Skill Demand Versus Skill Use

Comparing Job Posts With Individual Skill Use on the Job

Daly, Moira K.; Groes, Fane Naja; Jensen, Mathias Fjællegaard

Document Version
Final published version

Published in:
Labour Economics

DOI:
10.1016/j.labeco.2024.102661

Publication date:
2025

Labour Economics 92 (2025) 102661

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Skill demand versus skill use: Comparing job posts with individual skill use on the job☆

Moira Daly [a], Fane Groes [a], Mathias Fjællegaard Jensen [b],*

[a] Copenhagen Business School, Porcelænshaven 16A, Frederiksberg, Denmark
[b] University of Oxford, 10 Manor Rd, Oxford, OX1 3UQ, UK

ARTICLE INFO

JEL classification:
J24

Keywords:
Skills
Tasks
Job postings
Labour force survey

ABSTRACT

Skill requirements in a job post reflect an employer's "wish list," but do they also reflect skills used on the job by the hired worker? We compare skill measures derived from the text of online job posts with individual-level data from the Danish Labour Force Survey (LFS) in which participants report their main skills used on the job as free text. By identifying individual workers from the LFS who can be matched to a job post, we validate that the extensive margin skills measures derived from job postings data reflect main skills used on the job. Thus, using job postings data to analyze skill use on the job is generally a valid empirical strategy. However, we also show that heterogeneity in returns to skills is missed if only the extensive margin of skill demand is considered.

1. Introduction

An extensive literature has studied the dynamics of task-specific skills in Europe, and particularly, in the United States (e.g., Autor et al., 2003; Spitz-Oener, 2006; Black and Spitz-Oener, 2010; Goos et al., 2014; Arntz et al., 2016; Beaudry et al., 2016). This early literature generally relies on measures of skills at the occupation level, but in the last decade, a new data source has emerged from the online posting of job vacancies. Machine reading of job posts has allowed researchers to explore the variation in the demand for task-specific skills within firms and occupations, as well as across time. Task-specific skills refer to skills related to certain tasks, such as social and cognitive skills, and not education levels. A new and rapidly expanding literature has used text from job posts to, for instance, understand the variation in demand for skills within occupations and the effect of this variation on workers' pay (e.g., Modestino et al., 2016; Deming and Kahn, 2018; Hershbein and Kahn, 2018; Marinescu and Rathelot, 2018; Grinis, 2019; Atalay et al., 2020; Blair and Deming, 2020; Deming and Noray, 2020; Modestino et al., 2020; Alekseeva et al., 2021; Daly et al., 2022; Braxton and Taska, 2023).[1]

This nascent literature relies on proprietary job vacancy data, which, by its nature, capture skill demand and not necessarily skill use. Skills listed in a job vacancy may be an employer's "wish list", but not necessarily reflect skills used on the job by the hired worker. Still, much of the literature relies on the implicit assumption that demanded skills reflect skills actually used on the job. Although validation exercises have been performed at the occupational level (Hershbein and Kahn, 2018), the degree to which skill demand advertised by firms captures the skills used by workers at the individual level is still unclear, no doubt due to a lack of data. At the same time, recent papers match vacancy data and administrative data at increasingly granular levels (see e.g., Kettemann et al. (2018), Jensen (2024), Bagger et al. (2022) and Fluchtmann et al. (2022); see also Kircher (2022), for a review of studies linking job seekers to vacancies). Thus, validations of the data at more granular levels are also warranted. By linking job posts with self-reported skill use of workers hired for the posted jobs, we can describe both the relationship between skills demanded by employers and skills used on the job, and each of their effects on wages. We believe that our paper is the first to empirically test the assumption that skills demanded in job posts reflect skill use on the job at the individual level.

To capture skill use on the job, we use the Danish Labour Force Survey (LFS), the country-specific version of the widely available non-proprietary European Union Labour Force Survey (EU LFS). Survey

☆ We are grateful for the help and comments of Lisa Kahn and Ning Zhang. We would also like to thank our research assistant at the time, Oliver-Alexander Press. This work was supported by the Novo Nordisk Foundation grant number NNF16OC0021056 and the Kraks Fond project number 301061. Jensen thanks the Independent Research Fund Denmark for financial support through grant number 1058-00011B.

* Corresponding author.
E-mail addresses: moda.eco@cbs.dk (M. Daly), fg.eco@cbs.dk (F. Groes), mathias.jensen@economics.ox.ac.uk (M.F. Jensen).

