

De Vera, Micole; Garcia-Brazales, Javier

Working Paper

Firm Size and the Task Content of Jobs: Evidence from 47 Countries

Suggested Citation: De Vera, Micole; Garcia-Brazales, Javier (2021) : Firm Size and the Task Content of Jobs: Evidence from 47 Countries, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/246591>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Firm Size and the Task Content of Jobs: Evidence from 47 Countries*

Micole De Vera[†] Javier Garcia-Brazales[‡]

November 15, 2021

Abstract

Using a mix of household- and employer-based survey data from 47 countries, we provide novel evidence that workers in larger firms perform more non-routine analytical and interpersonal tasks, even within narrowly defined occupations. Moreover, workers in larger firms rely more on the use of information and communications technologies (ICT) to perform these tasks. We also document a 17% wage premium that workers in larger firms enjoy relative to their counterparts in smaller firms. We find evidence that the firm size gradient in the task content of jobs accounts for around 10% of the large firm wage premium.

JEL CODES: J24, J31, L25

KEYWORDS: tasks, occupations, firm size, cross-country evidence, wage differential

*We thank Jaime Arellano-Bover for valuable comments and suggestions. We also thank Andrés Carmona for his excellent research assistance. De Vera acknowledges funding from Spain's Ministerio de Ciencia, Innovación y Universidades (PRE2018-084485) and Ministerio de Economía, Industria y Competitividad (María de Maeztu Programme for Units of Excellence in R&D, MDM-2016-0684). Garcia-Brazales acknowledges funding received from the Fundación Ramón Areces.

[†]CEMFI: micole.devera@cemfi.edu.es

[‡]CEMFI: javier.garcia@cemfi.edu.es

1 Introduction

Empirical studies show that the task composition of jobs can explain a large part of wage dispersion in the economy over time (Autor et al., 2003; Acemoglu and Autor, 2011; Autor and Handel, 2013; Goos et al., 2014; Acemoglu and Restrepo, 2021). Most studies are, however, limited by the fact that direct measures of the task content of jobs are not available in standard data sets. To overcome this limitation, authors typically resort to imputing the task content of occupations by means of alternative data sets such as the Occupation Information Network (O*NET) or the European Working Conditions Survey (EWCS). An implied assumption of this approach is that the task composition of jobs under the same occupation is homogeneous.

We show that this assumption does not hold empirically when we compare workers employed in firms of different sizes. Workers in larger firms perform more non-routine analytical and interpersonal tasks, and work with more information and communications technologies (ICT), even *within* narrowly defined occupations. This empirical pattern is robust across the 47 high-, middle-, and low-income countries we study combining the OECD Survey of Adult Skills and the World Bank Skills Measurement Surveys. We provide evidence that our finding is not only true for the average worker. More specifically, the distributions of the intensity of performed tasks (in non-routine analytical, non-routine interpersonal, and ICT) in larger firms are all shifted to the right relative to smaller firms.

To the best of our knowledge, this is the first paper to document in detail that tasks performed by workers differ between large and small firms. This result contributes to the growing literature documenting the heterogeneity in the task content of occupations, which has emphasized heterogeneity along dimensions other than firm size. For instance, based on reported job tasks, Autor and Handel (2013) and Stinebrickner et al. (2019) uncover differences in the job task composition within occupations. Based on job postings, Deming and Kahn (2018) find that there is substantial variation in skill requirements, even within narrowly defined occupations. Atalay et al. (2021) show that this variation is correlated with city size such that larger cities have higher intensity of analytical and interactive tasks, more technological requirements, and increased task specialization. A number of papers also document that reported occupational task contents vary across countries (Dicarlo et al., 2016; Lewandowski et al., 2019; De La Rica et al., 2020; Caunedo et al., 2021).

We then move on to explore the implications of this firm-size gradient in occupational task intensities on wage determination. We do this in two steps. First, we document that large firms pay higher wages compared to smaller firms — the so-called large firm wage premium (LFWP). On average, we find that workers in larger firms earn about 17% more than their counterparts in smaller firms, after controlling for 2-digit occupation codes. Furthermore, we show that this is not driven only by a few workers in larger firms that are

paid disproportionately more. Rather, the distribution of wages in larger firms is shifted to the right compared to the distribution of wages in smaller firms. Our measured large firm wage premium is consistent with the existence of a large, economically significant LFWP found in other studies employing alternative data sets (Velenchik, 1997; Gerlach and Hübler, 1998; Schaffner, 1998; Troske, 1999; Winter-Ebmer, 2001; Dobbelaere, 2004; Söderbom et al., 2005; Lehmer and Möller, 2010; Bloom et al., 2018; Colonnelli et al., 2018; Reed and Thu, 2019; Lochner et al., 2020; Porcher et al., 2021).

Second, we conduct a simple mediation analysis to provide suggestive evidence on the sources of this large firm wage premium, including our novel finding that tasks compositions vary across firms of differing size. A number of explanations for the existence of the LFWP have been proposed (Brown and Medoff, 1989; Oi and Idson, 1999): (i) large firms hire more skilled workers (worker selection); (ii) large firms have worse working conditions (compensating differentials); (iii) large firms have market power and share rents with workers (productivity); (iv) large firms have higher costs of monitoring and pay efficiency wages; and (v) large firms pay higher because of threat of unionization. In this paper, we explore the firm size gradient in occupational task intensities as a novel source of the LFWP. We find that the firm size gradient to the task content of jobs is able to explain 10% of the raw LFWP. This proportion explained may vary from 5 to 20% depending on the country. This is an economically significant fraction of the LFWP and is comparable to the fraction explained by the sorting of higher educated individuals to larger firms.

Our results have implications for our understanding of the labor market and open new questions about its functioning. Most notably, the underlying sources of the firm size gradient in the task content of jobs remain unexplored. Answering this requires better understanding of the choice made by firms when bundling tasks into occupations needed for their production. Moreover, the ramifications of our results on dynamic wage determination also remain open. There is evidence that experience in doing jobs that are intensive in non-routine tasks are highly rewarded in the market (Stinebrickner et al., 2019). Our results, then, suggest an additional mechanism driving overall wage inequality in the economy. We urge future research to probe along these lines.

Outline of the paper. The remainder of this paper proceeds as follows. In Section 2, we describe the main data sets and detail the measures of task content used in the analysis. In Section 3, we document novel facts on the heterogeneity of occupational task contents across firms of differing size. In Section 4, we measure the large-firm wage premium and study a number of explanations for its existence, including the firm-size gradient in task intensity. Finally, we conclude in Section 5 with a summary of the findings and discussion of future directions of work. An appendix contains additional results.

2 Data and measurement

2.1 Data sources

We take advantage of the availability of cross-country harmonized surveys reporting the tasks performed by individuals in their work to construct a rich dataset covering working (not self-employed) individuals aged 16–65 across 47 countries. We combine two main data sets.

OECD Survey of Adult Skills. The Survey of Adult Skills is conducted under the OECD’s Programme for International Assessment of Adult Competencies (PIAAC). This survey aims to measure cognitive skills (literacy, numeracy, and problem solving in technology-rich environments), as well as skills used both at work and in other contexts. It is representative of the country’s adult population aged 16–65, with around 5,000 individuals participating in each country.¹ There have been three rounds of data collection (2008–2013, 2012–2016, and 2016–2019). We focus on the surveys collected from the following 31 countries: Belgium, Chile, Cyprus, Czech Republic, Denmark, Ecuador, France, Greece, Hungary, Ireland, Israel, Italy, Japan, Kazakhstan, South Korea, Lithuania, Mexico, Netherlands, New Zealand, Norway, Peru, Poland, Russian Federation, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Turkey, United Kingdom, and United States.

Full earnings information are not available in the public-use files of New Zealand, Singapore, and USA. Instead, earnings are only reported in deciles. These countries are still employed in the analysis of tasks.

World Bank Skills Measurement Surveys. Our second main data source is the World Bank’s STEP Skills Measurement Program surveys. They aim at measuring the demand and supply of skills in urban areas of low- and middle-income countries. It is a counterpart to the OECD’s Survey of Adult Skills under PIAAC. There are two types of surveys in the program: household-based and employer-based.

The household-based survey interviews a randomly-selected household member (aged 15 to 64) about their personal education and training history, work status and history, skills used in their jobs, earnings, individual competencies, and non-cognitive abilities (e.g. personality, behavior, risk preferences). Sample sizes varied from 3,000 to 4,000 individuals. We focus on the surveys which contain consistent questions regarding tasks and skills, corresponding to the following 11 countries: Armenia, Bolivia, China (Yunnan Province), Colombia, Georgia, Kenya, Laos, Macedonia, Sri Lanka, Ukraine and Vietnam.²

¹The United States conducted a second round of data collection to get more reliable estimates for certain subgroups.

²Though Ghana has a household-based survey that contains the relevant variables, we exclude it because

Additionally, in some countries, firms were also surveyed using the employer-based survey. In this survey, an informed respondent from around 300 to 500 firms per country reported the worker composition of the firm, skills required of workers, and provision of training. We use the employer-based survey of the following 9 countries: Albania, Armenia, Azerbaijan, Bosnia-Herzegovina, Georgia, Kenya, Kosovo, Serbia, and Vietnam. Only 4 of these countries also have conducted the household-based survey.

Strengths and limitations of data used. The main virtue of these datasets is the availability of information about the tasks performed by individuals in their own work that are comparable across a wide range of countries. The main limitation is that they are cross-sectional. In the absence of a panel, we are limited in the mechanisms that we explore; for instance, we cannot control for additional individual heterogeneity outside the characteristics we observe.

2.2 Measuring firm size, occupational task content, and wages

Firm size and the presence of large firm gaps. The surveys provide a measure of firm size based on the number of employees, reported in bins. The survey questions refer more precisely to workplace or establishment but we follow the past literature and use the term “firm” interchangeably. We define firms that have at least 50 employees as large.³ We mainly report large firm gaps that compare workers in firms that have at least 50 employees to workers in firms with less than 50 employees.

Task content of occupations. We follow a comparable approach to [Autor et al. \(2003\)](#) and [Spitz-Oener \(2006\)](#) and distinguish three task components of occupations: non-routine analytical, non-routine interpersonal, and routine-manual. Non-routine analytical tasks are those that involve analyzing data, thinking creatively, or interpreting information. Non-routine interpersonal tasks involve establishing and maintaining relationships, coordinating and directing subordinated, and interacting with others (e.g., through advising, negotiating, teaching). Lastly, occupations that are intensive in routine-manual tasks involve doing repetitive tasks, controlling machines or equipment, and using physical power or manual dexterity.

For the World Bank STEP data, we follow [Dicarlo et al. \(2016\)](#) on how to match the survey questions to the three task groups.⁴ We map the questions in the PIAAC survey to

of the small sample size that remains after sample selection.

³The OECD classifies micro enterprises as those with fewer than 10 employees, small enterprises as those with between 10 to 49 employees, medium-sized enterprises as those with between 50 to 249 employees, and large enterprises as those with 250 or more workers. See <https://data.oecd.org/entrepreneur/enterprises-by-business-size.htm>.

⁴[De La Rica et al. \(2020\)](#) and [Dicarlo et al. \(2016\)](#) study the occupation-level correlation of similar measures

closely mirror in spirit the mapping in the STEP survey. Appendix Table A1 summarizes this mapping. We create an individual index measuring the intensity of a particular task category in two steps. First, we standardize the responses to each task variable to have a within-country mean of zero and standard deviation of unity. Second, to obtain the index for a skill category, we add the standardized responses to the task variables and re-standardize the result to again have a within-country mean of zero and standard deviation of one. By construction, these measures are interpreted as intensities in units of standard deviations relative to the country mean.

Usage of ICT. A particular focus of our paper is on documenting the intensity by which workers use technologies such as computers and specific software as part of their work. The usage of ICT is not in itself mutually exclusive to the task dimensions mentioned above. Rather, we interpret the use of ICT as a *means* through which the above task dimensions are performed. In Table A1, we show the questions in the survey that are relevant in measuring the use of ICT. We use the responses to create an index in a similar manner as for the above measures of task content.

Wages. To quantify the large firm wage gap, we focus on hourly wages in non-self-employment work. We deflate the values to 2018 local currency and use 2018 exchange rates to US dollar to convert earnings to real 2018 USD. In Section 4, we explore sources of the LFWP and show that the gap remains even after accounting for worker sorting, a leading explanation put forward in the literature.

Demographics and additional individual controls. To increase the comparability of demographic characteristics across surveys, we consider the following standard controls: gender, age block (10-year groups starting from age 16 and ending at 65), and three education categories based on ISCED 2008 — (i) Primary or less (ISCED 1); (ii) up to a professional tertiary education degree (ISCED 5), and (iii) bachelor’s degree and above (ISCED 5A and beyond). We also aim to better account for individual heterogeneity by controlling for cognitive ability. For STEP countries we standardize, at the country level, the proportion of correct responses over the total number of questions in three different linguistic tests (vocabulary, sentence, and passage). For PIAAC countries we apply item response theory both for their numeracy and their literacy competences (Khorramdel et al., 2020). Additionally, other non-cognitive and preference-related dimensions are available in STEP. We use pre-constructed measures for the following traits: extroversion, conscientiousness, openness,

in the PIAAC and STEP data, respectively, to measures computed in O*NET. They find a positive, but not perfect, correlation — this is consistent with heterogeneity in the task content of occupations across countries they document. In Appendix C.1, we consider an alternative construction of task intensity indices using multiple correspondence analysis.

stability, agreeableness, grit and risk aversion. Finally, we consider the following sectoral classification for STEP countries: (i) agriculture, fishing, mining; (ii) manufacturing and construction; (iii) commerce, or (iv) other services in STEP countries. For PIAAC we use more detailed information encompassing twenty-one different industries.

2.3 Summary Statistics

Appendix Table A2 reports summary statistics for the 36 countries for which we have a continuous measure of wages. To ensure that our results are not driven by extrapolation, we also impose that each firm size and 2-digit occupation code cell has at least 5 observations. The number of observations after focusing on working-age individuals that are not self-employed varies from 858 (Greece) to 3,953 (United Kingdom) among PIAAC countries and from 347 (Laos) to 1,260 (Vietnam) among STEP countries. In general, small firms are more dominant but there is significant cross-country variation. For instance, in Belgium or Netherlands 50% of the firms are large, while it is 20% in Ecuador or Greece.

