I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18355

Formal and Informal Assets in the Italian Labour Market

Francesca Borgonovi
Daniele Checchi
Valentina Gualtieri

DECEMBER 2025

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18355

Formal and Informal Assets in the Italian Labour Market

Francesca Borgonovi
UCL and OECD

Daniele Checchi
University of Milan and IZA

Valentina Gualtieri
INAPP

DECEMBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Formal and Informal Assets in the Italian Labour Market

This paper estimates labour-market returns to formal and informal human capital in Italy using data from the first cycle of PIAAC. We distinguish formal inputs (years of schooling) from directly assessed skills (literacy and numeracy) which we interpret as distinct forms of human capital that are shaped by school quality and by non-formal and informal learning. To address non-random employment and joint endogeneity of schooling and skills, we combine a Heckman selection model with instrumental variables. Schooling is instrumented using cohort exposure to the 1971 introduction of full-day primary schooling and the 1999–2001 Bologna '3+2' university reform; skills are instrumented using gender- and cohort-specific municipal illiteracy rates from population censuses, matched by birthplace. Results show that ignoring selection and endogeneity overstates the returns to schooling. After correction, numeracy yields the main wage premium, while formal credentials contribute little once skills are accounted for. The findings highlight the role of early cultural environments and skill accumulation for Italian wage inequality in Italy.

JEL Classification: J24, I26
Keywords: skill, schooling, credentials, Italy

Corresponding author:
Daniele Checchi
Università degli Studi di Milano
Via Festa del Perdono 7
Milan, 20122
Italy
E-mail: daniele.checchi@unimi.it

1. Introduction

Human capital, which comprises knowledge and skills, has long been recognised as a key driver of economic growth at the societal level, and of labour market success at the individual level. Education and training are key to turn individuals' human capital potential into increased productivity (Becker, 1962; Lucas, 1988). However, measures of the quantity of schooling (such as years of schooling or educational qualifications) which have typically been used as measures of human capital in Mincerian wage regressions (Mincer, 1970) are poor proxies of human capital because they do not capture differences in the quality of educational inputs and ignore the role played by human capital development opportunities that occur outside of formal education settings. Education is just one of the channels through which human capital can be acquired: skills and knowledge can in fact be gained at work, through on-the-job training, and other informal or non-formal opportunities. Two workers with the same years of schooling may therefore have very different levels of human capital if the quality of their education differs, or if one accumulated substantially more informal or non-formal learning than the other. As a result, educational attainment or years of schooling in wage equations can lead to measurement bias of the role of human capital, especially in cross-country comparisons because of variations across countries in the quality of educational inputs (Hampf, Wiederhold and Woessmann 2017) and in the prevalence of non-formal and informal learning (OECD, 2025a).

Data constraints determined how human capital was originally operationalised when estimating the determinants of wage differentials. This started to change because of the development of direct assessments of adult skills. The International Adult Literacy Survey (IALS, conducted between 1994 and 1998 over 22 countries) was the first large-scale international comparative assessment designed to identify and measure literacy skills. The Adult Literacy and Life Skills Survey (ALL, conducted between 2003 and 2008 over 12 countries), and the OECD's Programme for International Assessment of Adult Competencies (PIAAC, conducted in two cycles, between 2011 and 2018 in 39 countries and in 2022-23 covering 31 countries) have since extended the range of skills domains assessed as well as country coverage. What IALS, ALL and PIAAC have in common is that they administer instruments that allow to derive direct measures of skills and, as a result, to estimate of labour market outcomes as a function of direct measures of human capital as well as on proxies based on educational inputs.

A body of empirical studies has emerged using IALS, ALL, and PIAAC surveys, aiming to estimate the returns to formal education versus cognitive skills, which, when controlling for formal education, represent the share of human capital that is developed informally, either because of differences in the quality of schooling or because of nonformal and informal investments in learning. These studies typically estimate earnings or employment regressions including both schooling and skill measures and consistently finding that both education and skills are positively related to labour market outcomes (Leuven, et al., 2004; Green and Riddell, 2003; Blau and Kahn 2005; Paccagnella 2015; Hanushek et al., 2015; Cappellari et al 2017). Importantly, the cross-country standardised measures available in these studies allow to estimate how the returns to skills and education differ across institutional contexts. For example, evidence from PIAAC suggests that there is a large cross-country variation in the wage returns to numeracy skills: a one SD difference in numeracy is associated with around 10% wage premium in countries Italy and Spain but a 25% premium in the United States and 40% in Singapore (Hanushek et al., 2015). Generally, countries with more flexible labour markets and greater wage dispersion show higher skill returns, whereas those with more compressed wage structures or strong collective bargaining show lower skill-wage differentials (Blau and Kahn, 2005).

Although data collected through large-scale assessments allow to include better measures of human capital compared to standard surveys, identifying causal conclusions about the wage effects of investments in human capital requires identifying suitable analytical approaches given the underlying data structure and constraints. This is because individuals self-select into education and other skills development opportunities and individuals with higher abstract thinking potential are not only more likely to attain more education but are also more likely to invest in skills development outside formal education settings and to earn more. Similarly, employers might pay workers with higher levels of human capital more but the set of jobs that pay more is often the same set that allows individuals to hone their skills at work (Deming and Silliman, 2025). A rich literature has identified causality with respect to formal education using exogenous variations in schooling using natural experiments such as changes in compulsory schooling legislation (Brunello et al, 2009; Clay et al, 2021), child labour laws (Del Rey et al, 2018), differences in the geographical proximity of different educational institutions (Card, 1993), the timing of birth affecting school start age (Angrist and Krueger, 1991) or siblings and twin studies (Bound and Solon, 1999). A recent survey is Patrinos and Psacharopoulos (2025). This literature concludes that while there is some variation by context and methodology, schooling has a strong causal effect on earnings (Card, 1999; Heckman et al, 2006) and that estimates of causal effects are just as large, if not larger than estimates based on correlational approaches.

Whereas many studies have estimated the causal effect of schooling, far fewer have identified the causal effects of skills. One strategy that has been adopted in the literature exploits early-life factors as sources of exogenous variation in adult skills, for example using individuals' years of schooling and their parents' education as instruments for their skills measured when they were adults in PIAAC, or with exposure to changes in compulsory schooling laws (Hampf et al, 2017). Another strategy used to estimate the effect of skills on wages leverages natural experiments that shift specific skill endowments. Using German PIAAC data, for instance, geographical variation in broadband availability has been used to estimate the labour market returns of skills related to digital use (problem solving in technology rich environments) (Hampf et al. 2017). Cappellari et al. (2017) address the joint endogeneity of schooling and achievements, using the timing of educational reforms as instrument. Their estimates suggest the existence of a recursive structure, whereby skills influence wages conditional on years of schooling.

A second empirical challenge in estimations of the wage returns to formal and informal human capital investments is selection into employment. Wage equations are typically estimated only for individuals who are employed and report earnings. However, if people who are employed and those who are not differ in terms of human capital investments, then estimates of the returns to human capital may be biased because of sample selection. This concern is particularly salient in cross-country contexts because selection mechanisms may differ across labour markets, and in analyses of gender differences in the returns to human capital, because women are generally less likely to work than men and selection into employment is not equally likely to reflect differences in human capital investments among men and women (Heckman, 1979; Olivetti and Petrongolo, 2008).

Women often have lower labour force participation than men, because they generally have greater caring responsibilities and combining such responsibilities with employment can be difficult to achieve in the absence of supportive family policies. As a result, women observed in full-time employment may be positively selected in terms of education, attachment to the labour market, and career orientation, which can attenuate the observed gender wage gap relative to a counterfactual in which women's employment decisions faced the same constraints as men's (Olivetti and Petrongolo, 2008). A standard approach to address the issue of selection is Heckman's two-step selection model, which entails first modelling the probability of employment and then incorporating a correction term into the wage equation to adjust for the non-random selection of workers (Heckman, 1979).

