MERITOCRACY ACROSS COUNTRIES: SUPPLEMENTARY APPENDIX

Oriana Bandiera Ananya Kotia Ilse Lindenlaub Christian Moser
Andrea Prat

October 8, 2025

Throughout this Supplementary Appendix, we indicate figures, tables, and equations within this appendix by SA.# ('S' for Supplementary). In turn, figures, tables, and equations from the main paper are denoted by just 1,2,.... Figures, tables, and equations from the main appendix are denoted by A.#.

SA Data Appendix

SA.1 Details on the PIAAC Dataset

Our main data source is the Survey of Adult Skills of the PIAAC administered by the OECD between 2012 and 2018. It comprises a representative sample of the working-age population in each participating country. The target population for the survey is all non-institutionalized adults between the ages of 16 and 65 who reside in the country at the time of data collection. The data comprise over 230,000 respondents, representing 815 million adults aged 16 to 65 surveyed in 38 countries. Of these, 30 countries have all the information (e.g., continuous wages) required for our study. The countries included in our study are Austria, Belgium, Chile, Cyprus, the Czech Republic, Denmark, Ecuador, Estonia, Finland, France, Germany, Greece, Ireland, Israel, Italy, Japan, Kazakhstan, Korea, Lithuania, Mexico, the Netherlands, New Zealand, Norway, Poland, Russia, Slovakia, Slovenia, Spain, the United Kingdom, and the US. The countries available in PIAAC but excluded from our study due to missing key variables are Australia, Canada, Hungary, Indonesia, Peru, Singapore, Sweden, and Turkey. Technical documentation for the PIAAC is available from OECD (2019), and the original microdata are available for download from OECD (2024).

SA.2 Key Variables

Hourly Wages. We harmonize individual hourly wages to a common 2012 purchasing-power-parity (PPP) basis in two steps. First, starting from the survey's hourly raw wage in local currency units (LCU), we deflate to 2012 prices using each country's consumer price index matched to the survey year. This yields an hourly wage in LCU at 2012 prices. Second, we convert the 2012 LCU wages into 2012 PPP wages in international dollars using the World Bank's PPP conversion factor for 2012 (local currency per PPP $). In two cases (Lithuania and Austria), inconsistencies in LCU (raw) wages lead us to replace this manually adjusted 2012 PPP wage variable with the survey's PPP hourly wage variable.

Worker Skills. PIAAC formally assesses three cognitive skill domains: literacy, numeracy, and information and communication technology (ICT). For our analysis, we use only numeracy and literacy skills, as ICT skill information is missing for several key countries in our sample. PIAAC reports scores from incentivized tests for numeracy and literacy skills on a scale from 0 to 500, which we use to construct country-specific skill distributions (see Section 2 of the paper for details). These scores can be interpreted as follows.

A numeracy score below 176 means that the respondent can perform simple processes such as counting, sorting, performing basic arithmetic operations, or understanding simple percentages; a score of 176–225 implies that the respondent is comfortable with basic mathematical processes in common, concrete contexts where the mathematical content is explicit; a score of 226–275 implies that the respondent can identify and act on mathematical information embedded in common contexts; scores in the range of 276–325 mean that the respondent understands mathematical information which may be less clear, whose representation may be more complex or in unfamiliar contexts; scores in the range 326–375 imply that the respondent understands complex mathematical information and ideas, which may involve chained processes; and finally, a score of 376–500 implies that the respondent understands complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts.

For literacy, scores below 176 imply that a respondent can only locate a single piece of information in short, simple texts. A score of 176–225 implies basic text processing where the requested information is explicit and easily matched; 226–275 requires comparing or integrating information across short texts; 276–325 means being able to interpret and evaluate longer, denser texts and digital formats; 326–375 implies understanding and critically engaging with complex or abstract texts;

and 376–500 reflects advanced skills in evaluating, comparing, and synthesizing multiple sources, often in digital contexts.

Job Skill Requirements. To measure occupational skill requirements, we proceed as follows. We first measure individual-level skill use in numeracy and literacy domains based on self-reported data from PIAAC on workers' skill use in their jobs. The survey asks workers to report whether they perform certain tasks in each skill domain. For numeracy, these tasks are (1) use a calculator, (2) calculate costs or budgets, (3) use or calculate fractions or percentages, (4) prepare charts graphs or tables, (5) use simple algebra or formulas, (6) use advanced math or statistic. For literacy, these are (1) read directions or instructions, (2) read letters memos or mails, (3) read newspapers or magazines, (4) read professional journals or publications, (5) read books, (6) read manuals or reference materials, (7) read financial statements, (8) read diagrams maps or schematics, (9) write letters memos or mails, (10) write articles, (11) write reports, and (12) fill in forms.