3 Firm size gradient in the task content of jobs

In this section, we document our novel finding that even within narrowly defined occupation groups, workers in large firms perform more intensively non-routine analytical and interpersonal tasks, and they use more information and communications technologies.

3.1 Task Content of Jobs in Large Firms

We first document the differences in the task content of jobs between large and small firms by estimating the following regression:

$$T_i = \beta \times \text{LF}_{j(i)} + \delta_{o(i)}^o + \delta_{c(i)}^c + \varepsilon_i, \quad (1)$$

where T_i is the measure of task content of the job of worker i , $\text{LF}_{j(i)}$ is an indicator of whether the firm $j(i)$ of individual i has at least 50 employees, δ^o are occupation-code fixed effects, and δ^c are country fixed effects. We focus on our four main task categories as outcomes: non-routine analytical tasks, non-routine interpersonal tasks, routine manual tasks, and the usage of ICT.

In Table 1, we report estimates of β in Equation 1 using the pooled sample of both PIAAC and STEP countries with varying specificity of occupation codes. In Column (1), we have the raw estimates, not controlling for occupation. Qualitatively, we find that workers in larger firms report to perform more non-routine analytical tasks, more non-routine interpersonal tasks, less routine-manual tasks, and more tasks involving information and

Table 1: Pooled estimates of firm size gradient in the task content of jobs

| | Occupation controls | | | |
|---------------------------|---------------------|--------------------|---------------------|--------------------|
| | (1) Raw | (2) 1-digit | (3) 2-digit | (4) 3-digit |
| <i>Outcome variable:</i> | | | | |
| Non-routine analytical | 0.3036 (0.0232) | 0.1471 (0.0148) | 0.1272 (0.0127) | 0.1166 (0.0147) |
| Non-routine interpersonal | 0.2501 (0.0206) | 0.1568 (0.0140) | 0.1575 (0.0127) | 0.1600 (0.0131) |
| Routine-manual | -0.1055 (0.0216) | 0.0042 (0.0147) | -0.0001 (0.0135) | 0.0023 (0.0149) |
| Usage of ICT | 0.3031 (0.0245) | 0.1368 (0.0149) | 0.1124 (0.0129) | 0.0998 (0.0150) |
| Sample Size | 88,901 | 87,987 | 82,901 | 57,055 |

Notes: Regression of task requirement intensity on an indicator of large firm (at least 50 employees) and country dummies in STEP-PIAAC pooled sample. Differences in sample size explained by (1) restriction of at least 5 observations per firm-size cross occupation cell, or (2) availability of ISCO-08 occupation codes. Reported standard errors are clustered at the country level.

communications technology. One may, however, be concerned that the occupational structure of large firms and small firms differs in a way that occupations that use more analytical tasks, interpersonal tasks and ICT, and less manual tasks are disproportionately hired by larger firms. In Columns (2)–(4) of Table 1, we estimate β in Equation 1 controlling for occupation fixed effects at the 1-digit, 2-digit and 3-digit level based on the ISCO-08 classification.⁵ The results consistently show that workers in larger firms perform more non-routine analytical and interpersonal tasks, and use more ICT. Once we account for the possible sorting of occupations in firms of various sizes, there is no evidence towards differences in the routine-manual task performed between workers in larger and smaller firms.

Our preferred specification controls for 2-digit ISCO-08 occupation codes, which balances the trade-off between the specificity of the occupations and sample size. In terms of economic magnitude, we find that the average worker in a larger firm performs 0.1272 of a standard deviation more non-routine analytic tasks than a worker in a smaller firm. The standard deviation is computed using the country-specific distribution of performed tasks. Workers in larger firms also perform 0.1575 of a standard deviation more non-routine interpersonal tasks and use 0.1124 of a standard deviation more computer technology compared

⁵As mentioned, to reduce the dependence of our results on extrapolation, we ensure there are at least 5 observations in each firm size-occupation cell.

to workers in smaller firms.

The qualitative results hold even when we take a more flexible approach to account for occupations. In Appendix Table B1, we estimate regressions such as Equation 1 conditioning on 1-digit occupation codes, but controlling for dummies of 2-digit occupation codes. We find that workers in larger firms across all 1-digit occupation codes perform more non-routine analytical and interpersonal tasks. Workers in larger firms also use more ICT in their work, except for skilled agriculture, fishery and forestry workers (though this result is based on a modestly-sized sample). There are, however, differences in the magnitude of the firm size gradient in the task content of jobs depending on which 1-digit occupation code we focus on. As expected, the largest gap in the intensity of non-routine analytical tasks and ICT use by firm size is seen among services and sales workers. For non-routine interpersonal tasks, the largest firm size gradient is seen among managers.

We further find that these differences in tasks performed are present at the beginning of the job tenure and early in the workers' career. In Appendix Table B2, we estimate Equation 1 conditioning first on workers having been in their current job for a short time (up to 2 years) and then additionally on being young (less than 25 years) to discard cases of workers who have adopted more task-intensive work as they progressed in their careers. Among these young workers with short tenure, we find firm size gradients in task content of work that are of a similar magnitude to the full sample. This suggests that these differences do not arise from firms requiring workers to perform increasingly more non-routine tasks and ICT over their tenure in the firm. Rather, these differences are already present at the start of performing the job. This is consistent with results, discussed below, based on demand side information where we show that firms already expect new hires to perform more of non-routine tasks and use more ICT.

In Appendix C.2, we extend our analysis by also exploring the differences in the distribution of task intensity by firm size. We employ distributional regressions to show that the large firm differences in non-routine analytical and interpersonal tasks, and use of ICT are present at multiple thresholds throughout the support of the distribution. This is key in demonstrating that the mean differences we find are not driven by discrepancies in the upper tail of the distribution of tasks performed. Instead, the distributions of the intensity of performed tasks in larger firms are all shifted to the right compared to the distribution in smaller firms.

Cross-country comparisons. There are reasons why we might suspect that these differences in the task content of jobs by firm size would differ across countries: (1) differences in labor market institutions, and (2) differences in the relevance of firms with at least 50 employees. In Figure B1, we summarize the results from estimating a version of Equation 1 separately for each country, controlling for 2-digit occupation codes. We observe that

for most of the countries in our sample, the qualitative results from the pooled regression hold true: workers in larger firms perform more non-routine analytical tasks, non-routine interpersonal tasks, and use more ICT. As the pooled estimates suggest, there are still no differences in the routine-manual tasks performed by firm size, except in a few cases.

We explore the relationship between the firm size gradient in the task content of jobs and country-specific characteristics by studying partial correlations summarized in Appendix Figure B2. We highlight three empirical patterns. First, we find that the firm size differences in the use of analytical tasks, interpersonal tasks, and ICT are larger in countries with a larger services or industry sector. In contrast, the firm size gradient in the use of manual tasks is uncorrelated with the size of the services or industry sector. Second, we find that the firm size gradients across the four task dimensions are negatively correlated with log GDP per capita. That is, in richer economies, the firm size differences in the task content of jobs are smaller. Lastly, we find that in countries with a larger college-educated population, workers in larger firms disproportionately perform more analytical, interpersonal, and routine-manual tasks. The firm size gradient in the use of ICT does not seem to be correlated with the proportion of college-educated individuals.

Evidence from the demand side. We perform a demand-side analysis using employer surveys in the World Bank STEP Skill Surveys program. In these surveys, firms answer detailed questions on the skill requirements of occupations within the firm. To limit the burden on the survey respondent, the focus is placed on two occupations (randomly selected from nine categories). We discuss in detail the measurement of the task content of occupations using the employer-based survey in Appendix A.

Appendix Table B3 reports estimates of average differences in task requirement between large and small firms, within occupation categories, for the pooled sample of nine countries where the STEP Employer Survey is available. We find that large firms require more non-routine analytical and interpersonal tasks, and use more ICT. We then explore whether these differences in task requirements are reflected in the hiring decision and environment of firms. In particular, In Appendix Table B3, we explore whether larger firms have more difficulty in hiring along various dimensions (e.g., few applicants, applicants' salary expectations being too high, etc.). Overall, we find this to be the case. Most notably, they face a scarcity in applicants who possess the adequate skills to perform the jobs (i.e. an undersupply of qualified applicants).

4 Large firm wage premium and the role of individual selection, sectors, and tasks

In this section, we document the large firm wage premium and explore possible mechanisms to explain its existence, including our novel firm size gradient of task intensity. In Subsection 4.1, we measure the raw LFWP using the pooled PIAAC and STEP sample. In Subsection 4.2, we explore how much of this raw gap can be linearly explained by various mechanisms, including selection of individuals into occupations and differences in the task composition of occupations.

4.1 Raw large firm wage premium: Cross-country evidence

Table 2: Pooled estimates of the large firm wage premium

| | Mean Regressions | | Quantile Regressions | | | | |
|-------------|--------------------|--------------------|----------------------|--------------------|--------------------|--------------------|--------------------|
| | (1) Raw | (2) 2d occ. | (3) p10 | (4) p25 | (5) p50 | (6) p75 | (7) p90 |
| LFWP | 0.2427 (0.0130) | 0.1556 (0.0103) | 0.1532 (0.0087) | 0.1378 (0.0099) | 0.1387 (0.0102) | 0.1413 (0.0114) | 0.1258 (0.0103) |
| 2-d ISCO-o8 | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Country FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sample Size | 72,282 | 66,970 | 66,970 | 66,970 | 66,970 | 66,970 | 66,970 |

Notes: Regression of log hourly wages (in 2018 USD) on an indicator of large firm (at least 50 employees) and country dummies in STEP-PIAAC pooled sample. Differences in sample size explained by restriction of at least 5 observations per firm-size cross occupation cell. Countries for which continuous wage data are not available are excluded. Reported standard errors are clustered at the country level.

Similar to how we documented the large firm gap in the task content of jobs in Section 3.1, in this subsection we explore the large firm gap in wages. We estimate the following regression:

$$\ln w_i = \beta \times \text{LF}_{j(i)} + \delta_{o(i)}^o + \delta_{c(i)}^c + \varepsilon_i, \quad (2)$$

where $\ln w_i$ is log real hourly wages in USD of individual i , $\text{LF}_{j(i)}$ is an indicator of whether the firm $j(i)$ of individual i has at least 50 employees, δ^o are occupation-code fixed effects, and δ^c are country fixed effects. We interpret β as a measure of the LFWP, which is how much more workers in firms with at least 50 employees are paid (in log-points) relative to smaller firms within the same occupation and country. In the first two columns of Table 2, we report the coefficient β not controlling for occupation and controlling for 2-digit ISCO-o8 occupation codes, respectively. The estimated LFWP without accounting for occupations

is 0.2427, which means that an average worker in a larger firm earns about 27% more than an average worker in a smaller firm.⁶ One might be concerned that larger firms may utilize relatively more occupations for which the market wages are higher, suggesting that the large firm wage premium could be explained by differences in the occupational structure of firms. However, as we see in Column (2) of Table 2, even after controlling for 2-digit occupation codes, the log-point difference in wages paid to an average worker in the larger firm relative to a smaller firm is 0.1556, corresponding to around a 17% difference in wage levels. These estimates are not readily comparable to estimates in the literature which are usually reported as elasticities. Still, the magnitudes we estimate are reasonable given estimates in the literature such as in Reed and Thu (2019).

In Appendix Table B4, we account for occupations more flexibly by estimating Equation 2 conditioning on workers being in specific 1-digit occupation codes, and controlling for 2-digit occupation codes. We find that workers in larger firms are paid more, with the possible exception of skilled agricultural workers for which we do not have enough statistical power to make firm conclusions. We find the smallest LFWP among workers in elementary occupations, where the average worker in larger firms earns 0.0544 log points more than their counterparts in smaller firms. The LFWP is largest among managers where workers in larger firms earn 0.2676 log points — 31% more in levels.

This LFWP is not only present when comparing average wages between larger and smaller firms. Instead, we document that the entire wage distribution of large firms is shifted to the right compared to smaller firms, even within occupations. Columns (3)–(7) in Table 2 report the β coefficients in a quantile regression version of Equation 2, for various quantiles, controlling for 2-digit occupation codes. These results compare the wage distribution of larger and smaller firms. Looking within occupations, the worker at the 10th percentile in larger firms is paid 0.1532 log-points more than the worker at the 10th percentile in smaller firms. Comparing the respective median workers, the difference is about 0.1387 log-points and it is 0.1258 log-points at the 90th percentile.

Cross-country comparisons. In Figure B3, we summarize estimates of the LFWP by country. Most of the LFWP estimates without controlling for occupation lie between 0.2 and 0.4 log points, corresponding to approximately 20–50% average real hourly wage differences between workers in larger firms relative to those in smaller firms. After adjusting these estimates for occupational composition using 2-digit ISCO-08 occupation codes, the LFWP estimates lie roughly between 0.10–0.25 log points corresponding to about 10–28% wage differentials. In Appendix Figure B4, we report the differences in the distribution of wages across firm size by country. We find that in almost all the countries we study, the wage distribution of workers in larger firms is shifted to the right relative to the wage distribution

⁶Note that $\exp(0.2427) - 1 \approx 0.2747$.

of workers in smaller firms. These results echo the qualitative results we obtained using the pooled data.

For privacy reasons, not all countries in the PIAAC report information on hourly wages in the public use files. In such cases, only the decile of the wage distribution the person is located in is provided. We utilize this information and estimate linear probability models where the outcome is an indicator taking the value of 1 if the worker is *at least* in a certain wage decile. Appendix Figure B5 summarizes the coefficient of this linear probability model controlling for 2-digit occupation codes. The results complement what we learn from the quantile regressions: not only are workers in larger firms more likely to have wages in the last decile, these workers are also more likely to have wages that are at least above the second and fifth decile. This is further evidence towards the wage distribution of large firms being shifted to the right compared to smaller firms, even within narrowly defined occupation groups.

4.2 Sources of the large firm wage premium

There are a number of plausible reasons for the existence of the large firm wage premium. In this subsection, we explore the roles of (1) sorting by individual characteristics, (2) industry characteristics, and (3) differential task content of jobs in wage determination.

