills β_1 toward zero. Moreover, higher earnings may actually lead to improvements in skills, giving rise to the problem of reverse causality. Higher-paying jobs may more likely require and reinforce skills or they may provide the resources to invest in adult education and training. Reverse causality will likely lead to an upward bias of the returns-to-skills estimate β_1. Finally, omitted-variable bias may arise because unobserved variables like non-cognitive skills, personality traits, family background, or health status could directly influence earnings or employment prospects and may also be related to cognitive skills. A positive (negative) correlation of skills measured in PIAAC with other unobserved variables that are valued on the labor market would bias the least squares estimate β_1 upward (downward).

Our main approach to address these endogeneity problems is instrumental-variable (IV) estimation (see Stock and Watson, 2007, for a textbook treatment). This approach allows for consistent estimation even when the explanatory variable in a regression model (here: cognitive skills) is endogenous, that is, when it is correlated with the error

[20] The PIAAC Public Use File reports hourly wages only in deciles for Austria, Canada, Germany, Sweden, and the United State in the first round, as well as for Singapore and Turkey in the second round. For Germany, we obtained the Scientific Use File, which contains continuous wage information. For the other countries, we assign the median wage of each decile of the country-specific wage distribution (obtained from the OECD) to each person belonging to the respective decile. Hanushek et al. (2015) show that using decile medians has no substantive impact on estimated returns to skills for those countries with continuous wage data. To limit the influence of outliers, we trim the bottom and top one percent of the wage distribution in each country with continuous earnings information.

[21] In accordance with the International Labour Organization (ILO), employment in the PIAAC survey is defined as having paid work for at least 1 h in the week before the survey.

[22] In the pooled estimation, we also add country fixed effects so that all estimates rely just on within-country variation.

[23] See Hanushek et al. (2015) for an extensive discussion of the problems of interpreting the coefficient on years of schooling in a wage regression that also contains cognitive skills.

[24] For ease of exposition, we frequently refer to β_1 simply as the "return to skill". It does not, however, correspond to a rate of return calculation because we have no indication of the cost of achieving any given level of skill (see also Heckman et al., 2006).

term and hence an OLS regression yields a biased estimate of the true coefficient. An instrument is a variable that is correlated with the endogenous regressor but has no independent association with the dependent variable of interest (here: wage, employment). In other words, the instrument neither has a direct effect on the outcome variable nor is it related to the outcome through a channel other than the endogenous regressor. Hence, an instrument allows isolating variation in the explanatory variable that is uncorrelated with the error term, eliminating any part of the variation that may suffer from endogeneity bias.[25]

Our Mincerian wage Eq. (1) is likely to yield biased estimates of the true effect of individual skills, β_1, because the skill variable is correlated with the error term, that is, $Cov(C_{in}, \varepsilon_{in}) \neq 0$. Hence, we need to find a valid instrument Z_{in} which satisfies two conditions, known as *instrument relevance* and *instrument exogeneity*. If an instrument is relevant, the variation in Z_{in} is linked to the variation in C_{in}, that is, $corr(Z_{in}, C_{in}) \neq 0$. As a rule of thumb, the F statistic testing the hypothesis that the coefficient on the instrument Z_{in} in an equation that regresses C_{in} on the instrument Z_{in} and exogenous regressors ("first stage") is zero is supposed to be larger than 10.[26] In addition, the instrument has to be uncorrelated with the error term in the original estimation equation, that is, $corr(Z_{in}, \varepsilon_{in}) = 0$. This exogeneity condition cannot be directly tested due to missing unbiased estimates for ε_{in} and requires making a judgement based on personal knowledge and common sense.

Typically, the IV model is implemented using a two stage least squares (2SLS) estimator. This estimator is calculated in two steps, the first stage and second stage. In the first-stage estimation, the endogenous regressor from Eq. (1), C_{in}, is regressed on the instrument Z_{in} and all exogenous regressors captured in the $\boldsymbol{X}$ vector:

$$C_{in} = \pi_0 + \pi_1 Z_{in} + \boldsymbol{X_{in}\pi_2} + \upsilon_{in}. \tag{2}$$

The key idea is that the first stage isolates a part of the variation in C_{in} that is uncorrelated with ε_{in}, thereby overcoming problems such as reverse causality and omitted variables and achieving consistent estimation. The causal effect of C on y is obtained from the second stage of the 2SLS model, where y is regressed on the predicted values (here: predicted skills) from the first-stage estimation of C_{in}, denoted by $\hat{C}_{in}$, and control variables:

$$\log y_{in} = \beta_0^{2SLS} + \beta_1^{2SLS}\hat{C}_{in} + \boldsymbol{X_{in}\beta_2^{2SLS}} + \omega_{in}. \tag{3}$$

After having outlined the basic idea of the IV approach, we now describe how we use this model to address the sources of potential bias in the returns-to-skills estimation.

Returns to general skills: explorations into causality

We start by exploring issues of causality in estimating returns to general cognitive skills across the 32 PIAAC countries. The analysis focuses on numeracy skills, which we deem most comparable across countries, as, e.g., skill tests are less affected by cross-country differences in language complexity than literacy skills.[27]

[25] See Schlotter et al. (2011) for a non-technical discussion of IV estimation and example applications in the field of education.

[26] For a discussion, see Staiger and Stock (1997) and Stock et al. (2002).

[27] Hanushek et al. (2015) find that results are generally quite similar for literacy skills.

Evidence addressing different potential biases in the international sample

Table 1 reports results on the returns to numeracy skills in terms of hourly wages in the 32 PIAAC countries, using different specifications to address potential bias from measurement error, reverse causality, and omitted variables, respectively. Each cell in Table 1 reports the coefficient on numeracy skills from a separate regression. Row (1) provides the baseline least squares estimate on the returns to numeracy skills without any correction for sources of possible bias. We find that a one SD increase in numeracy skills is associated with an increase in wages of 20% in the pooled country sample. But the estimated returns vary substantially across countries, ranging from 10% in Greece to 47% in Singapore.[28] Despite these cross-country differences in the returns to skills, we observe that skills are significantly rewarded in *all* countries participating in PIAAC.[29]

Similarly, Table 2 shows how numeracy skills are related to the probability of being employed. One reason why skills would affect employment is that individuals with higher earnings potential (due to higher skills) are more likely to choose to participate in the labor market. Another reason would be that low-skilled people are less likely to find a job in labor markets with effective minimum wages. In the baseline specification [row (1)] for the pooled country sample, the probability of being employed increases by 7.9% points when numeracy skills increase by one SD. Estimated returns in terms of employment range from 2.4% points in Indonesia to 14% points in the Slovak Republic and Spain. One potential reason for the strong association between skills and employment in the latter countries could be their currently high rates of non-employment. Despite this country heterogeneity, the association between skills and employment prospects is again significant in each country.

However, as discussed above, these returns-to-skills estimates are unlikely to reflect the causal effect of skills on labor-market outcomes. In rows (2)–(7) of Tables 1 and 2, we deal with the different sources of possible bias consecutively.

Measurement error

As is well known, tests differ in how reliably they measure underlying domains of cognitive skills, and the implied errors can bias the estimates of the returns to skills. Perhaps the most straightforward way to address possible attenuation bias arising from errors in the measurement of skills is to use two measures of the same concept in an IV approach. In the PIAAC setting with multiple tests, we can use literacy skills as an instrument for numeracy skills. This approach essentially takes the variation that is common to both skill measures as the relevant cognitive dimension.

[28] Part of this country heterogeneity in estimated returns can be attributed to a country's institutional environment reflected by union density, strictness of employment protection, and the size of the public sector (Hanushek et al., 2015). In addition, Hanushek et al. (2017b) show that returns to skill are larger in countries with faster prior economic growth, consistent with models where skills are particularly important for adaptation to dynamic economic change.

[29] Note that the US estimate differs slightly from the estimate in Table 2 in Hanushek et al. (2015) who show results for continuous earnings after wage trimming (obtained from the US National Center for Education Statistics). Moreover, estimated returns in PIAAC round 2 countries are not precisely comparable to those reported in Hanushek et al. (2017b) because they estimated the skill gradient using age instead of actual labor market experience. However, results with either approach are very similar.