[1] A number of other papers also analyze job vacancy data, but focus less on skill demand, see e.g., Adams et al. (2020), Azar et al. (2020), Clemens et al. (2020), Forsythe et al. (2020), Javorcik et al. (2020) and Bagger et al. (2022).

https://doi.org/10.1016/j.labeco.2024.102661
Received 6 November 2023; Received in revised form 25 November 2024; Accepted 26 November 2024
Available online 12 December 2024

respondents are asked about the main tasks they perform on the job, and their free-text answers are recorded.[2] We extract these free-text answers from the LFS and, using a similar approach to that of Deming and Kahn (2018), categorize keywords from the text into 9 categories of task-specific skills: Cognitive, Social, Management, Financial, Computer (general), Computer (specific), Writing/Language, Customer Service, Character.[3] We interpret a reported job task as the utilization of a task-specific skill, and thus, in order to be consistent in our terminology, we refer to the measures extracted from the LFS as the main skills used on the job.

Next, we apply the same categorization to the text of online job posts from Denmark. We compare the measures of task-specific skills extracted from job postings with the individual task-specific skills reported in the LFS.[4] To our knowledge, we are both the first to use the reported task-specific skills from any of the EU LFS surveys and to link these to skills demanded in job postings. As the current literature using job postings data focuses primarily on the US, we believe that, given that such an exercise is not possible in the US, Denmark provides an ideal environment to conduct this exercise, as many characteristics of the Danish labor market resemble those of the US labor market (see e.g., Botero et al., 2004; Groes et al., 2015; Heckman and Landersø, 2022). Particularly relevant for this analysis is the fact that Denmark and the US share very similar levels of labor market turnover rates, employment protection, and economic freedom (see e.g., Kreiner and Svarer, 2022). At the same time, data derived from online job postings are becoming increasingly available across countries, including many European and OECD countries, which makes a validation of the data increasingly relevant.

We find that a significant proportion of workers report using only a single main skill that falls under one of the commonly cited skill categories in the job postings literature.[5] On the other hand, employers demand skills from six different skill categories on average. Given that there is no limit on the heterogeneity of skills that employees report when describing their main skills, one interpretation of these findings is that workers are more specialized than what the skill indicators derived from job postings may suggest: whereas the job posting skills may capture the extensive margin of skills used on the job, the LFS measures capture the intensity of skills used on the job by only including the most important skills. We cannot separately distinguish between concepts of frequency of skill use and importance of skill use. Some workers may deem that their main skill is the task they perform most frequently, while others may consider it to be the task that they think is most important for their work. Nonetheless, we believe that understanding how a measure of individual-level, intensive skill use relates to existing skills measures derived from job postings is a valuable addition to the literature.

Next, we show the degree to which measures of task-specific skills derived from job postings correspond with skills used on the job as reported by workers in the LFS. Workers who report a particular main skill are extremely likely to be in a job that advertised for that particular skill. In addition, we find positive and significant correlations between job post skills and self-reported main skills from the LFS, the only exception being character skills. We also find that about one-tenth of workers report mainly using a skill type that their employer did not include in the job posting. An investigation of the relative employer-employee match quality of this group finds no evidence of shorter match durations or negative wage effects suggesting that mismatch among skill supply and demand is not substantial. Based on this evidence, we continue our analysis under the assumption that skill demand as captured by job posts and main skills reported on the LFS can be interpreted as extensive and intensive measures of skill use, respectively.

We then estimate standard wage regressions and explore the returns to skills on the extensive margin with and without including on-the-job measures of skill intensity (as captured by the LFS). Including measures of on-the-job skill use greatly increases the model's ability to explain the variation in wages when individual controls are not included. The inclusion of skill intensity measures does not qualitatively affect the estimates of the extensive margin return to advertised skills. On the other hand, much of the variation in wages explained by the intensity of skill use is absorbed once individual controls are included in the regression. Taken together, these results suggest that the precision of extensive margin skill return estimates can increase noticeably if intensive skill measures are included, but their inclusion is less necessary when sufficient individual controls are available.