We conduct a simple mediation analysis adopting the conditional decomposition developed in Gelbach (2016). We begin with a raw estimate of the LFWP, β^{raw} , from the regression:

$$\ln w_i = \beta^{\text{raw}} \times LF_{j(i)} + \delta_{o(i)}^{o,\text{raw}} + \delta_{c(i)}^{c,\text{raw}} + \varepsilon_i^{\text{raw}}, \quad (3)$$

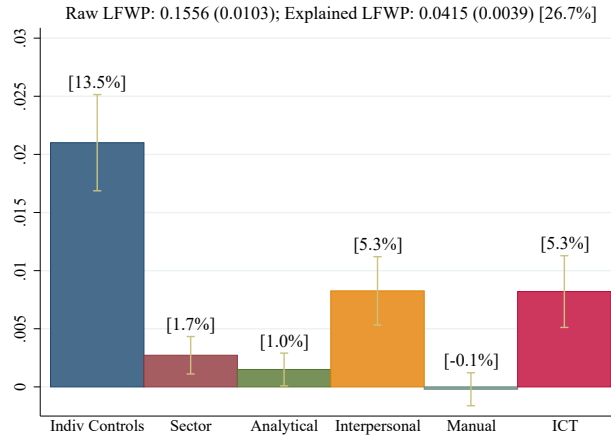
where $\ln w_i$ is log real hourly wages, $LF_{j(i)}$ is the indicator for worker i being in a large firm, δ^o are occupation fixed effects (2-digit ISCO code in the baseline), and δ_c are country fixed effects. This raw LFWP estimate coincides with the estimate in Column (2) of Table 2. Then, we include a set of individual controls X_i which we believe mediate the LFWP:

$$\ln w_i = \beta^{\text{full}} \times LF_{j(i)} + X_i' \gamma + \delta_{o(i)}^{o,\text{full}} + \delta_{c(i)}^{c,\text{full}} + \varepsilon_i^{\text{full}}. \quad (4)$$

In our case, the variables incorporated in X are (1) individual characteristics including sex, age and education, (2) sector dummies, and (3) the task content of jobs reported by workers. The difference $\beta^{\text{raw}} - \beta^{\text{full}}$ is interpreted as the part of the LFWP that we are able to explain by controlling for X . Gelbach (2016) proposes a methodology to apportion the explained part of the LFWP to the component variables of X .

The results are graphically summarized in Figure 1. The controls we consider are able to explain about 26% of the raw LFWP. The basic controls of age, sex, and education can explain a significant portion of the LFWP, around 13.5%. This is suggestive that large firms

Figure 1: Gelbach decomposition of LFWP, pooled



Notes: Pooled sample. Numbers in brackets indicate percentages of the raw LFWP. Reported confidence intervals at 95% confidence level. Standard errors are clustered at the country level.

pay more on average because they hire workers who are older and more educated. This sorting pattern of workers to larger firms has been documented by [Arellano-Bover \(2021\)](#). Though human capital sorting (by occupation or education) explains a large portion, it is unable to fully explain the existence of the LFWP.

The last four bars in Figure 1 are the fractions of the raw LFWP that are explained by the differences in the performed tasks of the workers. As we have comprehensively documented in Section 3, workers in larger firms perform more non-routine analytical and interpersonal tasks, and use more ICT. We examine how much these differences are able to explain the LFWP in various countries. The firm size gradient in the performance of non-routine interpersonal tasks explains about 5.3% of the raw large firm wage premium. Moreover, differences in the use of ICT between workers in larger and smaller firms explains an additional 5.3% of the raw LFWP. These results reflect not only the disproportionately higher intensity that workers perform interpersonal tasks and use ICT in larger firms, but also the growing importance of interpersonal ([Deming, 2017](#)) and computer skills ([Alekseeva et al., 2021](#)) in the labor market. Combined, these two task components explain a total of 10.6% of the raw LFWP, a magnitude comparable to that explained jointly by education, age, and sex.

Surprisingly, the differences in the performance of analytical tasks between workers in larger and smaller firms appear not to be reflected in differences in wages. This component explains only around 1% of the raw LFWP and is not statistically distinguishable from zero. This result and the fact that analytical task intensity and use of ICT are highly correlated

(as we show in Appendix C.3) would suggest that the role of analytical tasks in wage determination is mediated by how much of these analytical tasks are performed with the use of ICT.

Unsurprisingly, differences in the performance of routine-manual tasks do not translate to wage differences by firm size as, on average, workers in larger firms do not disproportionately perform more or less routine-manual tasks.

A main concern in the analysis of the LFWP is that it may be driven by differences in firm productivity — in many models of the labor market, including rent-sharing models or search and matching models, more productive firms pay higher wages to its workers. Unfortunately, we do not observe measures of firm productivity of the firms that the workers work in. To partially address this issue, we control for the sector in which the worker works in with the aim of accounting for aggregate productivity differences across sectors. We find that the sectoral membership of the worker only partially explains the existence of the LFWP, about 1.7% of the raw LFWP.

Cross-country comparisons. The results of the decomposition exercise by country are graphically summarized in Figure B6, focusing on countries for which both the LFWP and the explained portion of the LFWP are statistically significant. We find that the proportion of the raw LFWP explained by the controls we consider varies between 20% and 60%. It is consistent that basic individual controls such as age, sex, and education explain a significant portion of the raw LFWP (between 10–25%). Sectoral membership is only statistically and economically significant in a handful of countries. In countries where this component explains a statistically significant portion, sectors explains around 10–20% of the raw LFWP.

The firm size gradient in the performance of non-routine interpersonal tasks and the use of ICT explains, in general, a total of about 5–20% of the raw LFWP. The firm size gradient in the performance of non-routine interpersonal tasks explains, in many cases, between 3–10% of the raw large firm wage premium (up to 17.5% in the case of Kazajstan). Among the countries for which the ICT component is statistically significant, the estimates fall mostly between 5–9% of the raw LFWP. Consistent with the pooled regression, differences in the performance of non-routine analytical and routine-manual tasks by firm size do not translate to differences in wages.

Extensions. The main concern in the performed decomposition analysis is omitted variables. As mentioned above, firm performance is likely the most relevant one. However, in addition, it may be that there are unobserved individual characteristics that lead workers to sort into large and small firms, and that are also important in wage determination. In Appendix Figures B7 and B8, we extend our decomposition analysis by including more

individual-level controls. This approach leads to our results being estimated more imprecisely, especially for STEP countries, which prevents us from drawing firm conclusions.

In Appendix Figure B7, we included variables that capture cognitive measures of the workers. For the PIAAC countries, we include measures of numeracy and literacy while we only include a measure of literacy in STEP countries (as numeracy is not measured). It is still consistent that individual controls (sex, education and age) play economically significant roles in explaining the LFWP, capturing around 12% of the raw LFWP. The added cognitive measures explain little of the raw LFWP — about 3% of the raw LFWP in PIAAC countries and close to zero in STEP countries. Among the measures of task content, the use of ICT is consistently an important determinant of the LFWP, explaining around 5% in PIAAC countries and 7% in STEP countries. Differences in the performance of non-routine interpersonal tasks explain about 5% of the LFWP in PIAAC countries.

For the STEP countries, we have access to an even richer set of individual variables. In Appendix Figure B8, we include standard measures of non-cognitive skills (e.g., extroversion, stability, agreeableness, grit) on top of the measure of literacy. Individual controls still play an economically significant role. The intensity in the use of ICT explains around 5% of the raw LFWP.

5 Conclusion

In this paper we have documented novel facts about the heterogeneity in occupational task intensity across firms. We find that individuals working in larger firms perform more non-routine analytical, non-routine interpersonal, and ICT tasks, even within closely defined occupations. We complement this fact by showing that larger firms also require workers to perform more analytical and ICT intensive tasks.

Moreover, we document the existence of an economically significant large firm wage premium of about 17%. We provide suggestive evidence of the role of task heterogeneity in explaining this LFWP. By controlling for individual characteristics (age and education), sector, and the task content of jobs, we are able to explain about 26% of the raw LFWP. Differences in task content of jobs are able to account for 10% of the raw LFWP.

We consider our work to open two natural avenues for future research. First, an open question is the source of this firm size gradient in occupational task content. For instance, we are unable to determine whether firms that require more analytical, interpersonal and ICT tasks are the firms that grow larger, or if firms adopt technologies that require more analytical or ICT tasks from their worker as they grow. Answering this question requires panel data of firms and is, therefore, not suitable to be answered with the cross-sectional surveys used in this paper.

Second, we leave for further study other implications of the firm size gradient in occu-

pational content on labor markets. We have suggestive evidence of its role in static wage determination but lack exogenous identifying conditions to argue their causal nature. The implications of our results on dynamic wage determination remain unexplored. More specifically, our results may serve as a nexus to two seemingly parallel strands of the literature. First, a number of studies shows that having experience in certain tasks has different returns in the market: interpersonal and use of ICT have been found to have high market returns, especially in recent years (Stinebrickner et al., 2019; Deming, 2017; Alekseeva et al., 2021). Second, there is evidence that experience in large firms also has higher return in the market (Arellano-Bover, 2020). Our results would suggest a plausible mechanism for the larger dynamic returns to working in larger firms — workers in larger firms gain more experience in non-routine and ICT-intensive tasks, which are highly valued in the labor market.

References

- Acemoglu, Daron and David Autor (2011) "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," 4 of Handbook of Labor Economics, 1043–1171: Elsevier, [https://doi.org/10.1016/S0169-7218\(11\)02410-5](https://doi.org/10.1016/S0169-7218(11)02410-5). (cited in page 1)
- Acemoglu, Daron and Pascual Restrepo (2021) "Tasks, automation, and the rise in US wage inequality," Working Paper. (cited in page 1)
- Alekseeva, Liudmila, José Azar, Mireia Giné, Sampsa Samila, and Bledi Taska (2021) "The demand for AI skills in the labor market," *Labour Economics*, 71, 102002, <https://doi.org/10.1016/j.labeco.2021.102002>. (cited in page 13, 16)
- Arellano-Bover, Jaime (2020) "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size," Working Paper. (cited in page 16)
- (2021) "Who Gets Their First Job at a Large Firm? The Distinct Roles of Education and Skills," *AEA Papers and Proceedings*, 111, 465–69, [10.1257/pandp.20211009](https://doi.org/10.1257/pandp.20211009). (cited in page 13)
- Atalay, Engin, Sebastian Sotelo, and Daniel Tannenbaum (2021) "The geography of job tasks," https://enginataalay.github.io/city_task.pdf, Working Paper. (cited in page 1)
- Autor, David H. and Michael J. Handel (2013) "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages," *Journal of Labor Economics*, 31 (S1), S59–S96, [10.1086/669332](https://doi.org/10.1086/669332). (cited in page 1)
- Autor, David H., Frank Levy, and Richard J. Murnane (2003) "The Skill Content of Recent Technological Change: An Empirical Exploration*," *The Quarterly Journal of Economics*, 118 (4), 1279–1333, [10.1162/003355303322552801](https://doi.org/10.1162/003355303322552801). (cited in page 1, 4)
- Bloom, Nicholas, Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter (2018) "The Disappearing Large-Firm Wage Premium," *AEA Papers and Proceedings*, 108, 317–22, [10.1257/pandp.20181066](https://doi.org/10.1257/pandp.20181066). (cited in page 2)
- Brown, Charles and James Medoff (1989) "The Employer Size-Wage Effect," *Journal of Political Economy*, 97 (5), 1027–1059, [10.1086/261642](https://doi.org/10.1086/261642). (cited in page 2)
- Caunedo, Julieta, Elisa Keller, and Yongseok Shin (2021) "Technology and the task content of jobs across the development spectrum," Working Paper. (cited in page 1)
- Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly (2013) "Inference on Counterfactual Distributions," *Econometrica*, 81 (6), 2205–2268, <https://doi.org/10.3982/ECTA10582>. (cited in page A17)
- Colonnelli, Emanuele, Joacim Tåg, Michael Webb, and Stefanie Wolter (2018) "A Cross-Country Comparison of Dynamics in the Large Firm Wage Premium," *AEA Papers and Proceedings*, 108, 323–27, [10.1257/pandp.20181067](https://doi.org/10.1257/pandp.20181067). (cited in page 2)

- De La Rica, Sara, Lucas Gortazar, and Piotr Lewandowski (2020) "Job Tasks and Wages in Developed Countries: Evidence from PIAAC," *Labour Economics*, 65, 101845, <https://doi.org/10.1016/j.labeco.2020.101845>. (cited in page 1, 4)
- Deming, David J. (2017) "The Growing Importance of Social Skills in the Labor Market*," *The Quarterly Journal of Economics*, 132 (4), 1593–1640, [10.1093/qje/qjx022](https://doi.org/10.1093/qje/qjx022). (cited in page 13, 16)
- Deming, David and Lisa B. Kahn (2018) "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals," *Journal of Labor Economics*, 36 (S1), S337–S369, [10.1086/694106](https://doi.org/10.1086/694106). (cited in page 1)
- Dicarlo, Emanuele, Salvatore Lo Bello, Sebastian Monroy-Taborda, Ana Maria Oviedo, Maria Laura Sanchez Puerta, and Indhira Vanessa Santos (2016) "The skill content of occupations across low and middle income countries: evidence from harmonized data," Working Paper. (cited in page 1, 4)
- Dobbelaere, Sabien (2004) "Ownership, firm size and rent sharing in Bulgaria," *Labour Economics*, 11 (2), 165–189, <https://doi.org/10.1016/j.labeco.2003.07.001>. (cited in page 2)
- Gelbach, Jonah B. (2016) "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics*, 34 (2), 509–543, [10.1086/683668](https://doi.org/10.1086/683668). (cited in page 12)
- Gerlach, Knut and Olaf Hübler (1998) "Firm Size and Wages in Germany – Trends and Impacts of Mobility," *Empirica*, 25 (3), 245–261, [10.1023/A:1006904430572](https://doi.org/10.1023/A:1006904430572). (cited in page 2)
- Goos, Maarten, Alan Manning, and Anna Salomons (2014) "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 104 (8), 2509–26, [10.1257/aer.104.8.2509](https://doi.org/10.1257/aer.104.8.2509). (cited in page 1)
- Khorramdel, Lale, Matthias von Davier, Eugenio Gonzalez, and Kentaro Yamamoto (2020) *Plausible Values: Principles of Item Response Theory and Multiple Imputations*, 27–47, Cham: Springer International Publishing, [10.1007/978-3-030-47515-4_3](https://doi.org/10.1007/978-3-030-47515-4_3). (cited in page 5)
- Lehmer, Florian and Joachim Möller (2010) "Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort analysis for Germany," *The Annals of Regional Science*, 45 (1), 31–53, [10.1007/s00168-009-0290-y](https://doi.org/10.1007/s00168-009-0290-y). (cited in page 2)
- Lewandowski, Piotr, Albert Park, Wojciech Hardy, and Yang Du (2019) "Technology, Skills, and Globalization: Explaining International Differences in Routine and Nonroutine Work Using Survey Data," Working Paper. (cited in page 1)
- Lochner, Benjamin, Stefan Seth, and Stefanie Wolter (2020) "Decomposing the large firm wage premium in Germany," *Economics Letters*, 194, 109368, <https://doi.org/10.1016/j.econlet.2020.109368>. (cited in page 2)