To measure numeracy and literacy requirements of each individual in their job, we first map each of these tasks to a standardized proficiency scale. These proficiency levels are defined by the OECD (OECD, 2019) on a 1–5 scale and describe the kinds of tasks a worker at each level is expected to master. Numeracy proficiency level 1 corresponds to basic counting, sorting, or performing simple arithmetic operations; level 2 includes tasks such as simple measurement and spatial representation, and interpretation of relatively simple data and statistics in texts, tables and graphs; level 3 corresponds to recognizing and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; level 4 includes tasks which involve multi-step problem-solving strategies and complex reasoning; and level 5 involves understanding complex, abstract, and formal mathematical or statistical ideas. For literacy, proficiency level 1 includes reading directions or filling in forms, level 2 covers reading or writing short communications such as letters or emails, level 3 involves reading manuals, reference materials, or diagrams, level 4 includes writing reports and interpreting financial statements, and level 5 involves more complex tasks such as writing articles or reading professional journals.

We assign each PIAAC task the closest proficiency level based on the descriptions of both tasks and proficiency levels. In doing so, the six numeracy tasks are assigned the following proficiency levels: using a calculator (level 1), working with fractions or percentages (level 1.5), preparing figures, charts, or tables (level 1.5), calculating costs or budgets (level 3), using algebra or formulas (level 4), and using advanced mathematics or statistics (level 5). In turn, we assign the following proficiency levels to the twelve literacy tasks: reading directions or instructions (level 1), reading diagrams, maps, or schematics (level 1.5), reading letters, memos, or emails (level 2), reading newspapers or magazines (level 2), reading reference manuals (level 3), reading books (level 4), reading financial statements (level 4), reading professional journals or publications (level 4), filling in forms (level 2), writing memos or emails (level 3), writing reports (level 5), and writing articles (level 5). For some robustness checks, we also use problem-solving skill requirements, and so we assign the following proficiency levels to the problem-solving tasks: performing simple computer operations (level 2) and performing complex computer operations (level 5).

For each worker, we code as 1 if they perform a task (0 otherwise), multiply it by the assigned proficiency level, and sum across all tasks within a domain. Dividing this weighted sum by the highest possible level of proficiency in the domain yields a normalized index in the range $[0, 1]$. These numeracy and literacy indices measure the effective difficulty of the tasks actually performed by the worker. Aggregating across workers within the same 2-digit ISCO occupation provides us with occupation-level skill requirements for numeracy and literacy in each country.