For several skills, we find large differences in the estimated returns to skills derived from job posts and from the LFS. In particular, individuals in jobs that advertise for cognitive and management skills are rewarded substantially more if they use these skills intensively as measured by the LFS. On the other hand, workers in jobs that advertise for writing/language and customer service are severely penalized for using these skills intensively. Our results highlight the fact that although estimates of the return to skill on the extensive margin can accurately describe average returns to skills, workers who intensively use these skills can substantially benefit or suffer, depending on the skill considered.

We consider the robustness of our results across various matching strategies between the LFS and job postings data. We find that our results are robust to varying matching windows between the dates of job posting and job start as well as across various subsamples of new job starters. In addition, we consider different approaches to skill categorization by harmonizing the lexicons across the text data sources and by conditioning on informative LFS survey responses (i.e., those including at least one keyword indicative of a skill). These robustness checks in combination with our results on mismatch suggest that although measurement error affects our two skill measures, we find no systematic relationship in the measurement error between these two skill measures.

As researchers generally will not have access to skill intensity measures from additional data sources such as the LFS, we next extend our results by exploring potential alternative approaches to derive skill measures from job postings that capture some of the variation explained by the LFS skill intensity measures. In our first extension, we consider whether interactions between skills in job postings can explain variation in wages and absorb the explanatory power of LFS skill variables. Our results suggest that, while including interaction terms marginally increases explanatory power, these interactions do not significantly alter the influence of the LFS skill intensity indicators. Second, we examine how job complexity, defined by the number of different skill types required in a job, affects wages and interacts with the main skill used in the job, as indicated by the LFS indicators. This analysis shows that job complexity has a U-shaped relationship with wages but does not significantly affect the explanatory power of LFS skill indicators, suggesting that LFS measures capture aspects of skill importance not highly correlated with job complexity. Third, we assess several candidate skill intensity measures derived directly from the job postings data. Again, these measures, when compared to the LFS indicators, offer some explanatory power but do not fully capture the variation explained by the LFS intensity measures. The findings indicate that the LFS skill intensity measures capture aspects of skill importance that are not only reflected in the frequency of skill mentions

[2] These data, collected from all EU member states, 4 candidate countries, and 3 countries of the European Free Trade Association, have the potential to serve as an important source of skills data.

[3] We list the most common keywords for each skill category in Tables A.2 and A.3.

[4] We focus on task-specific skills, meaning the type of skills that are associated with specific tasks, such as social skills, cognitive skills, and computer skills.

[5] Almost three-quarters of all workers report using a main skill that falls within the skill categories often used in the job posting literature.

Table 1
Summary statistics of skill categories, estimation sample.

	(1) JP population matched	(2) LFS	(3) JP-LFS Matched	(4)
	Derived from JP	Derived from LFS	Derived from JP	Derived from LFS
Panel (A)				
Skill category				
Cognitive	0.56	0.05	0.61	0.06
Social	0.90	0.05	0.90	0.05
Management	0.70	0.23	0.74	0.22
Financial	0.57	0.12	0.63	0.10
Computer, General	0.55	0.06	0.59	0.06
Computer, Specific	0.35	0.04	0.39	0.05
Writing/Language	0.67	0.03	0.71	0.02
Customer service	0.88	0.31	0.87	0.34
Character	0.97	0.04	0.97	0.05
Panel (B)				
At least 1 skill in a category	0.99	0.69	0.99	0.72
Conditional on at least 1 skill in a category,fraction with skills falling across:				
Only one skill category	0.03	0.32	0.03	0.72
Two different skill categories	0.04	0.49	0.03	0.23
Three or more different skill categories	0.93	0.19	0.94	0.05
Observations	**499,645**	**13,138**	**2750**	

Notes: In Column 1, we consider employees in the employment register data that can be matched to the JP data. We report the fraction of employees that are in jobs requiring each of the 9 skills as captured by the corresponding job post(s) for that job. Column 2 reports the same fractions but where skills are observed in the Labour Force Survey by employees that can also be matched to the employment register data. In both Columns 1 and 2, we match to register data covering employees in the first year of employment spells. In Column 3, we report the fraction of workers in our estimation sample that are in jobs that require each of the 9 skills as captured by the job posting for that job. In Column 4, among the same group of workers, we report the fraction who report that they use a main skill in each of the 9 categories. See Appendix Table A.1 for more details on the estimation sample.

in job postings, highlighting the complementary nature of job posting and LFS data to understand wage determinants. Based on this analysis, we recommend that researchers using only job postings data should consider including both intensive and extensive measures of skill along with a job complexity measure; in our case, such a model yields the most precise estimates.