- Oi, Walter Y. and Todd L. Idson (1999) "Chapter 33 Firm size and wages," 3B of Handbook of Labor Economics, 2165–2214: Elsevier, [https://doi.org/10.1016/S1573-4463\(99\)30019-5](https://doi.org/10.1016/S1573-4463(99)30019-5). (cited in page 2)
- Porcher, Charly, Hannah Rubinton, and Clara Santamaria (2021) "The role of establishment size in the city-size earnings premium," https://drive.google.com/file/d/1y-ZIUJj9_NRGfvyN-wnaQ2TUU9RzdrMZ/view, Working Paper. (cited in page 2)
- Reed, Tristan and Trang Thu (2019) "The Large-Firm Wage Premium in Developing Economies," <https://openknowledge.worldbank.org/handle/10986/32371>, Working Paper. (cited in page 2, 11)
- Schaffner, Julie Anderson (1998) "Premiums to employment in larger establishments: evidence from Peru," *Journal of Development Economics*, 55 (1), 81–113, [https://doi.org/10.1016/S0304-3878\(98\)00038-8](https://doi.org/10.1016/S0304-3878(98)00038-8). (cited in page 2)
- Söderbom, Måns, Francis Teal, and Anthony Wambugu (2005) "Unobserved heterogeneity and the relation between earnings and firm size: evidence from two developing countries," *Economics Letters*, 87 (2), 153–159, <https://doi.org/10.1016/j.econlet.2004.09.012>. (cited in page 2)
- Spitz-Oener, Alexandra (2006) "Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure," *Journal of Labor Economics*, 24 (2), 235–270, [10.1086/499972](https://doi.org/10.1086/499972). (cited in page 4)
- Stinebrickner, Ralph, Todd Stinebrickner, and Paul Sullivan (2019) "Job Tasks, Time Allocation, and Wages," *Journal of Labor Economics*, 37 (2), 399–433, [10.1086/700186](https://doi.org/10.1086/700186). (cited in page 1, 2, 16)
- Troske, Kenneth R. (1999) "Evidence on the Employer Size-Wage Premium from Worker-Establishment Matched Data," *The Review of Economics and Statistics*, 81 (1), 15–26, [10.1162/003465399557950](https://doi.org/10.1162/003465399557950). (cited in page 2)
- Velenchik, Ann D. (1997) "Government intervention, efficiency wages, and the employer size wage effect in Zimbabwe," *Journal of Development Economics*, 53 (2), 305–338, [https://doi.org/10.1016/S0304-3878\(97\)00019-9](https://doi.org/10.1016/S0304-3878(97)00019-9). (cited in page 2)
- Winter-Ebmer, Rudolf (2001) "Firm Size, Earnings, and Displacement Risk," *Economic Inquiry*, 39 (3), 474–486, <https://doi.org/10.1093/ei/39.3.474>. (cited in page 2)

A Appendix: Data measurement and summary statistics

Table A1 summarizes the mapping of the questions in PIAAC and STEP to the different dimensions of work we are interested in: non-routine analytical tasks, non-routine interpersonal tasks, routine manual tasks, and the usage of ICT. The following exceptions to the table are noted:

- "Type of document read" uses up to m5a_q05_7 for Armenia, Kenya, and Macedonia and up to m5a_q05_5 for China, Sri Lanka, and Ukraine
- For Bolivia, China, Colombia, Laos, Sri Lanka, Ukraine, and Vietnam:
 - "Thinking for over 30 minutes" is obtained from m5b_q09
 - "Learning new things" is obtained from m4b_q17
 - They do not have a measure of collaboration with coworkers
 - "Supervising coworkers" is obtained from m5b_q11
 - "Making presentations" is obtained from m5b_q10
 - "Contact clients" is obtained from m5b_q04 and m5b_q05
 - "Autonomy" is obtained from m5b_q12
 - "Repetitiveness" is obtained from m5b_q14
 - "Operate equipment" is obtained from m5b_q08
 - "Driving" is obtained from m5b_q06
 - "Repair" is obtained from m5b_q07

Construction of task content measures using employer-based surveys. In the employers questionnaire, a knowledgeable person within the firm was asked about the task needs from two randomly selected occupations. One of them would be chosen among the following ones: manager/professional/technician while the second one would be selected from: clerk/services/sales/crafting/operator/elementary occupation. We obtain a measure of firm's demand for each of the following dimensions: analytical, interpersonal and ICT. They are standardized measures obtained from the sum of the standardized score in each of the following questions (possible answers were yes/no):

- Analytical: (a) does the job involve reading?; (b) does the job involve writing; (c) does the job involve math?; (d) does the job involve problem solving?, and (e) does the job involve speaking other languages?
- Interpersonal: (a) does the job involve making presentations?, and (b) does the job involve interacting with co-workers?

- ICT: (a) what is the highest level of computer use in the job? (possible responses were: none, straightforward, moderate, complex, and specialized).

Table A1: Mapping of survey questionnaires to task categories

| Task category | STEP surveys | | PIAAC surveys | |
|---------------------------|------------------------------|-------------------------------------|------------------------------|--------------------------|
| | Item description | Item nos. | Item description | Item nos. |
| Non-routine analytical | Type of document read | m5a_q04, m5a_q05(1-6) | Type of document read | G_Q01(a-h) |
| | Length of document read | m5a_q04, m5a_q06 | Length of document written | G_Q02(a-d) |
| | Length of document written | m5a_q11, m5a_q12, m5a_q13 | Math tasks | G_Q03(b-h) |
| | Math tasks | m5a_q18(1-5) | Problem solving | F_Q05(b) |
| | Thinking >30 minutes | m5b_q10 | Learning new things | D_Q13(a-c) |
| | Learning new things | m4b_q17 | | |
| | | | | |
| Non-routine interpersonal | Supervising coworkers | m5b_q13 | Supervising coworkers | D_Q08(a-b) |
| | Making presentations | m5b_q12 | Making presentations | F_Q02(c) |
| | Contact with clients | m5b_q05, m5b_q06 | Contact with clients | F_Q02(b,d,e), F_Q04(a-b) |
| | Collaborating with coworkers | m5b_q04 | Collaborating with coworkers | F_Q01(b), F_Q02(a) |
| Routine manual | Autonomy | m5b_q14 | Autonomy | D_Q11(a-c), F_Q03(a,c) |
| | Repetitiveness | m5b_q16 | Physical demand | F_Q06(b) |
| | Operate eqpt | m5b_q09 | Hand use | F_Q06(c) |
| | Driving | m5b_q07 | | |
| | Repair | m5b_q08 | | |
| | Physical demand | m5b_q03 | | |
| Usage of ICT | Computer use | m5b_q18, m5b_q19 | Internet use | G_Q05(a,c,d) |
| | Computer type | m5b_q18, m5b_q20(1-5), m5b_q22(1-8) | Computer use | G_Q05(e-h) |

Table A2: Summary Statistics

| Variable | New Zealand | United Kingdom | Slovak Republic | Russia | Czech Republic | Mexico | Lithuania | Kazajstan | Israel |
|--|-----------------|----------------|-----------------|---------------|----------------|---------------|---------------|----------------|----------------|
| Large Firm | 0.368 (0.482) | 0.5 (0.5) | 0.434 (0.496) | 0.465 (0.499) | 0.376 (0.485) | 0.318 (0.466) | 0.456 (0.498) | 0.367 (0.482) | 0.367 (0.482) |
| Log Earnings | 2.811 (0.508) | 2.774 (0.566) | 1.579 (0.673) | 0.691 (0.775) | 1.702 (0.579) | 0.812 (0.873) | 2.594 (0.597) | 0.289 (0.593) | 2.49 (0.811) |
| Female | -0.223 - 10.675 | 0.387 - 6.784 | 0.209 - 7.617 | -3.63 - 5.49 | -2.573 - 6.793 | -2.051 - 5.01 | 0.334 - 9.092 | -3.185 - 4.362 | -0.489 - 8.849 |
| Age (16-24) | 0.575 (0.494) | 0.593 (0.491) | 0.506 (0.5) | 0.637 (0.481) | 0.514 (0.5) | 0.384 (0.486) | 0.609 (0.488) | 0.569 (0.495) | 0.522 (0.5) |
| Age (25-34) | 0.204 (0.403) | 0.127 (0.333) | 0.093 (0.29) | 0.222 (0.416) | 0.147 (0.354) | 0.217 (0.412) | 0.088 (0.284) | 0.106 (0.308) | 0.282 (0.45) |
| Age (35-44) | 0.191 (0.393) | 0.127 (0.333) | 0.255 (0.436) | 0.281 (0.45) | 0.292 (0.455) | 0.266 (0.442) | .208 (0.406) | 0.315 (0.465) | 0.33 (0.47) |
| Age (45-54) | 0.223 (0.416) | 0.254 (0.435) | 0.252 (0.434) | 0.198 (0.398) | 0.235 (0.424) | 0.257 (0.437) | 0.234 (0.424) | 0.287 (0.453) | 0.24 (0.427) |
| Age (above 55) | 0.198 (0.399) | 0.23 (0.421) | 0.256 (0.436) | 0.194 (0.396) | 0.185 (0.389) | 0.166 (0.372) | 0.273 (0.446) | 0.209 (0.407) | 0.1 (0.3) |
| Education (Primary or less) | 0.184 (0.387) | 0.152 (0.359) | 0.144 (0.351) | 0.105 (0.307) | 0.14 (0.348) | 0.094 (0.292) | 0.196 (0.397) | 0.082 (0.275) | 0.049 (0.215) |
| Education (up to professional tertiary education degree) | 0.013 (0.113) | 0.078 (0.269) | 0.002 (0.041) | 0.006 (0.076) | 0.001 (0.029) | 0.33 (0.414) | 0.004 (0.066) | 0.003 (0.055) | 0.048 (0.213) |
| Education (bachelor and above) | 0.579 (0.494) | 0.492 (0.5) | 0.766 (0.423) | 0.236 (0.425) | 0.753 (0.431) | 0.615 (0.487) | 0.605 (489) | 0.532 (0.499) | 0.473 (0.499) |
| Observations | 0.408 (0.492) | 0.43 (0.495) | 0.232 (0.422) | 0.758 (0.429) | 0.246 (0.431) | 0.615 (0.487) | 0.391 (0.488) | 0.465 (0.499) | 0.48 (0.5) |
| | 2,456 | 3,953 | 2,334 | 1,396 | 2,399 | 2,236 | 2,552 | 2,280 | 1,172 |

| Variable | Greece | Ecuador | Chile | Spain | Slovenia | Poland | Norway | Netherlands | Korea |
|--|---------------|----------------|----------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Large Firm | 0.223 (0.416) | 0.254 (0.435) | 0.361 (0.48) | 0.348 (0.476) | 0.525 (0.5) | 0.376 (0.485) | 0.446 (0.497) | 0.483 (0.5) | 0.378 (0.485) |
| Log Earnings | 2.056 (0.523) | 1.116 (0.65) | 1.519 (0.782) | 2.519 (0.569) | 2.223 (0.426) | 1.333 (0.602) | 3.381 (0.43) | 3.057 (0.616) | 4.778 (0.735) |
| Female | 0.324 - 3.892 | -0.002 - 4.066 | -0.397 - 6.467 | 0.392 - 9.418 | 0.967 - 4.277 | -1.148 - 10.1 | 0.495 - 6.559 | 0.46 - 12.948 | 1.783 - 8.332 |
| Age (16-24) | 0.569 (0.496) | 0.387 (0.487) | 0.493 (0.5) | 0.484 (0.5) | 0.517 (0.5) | 0.446 (0.497) | 0.517 (0.5) | 0.506 (0.5) | 0.46 (0.498) |
| Age (25-34) | 0.077 (0.267) | 0.201 (0.401) | 0.155 (0.362) | 0.076 (0.266) | 0.043 (0.202) | 0.398 (0.489) | 0.17 (0.376) | 0.173 (0.378) | 0.096 (0.294) |
| Age (35-44) | 0.258 (0.438) | 0.317 (0.466) | 0.282 (0.45) | 0.25 (0.433) | 0.246 (0.431) | 0.314 (0.464) | 0.193 (0.395) | 0.185 (0.388) | 0.265 (0.441) |
| Age (45-54) | 0.325 (0.469) | 0.236 (0.425) | 0.207 (0.405) | 0.304 (0.46) | 0.295 (0.456) | 0.119 (0.324) | 0.246 (0.431) | 0.233 (0.423) | 0.286 (0.452) |
| Age (above 55) | 0.276 (0.447) | 0.158 (0.365) | 0.229 (0.42) | 0.26 (0.439) | 0.33 (0.47) | 0.108 (0.311) | 0.231 (0.422) | 0.249 (0.432) | 0.239 (0.427) |
| Education (Primary or less) | 0.064 (0.245) | 0.089 (0.284) | 0.127 (0.334) | 0.11 (0.313) | 0.087 (0.282) | 0.062 (0.241) | 0.16 (0.367) | 0.161 (0.368) | 0.115 (0.319) |
| Education (up to professional tertiary education degree) | 0.054 (0.225) | 0.305 (0.461) | 0.076 (0.264) | 0.132 (0.339) | 0.004 (0.067) | 0.009 (0.096) | 0.001 (0.028) | 0.005 (0.07) | 0.05 (0.217) |
| Education (bachelor and above) | 0.534 (0.499) | 0.53 (0.499) | 0.625 (0.484) | 0.424 (0.484) | 0.606 (0.489) | 0.67 (0.47) | 0.555 (0.497) | 0.606 (0.489) | 0.464 (0.499) |
| Observations | 0.413 (0.493) | 0.165 (0.372) | 0.3 (0.458) | 0.444 (0.497) | 0.389 (0.488) | 0.321 (0.467) | 0.445 (0.497) | 0.344 (0.475) | 0.487 (0.5) |
| | 858 | 1,446 | 2,118 | 2,030 | 1,788 | 3,789 | 2,562 | 2,840 | 3,007 |

Notes: Baseline sample. All variables except log earnings are binary. Means are reported with standard deviations in parentheses. For log earnings, the min and max are reported on a separate row.