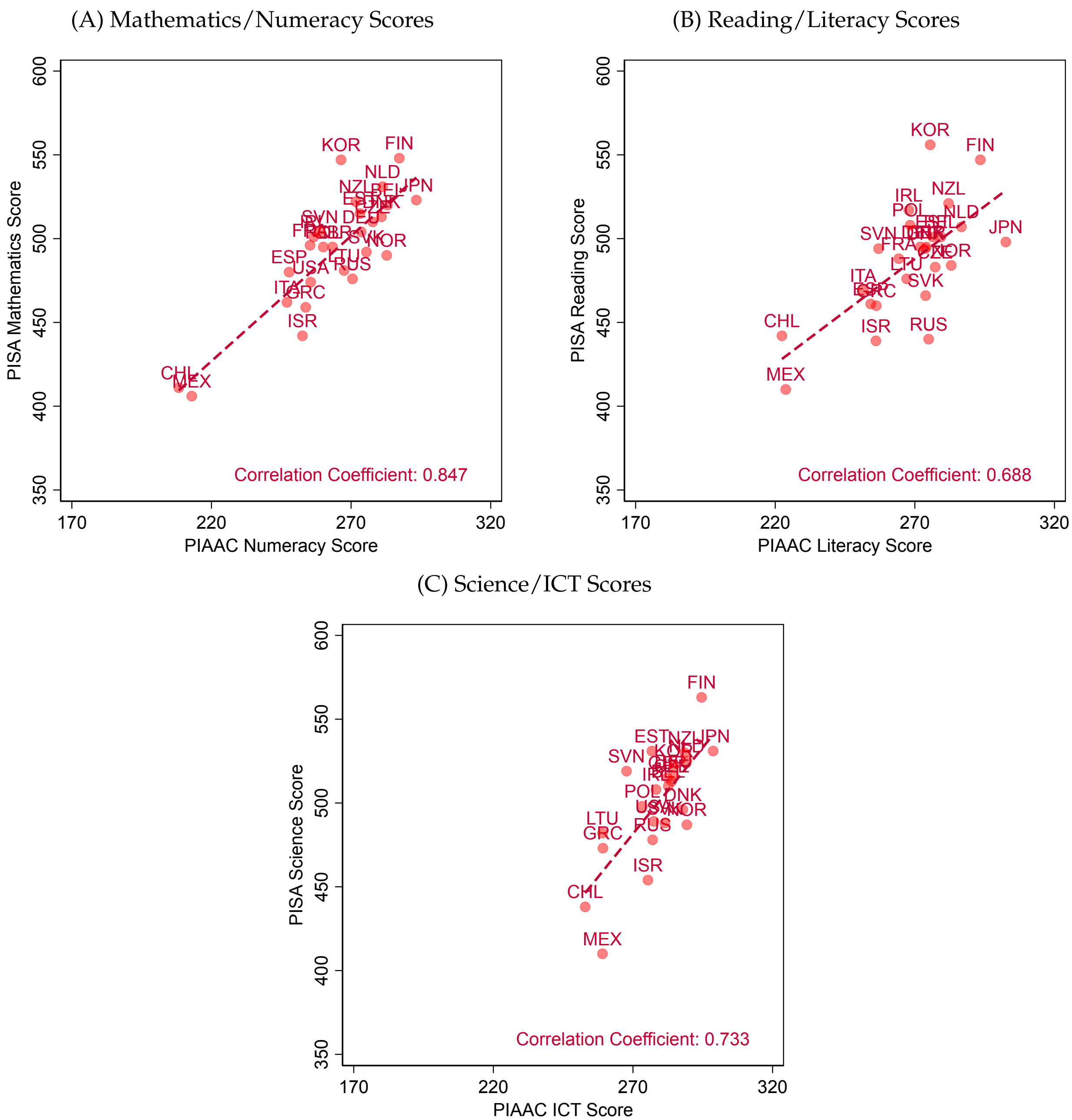
Demographics, Labor Markets, and Firms. In addition to data on worker skills and job skill requirements, we also use the PIAAC background questionnaire for data on the following variables: years of schooling, years of experience in the labor market, gender, age, occupation (4-digit ISCO codes), hourly wage, and employment status. We combine workers who are unemployed and out of the labor force so that each worker in our data is either employed or nonemployed. Finally, we use data on the size of the firm of a worker as a proxy for overall job productivity.

SA.3 Other Data Sources

We obtain GDP per capita (2012 $, PPP) from the World Development Indicators of the World Bank. We obtain data on labor shares from Our World in Data. We use 2012 as the reference year to merge these external data into our sample of PIAAC countries. Supplementary data on adolescents' test scores and institutional quality come from the PISA study and the World Development Indicators of the World Bank, respectively. For validation of the PIAAC labor market outcomes, we also use the Jobs of the World Database (JWD), a harmonized dataset on labor market outcomes and job characteristics in low- and middle-income countries (1990–2019), and the Luxembourg Income Study (LIS), a cross-national household survey database on income, employment, and demographics.

SA.4 Data Validation

Figure SA1: Correlation between Various PIAAC Test Scores and Corresponding PISA Test Scores