This paper focuses on Italy and uses the first cycle of the PIAAC survey which was administered in 2012 to estimate the wage returns to human capital, incorporating both measures of formal education (which we take to measure human capital produced in formal settings) and skills (which we take to measure human capital accumulated in nonformal and informal settings after conditioning on education). We make three original contributions. First, we estimate the causal effects of formal education and skills on the probability of employment and wages while explicitly addressing selection into employment and the joint endogeneity of schooling and skills. Our empirical strategy combines a Heckman selection framework with separate instruments for education and skills. For education, we exploit cohort exposure to two nationwide reforms: the 1999–2001 "3+2" Bologna reform of university degrees, and the introduction/expansion of "*tempo pieno*" (full-day primary schooling) in the 1970s. For skills, we use gender- and cohort-specific municipality illiteracy rates measured in population censuses and mapped to respondents by birthplace and decade of birth. We show that ignoring sample selection and endogeneity of education and skills leads to biased estimates of the returns to human capital investments. Second, we use this census-based variation not only for identification but also to reflect on a central feature of the Italian context: the persistently low level of adult skills documented by PIAAC. Our results support the view that Italy's skills gap is partly rooted in long-lasting differences in the cultural and educational environments in which existing cohorts of adults were raised, differences that are captured by historical illiteracy rates in individuals' municipalities of birth. In this sense, we provide evidence that the cross-cohort and cross-area dispersion in adult skills is not merely a reflection of contemporaneous labour-market conditions but is also linked to early-life local environments that shaped skill formation over the life course. Third, we contribute to a largely under-examined issue in the returns-to-skills literature: the role of income non-reporting in shaping estimated wage premia. In Italy, where self-employment is relatively prevalent, excluding individuals with missing earnings can generate a selected sample that differs systematically in both education and skills. We implement and compare alternative strategies to handle missing earnings and document how resulting estimates differ.

The paper is organised as follows: the next two sections illustrate the Italian peculiarity in the first cycle of PIAAC survey; section 4 presents our empirical strategy; section 5 contains descriptive evidence of our dataset; section 6 contains our main results dealing with sample self-selection and endogeneity of human capital measures; section 7 concludes.

2. Italy in PIAAC

Although the second cycle of PIAAC was released at the end of 2024 and contains similar data to those used in this work, we decided to rely on data collected in 2011-12 rather than 2023 for two reasons. First, part of our identification strategy relies on respondents' exposure to two Italian education reforms. The first affected cohorts starting from those who were born in 1966 cohort, and the second affected cohorts starting from those who were born in 1980. Because we focus on core working age adults (23-55-year-old), virtually all individuals observed in 2023 would be exposed to the first reform, reducing the variation needed for identification. Second, fieldwork for the second cycle of PIAAC was conducted in the aftermath of the COVID pandemic, which may have shaped individuals' employment and wages in ways that reflect a transitory shock rather than underlying structural conditions (OECD, 2021).

The OECD report based on PIAAC cycle 1 highlighted the low level of human capital of Italian adults, both in terms of formal education attainment and measured skills, and documented labour market returns to human capital, without offering a fully convincing explanation of Italy's lag (OECD 2013; ISFOL 2013). In the first round of PIAAC cycle 1 in which 24 countries participated, the share of adults at the lowest literacy proficiency levels (below Level 2[1]) ranges between 5.0% in Japan and 27.9% in Italy. In numeracy, the corresponding shares range between 8.3% in Japan and 31.9% in Italy. In addition to low average proficiency, Italy exhibits a relatively limited dispersion in skills (for example, the difference between 95th and 5th percentile in literacy is 146 points compared to 152 points on average across participating countries). In 2015 and 2018 two additional rounds of data collection were organised in the PIAAC programme. PIAAC was administered in 14 additional countries and results confirmed Italy's relatively low international standing in terms of skills proficiency (OECD 2019).

The Italian country report (ISFOL 2013) examined the national situation in a comparative perspective and emphasised regional divides as a potential explanation for the low average results in the PIAAC skills assessment. Whereas adults who resided in the Centre and North Easter regions of the country had levels of proficiency in both literacy and numeracy that were comparable to those of France, adults residing in the South

[1] Levels below 2 are the levels where adults are asked to complete simple tests, involving a few steps, handling a limited amount of information and involving simple basic cognitive operations.

and Islands performed at very low levels in the assessment, comparable to Kazakhstan. A second possible explanation pertains to Italy's comparatively low levels of educational attainment: among PIAAC participants, 54% of Italian adults did not complete upper secondary qualifications, 34% obtained an upper secondary qualification and just 12% obtained a tertiary degree. The respective figures for the OECD average were 27%, 43% and 29%. Results from second cycle of PIAAC that was administered in 2023 show little evidence of convergence: Italy remains among the lowest ranked countries in terms of skills proficiency, the geographical divide between North and South persists, and educational attainment is lower than in most other participating countries (OECD, 2025a; ISFOL 2024).

Skills use, the labour-market returns to human capital and skills mismatches have also been studied using Italian PIAAC data, often in a comparative perspective. Returns to measured skills are estimated to be among the lowest in Italy (Hanushek et al., 2015). A complementary line of work links these low returns to mismatch between workers' proficiency and the extent to which skills are used on the job. Using measures of mismatch based on the divergence between workplace skill use and individual proficiency, wages in Italy appear to be particularly associated with the underutilisation of literacy skills, while numeracy over-/underutilisation is typically not statistically significant (Allen et al., 2013). Further evidence distinguishes between skill mismatch and qualification mismatch and suggests that, in Italy, qualification mismatch is more prevalent than skill mismatch. This pattern is consistent with the interpretation that formal qualifications may be a relatively noisy signal of productive skills in the Italian labour market (Pellizzari and Fichen, 2017; Monti and Pellizzari, 2016).

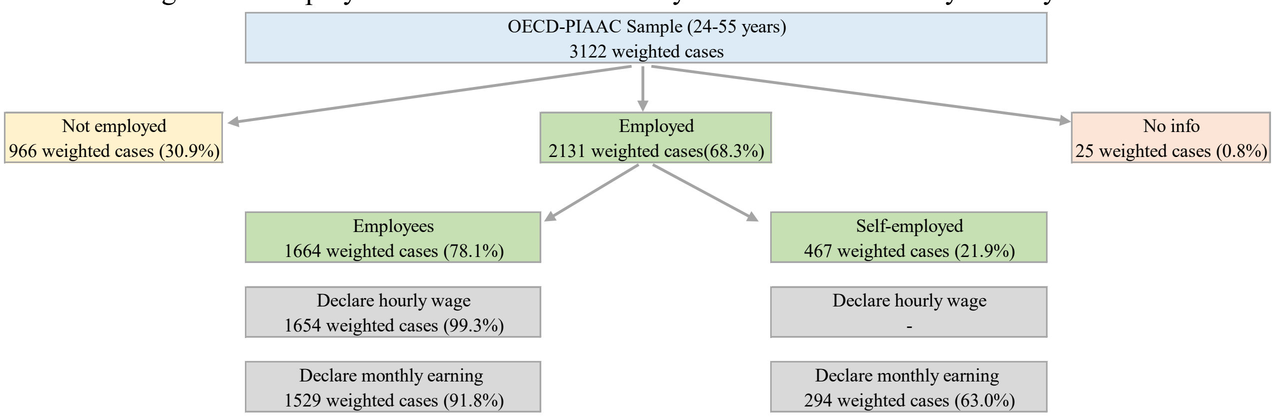
3. Data

Our primary source of data is the Italian sample of the first cycle of the Programme for the International Assessment of Adult Competencies (PIAAC). Fieldwork took place between 2011 and 2012 and involved a nationally representative sample of Italian adults aged 16-65. PIAAC combines a direct assessment of adults' information processing skills and a background questionnaire. Trained interviewers went to people's homes, where they first administered a background questionnaire using Computer Assisted Personal Interviewing (CAPI) and then handed them a computer-based assessment (or a paper-based assessment for those with low familiarity with digital tools or who opted to use paper-based instruments) that respondents were asked to complete.

In PIAAC cycle 1, the direct, low stakes and untimed assessment of literacy and numeracy was assessed in all countries whereas an assessment of problem solving in technology rich environments was administered in a selection of countries excluding Italy. The background questionnaire was administered before the assessment and was used to measure, through self-reports, individuals' socio-economic and demographic background, educational attainment and labour market outcomes, as well as skills use at work, at home, and in everyday life. In Italy a total of 4621 individuals took part in the first cycle of PIAAC and the response rate was 56% (OECD 2016, table 16.4a). Because we are interested in estimating the labour market returns of human capital investments, we focus on the subsample of 23- to 55-year-old PIAAC respondents, corresponding to 3122 individuals.