Table A2: Summary Statistics (cont.)

| Variable | Japan | Italy | Ireland | France | Denmark | Cyprus | Belgium | Vietnam | Ukraine |
|--|----------------|---------------|---------------|---------------|---------------|---------------|---------------|----------------|----------------|
| Large Firm | 0.429 (0.495) | 0.376 (0.484) | 0.405 (0.491) | 0.449 (0.497) | 0.458 (0.498) | 0.314 (0.464) | 0.533 (0.499) | 0.417 (0.493) | 0.545 (0.499) |
| Log Earnings | 2.638 (0.632) | 2.715 (0.517) | 3.052 (0.565) | 2.782 (0.445) | 3.407 (0.438) | 2.547 (0.571) | 3.122 (0.442) | -6.888 (3.159) | 0.009 (0.582) |
| Female | -3.016 - 7.841 | 0.743 - 5.607 | 0.179 - 7.333 | 0.387 - 6.784 | 0.125 - 6.604 | 0.683 - 5.836 | 0.565 - 6.895 | -60.477 - 1.59 | -2.287 - 3.012 |
| Age (16-24) | 0.489 (0.5) | 0.48 (0.5) | 0.592 (0.492) | 0.495 (0.5) | 0.507 (0.5) | 0.592 (0.492) | 0.497 (0.5) | 0.544 (0.498) | 0.643 (0.48) |
| Age (25-34) | 0.109 (0.312) | 0.051 (0.22) | 0.089 (0.284) | 0.089 (0.285) | 0.12 (0.325) | 0.084 (0.277) | 0.098 (0.297) | 0.114 (0.318) | 0.06 (0.238) |
| Age (35-44) | 0.205 (0.403) | 0.203 (0.402) | 0.267 (0.442) | 0.229 (0.42) | 0.131 (0.337) | 0.323 (0.468) | 0.249 (0.432) | 0.304 (0.46) | 0.28 (0.45) |
| Age (45-54) | 0.267 (0.442) | 0.344 (0.475) | 0.303 (0.459) | 0.262 (0.44) | 0.206 (0.404) | 0.264 (0.441) | 0.245 (0.43) | 0.297 (0.457) | 0.233 (0.423) |
| Age (above 55) | 0.223 (0.416) | 0.279 (0.449) | 0.204 (0.403) | 0.278 (0.448) | 0.23 (0.421) | 0.211 (0.408) | 0.296 (0.457) | 0.206 (0.405) | 0.25 (0.434) |
| Education (Primary or less) | 0.196 (0.397) | 0.123 (0.328) | 0.139 (0.346) | 0.141 (0.349) | 0.313 (0.464) | 0.118 (0.322) | 0.112 (0.316) | 0.079 (0.269) | 0.178 (0.383) |
| Education (up to professional tertiary education degree) | 0 (0.018) | 0.033 (0.18) | 0.034 (0.182) | 0.028 (0.166) | 0.004 (0.06) | 0.046 (0.209) | 0.02 (0.14) | 0.14 (0.347) | 0.003 (0.05) |
| Education (bachelor and above) | 0.479 (0.5) | 0.743 (0.437) | 0.486 (0.5) | 0.576 (0.494) | 0.538 (0.499) | 0.388 (0.487) | 0.539 (0.499) | 0.419 (0.494) | 0.378 (0.485) |
| Observations | 0.521 (0.5) | 0.224 (0.417) | 0.009 (0.09) | 0.395 (0.489) | 0.459 (0.498) | 0.566 (0.496) | 0.441 (0.497) | 0.441 (0.497) | 0.62 (0.486) |
| | 3,124 | 1,530 | 2,221 | 3,019 | 3,597 | 1,639 | 2,465 | 1,260 | 400 |

| Variable | Sri Lanka | Macedonia | Laos | Kenya | Georgia | Colombia | China | Bolivia | Armenia |
|--|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|----------------|
| Large Firm | 0.3 (0.458) | 0.438 (0.496) | 0.30 (0.462) | 0.211 (0.408) | 0.379 (0.486) | 0.344 (0.475) | 0.405 (0.491) | 0.285 (0.452) | 0.422 (0.494) |
| Log Earnings | -0.132 (0.912) | 0.512 (1.689) | -0.144 (0.724) | -0.007 (1.099) | 0.119 (0.815) | 0.493 (0.841) | 0.511 (0.689) | 0.711 (2.584) | -6.826 (0.746) |
| Female | -3.255 - 4.366 | -54.59 - 3.924 | -3.488 - 2.621 | -3.297 - 4.221 | -2.744 - 3.296 | -7.503 - 4.517 | -2.638 - 3.747 | -52.352 - 4.324 | -8.545 - 0.690 |
| Age (16-24) | 0.374 (0.484) | 0.481 (0.5) | 0.34 (0.474) | 0.398 (0.49) | 0.679 (0.467) | 0.428 (0.495) | 0.506 (0.5) | 0.54 (0.499) | 0.665 (0.472) |
| Age (25-34) | 0.1 (0.3) | 0.04 (0.196) | 0.127 (0.333) | 0.246 (0.431) | 0.077 (0.267) | 0.184 (0.388) | 0.064 (0.244) | 0.184 (0.388) | 0.095 (0.293) |
| Age (35-44) | 0.273 (0.446) | 0.273 (0.445) | 0.331 (0.471) | 0.447 (0.497) | 0.22 (0.414) | 0.356 (0.479) | 0.258 (0.438) | 0.354 (0.479) | 0.22 (0.415) |
| Age (45-54) | 0.323 (0.468) | 0.274 (0.446) | 0.308 (0.462) | 0.185 (0.389) | 0.26 (0.439) | 0.233 (0.423) | 0.398 (0.49) | 0.272 (0.445) | 0.222 (0.416) |
| Age (above 55) | 0.211 (0.409) | 0.244 (0.43) | 0.182 (0.386) | 0.088 (0.284) | 0.249 (0.433) | 0.168 (0.375) | 0.234 (0.424) | 0.121 (0.327) | 0.254 (0.436) |
| Education (Primary or less) | 0.093 (0.29) | 0.169 (0.375) | 0.052 (0.222) | 0.033 (0.18) | 0.194 (0.396) | 0.059 (0.235) | 0.046 (0.21) | 0.069 (0.254) | 0.209 (0.407) |
| Education (up to professional tertiary education degree) | 0.182 (0.386) | 0.008 (0.09) | 0.233 (0.424) | 0.33 (0.47) | 0.004 (0.06) | 0.21 (0.408) | 0.05 (0.218) | 0.079 (0.271) | 0.003 (0.056) |
| Education (bachelor and above) | 0.690 (0.463) | 0.648 (0.478) | 0.352 (0.478) | 0.482 (0.5) | 0.229 (0.421) | 0.44 (0.497) | 0.546 (0.498) | 0.322 (0.468) | 0.26 (0.439) |
| Observations | 0.128 (0.334) | 0.344 (0.475) | 0.415 (0.493) | 0.188 (0.391) | 0.767 (0.423) | 0.35 (0.477) | 0.404 (0.491) | 0.598 (0.491) | 0.736 (0.441) |
| | 681 | 1,222 | 347 | 1,019 | 546 | 748 | 802 | 478 | 645 |

Notes: Baseline sample. All variables except log earnings are binary. Means are reported with standard deviations in parentheses. For log earnings, the min and max are reported on a separate row.

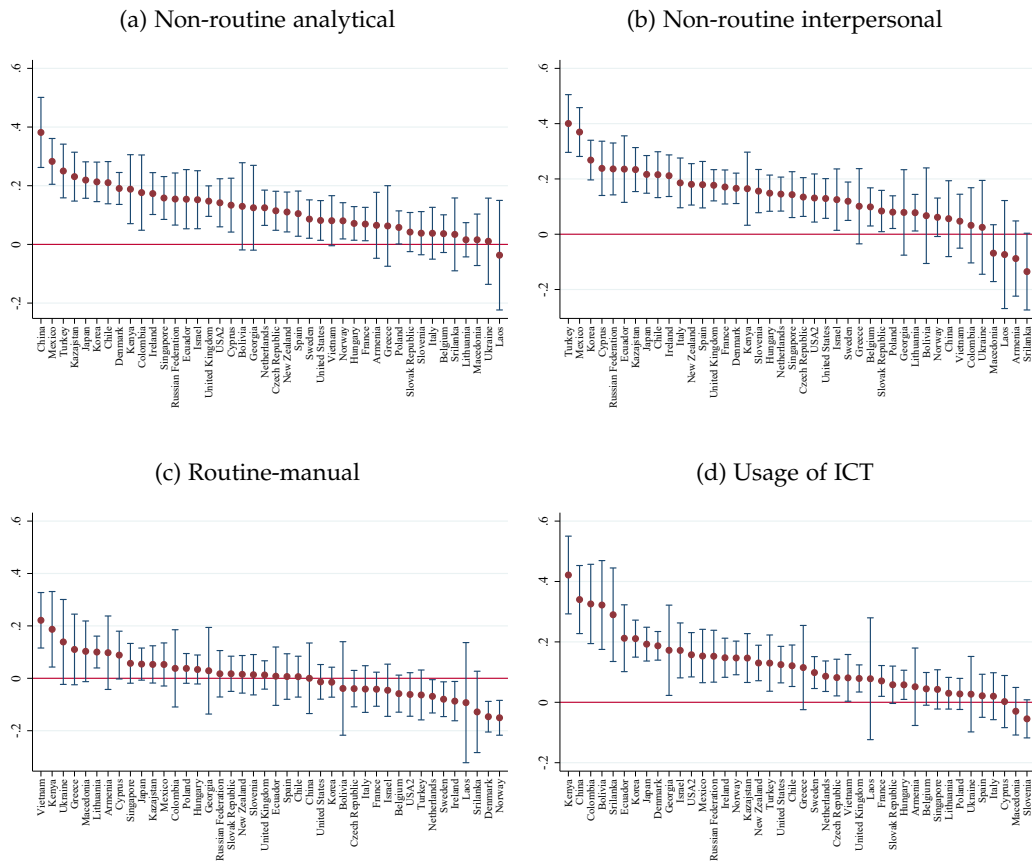
B Appendix: Additional tables and figures

Table B1: Firm size gradient in the task content of jobs, by occupation, pooled estimates

| 1-digit ISCO-08 Category | Task Category | | | | # Obs. |
|--|--------------------|--------------------|---------------------|---------------------|--------|
| | Analytical | Interpersonal | Manual | ICT | |
| Managers | 0.1374 (0.0233) | 0.3469 (0.0331) | -0.1704 (0.0353) | 0.0864 (0.0245) | 5,360 |
| Professionals | 0.1694 (0.0176) | 0.1657 (0.0173) | -0.0794 (0.0250) | 0.1721 (0.0192) | 17,658 |
| Technicians & associate professionals | 0.0893 (0.0202) | 0.1201 (0.0179) | -0.0147 (0.0256) | 0.0779 (0.0246) | 12,107 |
| Clerical support workers | 0.0914 (0.0331) | 0.1491 (0.0370) | 0.0243 (0.0204) | 0.0738 (0.0289) | 9,110 |
| Services & sales workers | 0.1866 (0.0249) | 0.2115 (0.0348) | 0.0444 (0.0238) | 0.1548 (0.0271) | 15,902 |
| Skilled agri, forestry & fishery workers | 0.1061 (0.0643) | 0.2635 (0.0940) | 0.1010 (0.0884) | -0.0096 (0.0442) | 524 |
| Craft & related trade workers | 0.0822 (0.032) | 0.0471 (0.034) | 0.0368 (0.0316) | 0.1129 (0.0292) | 7,969 |
| Plant & machine operators, & assemblers | 0.1147 (0.0304) | 0.0721 (0.0266) | 0.0636 (0.0260) | 0.0540 (0.0200) | 6,384 |
| Elementary occupations | 0.0836 (0.0231) | 0.1218 (0.0191) | 0.0968 (0.0190) | 0.0544 (0.0135) | 7,741 |