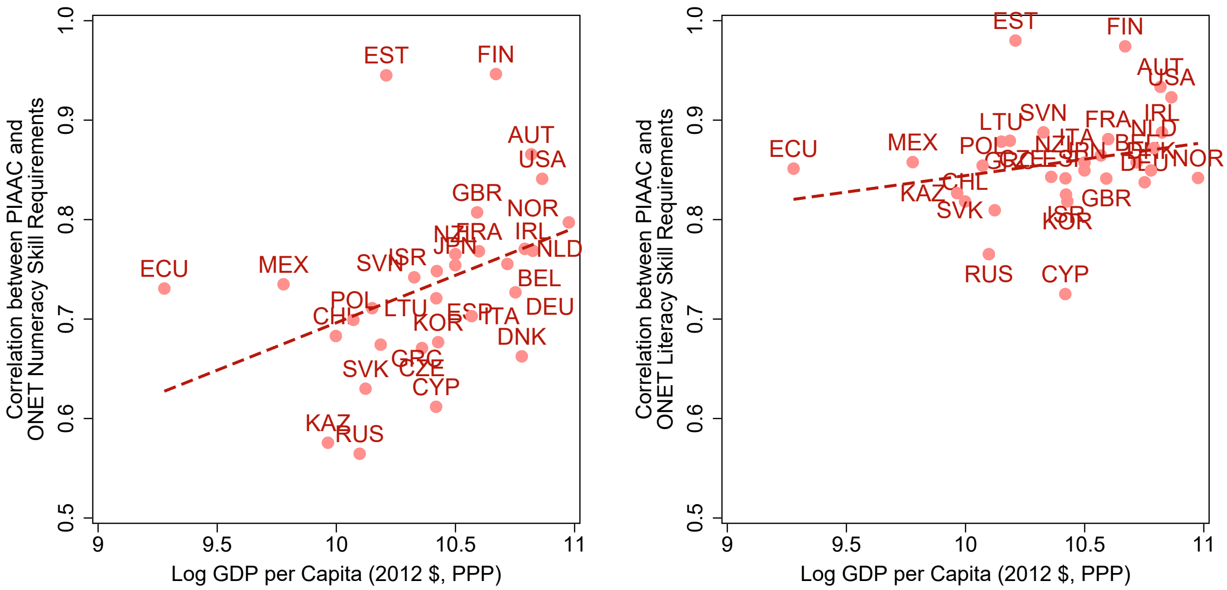


Notes: This figure shows scatter plots of the country means of various skill dimensions measured in the PISA (where tests were completed by 15-year-old students) and the corresponding skill dimensions in the PIAAC (where tests were completed by working-age adults) in the same countries for mathematics/numeracy skills (panel A), reading/literacy skills (panel B), and science/ICT scores (panel C). To account for the fact that individuals take PISA tests at a younger age, all PISA test scores are taken from the 2006 wave, except Lithuania, for which data is available only in the 2018 wave. *Source:* PISA and PIAAC.

Figure SA2: Correlation between PIAAC and O*NET Measures of Skill Requirements

(A) Numeracy Skill Requirements

(B) Literacy Skill Requirements



Notes: Both panels report the R^2 of country-specific regressions of the PIAAC skill requirement measures on the O*NET skill requirement measures at the occupational level plotted against log GDP per capita; panel A for numeracy skill requirements and panel B for literacy skill requirements. Occupational skill requirements from O*NET are computed as follows. We use three modules of the O*NET that provide numeracy and literacy task contents in the form of 128 skill requirements. These are characterized under the broad headings of (i) work activities, (ii) skills, and (iii) abilities for each of O*NET's 873 occupations under the *Standard Occupational Classification (SOC)* nomenclature. Using a crosswalk between O*NET SOC and PIAAC ISCO-08 occupation codes, we can assign O*NET numeracy and literacy skill requirement vectors to 441 4-digit ISCO-08 occupations in each country of the PIAAC. To reduce the dimensionality of the 128 skill requirements to just one numeracy and one literacy attribute, we follow Lise and Postel-Vinay (2020). First, we run Principal Component Analysis (PCA) on our large set of O*NET measures and keep the first two principal components. We then recover our numeracy and literacy skill requirement indices by recombining those two principal components, which by default are constructed to be orthonormal, in such a way that they satisfy the following two exclusion restrictions: (1) the "mathematics" score only reflects numeracy skill requirements, and (2) the "written comprehension" score only reflects literacy requirements. We rescale these O*NET skill requirements to lie in $[0, 1]$ by subtracting the minimum value and dividing by the range of values. *Sources:* PIAAC and O*NET.