Our key dependent variables are employment status and wages. Respondents were asked to report their employment status and were classified as being employed if, in the previous week, they worked for pay for at least one hour or if they were temporarily absent from a job of business they intended to return to (as a result, people on parental leave or on sick leave are still considered as employed).[2] Our earnings measure consists in the log of hourly earnings converted into PPP-adjusted US dollars. However the employment structure of the Italian workforce includes a large fraction of self-employed, who are more likely to abstain from reporting labour earnings, as shown in figure 1. Thus, when studying the economic returns of education and skills in Italy, one faces a double selection: the selection into being employed and conditional on it the selection into reporting positive earnings. While hourly wages are reported by 99% of employees and monthly earnings are declared by 92% of employees, just 65% of the self-employed report positive monthly earnings. The self-employed who do not report any earnings are not "randomly distributed": they are more likely to reside in Southern regions where employment rates are lower and self-employment is more frequent, possibly because self-employment lies in a grey area between participation in formal and informal market activities (see table 1). There is also evidence income under/not reporting being correlated to lower level of trust in others (see tables A.1 and A.2 of the online Appendix - on the concept of trust see Borgonovi and Burns 2015). However, we are not able to find a credible identification strategy associated to earnings reporting, and therefore in the sequel we will provide results referred to hourly wage (dependent employees only) and monthly earnings (all reporting employees).

Figure 1 – Employment status of PIAAC 1st cycle interviewees – Italy 23-55 year old



[2] For respondents in employment, PIAAC asks earnings in the format that is most convenient to the respondent (hourly, weekly, monthly, or yearly) and then harmonises such responses in hourly earnings. If a respondent declines to provide an amount, interviewers use a card showing national wage percentiles and record the respondent's decile location; actual wages are then imputed from the decile placement and covariates including age, gender, education, occupation, and assessed skills (see section 20.4 of OECD 2016).

Table 1 – Population by macro-regions and self-declared employment status (sample weights) - Italy PIAAC 2012

	dependent employee	self-employed	not employed	Total (excluding missing)	share of self-employment on total employment	non-employment rate	reporting monthly earnings: employees	reporting monthly earnings: self-employed
North West	522	119	179	821	0.186	0.218	0.910	0.732
North East	367	115	110	592	0.238	0.186	0.939	0.517
Centre	348	96	164	608	0.216	0.269	0.942	0.728
South	288	95	355	738	0.247	0.481	0.913	0.659
Islands	138	42	158	338	0.233	0.467	0.917	0.618
Total	1664	467	966	3097	0.219	0.312	0.924	0.654

Human capital is captured through educational attainment and directly assessed skills. Educational attainment is measured as the country-specific typical number of years required to complete the highest qualification held by the respondent (i.e., an imputed "years of schooling" measure consistent with the national education system). Skills are measured using PIAAC literacy and numeracy proficiency scores. Literacy reflects the "*ability to understand, evaluate, use and engage with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential*" and numeracy reflects the "*ability to access, use, interpret, and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life*" (OECD, 2012).

PIAAC reports literacy and numeracy proficiency as ten plausible values, i.e. random draws from each respondent's posterior proficiency distribution rather than a single deterministic score (Rhemtullaa and Savalei 2025). We used the Stata routine REPEST.ADO by Avvisati and Keslair (2014) for descriptive analyses. However, in the IV framework adopted in this paper, the objective is to isolate the component of skills that is predictably shifted by excluded instruments. For this purpose, the analysis uses the respondent-level mean across the ten plausible values as the main proficiency measure. This average can be interpreted as the expected posterior mean (EAP) of proficiency, which is the minimum-mean-squared-error summary of the latent skill given the respondent's item responses (and the conditioning information used in the assessment model) (Wu, 2005). Using the PV mean therefore "de-noises" proficiency by averaging out within-respondent imputation variability that is not informative about the systematic component exploited by IV. In practical terms, this choice yields a single, reproducible skill regressor and stabilises first-stage diagnostics, while preserving the economically relevant signal that is mapped into fitted values in 2SLS. Consistent with this interpretation, and anticipating results, we find that the estimated first-stage relationships and second-stage coefficients are very similar when the model is instead run separately on each plausible value and combined. This indicates that the substantive identifying variation is already contained in the PV-mean measure. PIAAC provides replicate weights to allow variance estimation that reflects complex sample features (stratification, clustering, unequal selection probabilities) as well as population weights. We incorporate both set of weights in analyses (OECD, 2016).

Whereas the PIAAC public use datafile does not contain information on where respondents live, we were able to access a restricted use dataset for Italy containing information on the municipality in which respondents were born. We use this information in our identification strategy. More specifically, we link information obtained from the population census of 1971 and 1981 on illiteracy rate in the municipality in which PIAAC respondents were born, separately for men and women. For each PIAAC participant we estimate the share of people of the same gender as the respondent and in the same municipality in which the respondent was born who were without formal education.

Two policy changes are used to identify exogenous variation in schooling. Both are implemented as intention-to-treat indicators based on cohort exposure. The first concerns the establishment of full-day primary schooling (*tempo pieno*). Law 820 of del 1971 introduced, for the first time, the possibility of full-day schooling at the primary level (grade 1 to 5, from 8.00 to 16.00), while previously only 4 hours were covered.[3] The aim of the reform was to enrich students' education through a more comprehensive and flexible school experience, expanding the role of the school to include not only cognitive aspects, but also relational, social, socialisation- and autonomy-related ones. We consider individuals born after 1966 as being potentially affected by this reform, even if we are unable to ascertain whether they were attending full-day classes. The second reform is the adoption of the "3+2" Bologna university system, which changed the structure of tertiary education. The Bologna process comprises a set of European wide agreements aimed at harmonising the architecture of the European higher education system. Italy ratified the agreement in 1999 (Law 509/1999). The reform changed the length of most tertiary education degrees, splitting the old duration (mostly four years) into 3-year courses (bachelor) followed by 2-year courses (master). The aim of the reform was to increase participation in tertiary education by shortening its duration, increasing the variety of contents and standardising course lengths through the adoption of ECTS (European Credit Transfer System). Given its gradual introduction within Italian universities, we assume that birth cohorts born after 1980 were potentially treated. In both cases individuals were potentially treated, and these reforms are to be interpreted as "intention to treat". As such they provide lower bound estimates of true effects.

4. Identification strategy

We intend to estimate the causal impact of schooling and skills on labour earnings in the Italian labour market. This estimation poses two problems: non-random selection into employment (and, for some outcomes, into earnings observability) and potential endogeneity of schooling and skills due to self-selection and omitted ability.[4] To address these issues we adopt an empirical strategy that combines an instrumental-variables

[3] The introduction is conditional on parental demand (which is correlated with female participation to the labour market), teacher availability and presence of refectories within the schools. This explains the large variability of full-day schooling across the country, with a national average of 33% of students spending 40 hours per week in schools and 37% spending less than 27 hours per week. On the advantages of longer hours in schools in deprived areas see Battistin and Meroni (2016).

[4] Hanushek et al (2015) discuss extensively the issue but they do not consider the joint endogeneity of schooling and numeracy (that they take as best proxy for skill).

approach with a Heckman-style selection correction following the control-function implementation proposed by Wooldridge (2010, pg.567 ss). Selection bias in the estimation of the returns to education (schooling and achievements) is therefore modelled as an omitted variable problem (Heckman 1979), where the omitted variable is defined as the inverse Mills ratio obtained from the selection equation that includes the instruments used to cope with potential endogeneity.

In symbols our model consists of three equations

$$y_{1i} = \mathbf{x}_{\mathbf{1}i}\boldsymbol{\delta}_{\mathbf{1}} + \boldsymbol{\beta}\mathbf{y}_{\mathbf{2}i} + v_{1i} \tag{1}$$

$$\mathbf{y}_{\mathbf{2}i} = [\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3]_{\boldsymbol{i}}\, \boldsymbol{\delta}_{\mathbf{2}} + v_{2i} \tag{2}$$

$$y_{3i} = 1([\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3]_{\boldsymbol{i}}\boldsymbol{\delta}_{\mathbf{3}} + v_{3i} > 0) \tag{3}$$

The first equation is the structural equation of interest (determinants of earnings), the second equation is a linear projection for the potentially endogenous variables $\mathbf{y}_{\mathbf{2}}$ (schooling and skills), and the third equation is the selection equation (presence in the labour market and/or reporting earnings). The proposed solution is estimating

$$y_{1i} = \mathbf{x}_{\mathbf{1}i}\boldsymbol{\delta}_{\mathbf{1}} + \boldsymbol{\beta}\hat{\mathbf{y}}_{\mathbf{2}i} + \gamma\hat{\lambda}_i + \epsilon_i \tag{4}$$

where the inverted Mills ratio $\hat{\lambda}_i = \frac{f(\mathbf{x}_i\hat{\boldsymbol{\delta}}_{\mathbf{3}})}{F(\mathbf{x}_i\hat{\boldsymbol{\delta}}_{\mathbf{3}})}$ is obtained from estimating equation (3) on the entire sample including the entire vector of $\mathbf{x}$'s in the regression, and $\hat{\mathbf{y}}_{\mathbf{2}}$ is the projection of the (potentially) endogenous variables obtained from equations (2) including the inverse Mills ratio among the regressors.