Table 1 Returns to skills: international evidence. *Data source:* PIAAC 2016

	Pooled	Australia	Austria	Belgium	Canada	Chile	Cyprus	Czech R.	Denmark	Estonia	Finland
Baseline (OLS)	.200***	.205***	.182***	.151***	.204***	.367***	.141***	.122***	.135***	.177***	.140***
	(.003)	(.016)	(.012)	(.009)	(.009)	(.033)	(.020)	(.018)	(.009)	(.013)	(.012)
Measurement error—literacy (2SLS)	.221***	.216***	.218***	.169***	.231***	.378***	.173***	.159***	.156***	.168***	.154***
	(.003)	(.014)	(.013)	(.011)	(.009)	(.029)	(.023)	(.022)	(.009)	(.015)	(.013)
	[61223]	[2824]	[2096]	[3141]	[7334]	[649]	[847]	[581]	[5153]	[3096]	[2812]
Measurement error—second PV (2SLS)	.232***	.250***	.211***	.184***	.242***	.402***	.182***	.149***	.162***	.209***	.164***
	(.003)	(.015)	(.012)	(.011)	(.009)	(.030)	(.024)	(.019)	(.010)	(.015)	(.014)
	[65811]	[2264]	[1859]	[2687]	[5054]	[2238]	[1142]	[1377]	[3346]	[2873]	[2062]
Reverse causality—years school (2SLS)	.517***	.492***	.471***	.322***	.489***	.673***	.752***	.313***	.370***	.450***	.477***
	(.007)	(.039)	(.036)	(.022)	(.023)	(.044)	(.083)	(.040)	(.023)	(.032)	(.036)
	[7569]	[197]	[179]	[324]	[553]	[342]	[93]	[149]	[382]	[369]	[170]
Reverse causality—parents (2SLS)	.461***	.516***	.384***	.315***	.338***	.635***	.502***	.255***	.286***	.487***	.364***
	(.012)	(.083)	(.045)	(.041)	(.032)	(.085)	(.117)	(.056)	(.040)	(.064)	(.080)
	[1862]	[48]	[69]	[79]	[199]	[47]	[38]	[41]	[94]	[95]	[24]
Omitted variables—parents (OLS)	.182***	.189***	.165***	.138***	.194***	.324***	.129***	.112***	.127***	.155***	.137***
	(.003)	(.017)	(.012)	(.010)	(.009)	(.034)	(.021)	(.019)	(.010)	(.015)	(.012)
Omitted variables—health (OLS)	.194***	.204***	.173***	.149***	.204***	.335***	.132***	.115***	.132***	.164***	.135***
	(.003)	(.016)	(.011)	(.009)	(.009)	(.035)	(.020)	(.017)	(.009)	(.014)	(.012)
Observations	45064	1433	1115	1221	7197	903	938	1066	1879	1763	1480

Table 1 continued

	France	Germany	Greece	Indonesia	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.
Baseline (OLS)	.175***	.239***	.101***	.245***	.230***	.273***	.135***	.181***	.210***	.172***	.180***
	(.009)	(.015)	(.019)	(.036)	(.023)	(.020)	(.018)	(.014)	(.019)	(.016)	(.011)
Measurement error—literacy (2SLS)	.179***	.268***	.106***	.331***	.259***	.321***	.154***	.162***	.235***	.169***	.217***
	(.010)	(.015)	(.024)	(.049)	(.023)	(.021)	(.020)	(.016)	(.017)	(.019)	(.014)
	[4326]	[2976]	[601]	[538]	[1583]	[1551]	[1515]	[2778]	[4204]	[1982]	[2252]
Measurement error—second PV (2SLS)	.200***	.274***	.137***	.306***	.280***	.320***	.145***	.221***	.249***	.198***	.215***
	(.010)	(.015)	(.024)	(.043)	(.021)	(.020)	(.019)	(.017)	(.018)	(.018)	(.014)
	[4940]	[2249]	[763]	[661]	[1152]	[2233]	[1535]	[1862]	[2211]	[1590]	[2082]
Reverse causality—years school (2SLS)	.361***	.456***	.455***	.747***	.512***	.572***	.423***	.478***	.563***	.600***	.456***
	(.019)	(.030)	(.068)	(.080)	(.042)	(.042)	(.045)	(.037)	(.040)	(.054)	(.039)
	[524]	[361]	[62]	[190]	[172]	[190]	[220]	[266]	[332]	[156]	[178]
Reverse causality—parents (2SLS)	.354***	.370***	.448***	1.381***	.509***	.562***	.466***	.717***	.540***	.401***	.567***
	(.037)	(.049)	(.146)	(.394)	(.074)	(.075)	(.093)	(.131)	(.068)	(.062)	(.096)
	[129]	[86]	[13]	[12]	[64]	[68]	[51]	[32]	[82]	[67]	[30]
Omitted variables—parents (OLS)	.157***	.228***	.091***	.210***	.214***	.251***	.114***	.166***	.187***	.150***	.173***
	(.010)	(.017)	(.019)	(.034)	(.025)	(.021)	(.020)	(.015)	(.019)	(.017)	(.012)
Omitted variables—health (OLS)	.172***	.234***	.099***	.231***	.230***	.264***	.135***	.180***	.208***	.158***	.176***
	(.009)	(.015)	(.019)	(.036)	(.023)	(.019)	(.018)	(.014)	(.019)	(.016)	(.011)
Observations	1715	1296	623	806	1033	908	1019	1319	1441	1260	1012

Table 1 continued

	New Zealand	Norway	Poland	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Baseline (OLS)	.191***	.132***	.190***	.466***	.188***	.186***	.225***	.130***	.208***	.220***	.254***
	(.013)	(.008)	(.018)	(.015)	(.021)	(.012)	(.017)	(.009)	(.035)	(.015)	(.028)
Measurement error—literacy (2SLS)	.206***	.131***	.250***	.444***	.184***	.222***	.261***	.139***	.273***	.262***	.278***
	(.013)	(.008)	(.021)	(.014)	(.023)	(.013)	(.020)	(.010)	(.036)	(.016)	(.022)
	[2848]	[4754]	[1483]	[8089]	[1782]	[3120]	[2265]	[2646]	[509]	[1978]	[2947]
Measurement error—second PV (2SLS)	.221***	.152***	.221***	.501***	.225***	.224***	.259***	.143***	.246***	.247***	.274***
	(.012)	(.009)	(.020)	(.015)	(.023)	(.013)	(.018)	(.009)	(.032)	(.015)	(.023)
	[3184]	[2990]	[1587]	[7153]	[1709]	[2484]	[2350]	[2489]	[703]	[2484]	[3389]
Reverse causality—years school (2SLS)	.460***	.330***	.642***	.723***	.742***	.506***	.653***	.230***	.804***	.606***	.487***
	(.036)	(.026)	(.055)	(.025)	(.071)	(.030)	(.040)	(.023)	(.096)	(.058)	(.033)
	[217]	[257]	[161]	[1055]	[142]	[405]	[322]	[272]	[87]	[117]	[473]
Reverse causality—parents (2SLS)	.363***	.278***	.740***	.747***	.654***	.461***	.676***	.201***	.649***	.408***	.443***
	(.063)	(.036)	(.133)	(.062)	(.091)	(.043)	(.117)	(.035)	(.149)	(.050)	(.052)
	[43]	[82]	[31]	[133]	[66]	[135]	[34]	[55]	[34]	[120]	[131]
Omitted variables—parents (OLS)	.182***	.120***	.168***	.445***	.148***	.157***	.206***	.127***	.193***	.188***	.213***
	(.014)	(.009)	(.018)	(.016)	(.020)	(.014)	(.019)	(.009)	(.036)	(.019)	(.030)
Omitted variables—health (OLS)	.186***	.128***	.183***	.456***	.182***	.178***	.224***	.127***	.208***	.214***	.236***
	(.013)	(.008)	(.018)	(.015)	(.021)	(.013)	(.018)	(.008)	(.035)	(.014)	(.027)
Observations	1204	1520	817	1507	1193	1306	1192	1316	674	1787	1121