Figure SA3: Average Numeracy Skill Requirements across Occupations and Countries

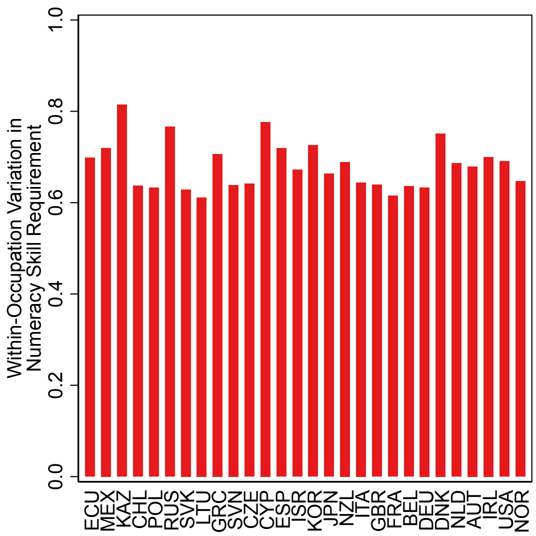
Note: This figure shows average numeracy skill requirements for each 1-digit occupation based on ISCO-08 occupation codes across countries. The occupational categories on the horizontal axis are the 1-digit ISCO-08 major groups. We exclude code 0, which corresponds to Armed Forces, and include codes 1-9: 1 = Managers, 2 = Professionals, 3 = Technicians and Associate Professionals, 4 = Clerical Support Workers, 5 = Services and Sales Workers, 6 = Skilled Agricultural/Forestry/Fishery Workers, 7 = Craft and Related Trades, 8 = Plant and Machine Operators/Assemblers, and 9 = Elementary Occupations. These ISCO major groups are ordered by skill level based on the ILO's framework (with managers/ professionals requiring higher complexity and training, and elementary occupations lower). To compute skill requirements, we sum the number of tasks a worker completes in each skill dimension—i.e., numeracy and literacy—and weigh them by their respective proficiency level; see Supplementary Appendix SA.2 for details. We then take the mean of this measure across all workers in a given 1-digit occupation in each country. *Source:* PIAAC.

Figure SA4: Average Literacy Skill Requirements across Occupations and Countries

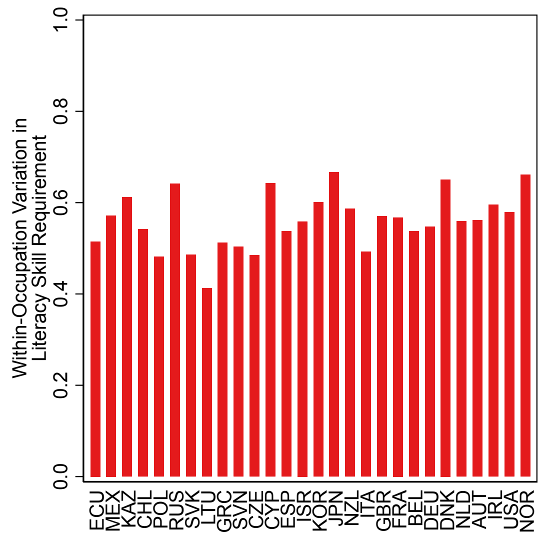
Note: This figure shows average literacy skill requirements for each 1-digit occupation based on ISCO-08 occupation codes across countries. The occupational categories on the horizontal axis are the 1-digit ISCO-08 major groups. We exclude code 0, which corresponds to Armed Forces, and include codes 1-9: 1 = Managers, 2 = Professionals, 3 = Technicians and Associate Professionals, 4 = Clerical Support Workers, 5 = Services and Sales Workers, 6 = Skilled Agricultural/Forestry/Fishery Workers, 7 = Craft and Related Trades, 8 = Plant and Machine Operators/Assemblers, and 9 = Elementary Occupations. These ISCO major groups are ordered by skill level based on the ILO's framework (with managers/ professionals requiring higher complexity and training, and elementary occupations lower). To compute skill requirements, we sum the number of tasks a worker completes in each skill dimension—i.e., numeracy and literacy—and weigh them by their respective proficiency level; see Supplementary Appendix SA.2 for details. We then take the mean of this measure across all workers in a given 1-digit occupation in each country. *Source:* PIAAC.