The vectors of regressors that we have chosen are the following:
$\mathbf{x}_1$ (demographics) = gender; age; age^2; foreign born; macro-regions of residence.
$\mathbf{x}_2$ (instruments for education) = illiteracy rate (of the corresponding gender) in the municipality of birth during the decade of birth (from population census); full day schooling (applies if born after 1966); separation of BA and MA (known as *Bologna process* – applies if born after 1980)
$\mathbf{x}_3$ (identification of selection equation) = spouse unemployed; number of children; having a child aged below 10; (subjective) trust in others.

5. Descriptive evidence

Descriptive statistics for the relevant variables are reported in table A.3 in the online Appendix. Consistent with prior evidence based on PIAAC data (Hanushek et al 2015), we find that hourly wages are positively correlated with both years of schooling and skills, with the strongest bivariate association observed for numeracy. Table 2 reports employment rates and mean monthly earnings by educational attainment and numeracy proficiency. Educational attainment reflects three categories (individuals who did not obtain an upper secondary qualification, those who obtained an upper secondary qualification, those who obtained a tertiary degree). For numeracy, adults are grouped depending on the PIAAC proficiency levels.[5] Table 2 indicates that both employment and earnings increase with schooling and with numeracy: individuals with higher educational attainment and higher numeracy proficiency display higher employment probabilities and higher mean monthly earnings. For example, the employment rate of individuals who completed post-secondary and tertiary qualifications and achieve at proficiency level 4 in the PIAAC numeracy assessment proficiency is 87% whereas it is only 47% among those who completed, at most, lower secondary qualifications and achieve at below Level 1 in the PIAAC numeracy assessment.

Table 2 – Employment and monthly earnings (sample weighed means)
– Italy PIAAC 2012 – population aged 23-55

EMPLOYMENT RATE	level of numeracy					
Maximal educational attainment	0	1	2	3	4	Total
Lower secondary or less (ISCED 1, 2, 3C)	0.47	0.59	0.67	0.81	0.85	0.64
Upper secondary (ISCED 3A-B, C long)	0.70	0.66	0.68	0.79	0.90	0.73
Postsec (ISCED 4) and tertiary (ISCED 5-6)	0.87	0.75	0.75	0.82	0.87	0.80
Total	0.51	0.62	0.68	0.80	0.88	0.70
MONTHLY EARNINGS	level of numeracy					
Maximal educational attainment	0	1	2	3	4	Total
Lower secondary or less (ISCED 1, 2, 3C)	1524	1834	1654	1795	2171	1726
Upper secondary (ISCED 3A-B, C long)	1602	1796	1822	2074	2362	1968
Postsec (ISCED 4) and tertiary (ISCED 5-6)	1756	3755	2406	2596	3322	2738
Total	1545	1959	1823	2156	2755	2012

While table 2 reports unconditional means, in table 3 we report the conditional correlations between earnings, schooling and skills (full model results are available in table A.4 of the online Appendix). In these specifications, the estimated return to years of schooling is around 4%, which is modest relative to many other countries and suggests limited wage differentiation by formal educational attainment in the Italian labour market.[6] Age, which partly proxies labour-market experience, accounts for an additional component of earnings variation. When skills are included (either proxied by numeracy or literacy, since their joint inclusion is never statistically significant), they exhibit a positive ad non negligible correlation: one standard deviation increase in skill is associated to a 5% increase, mostly among dependent employees. Conditional on observed characteristics, women and the foreign-born exhibit lower earnings, and the largest penalty in terms of wages and earnings is associated with living and working in Southern Italy.

[5] According to OECD (2016), level 0 corresponds to a score <175, level 1 to the interval (175,225], level 2 to the interval (225,275], level 3 to the interval (275,325] and level 4 to a score >325. The assignment to an interval has been based on the individual mean of ten plausible values of numeracy.
[6] The estimated return to education is similar when this is estimated using other datasets, like the Bank of Italy survey. See, for example, Bussolo et al. (2023).

Table 3 - Conditional correlation with hourly wages (dep.employees) or monthly earnings (all employed) – OLS – Italy PIAAC 2012 – population aged 23-55 – 10 plausible values with 80 replication weights

Dep. Variable	log hourly wage (dep.employees)	log hourly wage (dep.employees)	log hourly wage (dep.employees)	log monthly earnings (dep.+self-employed)	log monthly earnings (dep.+self-employed)	log monthly earnings (dep.+self-employed)
numeracy	0.001***		0.001	0.001		0.001
	[0.000]		[0.001]	[0.000]		[0.001]
literacy		0.001**	0.000		0.000	0.000
		[0.000]	[0.001]		[0.001]	[0.001]
years of schooling	0.037***	0.038***	0.037***	0.045***	0.046***	0.045***
	[0.003]	[0.003]	[0.003]	[0.006]	[0.005]	[0.005]

Gender, age, citizenship, region of residence and constant included; monthly earnings also includes log of worked hours (see table A.4 of the online Appendix) - 80 replications over 10 plausible values - *** p<0.01, ** p<0.05, * p<0.1

6. Self-selection and endogeneity

The OLS results reported in table 3 cannot be interpreted as causal. Estimates may in fact be biased by non-random selection into employment (particularly salient given the North-South divide), endogeneity of educational attainment and skills (given the absence of controls for unobservable ability), and measurement error (in the skills assessment). To address these issues, we implement the selection-corrected IV strategy outlined above.

The first step estimates a probit model for the probability of being employed. A key practical choice concerns the definition of "employment" for the selection equation. There are three possibilities: a broad definition based on self-declared employment; an intermediate definition based on being employed and reporting positive labour earnings; and a restrictive definition based on dependent employment with reported earnings. In the main text we report in analyses that adopt the intermediate definition, as it aligns the selection model with the population used in the earnings regressions. Results for alternative definitions are reported in Table A.5 of the online Appendix. Table 4 reports the corresponding selection equation for the preferred definition.

Table 4 – Self-selection into employment for those reporting positive wages or earnings – probit (sample weights) – Italy PIAAC 2012 - population aged 23-55

VARIABLES	non missing earnings (includes self-employed, excludes non-reporting)	non missing earnings (includes self-employed, excludes non-reporting)
Outcome: employment probability	0.644	0.644
female	-0.575***	-0.586***
	[0.061]	[0.062]
age	0.206***	0.200***
	[0.059]	[0.060]
age²	-0.003***	-0.003***
	[0.001]	[0.001]
foreign born	-0.055	-0.086
	[0.102]	[0.103]
region of residence=north east	-0.005	0.004
	[0.089]	[0.090]
region of residence=centre	-0.035	-0.028
	[0.090]	[0.090]
region of residence=south	-0.562***	-0.566***
	[0.090]	[0.090]
region of residence=islands	-0.479***	-0.472***
	[0.113]	[0.113]
spouse unemployed	-0.174	-0.201*
	[0.115]	[0.115]
number of children	-0.103***	
	[0.026]	
having a child younger than 10		-0.155**
		[0.071]
trust in others	0.105***	0.104***
	[0.029]	[0.029]
Observations	3,033	3,033
Pseudo R²	0.0941	0.0904

Robust standard errors in brackets – Illiteracy rates, educational reforms and constant included – *** p<0.01, ** p<0.05, * p<0.1

The exclusion restrictions used to identify the selection equation are based on household composition (spouse's employment status, number of children, the presence of a child under the age of 10) and on a measure of interpersonal trust. These variables are expected to shift labour supply and employability (and, in the Italian context, the likelihood of being observed with reportable earnings), but we expect these to be less directly related to hourly pay conditional on human capital and region of residence. Trust is included because it plausibly captures features of the social environment that are associated with compliance and reporting behaviour in Italy, like, for example, lower social capital and higher tax evasion in some areas (Guiso et al., 2004) and the greater prevalence of informal economic activity (Barra and Papaccio, 2024). Since regional differences are already absorbed by region-of-residence dummies, the residual variation in trust is interpreted as an individual attitude that may correlate with employability and reporting propensity rather than as an independent structural determinant of wages. The estimated probit is then used to compute the inverse Mills ratio, which enters subsequent stages of the estimation.

Table 4 documents large differences in selection into the earnings sample. Conditional on other covariates, women have a considerably lower probability of having non missing earnings than men, which is consistent with a large gender gap in employment in Italy. A strong geographical gradient also emerges: residence in the South and Islands is associated with markedly lower employment probabilities relative to the North West. Parenthood is negatively associated with having non missing earnings, and the parenthood penalty is larger when a child below age 10 is present. Finally, higher reported trust is positively associated with selection into the earnings sample, even after controlling for region of residence. The estimates presented in table 4 correspond to equation (3) and are used to construct the Mill's ratio which is included in the selection-corrected earnings equation (4).