Dependent variable: log gross hourly wage. Each cell reports the coefficient on numeracy skills from a separate regression. Regressions weighted by sampling weights. Model names indicate bias addressed by the specification; "OLS" refers to ordinary least squares estimations, "2SLS" refers to two-stage least squares estimations. Least squares estimates (except for pooled specification) take into account all 10 plausible values of numeracy skills. Sample: full-time employees aged 35–54 years. All regressions control for a quadratic polynomial in actual work experience and gender. *Baseline* standard model controlling for work experience and gender. *Measurement error—literacy* numeracy skills are instrumented by literacy skills. *Measurement error—second PV* numeracy skills are instrumented by second plausible value of numeracy skills. *Reverse causality—years school* numeracy skills are instrumented by years of schooling. *Reverse causality—parents* numeracy skills are instrumented by parental education (*1* neither parent attained upper secondary education, *2* at least one parent attained upper secondary education, *3* at least one parent attained tertiary education). *Omitted variables—parents* adds control for parental education. *Omitted variables—health* adds control for respondent's health (*1* poor, *2* fair, *3* good, *4* very good, *5* excellent). Observations refer to baseline specification. Numeracy skills standardized to SD 1 within country. Pooled specification includes country fixed effects and gives same weight to each country. Robust standard errors (jacknifed in the OLS models due to the complex sampling design) in parentheses. F-statistics of excluded instrument in the first stage of 2SLS models in brackets. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2 Employment effects of skills: international evidence. *Data source:* PIAAC 2016

	Pooled	Australia	Austria	Belgium	Canada	Chile	Cyprus	Czech R.	Denmark	Estonia	Finland
Baseline (OLS)	.079*** (.002)	.081*** (.011)	.042*** (.011)	.057*** (.009)	.083*** (.007)	.103*** (.016)	.080*** (.012)	.048*** (.016)	.105*** (.010)	.087*** (.009)	.080*** (.012)
Measurement error—literacy (2SLS)	.079*** (.002) [131508]	.078*** (.011) [6472]	.051*** (.011) [4747]	.065*** (.009) [6473]	.080*** (.008) [12005]	.099*** (.016) [1398]	.069*** (.014) [2108]	.056*** (.021) [1210]	.116*** (.010) [9153]	.089*** (.009) [6076]	.084*** (.012) [4372]
Measurement error—second PV (2SLS)	.092*** (.002) [135446]	.101*** (.011) [6105]	.047*** (.012) [4548]	.067*** (.009) [5839]	.094*** (.008) [9848]	.116*** (.015) [3360]	.088*** (.013) [2840]	.056*** (.018) [2356]	.121*** (.010) [5725]	.092*** (.009) [5517]	.099*** (.013) [3123]
Reverse causality—years school (2SLS)	.171*** (.004) [15746]	.158*** (.021) [484]	.080*** (.024) [375]	.141*** (.015) [676]	.159*** (.014) [932]	.147*** (.019) [713]	.243*** (.028) [296]	.117*** (.029) [248]	.179*** (.018) [675]	.191*** (.017) [715]	.147*** (.021) [282]
Reverse causality—parents (2SLS)	.095*** (.007) [3873]	.091** (.042) [129]	.029 (.038) [127]	.116*** (.023) [213]	.141*** (.023) [283]	.136*** (.041) [73]	.132*** (.051) [83]	−.066 (.060) [72]	.061* (.034) [155]	.118*** (.036) [145]	−.056 (.070) [29]
Omitted variables—parents (OLS)	.075*** (.002)	.074*** (.012)	.040*** (.011)	.051*** (.009)	.071*** (.008)	.105*** (.019)	.078*** (.012)	.047*** (.015)	.106*** (.010)	.084*** (.009)	.079*** (.012)
Omitted variables—health (OLS)	.066*** (.002)	.068*** (.011)	.021** (.011)	.037*** (.008)	.083*** (.007)	.083*** (.017)	.070*** (.012)	.032** (.016)	.080*** (.009)	.072*** (.009)	.064*** (.011)
Observations	85800	3214	2248	2149	11,583	2018	1810	1865	2747	3091	2044

Table 2 continued

	France	Germany	Greece	Indonesia	Ireland	Israel	Italy	Japan	Korea	Lithuania	Netherl.
Baseline (OLS)	.082***	.104***	.032***	.024**	.105***	.104***	.098***	.033***	.029***	.091***	.088***
	(.008)	(.010)	(.012)	(.011)	(.011)	(.010)	(.014)	(.009)	(.010)	(.012)	(.010)
Measurement error—literacy (2SLS)	.073***	.108***	−.004	.043***	.103***	.120***	.095***	.006	.018*	.060***	.094***
	(.010)	(.010)	(.013)	(.012)	(.012)	(.012)	(.013)	(.012)	(.010)	(.014)	(.011)
	[8142]	[5965]	[2848]	[1771]	[4951]	[4057]	[3385]	[4755]	[7949]	[3103]	[5077]
Measurement error—second PV (2SLS)	.088***	.115***	.050***	.031***	.121***	.113***	.119***	.042***	.037***	.108***	.095***
	(.009)	(.010)	(.014)	(.011)	(.012)	(.011)	(.013)	(.012)	(.011)	(.014)	(.012)
	[8099]	[6354]	[2309]	[2338]	[3585]	[5501]	[3807]	[3461]	[4880]	[3291]	[4646]
Reverse causality—years school (2SLS)	.137***	.120***	.190***	.050***	.254***	.241***	.248***	.033*	.031*	.218***	.179***
	(.015)	(.016)	(.029)	(.019)	(.022)	(.019)	(.023)	(.019)	(.019)	(.025)	(.019)
	[1178]	[829]	[325]	[576]	[400]	[549]	[458]	[415]	[611]	[312]	[438]
Reverse causality—parents (2SLS)	.087***	.070**	.084*	.160**	.177***	.222***	.156***	−.055	−.032	.139***	.081**
	(.027)	(.029)	(.049)	(.062)	(.034)	(.032)	(.040)	(.049)	(.040)	(.040)	(.038)
	[291]	[188]	[80]	[36]	[199]	[216]	[120]	[71]	[110]	[116]	[112]
Omitted variables—parents (OLS)	.078***	.105***	.029**	.023*	.101***	.088***	.093***	.036***	.031***	.083***	.086***
	(.009)	(.011)	(.013)	(.012)	(.012)	(.011)	(.014)	(.011)	(.011)	(.014)	(.012)
Omitted variables—health (OLS)	.068***	.087***	.028**	.021*	.085***	.089***	.094***	.032***	.024**	.081***	.071***
	(.008)	(.010)	(.012)	(.012)	(.010)	(.011)	(.013)	(.010)	(.010)	(.013)	(.010)
Observations	2919	2384	2356	3408	2655	1901	2243	2224	3066	2134	2205

Table 2 continued

	New Zealand	Norway	Poland	Singapore	Slovak R.	Slovenia	Spain	Sweden	Turkey	U.K.	U.S.
Baseline (OLS)	.073*** (.009)	.096*** (.011)	.079*** (.012)	.040*** (.010)	.141*** (.010)	.096*** (.011)	.142*** (.010)	.080*** (.011)	.052*** (.014)	.078*** (.009)	.097*** (.010)
Measurement error—literacy (2SLS)	.074*** (.010) [6147]	.094*** (.011) [7809]	.073*** (.013) [3374]	.034*** (.009) [13343]	.129*** (.012) [4395]	.092*** (.011) [4956]	.143*** (.011) [7128]	.087*** (.013) [5205]	.060*** (.013) [2753]	.077*** (.010) [4756]	.103*** (.011) [6419]
Measurement error—second PV (2SLS)	.089*** (.010) [6390]	.104*** (.011) [5635]	.093*** (.013) [3596]	.045*** (.009) [10787]	.168*** (.012) [4371]	.108*** (.011) [4138]	.158*** (.011) [5549]	.084*** (.013) [3936]	.060*** (.014) [2164]	.082*** (.011) [5262]	.104*** (.011) [5617]
Reverse causality—years school (2SLS)	.138*** (.020) [496]	.182*** (.020) [378]	.251*** (.025) [370]	.060*** (.013) [1618]	.272*** (.021) [455]	.195*** (.017) [743]	.288*** (.018) [1046]	.134*** (.023) [472]	.204*** (.025) [393]	.088*** (.026) [255]	.139*** (.016) [656]
Reverse causality—parents (2SLS)	.058 (.042) [92]	.068** (.032) [133]	.152*** (.043) [120]	.002 (.034) [166]	.212*** (.032) [195]	.132*** (.026) [260]	.167*** (.043) [148]	.045 (.038) [106]	.199*** (.073) [24]	.074** (.031) [194]	.098*** (.026) [261]
Omitted variables—parents (OLS)	.064*** (.011)	.098*** (.011)	.069*** (.013)	.041*** (.010)	.130*** (.012)	.088*** (.012)	.141*** (.011)	.075*** (.012)	.048*** (.015)	.078*** (.011)	.089*** (.011)
Omitted variables—health (OLS)	.060*** (.009)	.074*** (.010)	.060*** (.012)	.040*** (.010)	.121*** (.010)	.079*** (.011)	.134*** (.011)	.069*** (.011)	.052*** (.014)	.057*** (.010)	.066*** (.011)
Observations	2429	2101	1716	2277	2182	2254	2737	1762	2224	3864	1990