Figure SA5: Within-Occupation Variation in Skill Requirements

(A) Numeracy Skill Requirement



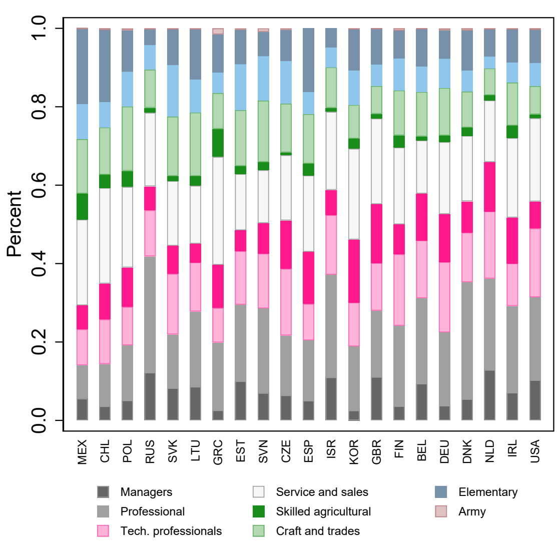
(B) Literacy Skill Requirement



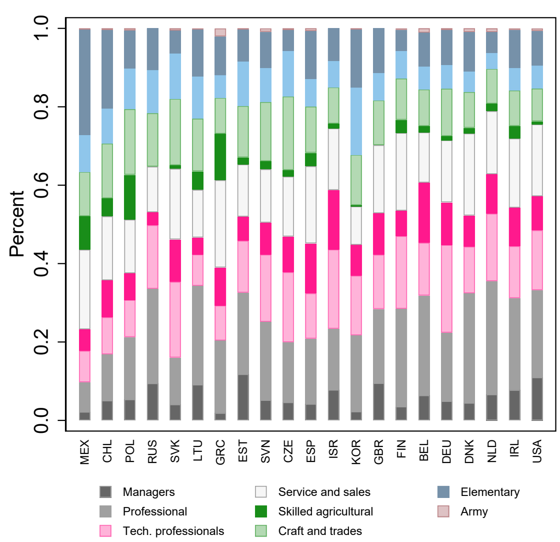
Notes: This figure shows the within-occupation (2-digit ISCO) variation in job skill requirements for each country, measured separately for numeracy (panel A) and literacy (panel B). The measure is calculated as the share of residual variance from a regression of *individual* job skill requirements on 2-digit occupation indicators. Countries are sorted by log GDP per capita (2012 $, PPP). *Source:* PIAAC.