While much of the literature has focused on the returns to skill demand on the extensive margin (e.g., Deming and Kahn, 2018), few studies have tried to quantify the intensive margin of skill demand, and fewer still have looked at both the extensive and intensive margin of skills, especially when the intensive margin is evaluated using self-reported skill use measures. We believe our ability to study both the intensive and extensive margin of task-specific skills at the job level is new to the literature.

The rest of the paper is organized as follows. We provide details on the data in Section 2, describe and discuss our results in Section 3, robustness exercises in Section 4, and extensions in Section 5. Finally, we conclude in Section 6.

2. Data

2.1. Danish job postings

The online job postings data (JP) from 2007–2017 are supplied by the consultancy firm HBS Economics (HBS) and cover the near universe of publicly accessible online job postings in Denmark. The Danish JP are generally analogous to the equivalent US data supplied by Burning Glass Technologies.[6] However, relative to the US, Denmark has a large public sector that is legally obligated to post all jobs online.[7] To facilitate comparison with the US job postings literature, we consider only positions advertised by private firms, and we concentrate on job posts for occupations that are well represented in both the job posts and the LFS data: professionals, technicians and associate professionals, clerical workers, and service and sales workers.[8]

The JP include keywords from each job post, a posting date, and an occupational code. In addition, the JP contain a firm identifier that allows us to match the data with Danish employer-employee matched registers provided by Statistics Denmark. By matching a job post to a firm, we can further match a specific job post with employees who recently started working within the same firm and within the corresponding occupation.

Like Deming and Kahn (2018), we focus on the extensive margin in the JP: indicators are created at the job-posting level capturing whether or not a posting contains a keyword in a particular skill category. In Panel A of Table 1, Column 1 presents the fraction of job posts with each of the categorized task-specific skills from all of the private-sector JP from the period.[9] Cognitive skills are one of the least-occurring skills, included in 56% of the job posts, whereas 88%–97% of the job posts include customer service, social, or character skills.

2.2. Labour force survey text data

Since the 1980s, EU member states have administered the Labour Force Survey based on common survey guidelines to enable cross-country comparisons. About 1.5 million people were surveyed quarterly in 2018.[10] One of the variables in the EU LFS is an occupational code, based on the ISCO standards. To classify individuals into the correct occupation, the national statistical institutes collect information on job titles, and more importantly, job tasks. Specifically, in the Danish LFS, respondents are asked to *"Describe the specific main tasks in your job"* to

[6] We have purchased the data through the Danish consultancy firm, HBS Economics. See Appendix A for more details.

[7] Approximately 30% of workers are in the public sector.

[8] We exclude blue collar occupations as in Deming and Kahn (2018). We also exclude managers (ISCO occupation 1) because these jobs are not well represented in the LFS data. We include 1-digit ISCO occupation codes of 2–5.

[9] This refers to the sample of job posts that can be matched to the population-level employment register data using the procedure described in Section 2.3.

[10] The EU LFS is found here: https://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey.

which they can provide open-ended responses.[11] Although the national statistical institutes collect and process data on job titles and job tasks, these free-text data tend not to be available to researchers. Uniquely, we have access to the Danish text data from the LFS from 2007–2017. Importantly, Statistics Denmark supplies this data with personal identifiers so it can be easily linked to Danish register data, allowing us to determine the firm for which a surveyed individual works, the wage they earn, and many other worker–firm characteristics that are not available in the LFS.

In order to compare how skills from the JP reflect on-the-job skill use as captured by the LFS, we consider workers in the LFS who started their job within the last year, as skill use may change with tenure. Column 2 of Table 1, Panel A, presents summary statistics of the categorized task-specific skills from the LFS data.[12] The lower incidence of task-specific skills is immediately clear — a consequence of the LFS capturing a measure of main skill(s) only. In contrast, the measures of task-specific skills derived from the job postings data demonstrate that employers almost always list at least one "character" and one "social" keyword.

2.3. Linking the JP and LFS

Each job posting in the JP is matched with one or more individuals in the registers, following the matching procedure described by Jensen (2024) and Daly et al. (2022). We identify individuals in the employer-employee linked register data who have recently started a new job (either in a new occupation or in a new firm). We match individuals to a job post in the same firm-occupation cell if the job was posted in the month in which they started their new job or a maximum of four months prior.[13] We call such a match a pseudo-individual match. This means that a person starting in a given firm-occupation cell can also be linked to multiple job posts if more than one job post has been posted in the same firm-occupation cell within the five-month window. If a person is matched to more than one job post, the indicator equals 1 if a skill is mentioned in any of the job posts.