Dependent variable is a binary indicator that equals 1 if the respondent reports that she is currently employed, and 0 otherwise. Each cell reports the coefficient on numeracy skills from a separate regression. Regressions weighted by sampling weights. Model names indicate bias addressed by the specification; "OLS" refers to ordinary least squares estimations, "2SLS" refers to two-stage least squares estimations. Least squares estimates (except for pooled specification) take into account all 10 plausible values of numeracy skills. Sample: persons aged 35–54 years. All regressions control for a quadratic polynomial in age and gender. *Baseline* standard model controlling for age and gender. *Measurement error* numeracy skills are instrumented by literacy skills. *Measurement error—second PV* numeracy skills are instrumented by second plausible value of numeracy skills. *Reverse causality—years school* numeracy skills are instrumented by years of schooling. *Reverse causality—parents* numeracy skills are instrumented by parental education (*1* neither parent attained upper secondary education, *2* at least one parent attained upper secondary education, *3* at least one parent attained tertiary education). *Omitted variables—parents* adds control for parental education. *Omitted variables—health* adds control for respondent's health (*1* poor, *2* fair, *3* good, *4* very good, *5* excellent). Observations refer to baseline specification. Numeracy skills standardized to SD 1 within country. Pooled specification includes country fixed effects and gives same weight to each country. Robust standard errors (jacknifed in the OLS models due to the complex sampling design) in parentheses. F-statistics of excluded instrument in the first stage of 2SLS models in brackets. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Pursuing this IV strategy, row (2) of Table 1 indicates that literacy is a very strong instrument for numeracy, with a point estimate of 0.85 in the first-stage estimation and an *F* statistics of over 61,000 in the pooled sample (shown in brackets at the bottom of each cell). In the second stage, the estimate on numeracy skills in predicting wages increases from 0.20 in the baseline OLS model to 0.22. That is, just by taking away domain-specific measurement error in the PIAAC test, the estimate of skill returns increases by 10% in the IV model, suggesting that downward bias from measurement error may indeed be an important issue in the analysis of returns to skills. A similar increase in estimated returns when correcting for measurement error can be observed in almost all countries; only in six countries (Estonia, Japan, Lithuania, Norway, Singapore, and the Slovak Republic) estimated returns decrease slightly.

While instrumenting numeracy skills by literacy skills addresses common concerns about test quality such as specific items on the numeracy test being a bad measure of skills relevant on the labor market, it ignores any earnings effects of domain-specific skills by considering only the returns to the skill component that is common to both skill domains. Therefore, we also pursue another approach to correct for measurement error in numeracy skills. As respondents received different item booklets, PIAAC reports 10 plausible values, or multiple imputations of proficiency values from the posterior distribution of a latent regression item response model, for each skill domain. As each plausible value provides an estimate of numeracy proficiency, we can use one plausible value of numeracy proficiency as an instrument for another plausible value of numeracy proficiency to correct for measurement error bias. Results of IV estimations that use the second plausible value of numeracy skills as an instrument for the first plausible value are shown in row (3) of Table 1. The *F* statistics of the excluded instrument in the first stage again indicate a very strong instrument. The second-stage estimate increase to 0.23 in the pooled country sample. In fact, the estimates of this IV model are larger than the OLS estimates in every single country in the sample.[30]

In the employment regressions, the coefficient on numeracy skills in the pooled sample remains unchanged when numeracy is instrumented with literacy and increases by 16% when the first plausible value of numeracy is instrumented with the second plausible value [rows (2) and (3) of Table 2]. Correcting for measurement error with the plausible-value based IV model leads to an increase in the estimated employment impact of skills in every single country. The literacy-based IV model also leads to significant effects in all but two countries, Greece and Japan. The reason for the weak relation between skills and employment in the literacy IV in these two countries is that better literacy skills are themselves not associated with a higher probability of being employed as literacy skills of unemployed and inactive adults are similar or even larger than literacy skills of their employed counterparts (OECD, 2016).[31] By considering only the returns to the skill component that is common to numeracy and literacy, the literacy IV thus fails to detect a significant association of skills with employment. This is also consistent with the

[30] Instrumenting the first plausible value of numeracy with any other plausible value or with all other plausible values simultaneously delivers almost identical results.

[31] As is discussed in OECD (2016), the differences in literacy skills between employed, unemployed, and inactive adults are small in most PIAAC countries. This can partly be attributed to the high rate of unemployment among young people (who tend to have higher literacy skills than older people) and the fact that many are inactive as they remain in education. Moreover, the difference in literacy skills between employed and unemployed adults is considerably larger when only the long-term (longer than 12 months) unemployed are used in the comparison.

fact that also in these two countries, the estimate increases compared to the OLS estimate when the second plausible value of numeracy is used as an instrument.

It is important to note that neither approach solves all possible measurement error issues. Errors common to both numeracy and literacy or common to all plausible values in the numeracy domain, ranging from the tested person having a bad testing day or the fact that the test measures may not be an encompassing measure of the underlying concept of human capital, are not eliminated (for an in-depth discussion, see section 2 in Hanushek et al., 2015).

Reverse causation

The second threat to causal identification is that people may have better skills *because* they have a better job. Such issues of reverse causation can be addressed by instrumental variables that are related to an individual's skills but observed before the start of the labor-market career. In line with this reasoning, school attainment could in fact serve as an instrument for skills, having been determined before entering the labor market. Similarly, family background potentially provides another instrument for skills that influences skill development but is predetermined with respect to labor-market experience. Indeed, in the literature on returns to school attainment, parental education has been used as an instrument for years of schooling (e.g., Ichino and Winter-Ebmer, 1999).

In rows (4) and (5) of Table 1, we use years of schooling and parental education, respectively, as instruments for numeracy skills in predicting wages. Both are strongly related to numeracy skills in the first stage, yielding strong instruments not only in the pooled sample of countries, but in fact in each country separately. In both cases, the second-stage estimate on the returns to skills increases substantially compared to the OLS estimate. In the pooled sample, estimated returns increase to 0.52 when using years of schooling as the instrument and to 0.46 when using parental education as the instrument. This pattern is again very similar across individual countries. The largest increases in estimated returns compared to the OLS results are observed in Indonesia, Poland, the Slovak Republic, Spain, and Turkey.

In the employment regressions, the positive effect of higher skills increases in both specifications compared to the baseline OLS estimate in the pooled sample of countries [rows (4) and (5) of Table 2]. While all previous results were surprisingly consistent across countries, using the parental education instrument in the employment regression yields statistically insignificant estimates in eight countries (Austria, Czech Republic, Finland, Japan, Korea, New Zealand, Singapore, and Sweden). One potential reason for this result is that parental education is reported in PIAAC only in three crude categories and there is little variation in this variable in the aforementioned countries. In particular, the share of adults with at least one tertiary-educated parent is rather low in these countries at atmost 15% (with the exception of Japan).[32]

While these IV estimates do not suffer from bias due to direct reverse causation, we shy away from interpreting them as causal effects. The main reasons for this are those discussed in the literature on returns to years of schooling: schooling is a choice variable, family background may exert direct effects on earnings and employment, and ability

[32] In Japan, the share of adults with at least one tertiary-educated parent is 29%.

may show intergenerational persistence (Card, 1999). Moreover, school attainment may proxy for some additional component of human capital that is relevant for earnings and employment—such as non-cognitive aspects of education that are not captured in the numeracy score. If any of these arguments hold true, the exclusion restriction that the instrument is related to wages and employment only through individuals' numeracy skills and not in any other way would be violated.