Figure SA6: Comparison between Occupational Structures in PIAAC and LIS Data

(A) PIAAC

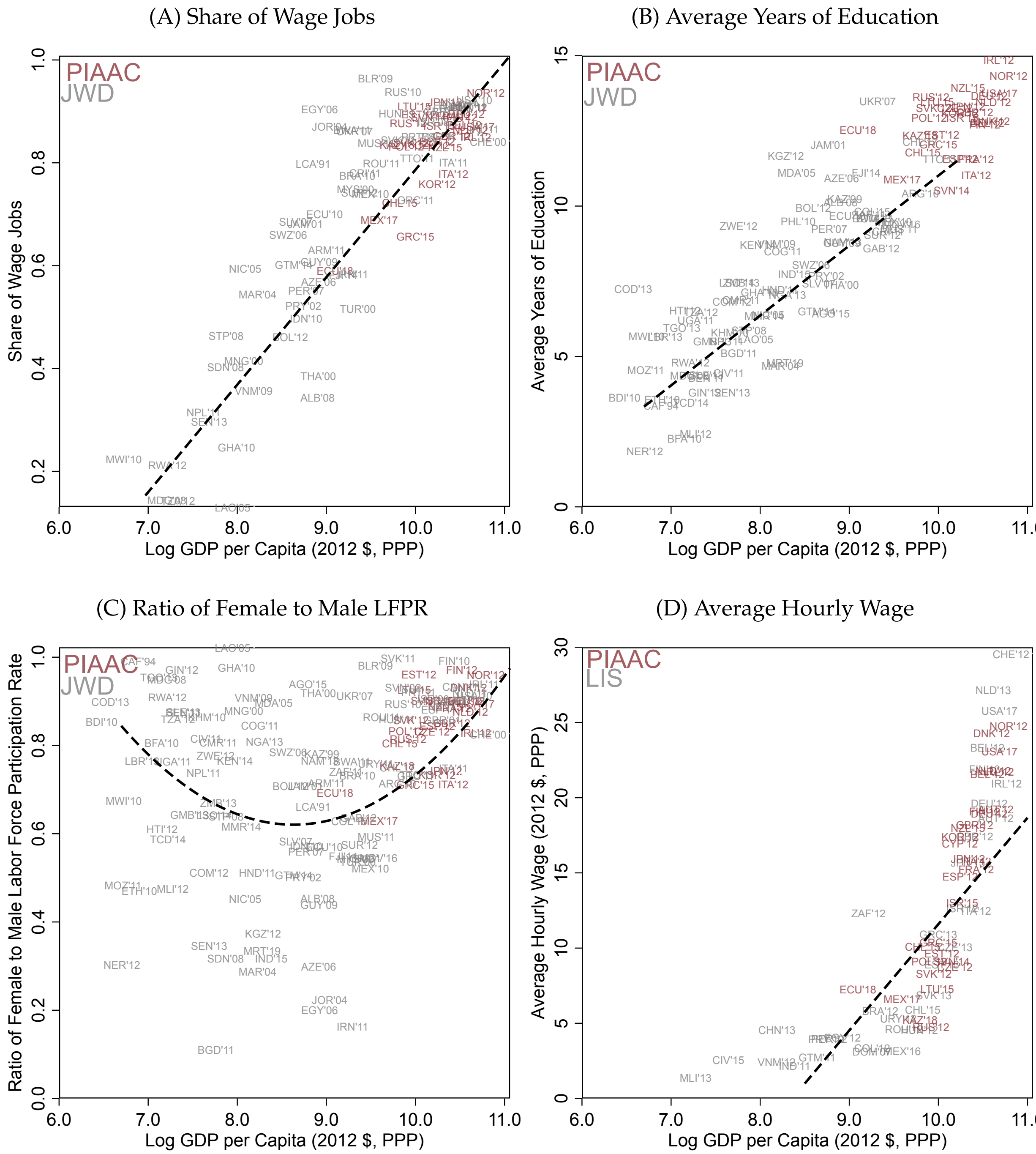


(B) LIS



Notes: This figure compares the occupational composition in the PIAAC (panel A) with the occupational composition in the LIS (panel B). Occupations are grouped into ten categories according to 1-digit ISCO-08 occupation codes. We report information for the 21 countries for which both PIAAC and LIS data are available. *Source:* PIAAC and LIS.

Figure SA7: Comparison of Cross-Country Labor Market Outcomes between PIAAC and Cross-Country Census as well as Survey Data



Notes: Panels A-C are based on cross-country data from PIAAC and The Jobs of the World Dataset (JWD); panel D is based on PIAAC and the Luxembourg Income Study (LIS). The JWD harmonizes cross-country census data on labor market outcomes from the Integrated Public Use Microdata Series (IPUMS) and the Demographic and Health Surveys (DHS). The year of the respective survey is indicated along the country labels. Panel A plots the share of workers in wage jobs (as opposed to being self-employed) for all countries in the JWD (gray) and PIAAC (red) data. Panel B plots the average years of education of the population in the JWD and PIAAC. Panel C plots the ratio of female to male labor force participation rate in the JWD and PIAAC. Finally, panel D compares average hourly wages in PIAAC (red) to average hourly wages of the same countries in the LIS (gray). *Source:* PIAAC, JWD, and LIS.

SB Estimation Appendix

SB.1 Homogeneity Properties

Proposition SA1 (Homogeneity Properties)**.** *Systematic wages $w(x,y)$, idiosyncratic wages $\tilde{w}(x_i,y_k)$, systematic profits $\nu(x,y)$, and systematic surplus $s(x,y)$ are all homogeneous of degree 1 in*

$$\mathcal{P} := (f(x,y), f_{x\emptyset}(x), f_{\emptyset y}(y), \sigma).$$

As a result, the matching frequencies $(\mu(x,y), \mu(\emptyset,y), \mu(x,\emptyset))$, the within-worker-job-cell share of log-wage dispersion out of overall log-wage dispersion, the economy-wide profit share, and the Matching Index $\mathcal{M}(\theta)$ are all homogeneous of degree 0 in $\mathcal{P}$.

Proof. Let $\lambda > 0$ be an arbitrary scalar. For now, assume that $(\mu(x,y), \mu(\emptyset,y), \mu(x,\emptyset))$ are homogeneous of degree 0 in $\mathcal{P}$. We will verify this property below.

Indexing the wage function by the set of parameters $\mathcal{P}$, $w(x,y;\mathcal{P})$, it follows that for all (x,y)

$$\begin{aligned} w(x,y;\lambda\mathcal{P}) &= (\lambda\sigma)\log\left(\frac{\mu(x,y)}{\mu(x,\emptyset)}\right) + (\lambda f_{x\emptyset}(x)) \\ &= \lambda w(x,y;\mathcal{P}), \end{aligned}$$

so the systematic wage function w is homogeneous of degree 1 in $\mathcal{P}$.

Similarly, for the systematic profit function, $\nu(x,y;\mathcal{P})$, we have for all (x,y)

$$\begin{aligned} \nu(x,y;\lambda\mathcal{P}) &= \lambda f(x,y) - w(x,y;\lambda\mathcal{P}) \\ &= \lambda\nu(x,y;\mathcal{P}), \end{aligned}$$

where the second equality follows from the previous result, so the systematic profit function, ν, is also homogeneous of degree 1 in $\mathcal{P}$.