We are able to match about one-fifth of all new jobs recorded in the Danish register data to job posts in the relevant occupation, reflecting the fact that many private companies do not advertise all new jobs, especially new jobs resulting from occupational changes within a firm. Our final sample links individuals with the firm and job posting to which they responded, and includes individual responses to the LFS. From the job postings, we derive skill indicators associated with each job match, and from the register data, we obtain other characteristics of the job during the respondent's first year in the job (e.g., wages and hours worked).

To understand the representativeness of the JP and LFS data, we report various summary statistics in Appendix Table A.1. We find that age and location vary slightly between the samples, so we control for these factors in our empirical specification. We find that the matched JP-LFS (henceforth the Estimation Sample) is representative of the overall population of new jobs, of the JP-population matched data, and of the full set of new jobs identified in the LFS. As we would expect given that the LFS is a representative sample of Danish workers, the match rate between the full population of new jobs to the JP is similar to the match rate between the subset of new jobs that are sampled in the LFS and the JP. In Table 1, Columns 3 and 4, we consider employees in the matched JP-LFS sample and their corresponding skills derived from the JP data and the Labour Force Survey, respectively. See Appendix A for more details on the data.

2.4. Interpretation of task-specific skill measures

It is important to recognize how the measures of task-specific skills differ depending on whether they are derived from the LFS or from the job postings data. To fix ideas, we consider a simplified setting in which some workers are hired into jobs that demand a particular skill and, for a subset of these workers, that skill is intensively used. Let s^{JP} and s^{LFS} be the observed skill measures captured by job postings data and the LFS, respectively:

$$s^{JP} = \mathbb{1}(s^D + e^{JP} > a)$$
$$s^{LFS} = \mathbb{1}(s^S + e^{LFS} > b)$$

where s^D corresponds to the true unobserved employer skill demand, s^S corresponds to the true unobserved skill supply of workers, and e^{JP} and e^{LFS} correspond to measurement error in both of these observed variables, respectively. This formulation highlights that there are generally three ways a potential misalignment of these two skill measures could occur.

First, there may be labor market frictions generating mismatch between skill supply and demand: firms may not be able to hire workers possessing the skills they demand, and instead hire workers with a different skill profile, i.e., $s^S \neq s^D$. Second, it is likely that $a < b$: workers are specifically asked to report only their main skills; in contrast, employers may list skills that are desirable but not necessary in addition to those that are necessary, given their small cost of doing so. Third, even in the case that there are no labor market frictions ($s^S = s^D = s^*$, where s^* is the skill level in equilibrium), and if the skill indicators capture the same intensity of the underlying skill ($a = b$), measurement error in either or both of the observed skill measures could lead to a misalignment between s^{JP} and s^{LFS}. Measurement error could arise on the demand side because certain skill categories are so fundamentally part of a job that explicitly stating the requirement may be unnecessary. On the supply side, measurement error could come from how a worker perceives a main task, or from the distinction between tasks and skills. Although this distinction is not often emphasized in the literature, in our context it may be more of an issue. For example, an employer may list "detail-oriented" as a skill required for the job, but when asked about skills use (tasks) on the job, an employee would likely not list "detail-oriented".

In the absence of mismatch and measurement error, we can write our skill measures as:

$$s^{JP} = \mathbb{1}(s^* > a) = s$$
$$s^{LFS} = \mathbb{1}(s^* > b) = s \cdot m$$

where s is an indicator equal to 1 if a job requires a skill of at least level a, and m is an indicator taking the value of 1 if that skill is particularly important, above level b. If we believe a is close to 0, we can interpret s as the extensive margin of the skill and m as a measure of skill intensity, conditional on the worker having a positive amount of the skill. In this case, the expected value of m corresponds to the fraction of workers

[11] In Danish, the specific questions on the survey include: (1) B2Stil: *"Hvad er din stillingsbetegnelse/titel?"* and (2) B2StilA: *"Beskriv de konkrete hoved-arbejdsopgaver i din stilling"*. The question included in the German Mikrozensus 2021 is: *"Please describe your current work in keywords"*. The question included in the UK LFS is: *"What did you mainly do in your job?"*, and in the Swedish LFS: *"What are your main tasks?"* See national LFS questionnaires by year here: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_labour_force_survey_-_documentation.