Omitted variables

The third—and presumably most daunting—source of bias in estimating skill returns is omitted variables that are related to both skills and labor-market success. For example, if family background is related to skill development and family networks help people find a better job, the association of skills with earnings and employment would not reflect just the causal effect of skills. In this sense, family background should be a control, rather than an instrument, in the estimation. As shown in row (6) of Table 1, controlling for parental education—which is indeed itself significantly associated with earnings (not shown)—does reduce the OLS estimate on numeracy skills in the wage regression in the pooled sample (from 0.200 to 0.182) and in all individual countries, suggesting that some (albeit small) part of the estimated returns to skills in the baseline least squares model may be attributable to family background.

Likewise, a person's health may positively affect both skill acquisition and labor-market outcomes. Controlling for the measure of self-assessed health status available in PIAAC, though, barely changes the estimate of skills on earnings [row (7) of Table 1]. Again, better health is itself positively associated with wages (not shown).

Similarly, including additional controls for family background and health somewhat reduces the estimated employment effects of skills, but better skills remain significantly related to higher employment probabilities in all countries [rows (6) and (7) of Table 2]. Thus, although somewhat less pronounced, the pattern of results in the employment regressions is again rather similar to what we observed for wages.[33] Of course, the available variables in PIAAC are obviously limited measures of the set of possible omitted traits. But gauging from these crude analyses, the empirical relevance of concerns from omitting family-background and health measures may be limited.

Robustness checks

In analyses not shown here, we have performed several additional robustness checks. Among others, we estimated the baseline model in Table 1 in a more encompassing set of workers that also includes part-time workers. Furthermore, we performed limited information maximum likelihood (LIML) estimates that are more robust to potentially weak instruments than two-stage least squares estimates. The pattern of results in these additional analyses is remarkably similar to our baseline estimates.[34]

Even though addressing several concerns regarding potential biases in the returns to skills estimations, neither of the aforementioned strategies provides an encompassing

[33] Accounting for the employment effects of skills in the wage equation—either by including the non-employed in the sample and assigning them a very low wage or by estimating Heckman selection models—yields returns to skills that are considerably larger than the baseline estimate (see Table 4 in Hanushek et al., 2015).

[34] Additional file 1: Tables S3 and S4 shows the results of these robustness checks for the pooled specification. Detailed country-by-country results are available from the authors upon request.

solution for all endogeneity problems. To make a further step towards causal analysis, we exploit variation in skills from two natural experiments. We do so by using changes in U.S. compulsory schooling laws as a source of exogenous variation in general skills ("Evidence from changes in compulsory schooling laws in the United States"). We then exploit technological peculiarities in broadband technology in Germany that affected the development of ICT skills ("Returns to ICT skills: evidence from peculiarities in broadband technology in Germany").

Evidence from changes in compulsory schooling laws in the United States

The biggest concern with the analysis in Tables 1 and 2 is that the instruments (i.e., other skill measures, years of schooling and parental background) are unlikely to capture exogenous variation in numeracy skills because they are themselves associated with higher wages. To provide more convincing evidence that the observed variation in cognitive skills is exogenous, Hanushek et al. (2015) exploit changes in U.S. compulsory schooling laws over time at the state level. The idea here is that schooling is one input into skill development and children who are forced to attend school longer should, ceteris paribus, build up more skills.[35] Since U.S. states changed compulsory schooling requirements at different points in time, our models can include state fixed effects that account for any (observed and unobserved) factors affecting skills and wages that remain constant over time within a state.

Table 3, replicated from Hanushek et al. (2015), shows returns-to-skills estimations using U.S. compulsory schooling laws as an instrument for numeracy skills. Column (1) starts with a model that includes state fixed effects and a quartic polynomial in age. In the first stage, each additional year of compulsory schooling is associated with 0.027 SD higher skills. A first-stage F statistic of 25.9 indicates a strong instrument. In the second stage, the part of the skill variation that is induced by changes in state compulsory schooling laws is significantly related to higher wages. The IV point estimate of 0.66 is substantially larger than the OLS point estimate of 0.25 [reported in row (1) of Table 1], although the relatively large standard errors do not allow distinguishing the coefficients at conventional levels of significance. The substantial increase in the IV estimate likely reflects that returns are higher for those who give rise to the identifying variation in this local average treatment effect (LATE), namely the population of compliers who are induced to get additional schooling because of the law changes. However, since PIAAC provides information only on the current state of residence, the estimated returns to skills in the IV model are potentially downward biased because interstate mobility would induce measurement error in the (state-level) compulsory schooling instrument.[36]

Column (2) replaces the quartic polynomial in age by a set of birth year fixed effects. The first stage estimate of the instrument remains strong (F statistic of 15.1), and the second-stage coefficient is even slightly higher. Columns (3) and (4) show this model separately for the samples of individuals with at most a high school degree and with more than a high

[35] Acemoglu and Angrist (2001) show that compulsory attendance requirements in the United States have generally been growing more restrictive, with the maximum enrollment age falling and the minimum dropout age rising.

[36] This measurement error is likely to be non-negligible because the United States is well known for the volume of internal migration. As shown in Hanushek et al. (2017a), more than 40% of a state's current working-age population (20–65 years) was not born in the same state. However, this share varies considerable between states, ranging from 22% in Louisiana to 84% in Nevada.

Table 3 Instrumental-variable models exploiting changes in compulsory schooling laws across US States. *Source:* adapted from Hanushek and Woessmann (2015)

	All levels of school attainment		At most high school	More than high school	At most high school	
	Age 35–54				Age 35–65	
	(1)	(2)	(3)	(4)	(5)	(6)
Second stage (dependent variable: log gross hourly wage)						
Numeracy	.661*** (.305)	.798*** (.246)	.659*** (.182)		.783*** (.273)	.633*** (.306)
State fixed effects	X	X	X	X	X	X
Age quartic	X					
Birth year fixed effects		X	X	X	X	X
State-specific trends						X
First stage (dependent variable: numeracy skills)						
Minimum school-leaving age	.027*** (.005)	.029*** (.008)	.061*** (.016)	−.000 (.006)	.047*** (.020)	.057*** (.022)
Instrument F statistic	25.9	15.1	14.5	0.0	5.7	6.5
Observations	932	932	369	563	520	520

Two-stage least squares regressions weighted by sampling weights. Sample: full-time employees in the United States. Second-stage coefficient is not displayed if the first-stage coefficient is insignificant. All regressions control for gender. Robust standard errors (adjusted for clustering at state level) in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

school degree, respectively. Reassuringly, the instrument of changes in compulsory schooling requirements affects only those with lower education levels and is not related to the skills of individuals with higher education, who should be unaffected by these laws.

Recently, Stephens and Yang (2014) have shown that identification from the timing of the law changes across US states over time can be very sensitive to the identifying assumption that there are no systematic state changes related to the variables of interest at the same time. To check whether results are driven by other variables changing simultaneously with compulsory schooling laws (e.g., school quality improvements), column (6) includes state-specific time trends. Even though the instrument becomes somewhat weaker in this highly demanding specification, estimated returns to skills remain statistically significant and sizeable.[37]

Returns to ICT skills: evidence from peculiarities in broadband technology in Germany

This section turns to estimating labor-market returns to one specific set of skills, namely skills to master information and communication technologies (ICT). This analysis exploits another natural experiment that specifically affected the development of ICT skills across German municipalities, leaving numeracy and literacy skills unchanged. This provides the unique opportunity to dig deeper into issues of causality in the estimation of returns to a domain-specific skill type, namely ICT skills—a skill domain that is commonly believed to be central in modern knowledge-based labor markets—and to isolate the wage effect of these ICT skills from skills in general.

Although there is the widespread belief that ICT skills matter for labor-market outcomes, the correlation between ICT skills and a person's general ability makes it hard

[37] In the trend estimation, the sample is extended to all workers with at most a high-school degree aged 35–65 so as to have enough variation over time. As a benchmark, column (5) of Table 3 provides the return-to-skills estimate for this sample without state-specific time trends.

to isolate the wage effect of ICT skills. For instance, an influential paper by DiNardo and Pischke (1997) shows that computer users at work possess unobserved skills which might have little to do with computers per se but which increase their productivity and wages. They strikingly demonstrate this by showing that positive wage effects can also be found for pencil use at work, being similar in magnitude to the wage effects of computer use. Based on this rather nonsensical finding, they conclude that returns to computer and pencil use at work must be biased due to unobserved skills of the users.

To isolate the wage effect of ICT skills from that of skills in general, Falck et al. (2016) exploit plausibly exogenous variation in ICT skills using technological peculiarities in broadband technology that led to uneven access in broadband Internet independent of individuals' (observed and unobserved) characteristics. Here, we summarize their identification strategy and the main results.