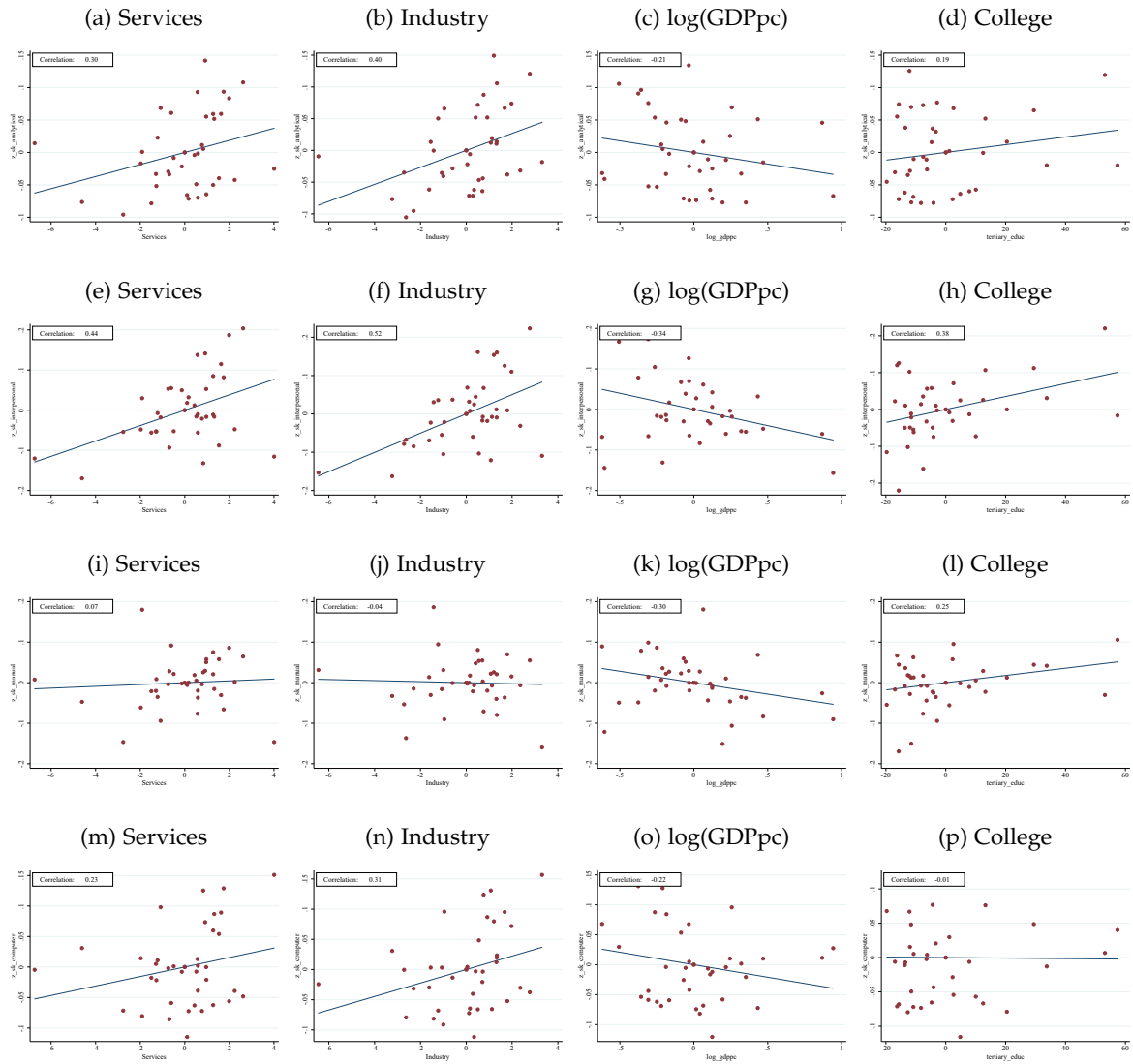
Notes: Pooled sample. Coefficient of an indicator for large firm in a regression of task content intensity on indicator of large firm (at least 50 employees) by 1-digit ISCO-08 occupation codes, controlling for 2-digit occupation codes and country fixed effects. We do not report one 1-digit category corresponding to armed forced occupations due to small sample size. Reported standard errors clustered at the country-level.

Figure B1: Firm size gradient in the task content of jobs by country, within 2-digit occupations



Notes: Baseline sample. Coefficient of an indicator for large firm in a regression of task content intensity on indicator of large firm (at least 50 employees) and dummies for 2-digit ISCO-o8 occupation codes. Countries ordered by decreasing point estimates. Regressions done for each country separately. Reported confidence intervals at 95% confidence level computed using heteroskedasticity-robust standard errors.

Figure B2: Partial correlations of the firm size gradient in task content



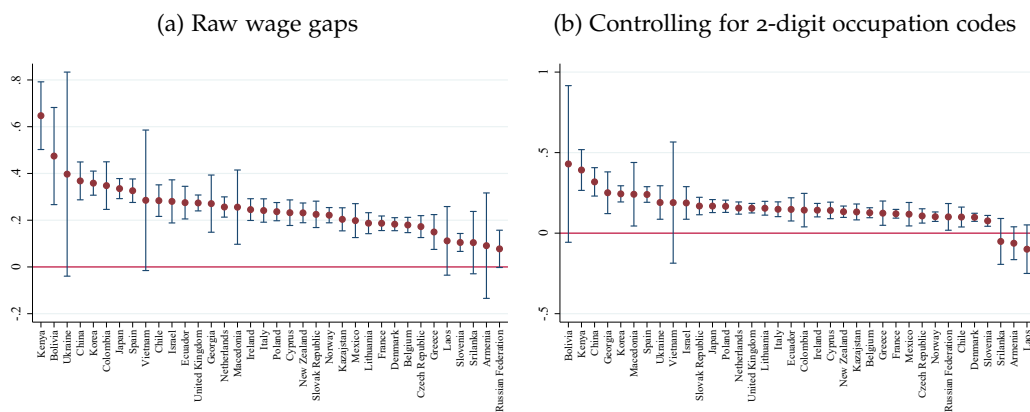
Notes: Partial correlations of the firm size gradient in task content on country-level covariates. Covariates considered are services share, industry share, log GDP per capita, and ratio of college educated. Rows of panels correspond to non-routine analytical, non-routine interpersonal, routine manual, and usage of ICT.

Table B2: Firm size gradient in the task content of jobs, pooled estimates, young workers with short tenure

| Task Category | PIAAC | | STEP | |
|---------------------------|---------------------|--------------------|--------------------|--------------------|
| | Short Tenure | + Age < 25 | Short Tenure | + Age < 25 |
| Non-routine Analytical | 0.1441 (0.0151) | 0.0866 (0.0282) | 0.1425 (0.0197) | 0.1684 (0.0383) |
| Non-routine Interpersonal | 0.1594 (0.0146) | 0.0580 (0.0232) | 0.0286 (0.0385) | 0.0835 (0.0781) |
| Routine Manual | -0.0040 (0.0154) | 0.0673 (0.0250) | 0.0571 (0.0383) | 0.2070 (0.0638) |
| Use of ICT | 0.1121 (0.0139) | 0.0814 (0.0369) | 0.2504 (0.0643) | 0.2392 (0.1127) |
| # of Observations | 32,802 | 6,642 | 2,844 | 750 |

Notes: Pooled sample. Coefficient of an indicator for large firm in a regression of task content intensity on indicator of large firm (at least 50 employees) controlling for 2-digit occupation codes and country fixed effects. The first and third columns restrict the sample to workers with short tenure. In PIAAC there is no direct question of tenure, so we proxy short-tenure by an individual having worked for multiple firms in the last five years. STEP does provide information on the months that the individual has worked for the firm. We are therefore able to define short tenure in a more demanding manner: having worked for the current employer for up to 24 months. The second and fourth columns additional require the worker to be up to 25 years Reported standard errors clustered at the country-level.

Figure B3: Estimated average large firm wage premium by country



Notes: Baseline Sample. Regression of log wages on an indicator of large firm (at least 50 employees). In (a) not controlling for anything, while in (b) controlling for 2-digit ISCO-o8 occupation codes. Countries ordered by decreasing point estimates. Regressions done for each country separately. Reported confidence intervals at 95% confidence level computed using heteroskedasticity-robust standard errors.

Table B3: Evidence from the demand side

(a) Task requirements

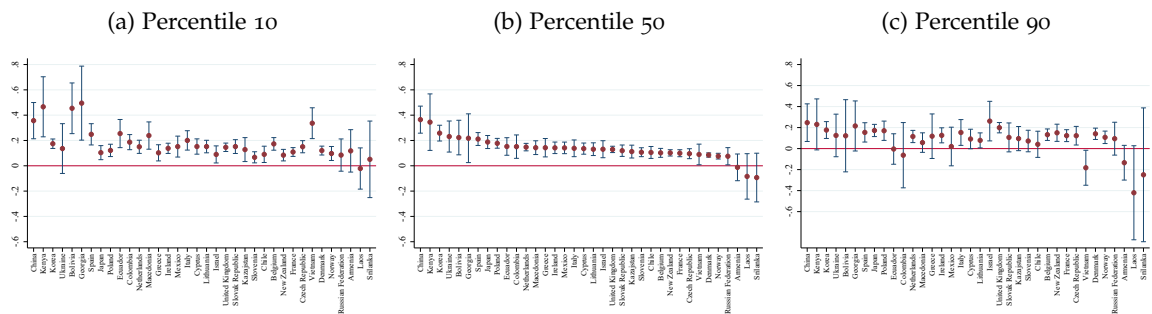
| Task category | LF estimate | # Obs. |
|---------------------------|-----------------|--------|
| Non-routine analytical | 0.1631 (0.0488) | 8,236 |
| Non-routine interpersonal | 0.1593 (0.0078) | 8,290 |
| Use of ICT | 0.1062 (0.0000) | 8,212 |

(b) Problems encountered in hiring

| Problems encountered | LF Estimate |
|---|-----------------|
| No or few applicants | 0.0236 (0.0567) |
| Applicants lack required skills | 0.0556 (0.0037) |
| Applicants expect higher wages than offered | 0.0376 (0.0273) |
| Applicants dislike working conditions | 0.0142 (0.0234) |
| # of observations | 43,590 |

Notes: Pooled sample, STEP employer surveys. Panel (a): Coefficient in a regression of task measure on an indicator of large firm, occupation fixed effects, and country fixed effects. In parenthesis, we report the p-values of the test that the effects are null using wild-bootstrapped standard errors clustered at the country level. Panel (b): Linear probability model of encountering problems in hiring on large firm indicator, occupation fixed effects, and country fixed effects. In parenthesis, we report the p-values of the test that the effects are null using wild-bootstrapped standard errors clustered at the country level.

Figure B4: Distribution of wages in large firms by country, within 2-digit occupations; percentiles 10, 50 and 90



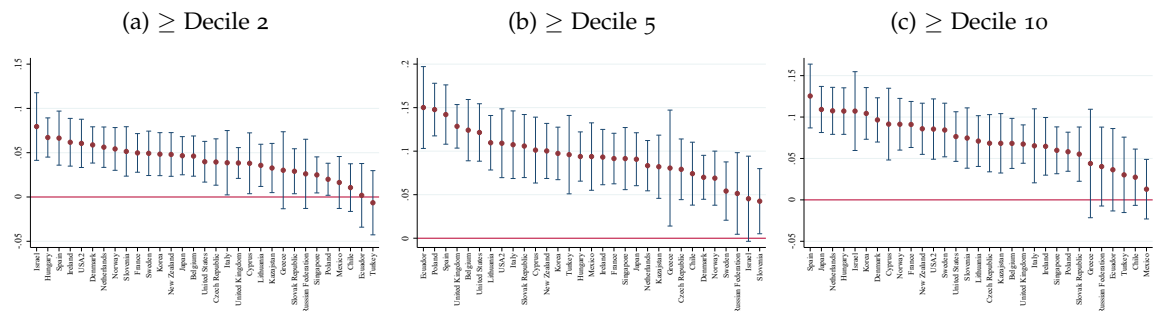
Notes: Baseline sample. Coefficient of an indicator for large firm in a quantile regression of task content intensity on indicator of large firm (at least 50 employees) and dummies for 2-digit ISCO-08 occupation codes. Countries ordered by decreasing point estimates in the median regression. Regressions done for each country separately. Reported confidence intervals at 95% confidence level computed using heteroskedasticity-robust standard errors.

Table B4: Large firm wage premium, by occupation, pooled sample

| 1-digit ISCO-08 Category | LFWP | # Obs. |
|--|--------------------|--------|
| Managers | 0.2676 (0.0537) | 4,237 |
| Professionals | 0.1600 (0.0194) | 14,369 |
| Technicians & associate professionals | 0.1077 (0.0198) | 9,383 |
| Clerical support workers | 0.1487 (0.0189) | 7,696 |
| Services & sales workers | 0.1860 (0.0200) | 12,920 |
| Skilled agri, forestry & fishery workers | 0.0541 (0.1253) | 430 |
| Craft & related trade workers | 0.1542 (0.0248) | 6,558 |
| Plant & machine operators, & assemblers | 0.1532 (0.0198) | 5,149 |
| Elementary occupations | 0.0544 (0.0202) | 6,158 |

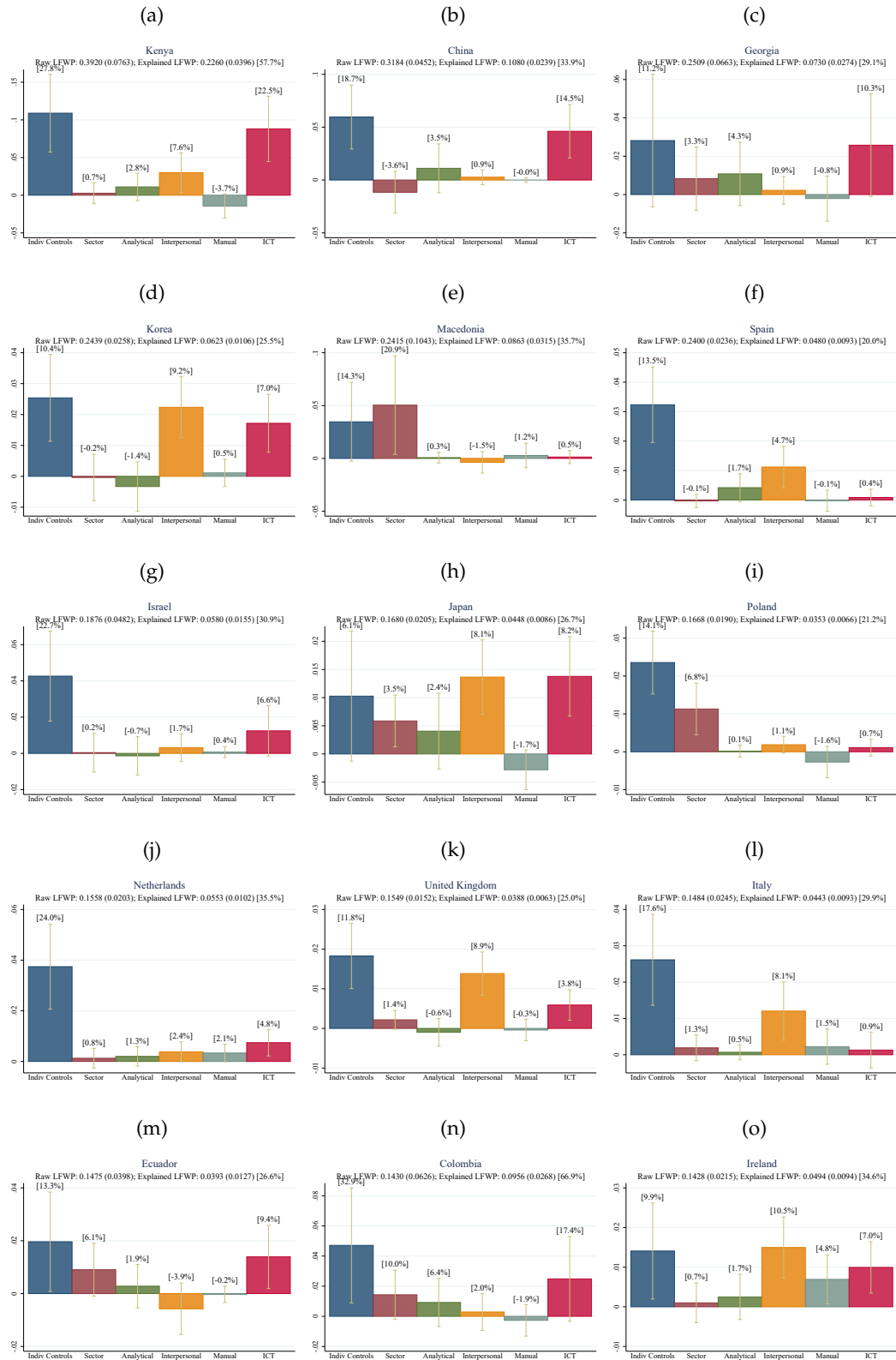
Notes: Pooled sample. Coefficient of an indicator for large firm in a regression of log real hourly wages on indicator of large firm (at least 50 employees) by 1-digit ISCO-08 occupation codes, controlling for 2-digit occupation codes and country fixed effects. We do not report one 1-digit category corresponding to armed forced occupations due to small sample size. Reported standard errors clustered at the country-level.