[12] This refers to the sample of respondents in the Labour Force Survey who can be matched to the population-level employment register data using the procedure described in Section 2.3.

[13] To match job posts to new job spells, we use detailed occupational codes at a level between 3- and 4-digit ISCO-codes. Statistics Denmark transitions from ISCO88-codes and adopt ISCO08-codes in 2010. We convert ISCO88/ISCO08-codes into 228 time-consistent occupational groups, which gives a level of detail between 3- and 4-digit ISCO-codes (see Appendix B.3.1 and Online Appendix D.2 in Jensen, 2024 for details on occupational codes and the conversion to time-consistent codes).

hired in jobs requiring skill s who use that skill intensively. If we regress s^{LFS} on s^{JP}, the slope should recover this fraction.[14]

The availability of a measure of on-the-job skill intensity allows us to investigate how sensitive wage regressions are to using just extensive margins defined from skills listed in job postings rather than, or in addition to, regressions that also include a measure of skill intensity. In this simplified environment, we can write the following population model:

$$w = \beta_0 + \beta_1 s + \beta_2 (s \cdot m) + u \tag{1}$$

where w is log wages, and under our simplifying assumptions, u captures all other unobserved determinants of wage such that $E(u \mid s, m) = 0$. β_1 is the return to the extensive margin of skill s, and β_2 is the additional return for using that skill intensively. Note that in this framework, there is no need to include m separately in the regression as individuals only intensively use the skill if they also use the skill at the extensive margin.

Deming and Kahn (2018) estimate a multi-skill version of Eq. (1) that implicitly restricts $\beta_2 = 0$. Our data uniquely enables us to also estimate the effect of intense skill use. By incorporating a measure of skill intensity, the interpretation of the our wage equation comes closer to a standard Mincerian-style wage equation in which increases in task-specific human capital skills yield additional wage returns (Mincer, 1974). As we cannot measure skill use continuously, β_2 measures the additional return awarded to those who not only posses the skill, but also intensively use it. We can interpret β_1 as the average return to a job that requires skill s, but does not require intense use of the skill, relative to jobs that do not require the skill at all.

The signs of β_1 and β_2 can vary. When both coefficients are positive, the returns of the skill are monotonic, in line with traditional models of human capital, where an increasing skill level leads to a higher wage. If $\beta_1 < 0$ and $\beta_2 < 0$, workers with a certain skill experience lower wages than workers without the skill, and for workers using the skill intensively, wages are even lower. One way we can understand this type of negative return is by acknowledging the presence of multiple skills: some skills are valued less than others, and as such investing in those skills generates a wage penalty. In both of these scenarios, we have monotonic returns to skills.

It may also be the case that the returns to skills are non-monotonic. Deming (2022) points out that not all skills can be interpreted as human capital where more of a skill is always better. He illustrates this with the skill conscientiousness. Conscientiousness positively predicts earnings, yet the opposite of conscientiousness (e.g., disruptive and aggressive behavior) has been shown to predict entrepreneurial success (Levine and Rubinstein, 2017; Papageorge et al., 2019). Our approach, including a flexible form of intensive skill use, allows us to show potential non-monotonicities in the returns to skills. If $\beta_1 > 0$ and $\beta_2 < 0$, there are positive returns to having a job that requires a given skill (at the extensive margin), but using the skill intensively reduces these otherwise positive returns. This type of skill has a satiation point — an optimal amount of how much a worker should use a skill to maximize returns. If $\beta_1 < 0$ and $\beta_2 > 0$, the return to intensively using a skill is positive, but only relative to an already negative return at the extensive margin. The scenario reflects that specializing or intensely using a skill may yield positive returns in terms of wages, but less specialized use of a skill having the opposite effect.

Finally, if the magnitude of the extensive margin return is significant, and that of the intensive margin return is close to zero, we may suspect that the skill is a certificate skill and/or has a type of signaling component following Spence (1978). In this case, intensely using a skill would yield no additional wage returns.

[14] More generally, $E(s^{JP}) = Pr(s^* > a)$ and $E(s^{LFS}) = Pr(s^* > b|s^* > a) * Pr(s^* > a)$. We thank an anonymous referee for this suggestion.