The underlying idea of the identification strategy is that ICT skills are developed through learning-by-doing for which Internet availability (which enables browsing the web, searching topics online, and receiving information through email) is a precondition. Since it is not random whether people have access to the Internet,[38] Falck et al. (2016) exploit the fact that the copper wires of the traditional voice-telephony network connecting households with the main distribution frame (MDF) were upgraded in many countries to provide fast Internet access by means of the so-called DSL technology (see Fig. 1). Indeed, the authors show that countries with a high fixed-line penetration before the introduction of DSL could roll out broadband earlier and reached a larger share of the population faster than countries lagging behind in fixed-line infrastructure.

This reliance of broadband rollout on traditional voice-telephony networks led to an uneven distribution of broadband Internet access within countries in the early years of the Internet era. Specifically, in West Germany, the general structure of the voice-telephony network dates back to the 1960s when the provision of telephone service was a state monopoly having the declared goal of providing universal telephone service to all German households.[39] While all households connected to an MDF enjoyed voice-telephony services in the same quality, only those households closer than 4200 m (2.6 miles) to their assigned MDF could gain access to broadband Internet when a DSLAM was installed.[40] Past this threshold, DSL technology was no longer feasible without replacing parts of the copper wire (typically placed between the MDF and the street cabinet) with fiber wire (see Fig. 1). Since this replacement involved costly earthworks that increased with the length of the bypass, certain West German municipalities were effectively excluded from early broadband Internet access.[41] Figure 2 shows that the share of house-

[38] For instance, people with better jobs are more likely to have the financial means to buy computers and equip their homes with Internet connections.

[39] Falck et al. (2016) exclude East Germany since it cannot be ruled out that location decisions for the MDFs in East Germany, which were made after Reunification in the 1990s, were partly determined by unobserved characteristics of the municipalities that are also correlated with individual wages (see Bauernschuster et al., 2014, for details). Berlin is also dropped from the analysis because DSL availability could not be distinguished between former West and East Berlin.

[40] The threshold value of 4200 m is a consequence of the DSL provision policy of the German telecommunication carrier, Deutsche Telekom, which marketed DSL subscriptions at the lowest downstream data transfer rate of 384 kbit/s only if the line loss was less than 55 decibel (dB). Since the copper cables connecting a household with the MDF usually had a diameter of 0.4 mm, a line loss of 55 dB was typically reached at about 4200 m. As the actual line loss depends on other factors as well, the 4200-m threshold is only a fuzzy threshold (Falck et al., 2014). This fuzziness in the technological threshold of DSL availability is substantially more severe in other countries, effectively limiting the use of the threshold identification to Germany.

[41] The costs of rolling out 1 km of fiber wire subsurface amount to 80,000 euro, with an additional 10,000 euro to install a new node where the remaining part of the copper wires is connected to the fiber wire (Falck et al., 2014).

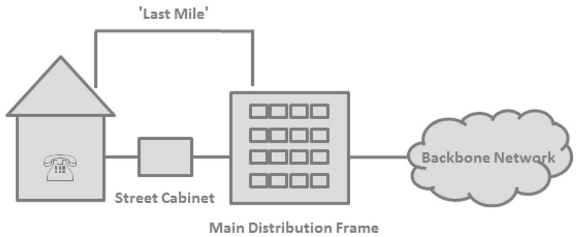

Fig. 1 The Structure of a DSL Network. The figure shows the structure of a DSL network that relies on the "last mile" of the preexisting fixed-line voice-telephony network. The "last mile" consists of copper wires connecting every household via the street cabinet to the main distribution frame (MDF). At the MDF, a DSLAM (Digital Subscriber Line Access Multiplexer) is installed that aggregates and redirects the voice and data traffic to the telecommunication operator's backbone network. Source: Falck et al. (2016)

holds with access to DSL is indeed substantially smaller in municipalities above the 4200-m threshold than in below-threshold municipalities.

This technological peculiarity can be exploited as a "natural experiment" in an IV analysis. In this analysis, being above or below the 4200-m threshold is used as an instrument for ICT skills in an earnings equation similar to Eq. (3). In particular, the threshold instrument is defined as a binary variable that equals 1 when the municipality in which an individual lives is more than 4200 m away from its MDF (lower probability of DSL availability) and 0 otherwise.[42]

Columns (3) and (4) of Table 4 present the results of this IV model in specifications with just municipality controls and with municipality and individual controls, respectively. The first-stage results in the lower panel of Table 4 provide support for the suggested learning-by-doing channel: persons in municipalities above the 4200 m threshold have 0.37 SD lower ICT skills than persons with an MDF within the 4200 m corridor in the model with all controls [column (4)]. The threshold instrument is significant at the 1% level and the first stage F statistic is 10.5, suggesting that a weak instrument bias is not a substantial concern in this context. In the second stage, a one SD increase in ICT skills attributable to the technical threshold leads to a wage increase of 31%.[43] The IV coefficients are about twice as large as the corresponding OLS results, shown in columns (1) and (2) of Table 4.[44] These higher returns in the IV specification likely reflect that returns are higher for the population of compliers that mainly consists of individuals with intermediate ICT skills.[45] Another reason for this difference is measurement error in ICT skills, biasing the OLS returns toward zero.

Column (5) shows that the threshold instrument is associated with no appreciable changes in numeracy skills—and in fact even has a positive coefficient—suggesting that

[42] This analysis extends the sample to persons aged 20–65 years and also includes part-time workers to have enough variation in the technical threshold across municipalities. First-generation immigrants are excluded because they often developed their ICT skills outside Germany and should therefore not be affected by the threshold instrument.

[43] Without controlling for individual characteristics, estimated returns to ICT skills are at 27%, significant at the 11% level [column (3)].

[44] The OLS results are based on variables aggregated at the municipality level, which provides the proper comparison to the IV results because the threshold instrument varies only at the municipality level. For the same reason, standard errors in Tables 4 and 5 are clustered at the municipality level.

[45] The OECD (2013) distinguishes three different ICT-proficiency levels: low (level 1 and below), intermediate (level 2), and high (level 3).

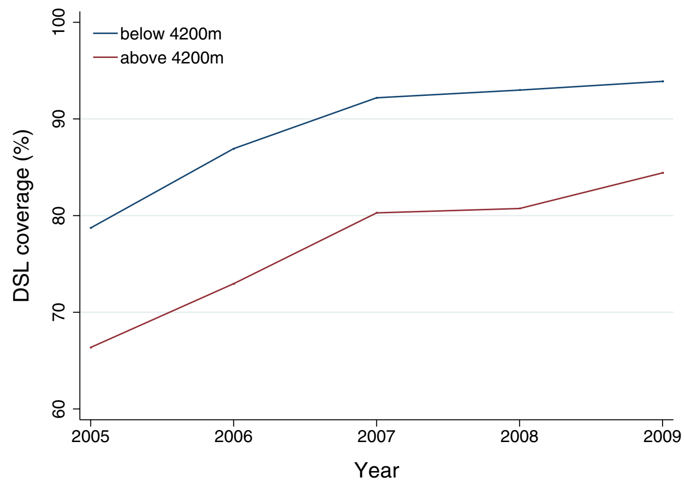

Fig. 2 DSL Coverage in Above-Threshold and Below-Threshold Municipalities. The figure shows the share of households with access to DSL in the period 2005–2009. The *blue* (*red*) *line* indicates municipalities that are less (more) than 4200 m away from their assigned main distribution frame (MDF). *Source:* Falck et al. (2016)

Table 4 Returns to ICT skills: instrumental-variable models exploiting technological peculiarities in broadband technology in West Germany. *Source:* adapted from Falck et al. (2016)

Dependent variable: log gross hourly wage					
	OLS (municipality level)		2SLS (second stage)		OLS
	(1)	(2)	(3)	(4)	(5)
ICT skills	.136*** (.025)	.148*** (.025)	.272 (.167)	.306*** (.151)	
Municipality characteristics	X	X	X	X	X
Individual characteristics		X		X	X
First stage					
Dependent variable:	**ICT skills**				**Numeracy skills**
Threshold			−.404*** (.102)	−.369*** (.114)	.044 (.054)
Instrument F statistic			15.8	10.5	
Individuals	–	–	1849	1849	1849
Municipalities	204	204	204	204	204