The second step provides valid exogenous variations for the potentially endogenous human capital measures - years of schooling, numeracy and literacy. Table 5 reports key first-stage coefficients of the determinants of skills and years of schooling, whereas the full model results is available in table A.6 in the online Appendix.[7] We hypothesise that adult skills partly reflect early-life local environments. Under the assumption that municipality of birth is informative about the environment in which early skill formation occurred (i.e. no migration in the early years), each respondent was matched to gender- and cohort-specific municipality illiteracy rates from the 1971 and 1981 population censuses.

Table 5 – First stage regression for schooling and skills (weighed OLS) –
– Italy PIAAC 2012 - population aged 23-55

VARIABLES	Numeracy (OLS PV average)	Literacy (OLS PV average)	Years of schooling (OLS)
Female	-20.675	-6.417	0.919
	[16.228]	[16.318]	[1.379]
Age	7.771	7.034	0.52
	[6.856]	[6.817]	[0.521]
Age²	-0.093	-0.085	-0.007
	[0.085]	[0.085]	[0.006]
foreign born	-39.105***	-41.880***	-1.290***
	[6.594]	[5.723]	[0.276]
illiteracy rate	-52.148**	-44.969**	-1.818
	[21.918]	[18.664]	[1.212]
Bologna process (3+2)	9.769	6.612	1.406***
	[7.926]	[7.423]	[0.506]
full day schooling	0.781	-0.927	-0.036
	[8.154]	[7.693]	[0.588]
Observations	1,637	1,637	1,637
R-squared	0.150	0.161	0.144

Robust standard errors in brackets – controls for macro-region of residence, spouse unemployed, number of children, youngest child below 10-year-old, trust in other and Mills ratio and constant included – numeracy and literacy are measured as averages across 10 plausible values – full model reported in Table A.6 in the online Appendix - - *** p<0.01, ** p<0.05, * p<0.1

[7] Table A.6 in the online Appendix also reports the same estimates using 80 replications for each of the 10 plausible values. While the coefficients are identical by construction, the corresponding standard errors tend to be larger, but the relevant coefficients still retain their significance.

The first-stage estimates presented in Table 5 confirm that the municipality-level illiteracy rate is strongly and negatively associated with adult numeracy and literacy, consistent with the interpretation that early-life local environments are predictive of later competencies. By contrast, years of schooling are primarily predicted by cohort exposure to education reforms, particularly the Bologna "3+2" reform, which is associated with a statistically significant increase in schooling. The full-day primary schooling reform exhibits the expected positive sign but is imprecisely estimated in the baseline specification[8]. These first-stage regressions correspond to equation (2) in the framework described above and provide the fitted values used in the selection-corrected IV estimates of equation (4), corresponding to the determinants of hourly wage in the Italian labour market.

Table 6 reports the selection-corrected IV estimates for log hourly wages among dependent employees. In the baseline specification, which uses the respondent-level mean across plausible values as the main proficiency measure, numeracy is positively and statistically significantly associated with hourly wages. By contrast, the estimate for years of schooling is not statistically distinguishable from zero once endogeneity is addressed and skills are included, and the point estimate is small and negative. The association between literacy and wages is negative but imprecisely estimated (results are not statistically significant in the specification that fitted using replicate weights). Given the high correlation between literacy and numeracy, this estimate is best interpreted as the partial association of literacy conditional on numeracy rather than as evidence that literacy is penalised in wage-setting. These patterns contrast with the OLS associations reported earlier, where schooling is strongly correlated with earnings, and suggests that the portion of the schooling-earnings gradient that survives the IV strategy operates primarily through competencies rather than formal attainment.

Table A.7 in the online Appendix provides additional evidence on the robustness of these findings to the treatment of plausible values. By running the model without replicate weights we are able to report standard IV diagnostics and we show that omitting replicate weights does not change identification or systematically shift point estimates although it yields marginally smaller standard errors. Estimating the model separately on each plausible value yields numeracy coefficients that are uniformly positive and of similar magnitude across the ten specifications, with most estimates statistically significant at conventional levels. Literacy coefficients are consistently negative but smaller in magnitude and less precisely estimated across plausible values, which is consistent with limited independent variation in literacy once numeracy is controlled for. By contrast, the schooling coefficient varies in sign across plausible values and is consistently imprecisely estimated, reinforcing the conclusion that the wage gradient attributed to schooling in OLS is not robust in the selection-corrected IV framework. Table A.7 reports first-stage F-statistics obtained from pv-by-pv estimation (as well as from the pv-mean model). These diagnostics are modest but stable across plausible values (roughly around 5-6 for numeracy and schooling, and somewhat lower for literacy).

[8] Other authors (Brunello and Miniaci 1999) have considered the 1962 extension of compulsory education, which hit cohorts born after 1952, thus not discriminating in our sample. However, we need three separate instruments if we aim to consider three potentially endogenous variables. We therefore retained this second reform despite its weak statistical significance.

Table 6 – Determinants of hourly wages – IV estimation –
Italy PIAAC 2012 - population aged 23-55

VARIABLES	PV avg with replicates	PV avg no replicates
Numeracy	0.026**	0.025**
	[0.012]	[0.010]
literacy	-0.024	-0.021*
	[0.015]	[0.011]
years of schooling	-0.006	-0.017
	[0.080]	[0.064]
female	0.334	0.282
	[0.293]	[0.216]
age	0.013	0.014***
	[0.005]	[0.004]
foreign born	-0.290	-0.21
	[0.290]	[0.213]
Observations	1637	1637
R^2	-1.06	-0.948

Robust standard errors in brackets – population weights – region of residence, Mills ratio and constant included -
column 1: average over plausible values – column 2: 80 replications over 10 plausible values – *** p<0.01, ** p<0.05, * p<0.1

7. Conclusions

This paper revisits the returns to human capital in Italy by distinguishing between formal educational attainment (years of schooling) and directly assessed adult competencies (literacy and numeracy) using data from the PIAAC first cycle. Descriptive patterns replicate the well-known finding that both schooling and skills are positively correlated with employment and earnings. However, these correlations are not readily interpretable as causal in the presence of non-random selection into observed earnings, endogenous schooling choices, and measurement uncertainty in assessed skills.

To address these issues, the analysis combines a Heckman-style selection correction with instrumental-variables identification of both years of schooling and skills. The first-stage results highlight two distinct sources of variation: cohort exposure to education reforms predicts years of schooling, whereas local cultural conditions in the place of birth, proxied by municipality-level illiteracy rates, strongly predict adult skills. In the second stage, once selection and joint endogeneity are accounted for, the earnings return to additional years of schooling becomes small and is statistically indistinguishable from zero. Numeracy remains the skill domain most consistently linked to wages in the Italian labour market. These findings suggest that, within the margin of variation identified by the instruments, labour-market rewards in Italy depend on numeracy skills but not formal educational attainment per se.

A possible interpretation of these findings is that educational qualifications may provide a relatively noisy signal of productive capabilities in the labour market, and, as a result, employers and wage-setting institutions place greater weight on skills that are revealed on the job or inferred through performance. This is consistent with previous findings that formal school assessment does not correspond to actual skills, especially in schools located in Southern regions (Argentin and Triventi, 2015; Iacus and Porro 2011). And broader evidence on mismatch in Italy, which indicates that qualification mismatch is relatively prevalent and that low measured skill returns coexist with limited skill use in some jobs (Hanushek et al., 2015; Allen et al., 2013; Pellizzari and Fichen, 2017; Monti and Pellizzari, 2016). It would also be consistent with evidence that, when educational credentials do not reliably map into skills, screening and pay may shift toward alternative signals of potential task performance.

These results are also relevant for the growing literature and policy interest on “skills-first” approaches - hiring and workforce practices that prioritise demonstrated skills over traditional proxies such as degrees or job titles (OECD, 2025b). There is a growing use of skills-first practices across OECD labour markets, including greater reliance on explicit skills requirements in online vacancies and the rising role of skills signalling through mechanisms such as online profiles and micro-credentials. Even where hiring may remain credential-oriented, our work suggests that wage premia within employment may be more strongly aligned with competencies than with additional schooling once endogeneity and selection are addressed. From a policy perspective, this combination points to the potential value of strengthening transparent, credible skill validation, so that skills-first practices do not simply replace one imperfect proxy with another and do not exacerbate bias or weaken occupational standards, risks the OECD highlights explicitly (OECD, 2025b).