Figure B5: LFWP based on wage deciles in PIAAC, controlling for 2-digit occupation codes



Notes: Baseline sample. Coefficient of an indicator of being at least in a certain wage decile in linear regression on indicator of large firm (at least 50 employees) and dummies for 2-digit ISCO-08 occupation codes. Countries ordered by decreasing point estimates. Regressions done for each country separately. Reported confidence intervals at 95% confidence level computed using heteroskedasticity-robust standard errors.

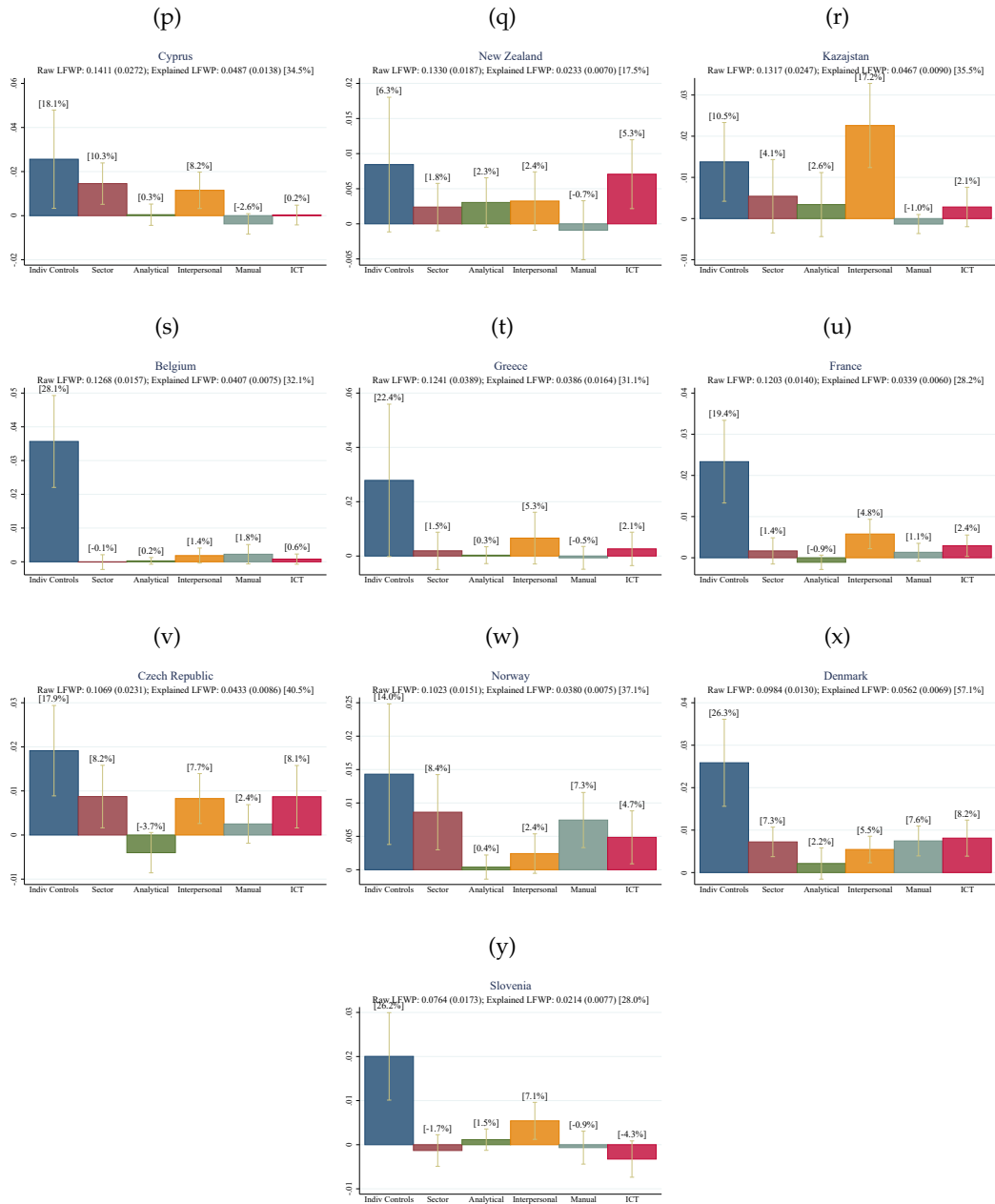
Figure B6: Gelbach decomposition of LFWP, by country



A11

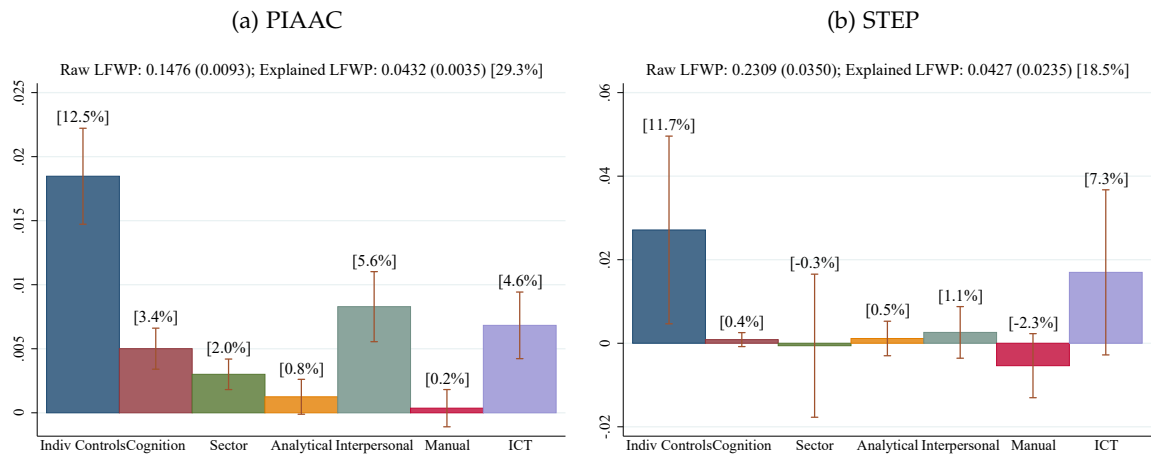
Notes: Only countries for which both the LFWP and explained portion of the LFWP are statistically significant at the 95% confidence level. Countries ordered in decreasing point estimate of the LFWP. Numbers in brackets indicate percentages of the raw LFWP.

Figure B6: Gelbach decomposition of LFWP, by country (cont.)



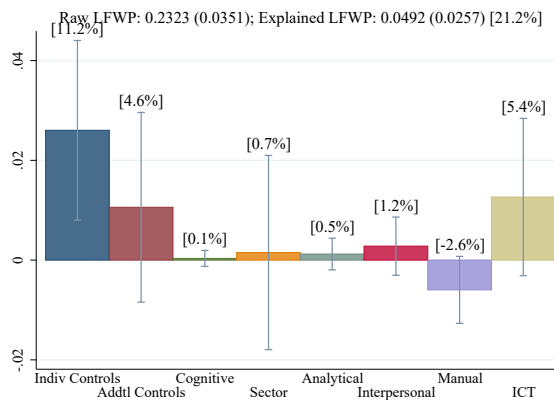
Notes: Only countries for which both the LFWP and explained portion of the LFWP are statistically significant at the 95% confidence level. Countries ordered in decreasing point estimate of the LFWP. Numbers in brackets indicate percentages of the raw LFWP.

Figure B7: Gelbach decomposition of LFWP (extended analysis), pooled



Notes: Numbers in brackets indicate percentages of the raw LFWP. Reported confidence intervals at 95% confidence level. Standard errors are clustered at the country level.

Figure B8: Gelbach decomposition of LFWP (extended analysis 2) for STEP countries, pooled



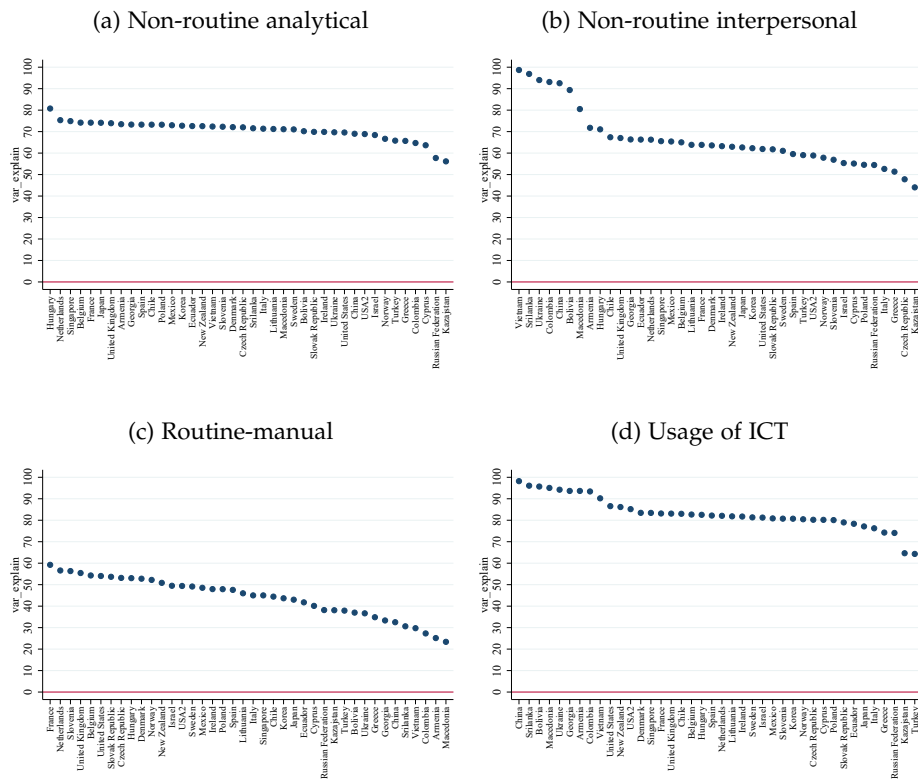
Notes: Numbers in brackets indicate percentages of the raw LFWP. Reported confidence intervals at 95% confidence level. Standard errors are clustered at the country level.

C Appendix: Additional analyses

C.1 Alternative construction of measures: Multiple correspondence analysis (MCA)

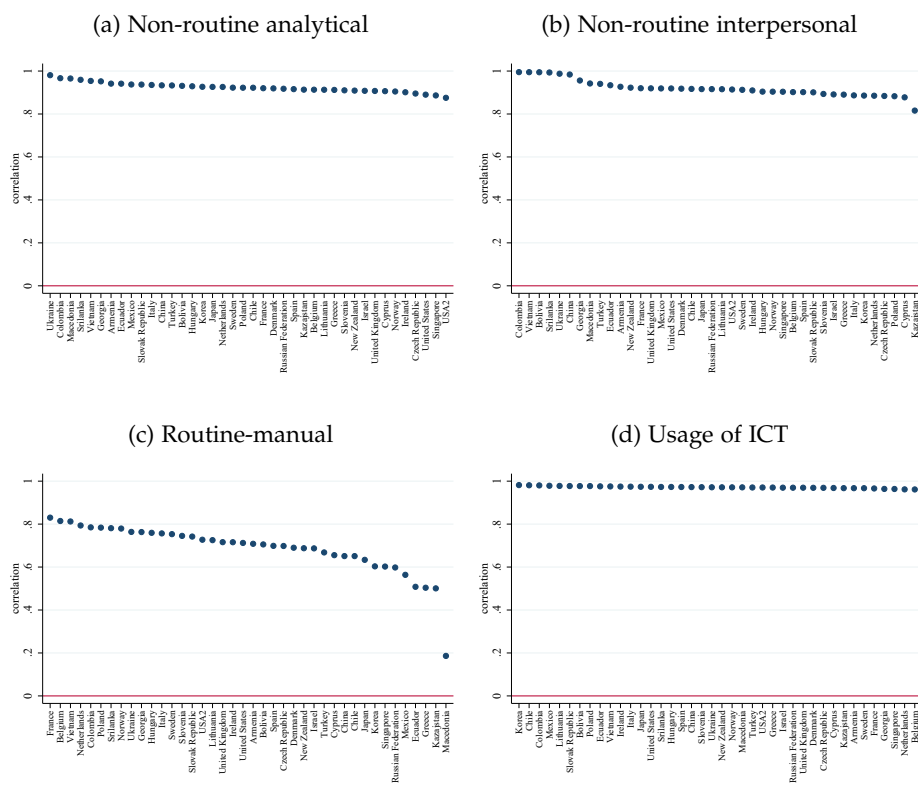
We assess the robustness of our qualitative results by performing some of the same analyses using a differently constructed measure of task content. To construct the measure, we use the same questions detailed in Appendix Table A1 but aggregate them in a different way. We choose to perform a multiple correspondence analysis (MCA) and choose the first dimension extracted which corresponds to the dimension that explains the largest variance in the data. MCA is a tool that could be thought of as the counterpart of principal component analysis for categorical (particularly, ordinal) data, such as the responses to the questions we have. We opt to perform the MCA by country and we standardize the resulting measure within the country. This means that they are again interpreted as standard deviations relative to the country mean.