Regressions weighted by sampling weights (giving same weight to each municipality). Least squares estimations with variables aggregated at the municipality level in columns (1)–(2); two-stage least squares estimations in columns (3)–(4); least squares estimations in column (5). Sample: West German employees aged 20–65 years, no first-generation immigrants. ICT and numeracy skills are standardized to SD 1 within country. *Threshold* binary variable indicating whether a municipality is more than 4200 m away from its MDF (*1* lower probability of DSL availability), and 0 otherwise. Distance calculations are based on municipalities' geographic centroid. Municipality characteristics are unemployment rate in 1999 (i.e., share of unemployed individuals in the working-age population aged 18–65 years) and population share of individuals older than 65 in 1999. Individual characteristics are quadratic polynomial in work experience and gender. Column (5) controls for ICT skills. Robust standard errors, adjusted for clustering at the municipality level, in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 Robustness of returns to ICT skills in sample of municipalities without own main distribution frames. *Source:* adapted from Falck et al. (2016)

Dependent variable: log gross hourly wage					
	OLS (municipality level)		2SLS (second stage)		OLS
	(1)	(2)	(3)	(4)	(5)
ICT skills	.209*** (.079)	.271*** (.087)	.405*** (.204)	.521*** (.213)	
Municipality characteristics	X	X	X	X	X
Individual characteristics		X		X	X
First stage					
Dependent variable:	**ICT skills**				**Numeracy skills**
Threshold			−.592*** (.126)	−.517*** (.153)	−.033 (.069)
Instrument F statistic			22.1	11.5	
Individuals	–	–	160	160	160
Municipalities	18	18	18	18	18

Regressions weighted by sampling weights (giving same weight to each municipality). Least squares estimations with variables aggregated at the municipality level in columns (1)–(2); two-stage least squares estimations in columns (3)–(4); least squares estimations in column (5). Sample: West German employees aged 20–65 years, no first-generation immigrants, only municipalities without an own main distribution frame (MDF). ICT and numeracy skills are standardized to SD 1 within country. *Threshold:* binary variable indicating whether a municipality is more than 4200 m away from its MDF (*1* lower probability of DSL availability), and 0 otherwise. Distance calculations are based on municipalities' geographic centroid. Municipality characteristics are unemployment rate in 1999 (i.e., share of unemployed individuals in the working-age population aged 18–65 years) and population share of individuals older than 65 in 1999. Individual characteristics are quadratic polynomial in work experience and gender. Column (5) controls for ICT skills. Robust standard errors, adjusted for clustering at the municipality level, in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the empirical strategy indeed isolates the effect of ICT skills (vis-à-vis generic skills or general ability).[46] Falck et al. (2016) also provide another placebo test to solidify the evidence for the existence of a learning-by-doing channel in the accumulation of ICT skills. They show that the threshold instrument is irrelevant for ICT skills in a sample of first-generation immigrants who are unlikely to have acquired their ICT skills in Germany.

One concern with the results in Table 4 is that they may partly reflect earnings differences between rural and urban areas. Densely populated municipalities always have at least one own MDF and households are typically below the 4200-m threshold; less agglomerated municipalities often share an MDF. Thus, Table 5 provides estimates based on regressions analogous to those underlying Table 4 in a sample of municipalities without an own MDF, leading to a more homogenous sample with respect to socioeconomic characteristics. Some municipalities, however, were (arguably randomly) lucky to be close enough to an MDF in another municipality to have access to broadband Internet. This provides variation in the instrument in this restricted sample.[47] Also in this sample, returns to ICT skills are economically and statistically significant, and even increase somewhat compared to the estimates in the full sample. Moreover, the threshold instrument is again not

[46] Estimates in column (5) condition on ICT skills to account for the high correlation between skill domains. Note that the threshold instrument continues to be a relevant predictor for ICT skills also when it is controlled for numeracy (or literacy) skills (not shown).

[47] However, sample size is considerably smaller than in the full sample because the sampling of municipalities in PIAAC was proportional to municipality size (Rammstedt, 2013).

significantly related to numeracy skills, indicating that the estimated returns to ICT skills are unlikely to be biased due to unobserved skills of PIAAC respondents.

To provide further evidence in favor of the validity of their identification strategy, Falck et al. (2016) show that potential direct productivity effects of broadband (e.g., the introduction of online job search channels increasing the quality of job matching) do not affect their results. Moreover, they show that households without broadband Internet access do not selectively relocate to regions where broadband is available.

Conclusion

The idea that human capital is crucial for future prosperity is widely accepted today. Policymakers regularly emphasize the importance of education for the economy's innovative capacity and ability to compete in the globalized world of the 21st century. In the words of former U.S. President Barack Obama, "Whether it's improving our health or harnessing clean energy, protecting our security or succeeding in the global economy, our future depends on reaffirming America's role as the world's engine of scientific discovery and technological innovation. And that leadership tomorrow depends on how we educate our students today".[48]

Existing research investigating the effects of human capital accumulation supports this view. Human capital has been shown to have substantial positive impacts not only on individuals' success in the labor market, but also on their general well-being. Moreover, a substantial amount of evidence suggests that the human capital of a population is a main driver of economic growth. However, the empirical literature on the labor-market effects of a person's cognitive skills, which have shown to be a more reliable proxy for effective human capital than years of schooling, is plagued by the apparent endogeneity of measured skills. For instance, different employment patterns could directly affect skills over the lifecycle, implying problems of reverse causation. Moreover, unobserved variables like family networks, health, or non-cognitive skills could directly influence earnings; if also related to skills, these could lead to omitted variable bias in the analysis of skills.

This paper aimed to address these sources of potential bias by estimating IV models that exploit variation in skills stemming from differences in family backgrounds and school attainment. In all participating PIAAC countries, we find larger returns to skills in these IV models than in standard least squares estimations, suggesting that the latter may in fact be biased downwards. This finding holds for wages and, albeit to a lesser degree, for employment. While information on family background and years of schooling is readily available in the PIAAC data, the issue remains that the variation in skills induced by these variables is not necessarily exogenous. We therefore complement the above analysis by two natural experiments that more credibly identify exogenous variation in skills that is independent of other influences such as family background or health limitations. These more convincing models similarly suggest that OLS returns provide a lower bound of the true returns to skills in the labor market.

Overall, our results show that modern knowledge-based economies highly reward skills. This puts the focus on policies for skill development at all levels—from the

[48] Office of the Press Secretary, White House Office, "Remarks by the president on the "Educate to Innovate" Campaign and Science Teaching and Mentoring Awards," January 6, 2010.

education provided before and in school to lifelong learning opportunities on and off the job—and on policies that ensure that skills are effectively retained and used. Our results emphasize that such policies are important to secure prosperity in the future.

Additional file

Additional file 1. Robustness checks.

Authors' contributions
All authors made substantial contribution to the conception and design, as well as to the analysis and interpretation of results. They were jointly responsible for drafting and revising the article. All authors read and approved the final manuscript.

Author details
[1] ifo Institute at the University of Munich, Poschingerstraße 5, 81679 Munich, Germany. [2] KU Eichstaett-Ingolstadt, ifo Institute, and CESifo, Auf der Schanz 49, 85049 Ingolstadt, Germany. [3] University of Munich, ifo Institute, CESifo, and IZA, Poschingerstraße 5, 81679 Munich, Germany.

Acknowledgements
We would like to thank Vanessa Denis and William Thorn from the OECD for access to and help with the international PIAAC data, as well as Dan McGrath, Eugene Owen, and Saida Mamedova from the National Center for Education Statistics and Anja Perry at GESIS for their assistance with the U.S. and German PIAAC data, respectively. We also thank the editor and two reviewers for their useful comments. We gratefully acknowledge financial support from the Leibniz Association through the project "Acquisition and Utilisation of Adult Skills—A Network for Analyzing, Developing and Disseminating PIAAC".

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
The main dataset employed in this paper uses the Public Use Files for all countries surveyed in the Programme for the International Assessment of Adult Competencies (PIAAC) other than Australia. Data are available at the OECD website (http://www.oecd.org/site/piaac/publicdataandanalysis.htm). The Public Use Files contain both responses to the background questionnaire (including wage data) and the cogni-tive assessment. The OECD website contains further information on the variables in the Public Use Files.