Next, note that for any random variable $\delta \sim$ EV Type I$(0,\sigma)$, the cumulative distribution function $s(\delta;\sigma)$ satisfies

$$\begin{aligned} s(\delta;\lambda\sigma) &= \exp\left(-\exp\left(-\delta/(\lambda\sigma)\right)\right) \\ &= s(\delta/\lambda;\sigma). \end{aligned}$$

Therefore, if $\delta \sim$ EV Type I$(0,\sigma)$, then $\lambda\delta \sim$ EV Type I$(0,\lambda\sigma)$.
From this, it follows that the idiosyncratic wage function, $\tilde{w}(x,y;\mathcal{P})$, has for all (x_i,y_k) the property

$$\tilde{w}(x_i,y_k;\lambda\mathcal{P}) = \lambda w(x,y) + \lambda\delta_{iy}\lambda\tilde{w}(x_i,y_k;\mathcal{P}),$$

so the idiosyncratic wage function, $\tilde{w}$, is homogeneous of degree 1 in $\mathcal{P}$.

Next, for all (x, y), we can write the systematic surplus, $s(x, y; \mathcal{P})$, as

$$\begin{aligned} s(x, y; \lambda\mathcal{P}) &= \lambda f(x, y) - \lambda f_{x\emptyset}(x) - \lambda f_{\emptyset y}(y) \\ &= \lambda s(x, y; \mathcal{P}), \end{aligned}$$

so the systematic surplus function s is homogeneous of degree 1 in $\mathcal{P}$.

To summarize, we have shown that systematic wages $w(x, y)$, idiosyncratic wages $\tilde{w}(x_i, y_k)$, systematic profits $v(x, y)$, and systematic surplus $s(x, y)$ are all homogeneous of degree 1 in $\mathcal{P}$.

Given the results above, it follows from equations (1)–(4) that the conditional choice probabilities of workers and jobs are all homogeneous of degree 0 in $\mathcal{P}$ and so are matching frequencies $(\mu(x, y), \mu(\emptyset, y), \mu(x, \emptyset))$, which verifies the claim made at the beginning of this proof.

Since $\tilde{w}(x_i, y_k)$ is homogeneous of degree 1 in $\mathcal{P}$, we have

$$\log \tilde{w}(x_i, y_k; \lambda\mathcal{P}) = \log \lambda + \log \tilde{w}(x_i, y_k; \mathcal{P}).$$

Therefore,

$$\begin{aligned} Var(\log \tilde{w}(x_i, y_k; \lambda\mathcal{P})) &= Var(\log \lambda + \log \tilde{w}(x_i, y_k; \mathcal{P})) \\ &= Var(\log \tilde{w}(x_i, y_k; \mathcal{P})), \end{aligned}$$

so the dispersion of log idiosyncratic wages, $Var(\log \tilde{w}(x_i, y_k; \mathcal{P}))$, is homogeneous of degree 0 in $\mathcal{P}$. An analogous argument shows that the coefficient of determination (R^2) from a regression of log wages, $\log \tilde{w}(x_i, y_k)$, on indicators for interacted worker and job types (x, y) is homogeneous of degree 0 in $\mathcal{P}$. Note that $1 - R^2$ in the above-referenced regression is simply the within-worker-job-cell share of log-wage dispersion out of overall log-wage dispersion, which is then also homogeneous of degree 0 in $\mathcal{P}$.

Due to the homogeneity properties of output and wages established above, it also follows that the job-level profit share (D.6) is homogeneous of degree 0 in $\mathcal{P}$, and so is the economy-wide average profit share.

Finally, that the Matching Index $\mathcal{M}(\theta; \mathcal{P})$ is homogeneous of degree 0 in $\mathcal{P}$ is a direct consequence of the fact that $f(x, y; \mathcal{P})$ is homogeneous of degree 1 in $\mathcal{P}$ while $\mu(x, y)$ is homogeneous of degree 0 in $\mathcal{P}$. $\square$

References

Lise, Jeremy and Fabien Postel-Vinay, "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, 2020, *110* (8), 2328–2376.

OECD, "The Survey of Adult Skills (PIAAC)," *Technical Report, 3rd Edition*, 2019.

_ , "Programme for the International Assessment of Adult Competencies (PIAAC)," https://www.oecd.org/skills/piaac/, accessed on March 17, 2024.