Our findings raise the issue of whether the Italian education system is able to efficiently translate additional schooling and educational qualifications into skills that are valued in the labour market. The evidence we present is consistent with the possibility that expansions in educational attainment, such as those induced by tertiary education reforms, may not translate into skills that promote higher productivity per se, at least for the marginal cohorts affected by the reform we have considered. This interpretation resonates with a broader literature emphasising that what matters for productivity and growth is not only the quantity of schooling but also quality of skills acquired (Hanushek and Woessmann, 2008; Hanushek and Woessmann, 2012). At the same time, this paper’s design identifies local average treatment effects tied to specific sources of variation.

A final note of caution pertains to the recognition of the broader benefits of education, which go well beyond the labour-market. A large literature documents that education has important social and civic benefits (Dee, 2004; Milligan, Moretti and Oreopoulos, 2004; Oreopoulos and Salvanes, 2011) and therefore education policy should be evaluated on the basis of a wider set of objectives than its effect on productivity alone. Based on this broader literature, the evidence we present on the earnings returns to numeracy and the fact that formal attainment is not linked to higher wages in the present IV framework underscores the importance of policies that raise the skill content of schooling but also strengthen pathways for lifelong learning and skill recognition, so that both economic and civic objectives can both be advanced.

References

Allen, J. P., Levels, M., and van der Velden, R. K. W. (2013). Skill mismatch and skill use in developed countries: Evidence from the PIAAC study. ROA. ROA Research Memoranda No. 017

Angrist, J. D., and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings?. *The Quarterly Journal of Economics*, 106(4): 979-1014.

Argentin, Gianluca and Moris Triventi. (2015). The North-South Divide in School Grading Standards: New Evidence from National Assessments of the Italian Student Population. *Italian Journal of Sociology of Education*, 7(2), 157-185.

Avvisati, Francesco and François Keslair, 2014. REPEST: Stata module to run estimations with weighted replicate samples and plausible values, Statistical Software Components S457918, Boston College Department of Economics, revised 11 Dec 2024.

Barra, Cristian and Anna Papaccio. 2024. Does Regulatory Quality Reduce Informal Economy ? A Theoretical and Empirical Framework. *Social Indicators Research* 172: 543–567

Battistin, Erich and Elena Claudia Meroni. (2016) Should we increase instruction time in low achieving schools? Evidence from Southern Italy. *Economics of Education Review* 55: 39–56

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5, Part 2): 9-49.

Blau, Francine and Kahn, Lawrence. (2005). Do cognitive test scores explain higher US wage inequality. *Review of Economics and Statistics*, 87(1): 184-193.

Borgonovi, F. and T. Burns (2015), The educational roots of trust, OECD Education Working Papers, No. 119, OECD Publishing, Paris. http://dx.doi.org/10.1787/5js1kv85dfvd-en

Bound, John and Gary Solon. (1999) Double trouble: on the value of twins-based estimation of the return to schooling. *Economics of Education Review* 18 (2): 169-182

Brunello, G. and Miniaci, R. (1999). The economic returns to schooling for Italian men. An evaluation based on instrumental variables. *Labour Economics*, *6*(4): 509-519.

Brunello, G., Fort, M. and Weber, G. (2009). Changes in compulsory schooling, education and the distribution of wages in Europe. *The Economic Journal*, 119(536): 516-539.

Bussolo, Maurizio, Daniele Checchi and Vito Peragine. 2023. The long term evolution of inequality of opportunity, *Journal of Economic Inequality* 21(2): 277-323

Cappellari, Lorenzo, Paolo Castelnovo, Daniele Checchi and Marco Leonardi. (2017). Skilled or educated? Educational reforms, human capital and earnings, *Research in Labor Economics* 45: 173-197

Card, David. (1993). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. NBER Working Papers 4483

Card, David. (1999). The causal effect of education on earnings. *Handbook of Labor economics*, *3*: 1801-1863.

Clay, Karen, Jeff Lingwall and Melvin Stephens Jr. (2021). Laws, educational outcomes, and returns to schooling evidence from the first wave of U.S. state compulsory attendance laws. *Labour Economics* 68, 101935

Dee, Thomas S. (2004). Are there civic returns to education? *Journal of Public Economics,* 88(9–10), 1697–1720.

Del Rey, Elena, Sergi Jimenez-Martin and Judit Vall Castello. (2018) Improving educational and labor outcomes through child labor regulation. *Economics of Education Review* 66: 51-66

Deming, David and Mikko Silliman. (2025). Skills and human capital in the labor market. Chpt.2 of Christian Dustmann and Thomas Lemieux (eds) *Handbook of Labor Economics* Volume 6: 115-157.

Green, D. A., and Riddell, W. C. (2003). Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics*, 10(2): 165-184.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2004). The Role of Social Capital in Financial. Development. *American Economic Review*, 94(3): 526–556

Hampf, F., Wiederhold, S. and Woessmann, L. (2017). Skills, earnings, and employment: exploring causality in the estimation of returns to skills. *Large-scale Assessments in Education*, 5(1), 12.

Hanushek, Eric A., and Woessmann, Ludger (2008). The role of cognitive skills in economic development. *Journal of Economic Literature,* 46(3), 607–668.

Hanushek, Eric A., and Woessmann, Ludger (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth,* 17*,* 267–321.

Hanushek, Eric A., Guido Schwerdt, Simon Wiederhold and Woessmann, Ludger. (2015). Returns to Skills around the World: Evidence from PIAAC. *European Economic Review*, 73(C):103-130.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. *Handbook of the Economics of Education*, *1*: 307-458.

Heckman, James. (1979). Sample Selection Bias as a Specification Error. *Econometrica* 47: 153–161

Iacus, Stefano M., and Giuseppe Porro. (2011). Teachers' Evaluations and Students' Achievement: A "Deviation from the Reference" Analysis. *Education Economics* 19(2):139–59.

ISFOL (2013). Le competenze per vivere e lavorare oggi – principali evidenze dall'indagine PIAAC (a cura di Gabriella Di Francesco). Isfol Research Paper n.9/2013

ISFOL (2024). *Le competenze cognitive in Italia nel contesto internazionale*. Rome

Leuven, E., Oosterbeek, H. and 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.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1): 3-42.

Milligan, Kevin, Moretti, Enrico, and Oreopoulos, Phillip. (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics,* 88(9–10), 1667–1695.

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.

Monti, Paola and Michele Pellizzari. (2016). Skill Mismatch and Labour Shortages in the Italian Labour Market. IGIER Bocconi Policy Brief n.2

OECD (2012), *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing, Paris, https://doi.org/10.1787/9789264128859-en.

OECD (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills,* OECD Publishing. http://dx.doi.org/10.1787/9789264204256-en

OECD (2016). Technical Report of the Survey of Adult Skills (PIAAC) (2nd Edition). Paris https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/piaac/technical-reports/cycle-1/PIAAC_Technical_Report_2016.pdf

OECD (2019). *OECD Skills Studies. Skills Matter. Additional results from the survey of adult skills,* OECD Publishing. https://www.oecd-ilibrary.org/education/skills-matter_1f029d8f-en

OECD (2021), COVID-19 and Well-being: Life in the Pandemic, OECD Publishing, Paris, https://doi.org/10.1787/1e1ecb53-en

OECD (2025a), *Getting Skills Right. Trends in Adult Learning. New Data from the 2023 Survey of Adult Skills*. OECD Publishing, Paris https://www.oecd.org/en/publications/trends-in-adult-learning_ec0624a6-en.html

OECD (2025b), *Empowering the workforce in the context of a skills-first approach*, OECD Skills Studies, OECD Publishing, Paris, https://doi.org/10.1787/345b6528-en.

Olivetti, C. and Petrongolo, B. (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics*, 26(4): 621-654.

Oreopoulos, Phillip and Salvanes, Kjell G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives,* 25(1), 159–184.

Paccagnella, M. (2015). Skills and Wage Inequality: Evidence from PIAAC. *OECD Education Working Papers*, No. 114, OECD Publishing, Paris, https://doi.org/10.1787/5js4xfgl4ks0-en.