As discussed in the paper, the estimation relies on some data that are not currently available in the public use file but that may be obtained by researchers through application to the respective countries. There are relatively minor differences in the main results from using the public use data for the United States and for Germany instead of the confidential data. These arise from having some data (including earnings) in categorical instead of continuous form. There is no public use file for the Australian data, requiring that full replication of our results requires individual researchers to obtain a data set directly from the Australians. The National Center for Education Statistics of the United States separately ran do files on their confidential data. The German municipality identifiers are available only via on-site use after an application process.

Ethics approval and consent to participate
We rely on data from the PIAAC Survey of Adult Skills which underlie ethics standards stated by the OECD. For further information see "PIAAC Technical Standards and Guidelines" (June 2014), http://www.oecd.org/skills/piaac/PIAAC-NPM(2014_06)PIAAC_Technical_Standards_and_Guidelines.pdf.

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

Received: 20 September 2016 Accepted: 4 April 2017
Published online: 26 April 2017

References
Acemoglu, D., & Angrist, J. D. (2001). How Large are Human-Capital Externalities? Evidence from compulsory-schooling laws. In B. S. Bernanke & K. Rogoff (Eds.) *NBER Macroeconomics Annual 2001* (pp. 9–74). Cambridge, MA: MIT Press.

Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4B, pp. 1043–1171). Amsterdam: North Holland.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics, 106*(4), 976–1014.

Ashenfelter, O., & Rouse, C. (1998). Income, schooling and ability: Evidence from a new sample of identical twins. *Quarterly Journal of Economics, 113*(1), 253–284.
Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics, 118*(4), 1279–1333.
Barrett, G. F. and Riddell, W. C. (2016). *Ageing and literacy skills: Evidence from IALS, ALL and PIAAC*. IZA Discussion Paper No. 10017.
Bauernschuster, S., Falck, O., & Woessmann, L. (2014). Surfing alone? The internet and social capital: Quasi-experimental evidence from an unforeseeable technological mistake. *Journal of Public Economics, 117,* 73–89.
Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy, 70*(5), 9–49.
Bishop, J. H. (1989). Is the test score decline responsible for the productivity growth decline? *American Economic Review, 79*(1), 178–197.
Bowles, S., Gintis, H., & Osborne, Melissa. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature, 39*(4), 1137–1176.
Card, D. (1999). The Causal Effect of Education on Earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3A, pp. 1801–1863). Amsterdam: North Holland.
Chetty, R., Friedman, J. N., Hilger, Nathaniel, Saez, Emmanuel, Schanzenbach, Diane Whitmore, & Yagan, Danny. (2011). How does your kindergarten classroom affect your earnings? Evidence from project STAR. *Quarterly Journal of Economics, 126*(4), 1593–1660.
Devereux, P. J., & Hart, R. A. (2010). Forced to be rich? Returns to compulsory schooling in Britain. *Economic Journal, 120*(549), 1345–1364.
DiNardo, J. E., & Pischke, J.-S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *Quarterly Journal of Economics, 112*(1), 291–303.
Falck, O., Gold, R., & Heblich, S. (2014). E-lections: Voting behavior and the internet. *American Economic Review, 104*(7), 2238–2265.
Falck, O., Heimisch, A., & Wiederhold S. (2016). *Returns to ICT Skills*. CESifo Working Paper No. 5720.
Goldin, C., & Katz, L. F. (2008). *The race between education and technology*. Cambridge: Harvard University Press.
Grenet, J. (2013). Is extending compulsory schooling alone enough to raise earnings? evidence from french and british compulsory schooling laws. *Scandinavian Journal of Economics, 115*(1), 176–210.
Hampf, F., & Woessmann L. (2016). *Vocational vs. general education and employment over the life-cycle: New evidence from PIAAC*. CESifo Working Paper No. 6116.
Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2014). *The value of smarter teachers: International evidence on teacher cognitive skills and student performance*. NBER Working Paper No. 20727.
Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics, 4,* 131–157.
Hanushek, E. A., Ruhose, J., & Woessmann, L. (2017a). Knowledge capital and aggregate income differences: Development accounting for U.S. States. *American Economic Journal: Macroeconomics*, forthcoming.
Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2017b). Coping with change: International differences in the returns to skills. *Economics Letters*, forthcoming.
Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2008). Returns to skills around the world: evidence from PIAAC. *European Economic Review, 73*(C), 103–130.
Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature, 46*(3), 607–668.
Hanushek, E. A., & Woessmann, L. (2011). The economics of international differences in educational achievement. In E. A. Hanushek, S. Machin & L. Woessmann (Eds.) *Handbook of the economics of education* (Vol. 3, pp. 89–200). Amsterdam: North Holland.
Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth, 17*(4), 267–321.
Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. Cambridge: MIT Press.
Hanushek, E. A., & Zhang, L. (2009). Quality-consistent estimates of international schooling and skill gradients. *Journal of Human Capital, 3*(2), 107–143.
Harmon, C., & Walker, I. (1995). Estimates of the economic return to schooling for the United Kingdom. *American Economic Review, 85*(5), 1278–1286.
Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: the mincer equation and beyond. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (Vol. 1, pp. 307–458). Amsterdam: North Holland.
Ichino, A., & Winter-Ebmer, R. (1999). Lower and upper bounds of returns to schooling: An exercise in IV estimation with different instruments. *European Economic Review, 43*(4), 889–901.
Lazear, E. P. (2003). Teacher incentives. *Swedish Economic Policy Review, 10*(3), 179–214.
Leuven, E., Oosterbeek, H., & van Ophem, H. (2004). Explaining international differences in male skill wage differentials by differences in demand and supply of skills. *Economic Journal, 114*(495), 466–486.
Levels, M., van der Velden, R., & Allen, Jim. (2014). Educational mismatches and skills: New empirical tests of old hypotheses. *Oxford Economic Papers, 66*(4), 959–982.
Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *The Review of Economic Studies, 72*(1), 189–221.
Lochner, L. (2011) Nonproduction benefits of education: Crime, health, and good citizenship. In S. Machin, E. A. Hanushek & L. Woessmann (Eds.) *Handbook of the Economics of Education* (Vol. 4, pp. 183–282). Amsterdam: North Holland
Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *The American Economic Review, 94*(1), 155–189.
Mincer, J. (1970). The distribution of labor incomes: A survey with special reference to the human capital approach. *Journal of Economic Literature, 8*(1), 1–26.

Mincer, J. (1974). *Schooling, experience, and earnings*. New York: NBER.
Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy, 107*(6), S184–S224.
Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management, 19*(4), 547–568.
Murnane, R. J., Willett, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics, 77*(2), 251–266.
Neal, D., & Johnson, W. R. (1996). The role of pre-market factors in Black–White differences. *Journal of Political Economy, 104*(5), 869–895.
OECD. (2013). *OECD skills outlook 2013: First results from the survey of adult skills*. Paris: Organisation for Economic Co-operation and Development.
OECD. (2016). *Skills matter: Further results from the survey of adult skills, OECD skills studies*. Paris: Organisation for Economic Co-operation and Development.
Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review, 96*(1), 152–175.
Oreopoulos, P., & Salvanes, Kjell G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives, 25*(1), 159–184.
Perry, A., Wiederhold, S., & Ackermann-Piek, D. (2014). How can skill mismatch be measured? New approaches with PIAAC. *Methods, Data, Analyses: A Journal for Quantitative Methods and Survey Methodology, 8*(2), 137–174.
Pischke, J.-S., & von Wachter, T. (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *Review of Economics and Statistics, 90*(3), 592–598.
Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics, 12*(2), 111–134.
Rammstedt, B. (Ed.). (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012*. Muenster: Waxmann.
Schlotter, M., Schwerdt, G., & Woessmann, L. (2011). Econometric methods for causal evaluation of education policies and practices: A non-technical guide. *Education Economics, 19*(2), 109–137.
Schultz, T. W. (1961). Investment in human capital. *The American Economic Review, 51*(1), 1–17.
Staiger, D., & Stock, J. H. (1997). Instrumental variables regressions with weak instruments. *Econometrica, 65*(3), 557–586.
Stephens, M., Jr., & Yang, D.-Y. (2014). Compulsory education and the benefits of schooling. *American Economic Review, 104*(6), 1777–1792.
Stock, J. H., & Watson, M. W. (2007). *Introduction to econometrics, second edition*. Boston: Pearson.
Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics, 20*(4), 518–529.
Woessmann, L. (2016). The economic case for education. *Education Economics, 24*(1), 3–32.