Figure C1: Explained variance by first dimension in MCA



Appendix Figure C1 summarizes the proportion of total variance explained by the first

Figure C2: Correlation of MCA measure with main measure



dimension in the MCA. For three of the task dimensions, the MCA measures constructed explain a large proportion of the variation in the responses to the underlying questions. For non-routine analytical-related questions, the MCA measure explains around 60–80% of the variation, depending on the country. This is around 50–100% for non-routine interpersonal and 70–100% for the usage of ICT. The variance explained by the first dimension in the MCA using the questions related to routine manual tasks is smaller — around 20–60%, depending on the country. The second dimension in MCA still explains a substantial portion of the variation. This would suggest that the questions we associate with routine manual tasks capture multiple dimensions which a single index could not fully capture. Further study with larger data sets is required to better understand this. Appendix Figure C2 reports the correlation between this MCA measure and the main measure we use in the main text. The correlation of our MCA measures are close to one when we look at non-routine analytical tasks, non-routine interpersonal tasks, and usage of ICT. Unsurprisingly, the correlation of the main measure we use for routine-manual is less correlated with the MCA measure (around 0.4–0.8, with 0.2 for Macedonia).

Table C1: Pooled estimates of firm size gradient in the task content of jobs, MCA measure

| | (1) Raw | Occupation controls | | |
|---------------------------|---------------------|---------------------|--------------------|--------------------|
| | | (2) 1-digit | (3) 2-digit | (4) 3-digit |
| <i>Outcome variable:</i> | | | | |
| Non-routine analytical | 0.3044 (0.0250) | 0.1333 (0.0153) | 0.1142 (0.0129) | 0.0981 (0.0144) |
| Non-routine interpersonal | 0.1961 (0.0234) | 0.0945 (0.0164) | 0.0950 (0.0138) | 0.0976 (0.0147) |
| Routine-manual | −0.0214 (0.0208) | 0.0637 (0.0158) | 0.0576 (0.0129) | 0.0492 (0.0135) |
| Usage of ICT | 0.3147 (0.0251) | 0.1439 (0.0152) | 0.1213 (0.0128) | 0.1051 (0.0145) |
| Sample Size | 88,901 | 87,987 | 82,901 | 57,055 |

Notes: Regression of task requirement intensity (MCA measure) on an indicator of large firm (at least 50 employees) and country dummies in STEP-PIAAC pooled sample. Differences in sample size explained by (1) restriction of at least 5 observations per firm-size cross occupation cell, or (2) availability of ISCO-08 occupation codes. Reported standard errors are clustered at the country level.

In Appendix Table C1, we replicate the results reported in Table 1 using the MCA measure we constructed. In general, we find qualitatively similar results for the main results we present in the main text — workers in larger firms tend to do more non-routine

analytical tasks, non-routine interpersonal tasks, and use more ICT. The differences in the performance non-routine interpersonal tasks is more modest using the MCA measure. Using the MCA measure, there is some evidence towards workers in larger firms also performing more routine-manual tasks. As discussed previously, this may reflect the MCA measure capturing a specific dimension of routine-manual tasks.

C.2 Distributional differences in the task content of jobs by firm size

The previous analyses have shown that the average worker in a large firm performs more non-routine analytical and interpersonal tasks, and use more ICT, even within narrowly defined occupations. To complement this finding, we run various distribution regressions which models the conditional distribution of the outcome Chernozhukov et al. (2013).² We approximate the probability that the task intensity performed T_i is greater than a particular value $t \in \mathcal{T} \subset \text{Supp}(T_i)$ as a linear probability model:

$$\Pr(T_i > t \mid LF_{j(i)}, o(i), c(i)) = \beta \times LF_{j(i)} + \delta_{o(i)}^o + \delta_{c(i)}^c + \varepsilon_i \quad (5)$$

where $LF_{j(i)}$ is the indicator for worker i being in a large firm, δ^o are occupation fixed effects, and δ^c are country fixed effects. We report the estimates for β for a set of support points \mathcal{T} . We have also tried using a probit specification and the qualitative results do not change.

Appendix Table C2 summarizes the estimated coefficients in the distribution regression of the task content of jobs on an indicator of being in a large firm controlling for 2-digit ISCO-o8 occupation codes. We consider the points on the support $\{-0.75, -0.5, 0, 0.5, 0.75\}$ as thresholds. We find that the coefficient on firm size for non-routine analytical, non-routine interpersonal and use of ICT is positive for all five support points. This suggests that our result that the average worker in a large firm that performs more of non-routine analytical or interpersonal tasks, or use more ICT relative to their counterparts in smaller firms is not driven by a few workers who perform these tasks more intensively — the difference in means is not because of differences in the tails but that the entire tasks intensity distribution in large firms is shifted to the right relative to smaller firms.

C.3 Correlations in firm size gradient of tasks

In Figure C3, we correlate the firm size gradients in the different task categories. We find interesting patterns. Panel (e) in the figure suggests that the firm size gradient in the performance of non-routine interpersonal tasks and that of tasks involving ICT seem to be

²A complementary approach is quantile regression. However, since the task content measures inherit the discrete nature of the responses to the survey questions, quantile regressions may not be appropriate while distribution regression remains valid.

Table C2: Distribution regression estimates of firm size gradient in the task content of jobs, pooled sample

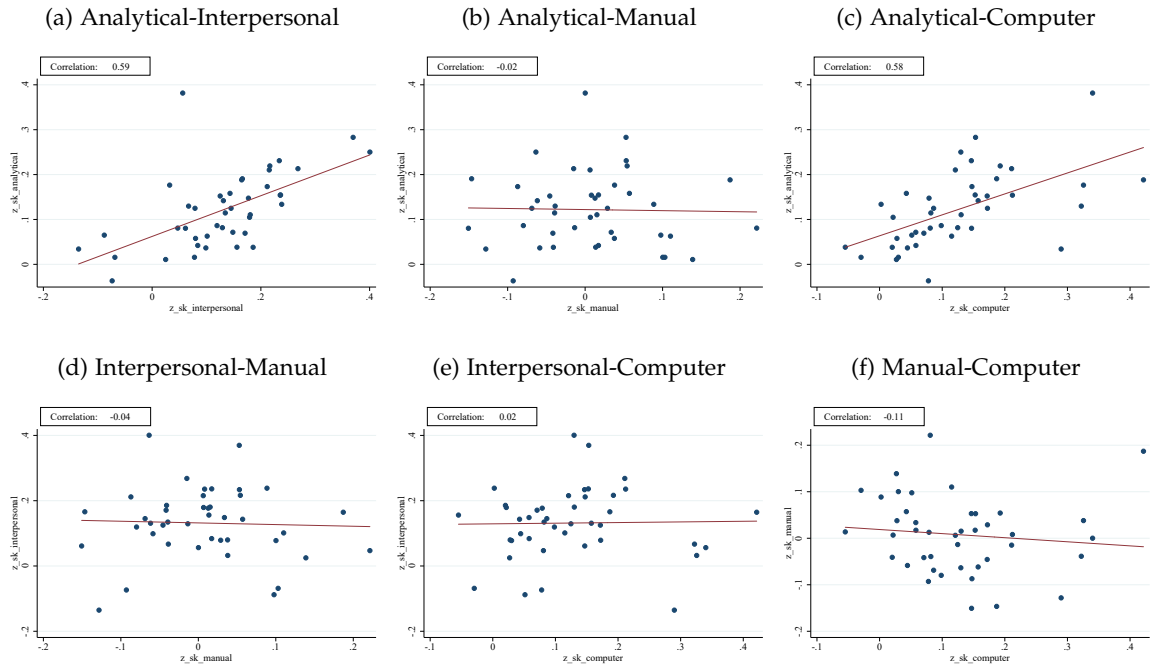
| | Support points | | | | |
|---------------------------|---------------------|---------------------|---------------------|--------------------|--------------------|
| | (1) | (2) | (3) | (4) | (5) |
| | -0.75 | -0.50 | 0.00 | 0.50 | 0.75 |
| <i>Outcome variable:</i> | | | | | |
| Non-routine analytical | 0.0358 (0.0040) | 0.0452 (0.0047) | 0.0534 (0.0059) | 0.0511 (0.0062) | 0.0438 (0.0057) |
| Non-routine interpersonal | 0.0326 (0.0044) | 0.0377 (0.0049) | 0.0362 (0.0052) | 0.0491 (0.0054) | 0.0502 (0.0048) |
| Routine-manual | -0.0111 (0.0054) | -0.0111 (0.0053) | -0.0040 (0.0059) | 0.0144 (0.0052) | 0.0182 (0.0044) |
| Usage of ICT | 0.0440 (0.0058) | 0.0562 (0.0053) | 0.0558 (0.0059) | 0.0454 (0.0059) | 0.0312 (0.0063) |
| Sample Size | 82,901 | 82,901 | 82,901 | 82,901 | 82,901 |

Notes: Pooled sample. Coefficient of an indicator for large firm in a distribution regression of task content intensity on indicator of large firm (at least 50 employees) and dummies for 2-digit ISCO-08 occupation codes and countries by a linear probability model. Reported standard errors are clustered at the country level.

uncorrelated across the countries. However, countries that have a larger firm size gradient in either the performance of interpersonal tasks or the use of ICT are also the countries for which non-routine tasks are disproportionately performed by workers in large firms (see Panels (a) and (c)).

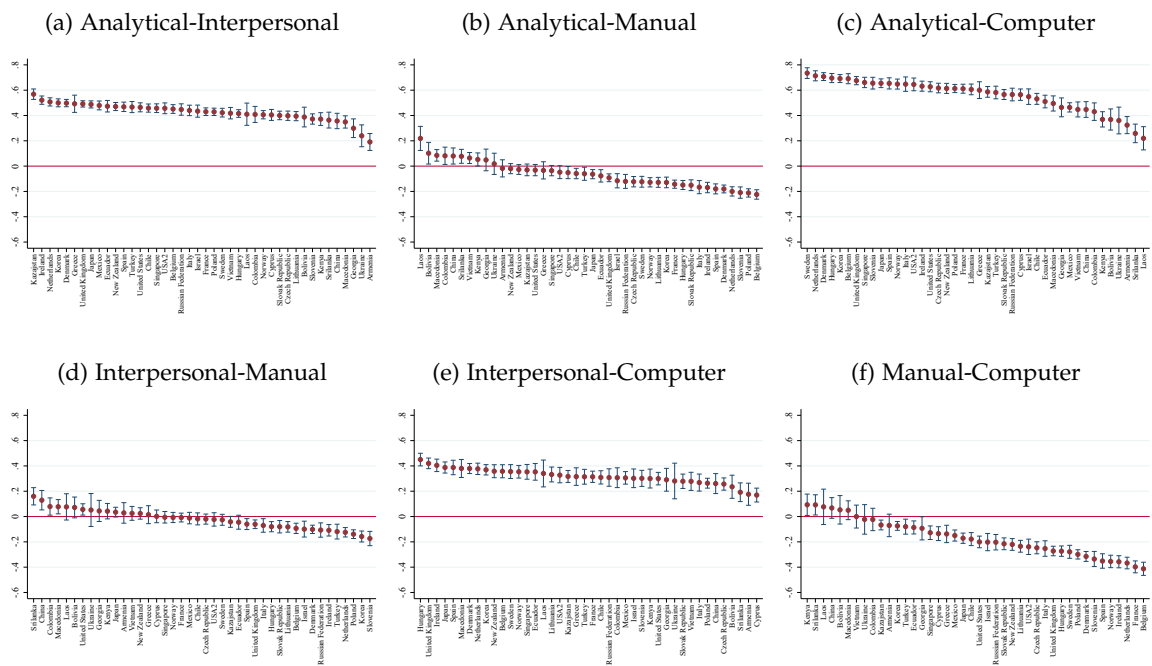
The above result is not informative of the within-country correlation of tasks — that is, whether the larger firms that perform more non-routine analytical tasks are also the firms that perform more non-routine interpersonal tasks or use more ICT. With individual data, we are able to speak towards this and to show that workers that perform more non-routine analytical tasks are the same workers who perform more intensively non-routine interpersonal tasks and use more ICT. Figure C4 reports the coefficients from regressing one measure of performed tasks against another at the individual level, controlling for 2-digit ISCO-08 occupation codes. Two main patterns arise. First, the analytical, interpersonal, and IT content of jobs seem to be complements: they are all strongly positively correlated throughout the countries in our sample. In terms of magnitude, these correlations are frequently around 0.4, and they are noticeably higher for analytical-computer, reaching levels of around 0.75. This last point will be of relevance in our decomposition of drivers of the LFWP. Second, there tends to be a negative correlation between the analytical/interpersonal/IT and the manual content of the job. This should be, however, qualified.

Figure C3: Correlation of firm size gradients in task content, between countries



For analytical and interpersonal, the correlation is slightly positive for a handful of countries. These are all STEP countries. In principle, this could be due to (i) the fact that STEP countries are less developed than their PIAAC counterparts, or (ii) the manual content of a task is relatively harder to capture through our survey measures (further discussed in Appendix C.1). This, together with the fact that STEP measures are coarser could make these countries to display a positive, albeit small, positive correlation between manual and any other form of content. Although it is hard to be conclusive, reason (i) seems less plausible since PIAAC also includes middle-income countries for which we do not observe such a positive correlation.

Figure C4: Within-job correlation of task content



Notes: Baseline sample. By country regression at the individual level of one measure of task content against another controlling for dummies for 2-digit ISCO-o8 occupation codes. Reported confidence intervals at 95% confidence level computed using heteroskedasticity-robust standard errors.