Patrinos, Harry Anthony and George Psacharopoulos. (2025). Causal Returns to Education. IZA DP No. 18077

Pellizzari, Michele and Anne Fichen. (2017). A new measure of skill mismatch: theory and evidence from PIAAC. *IZA Journal of Labor Economics* 6: 1

Rhemtullaa, Mijke and Victoria Savalei. 2025. Estimated Factor Scores Are Not True Factor Scores. *Multivariate Behavioral Research* 60(3): 598–619

Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data.* MIT Press

Wu, M. (2005). The Role of Plausible Values in Large-Scale Surveys. *Studies in Educational Evaluation* 31: 114-128

Appendix (online)

Table A.1 – Trust in others by macro-regions (sample weights) - Italy PIAAC 2012

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree	Total
	There are only a few people you can trust completely					
North West	41.79	38.35	9.58	6.73	3.54	100
North East	35.53	40.86	14.14	8.08	1.39	100
Centre	37.36	46.86	7.00	7.02	1.76	100
South	47.88	34.01	10.15	5.64	2.32	100
Islands	58.41	29.11	7.27	3.72	1.50	100
Total	42.99	38.46	9.83	6.45	2.27	100
	If you are not careful, other people will take advantage of you					
North West	37.70	39.25	14.48	5.44	3.12	100
North East	37.06	41.58	15.38	5.35	0.63	100
Centre	37.73	43.70	11.99	5.11	1.47	100
South	52.72	33.93	5.82	5.27	2.26	100
Islands	60.75	28.95	4.62	4.12	1.55	100
Total	43.68	38.18	11.02	5.17	1.95	100

Table A.2 – Fraction of population reporting earnings information by macro-regions and trust (sample weights) - Italy PIAAC 2012

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree	Total
	If you are not careful, other people will take advantage of you					
North West	0.669	0.687	0.692	0.814	0.626	0.686
North East	0.639	0.685	0.750	0.762	0.794	0.683
Centre	0.601	0.671	0.657	0.787	1.000	0.654
South	0.392	0.503	0.309	0.556	0.757	0.442
Islands	0.463	0.381	0.618	0.536	0.609	0.452
Total	0.542	0.619	0.649	0.712	0.727	0.595

Table A.3 – Descriptive statistics (sample weights) – Italy PIAAC 2012 – population aged 23-55

Variable	Obs	Mean	Std.dev.	Min	Max
demographics					
female	3069	0.502	0.500	0	1
foreign born	3069	0.090	0.286	0	1
age	3069	39.675	8.274	23	55
residence=north west	3069	0.265	0.442	0	1
residence=north east	3069	0.191	0.393	0	1
residence=centre	3069	0.197	0.398	0	1
residence=south	3069	0.239	0.427	0	1
residence=islands	3069	0.107	0.309	0	1
labour market outcomes					
employed	3069	0.702	0.458	0	1
worked hours	2315	38.91	12.05	1	125
hourly wage	1650	11.71	7.25	0.06	89.92
monthly earnings	1866	2011.55	1839.36	0.00	41666.67
schooling and skills					
years schooling	3069	11.28	3.83	5	21
numeracy (80 replications)	3069	252.07	49.40	64.06	411.29
literacy (80 replications)	3069	253.89	44.10	81.76	399.51
instrumental variables					
illiteracy rate 1970-80	3033	0.146	0.125	0	0.620
3+2 university reform	3069	0.205	0.404	0	1
full-day school reform	3069	0.716	0.451	0	1
unemployed spouse	3069	0.061	0.240	0	1
number children	3069	1.098	1.171	0	19
at least one child <10 year old	3069	0.198	0.398	0	1
trust in others	3068	1.838	0.952	1	5

Table A.4 - Conditional correlation with hourly wages (dep.employees) or monthly earnings (all employed) – OLS – Italy PIAAC 2012

	1	2	3	4	5	6
VARIABLES	log hourly wage	log hourly wage	log hourly wage	log monthly earnings	log monthly earnings	log monthly earnings
numeracy	0.001***		0.001	0.001		0.001
	[0.000]		[0.001]	[0.000]		[0.001]
literacy		0.001**	0.000		0.000	0.000
		[0.000]	[0.001]		[0.001]	[0.001]
years of schooling	0.037***	0.038***	0.037***	0.045***	0.046***	0.045***
	[0.003]	[0.003]	[0.003]	[0.006]	[0.005]	[0.005]
Female	-0.150***	-0.158***	-0.150***	-0.225***	-0.231***	-0.223***
	[0.025]	[0.025]	[0.026]	[0.046]	[0.045]	[0.047]
Age	0.015***	0.015***	0.015***	0.018***	0.018***	0.019***
	[0.001]	[0.001]	[0.001]	[0.003]	[0.002]	[0.003]
foreign born	-0.202***	-0.202***	-0.201***	-0.238***	-0.243***	-0.242***
	[0.040]	[0.042]	[0.042]	[0.062]	[0.062]	[0.063]
region of residence=North East	-0.078**	-0.072**	-0.078**	-0.119**	-0.112**	-0.118**
	[0.032]	[0.032]	[0.032]	[0.047]	[0.048]	[0.048]
region of residence=Centre	-0.097***	-0.091***	-0.097***	-0.238***	-0.233***	-0.239***
	[0.035]	[0.035]	[0.035]	[0.047]	[0.047]	[0.047]
region of residence=South	-0.183***	-0.185***	-0.183***	-0.295***	-0.297***	-0.295***
	[0.045]	[0.046]	[0.045]	[0.061]	[0.060]	[0.060]
region of residence=Islands	-0.192***	-0.193***	-0.192***	-0.408***	-0.409***	-0.409***
	[0.052]	[0.052]	[0.052]	[0.084]	[0.084]	[0.085]
(log of) worked hours				0.642***	0.643***	0.642***
				[0.071]	[0.071]	[0.071]
Observations	1,650	1,650	1,650	1,833	1,833	1,833
R^2	0.209	0.206	0.209	0.314	0.314	0.315

Robust standard errors in brackets – constant included - 80 replications over 10 plausible values -
*** p<0.01, ** p<0.05, * p<0.1

Table A.5 – Self-selection into employment for those reporting positive wages or earnings – probit (sample weights) – Italy PIAAC 2012 - population aged 23-55

VARIABLES	self declared employed	self declared employed	non missing earnings (includes self-employed, excludes non-reporting)	non missing earnings (includes self-employed and non-reporting)	non missing hourly wage (excludes self-employed and non-reporting)	non missing hourly wage (excludes self-employed but includes non-reporting)
Outcome: employment probability	0.758	0.758	0.644	0.644	0.537	0.537
female	-0.858***	-0.869***	-0.575***	-0.586***	-0.339***	-0.353***
	[0.070]	[0.071]	[0.061]	[0.062]	[0.060]	[0.060]
age	0.232***	0.230***	0.206***	0.200***	0.132**	0.131**
	[0.063]	[0.064]	[0.059]	[0.060]	[0.058]	[0.058]
age^2	-0.003***	-0.003***	-0.003***	-0.003***	-0.002**	-0.002**
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
foreign born	-0.257**	-0.285***	-0.055	-0.086	-0.035	-0.069
	[0.105]	[0.106]	[0.102]	[0.103]	[0.099]	[0.099]
region of residence=north east	0.190*	0.200*	-0.005	0.004	0.007	0.017
	[0.104]	[0.104]	[0.089]	[0.090]	[0.084]	[0.085]
region of residence=centre	-0.124	-0.114	-0.035	-0.028	-0.07	-0.061
	[0.102]	[0.103]	[0.090]	[0.090]	[0.086]	[0.086]
region of residence=south	-0.640***	-0.641***	-0.562***	-0.566***	-0.515***	-0.517***
	[0.102]	[0.102]	[0.090]	[0.090]	[0.087]	[0.087]
region of residence=islands	-0.556***	-0.539***	-0.479***	-0.472***	-0.425***	-0.415***
	[0.124]	[0.124]	[0.113]	[0.113]	[0.111]	[0.110]
spouse unemployed	-0.13	-0.154	-0.174	-0.201*	-0.149	-0.176
	[0.126]	[0.124]	[0.115]	[0.115]	[0.115]	[0.115]
number of children	-0.100***		-0.103***		-0.120***	
	[0.029]		[0.026]		[0.026]	
having a child younger than 10		-0.210***		-0.155**		-0.234***
		[0.077]		[0.071]		[0.069]
trust in others	0.111***	0.110***	0.105***	0.104***	0.098***	0.097***
	[0.032]	[0.032]	[0.029]	[0.029]	[0.028]	[0.028]
illiteracy rate	-0.547	-0.655*	-0.152	-0.265	0.035	-0.089
	[0.350]	[0.354]	[0.321]	[0.323]	[0.314]	[0.315]
Bologna process (3+2)	0.099	0.093	0.041	0.04	0.035	0.032
	[0.155]	[0.156]	[0.147]	[0.147]	[0.145]	[0.145]
full day schooling	-0.051	-0.013	-0.169	-0.142	-0.03	0.005
	[0.156]	[0.158]	[0.136]	[0.137]	[0.131]	[0.131]
Observations	3,033	3,033	3,033	3,033	3,033	3,033
Pseudo R^2	0.164	0.162	0.0941	0.0904	0.0572	0.0541

Robust standard errors in brackets – constant included - *** p<0.01, ** p<0.05, * p<0.1

Table A.6 – First stage regression for schooling and skills (OLS) –
– Italy PIAAC 2012 - population aged 23-55

VARIABLES	Numeracy (80 replications)	Literacy (80 replications)	Years of schooling (80 replications	Numeracy (OLS on PV average)	Literacy (OLS on PV average)	Years of schooling (OLS)
female	-20.675	-6.417	0.919	-20.675	-6.417	0.919
	[16.228]	[16.318]	[1.379]	[13.169]	[12.248]	[1.219]
age	7.771	7.034	0.52	7.771	7.034	0.52
	[6.856]	[6.817]	[0.521]	[5.622]	[5.299]	[0.484]
age²	-0.093	-0.085	-0.007	-0.093	-0.085	-0.007
	[0.085]	[0.085]	[0.006]	[0.070]	[0.066]	[0.006]
foreign born	-39.105***	-41.880***	-1.290***	-39.105***	-41.880***	-1.290***
	[6.594]	[5.723]	[0.276]	[5.347]	[4.786]	[0.356]
region of residence=north east	21.840***	17.153***	0.343	21.840***	17.153***	0.343
	[5.018]	[4.016]	[0.285]	[3.782]	[3.351]	[0.281]
region of residence=centre	15.885***	10.856***	0.353	15.885***	10.856***	0.353
	[4.628]	[3.933]	[0.262]	[3.869]	[3.498]	[0.301]
region of residence=south	-23.324	-15.192	-0.215	-23.324*	-15.192	-0.215
	[16.377]	[16.287]	[1.353]	[13.551]	[12.981]	[1.260]
region of residence=islands	-16.937	-13.213	0.558	-16.937	-13.213	0.558
	[13.841]	[13.915]	[1.184]	[11.711]	[10.953]	[1.085]
illiteracy rate	-52.148**	-44.969**	-1.818	-52.148***	-44.969***	-1.818
	[21.918]	[18.664]	[1.212]	[17.152]	[15.203]	[1.259]
Bologna process (3+2)	9.769	6.612	1.406***	9.769	6.612	1.406**
	[7.926]	[7.423]	[0.506]	[6.663]	[6.513]	[0.557]
full day schooling	0.781	-0.927	-0.036	0.781	-0.927	-0.036
	[8.154]	[7.693]	[0.588]	[7.436]	[6.543]	[0.571]
spouse unemployed	-10.974	-14.345*	-1.126*	-10.974*	-14.345**	-1.126**
	[8.159]	[8.045]	[0.592]	[6.613]	[6.486]	[0.562]
number of children	-6.074	-7.260*	-0.493*	-6.074*	-7.260**	-0.493*
	[4.067]	[3.883]	[0.265]	[3.333]	[3.006]	[0.265]
having a child younger than 10	7.088	6.74	0.492	7.088	6.74	0.492
	[5.561]	[5.558]	[0.385]	[4.645]	[4.225]	[0.375]
trust in others	11.136***	9.268***	0.912***	11.136***	9.268***	0.912***
	[3.107]	[2.981]	[0.251]	[2.608]	[2.388]	[0.239]
Mills ratio (reporting income)	67.104	48.625	2.143	67.104	48.625	2.143
	[49.360]	[50.469]	[4.242]	[41.110]	[38.658]	[3.827]
Constant	60.963	86.38	-0.999	60.963	86.38	-0.999
	[148.453]	[148.154]	[11.394]	[120.716]	[114.454]	[10.513]
Observations	1,637	1,637	1,637	1,637	1,637	1,637
R²	0.150	0.161	0.144	0.168	0.181	0.144

Robust standard errors in brackets – constant included – column 1-2-3: 80 replications over 10 plausible values – column 4-5-6: weighed OLS using averages across 10 plausible values - *** p<0.01, ** p<0.05, * p<0.1

Table A.7 – Determinants of hourly wages – IV estimation – Italy PIAAC 2012 - population aged 23-55

VARIABLES	PV1	PV2	PV3	PV4	PV5	PV6	PV7	PV8	PV9	PV10	PV avg
numeracy	0.019**	0.023*	0.014***	0.019**	0.012*	0.013***	0.016*	0.020***	0.031*	0.025**	0.025**
	[0.009]	[0.013]	[0.005]	[0.008]	[0.007]	[0.005]	[0.008]	[0.007]	[0.017]	[0.010]	[0.010]
literacy	-0.012	-0.016	-0.012*	-0.012	-0.015	-0.01	-0.006	-0.022**	-0.016	-0.020*	-0.021*
	[0.008]	[0.015]	[0.007]	[0.008]	[0.012]	[0.008]	[0.009]	[0.011]	[0.013]	[0.011]	[0.011]
years of schooling	-0.057	-0.037	0.012	-0.019	0.073	0.002	-0.053	0.068	-0.082	-0.012	-0.017
	[0.069]	[0.090]	[0.043]	[0.054]	[0.071]	[0.054]	[0.069]	[0.062]	[0.111]	[0.080]	[0.064]
female	0.289	0.294	0.087	0.098	-0.041	0.073	0.102	0.041	0.468	0.29	0.282
	[0.259]	[0.306]	[0.140]	[0.171]	[0.143]	[0.133]	[0.187]	[0.160]	[0.425]	[0.249]	[0.216]
age	0.012***	0.012**	0.014***	0.013***	0.017***	0.014***	0.014***	0.022***	0.014**	0.014***	0.014***
	[0.004]	[0.005]	[0.003]	[0.004]	[0.004]	[0.003]	[0.004]	[0.005]	[0.006]	[0.005]	[0.004]
foreign born	-0.174	-0.131	-0.254	-0.098	-0.322	-0.199	0.028	-0.329*	0.203	-0.112	-0.21
	[0.188]	[0.243]	[0.177]	[0.163]	[0.253]	[0.223]	[0.205]	[0.192]	[0.364]	[0.252]	[0.213]
region of residence=north east	-0.264**	-0.205	-0.160*	-0.216**	-0.065	-0.178	-0.280**	-0.122	-0.524*	-0.320*	-0.233**
	[0.127]	[0.143]	[0.082]	[0.101]	[0.118]	[0.108]	[0.128]	[0.094]	[0.290]	[0.164]	[0.118]
region of residence=centre	-0.215**	-0.275**	-0.150**	-0.212***	-0.119*	-0.175**	-0.226***	-0.177**	-0.409**	-0.283**	-0.232***
	[0.089]	[0.124]	[0.062]	[0.081]	[0.071]	[0.076]	[0.086]	[0.074]	[0.204]	[0.122]	[0.089]
region of residence=south	0.106	0.053	-0.001	-0.078	-0.058	-0.076	-0.053	-0.098	0.186	0.053	0.07
	[0.178]	[0.188]	[0.113]	[0.122]	[0.126]	[0.098]	[0.129]	[0.127]	[0.278]	[0.170]	[0.147]
region of residence=islands	0.031	0.002	-0.066	-0.089	-0.115	-0.127	-0.004	-0.134	0.069	0.023	0.002
	[0.162]	[0.187]	[0.110]	[0.130]	[0.132]	[0.118]	[0.151]	[0.154]	[0.235]	[0.189]	[0.148]
Mills ratio (reporting income)	-0.587	-0.403	-0.328	-0.145	-0.213	-0.166	-0.195	-0.192	-0.674	-0.600	-0.500
	[0.395]	[0.392]	[0.248]	[0.243]	[0.263]	[0.201]	[0.282]	[0.263]	[0.567]	[0.396]	[0.322]
Constant	1.015	0.69	1.219	0.597	1.738	0.974	-0.037	1.339	-0.835	0.841	1.138
	[1.011]	[1.276]	[0.846]	[0.875]	[1.283]	[1.060]	[1.041]	[0.927]	[1.864]	[1.329]	[1.068]
Observations	1637	1637	1637	1637	1637	1637	1637	1637	1637	1637	1637
R^2	-1.035	-1.649	-0.39	-0.821	-0.42	-0.342	-0.857	-1.223	-3.093	-1.759	-0.948
F-test for first stage regressions											
Numeracy	5.37	5.61	5.39	4.62	5.91	6.15	5.28	5.54	3.61	4.58	5.79
Literacy	3.56	5.14	4.21	3.81	3.93	2.95	4.42	4.74	3.36	4.25	4.50
Years of schooling	6.53	6.53	6.53	6.53	6.53	6.53	6.53	6.53	6.53	6.53	6.53

Robust standard errors in brackets – population weights – column 1-10: individual specific plausible values – column 11: average over plausible values – *** p<0.01, ** p<0.05, * p<0.1