To understand the relationship between wages, indicators of advertised skills, and main skills empirically, we first estimate Eq. (1) separately for each skill s_k with $k = 1 \ldots 9$ using the following regression model:

$$wage_i = \gamma_0^1 + \gamma_k^1 s_{ki}^{JP} + \theta_k^1 s_{ki}^{LFS} + x_i \delta + \epsilon_i \tag{2}$$

where $wage_i$ is the natural logarithm of the hourly wage of worker i. γ_k^1 and θ_k^1 can respectively be interpreted as β_1 and β_2 from Eq. (1) for skill k. Finally, x_i always contains year and municipality indicators. We also estimate specifications in which individual controls are included: age, age^2, experience, experience^2, years of education fixed effects, and a gender dummy.[15]

We then proceed to estimate a more general form of Eq. (1) that allows each of the 9 possible main skills to be correlated with multiple skills advertised in the JP with the specification:

$$wage_i = \gamma_0 + \sum_{k=1}^{9} (\gamma_k s_{ki}^{JP} + \theta_k s_{ki}^{LFS}) + x_i \delta + \epsilon_i \tag{3}$$

where $wage_i$, γ_k, and θ_k can be interpreted as above. γ_0 captures average earnings for those who work in jobs that require skills not captured by our categorization and who mainly use skills not captured by our categorization.[16] Specifically, γ_k captures the average wage premium (or penalty) of those who work in jobs that advertised for skill k, conditional on using the other $k - 1$ skills. In the same way, the θ_k captures the average wage premium (or penalty) of those who mainly use skill k. The term x_i contains the same controls as in Eq. (2).

In the data, there are individuals who use a skill intensively and work in a job that did not advertise for that skill ($s^{JP} = 0$ and $s^{LFS} = 1$). We conjecture that if this misalignment is mainly being driven mismatch, we should see evidence of negative effects of this misalignment on wages and on tenure as firms find workers with whom they are a better match. We explore this possibility further in the results section and find no evidence of either. Thus, we conclude that the relatively small misalignment is due to measurement error.

3. Results

3.1. Validation of skill measures

Column 3 of Panel B, Table 1, reports the frequency of having only one, two, and three different types of skills mentioned in a job post in the JP-LFS matched sample. The same frequencies, but now referring to the main skills reported in the LFS, are shown in Column 4. Almost all job posts contain at least one skill that falls within the categories we consider. On the other hand, in the LFS, almost three-quarters of the workers report at least one main skill that fall within one of our 9 skill categories. As workers are free to report whatever they consider their main skill used a certain degree of measurement error is expected. For example, a phlebotomist may report that their main skill is to "take blood samples", something not captured by our skills measures. Upon closer inspection, this type of measurement error appears to be

[15] In addition, one could consider controlling for occupation fixed effects, but such controls would change the interpretation of returns to skills (the correlation between the skill indicator and wage) to how much a person uses the skill relative to the occupation they are in, rather than whether they use a skill and how much they use the skill. Jensen (2024) includes a thorough analysis with inclusion of occupation fixed effects in the larger JP sample and finds that the returns to the nine categories of skills decrease after controlling for occupation fixed effects. We believe both measures are interesting, but refer to Jensen (2024) for a more in-depth discussion of wage returns using the larger JP sample only, making it possible to condition on detailed occupational fixed effects as well as on firm fixed effects.

[16] We follow Deming and Kahn (2018) in our skill categorization. See Section 2 and Appendix A for further discussion of our skill categorization.

Table 2
Conditional probabilities.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Conditional on LFS (main) skill, probability that job post has the same skill											
	Cognitive	Social	Manage-ment	Financial	Computer, General	Computer, Specific	Writing/ Language	Customer service	Character	Avg. match	Any match
Mean	0.849	0.915	0.848	0.853	0.872	0.683	0.808	0.932	0.976	0.887	0.916
SD	(0.359)	(0.280)	(0.360)	(0.355)	(0.335)	(0.467)	(0.398)	(0.251)	(0.153)	(0.293)	(0.277)
N	166	141	617	278	172	142	52	948	126	1968	1968
Panel B: Conditional on job post (required) skill, probability that LFS has the same skill											
	Cognitive	Social	Manage-ment	Financial	Computer, General	Computer, Specific	Writing/ Language	Customer service	Character	Avg. match	Any match
Mean	0.085	0.052	0.258	0.137	0.093	0.090	0.022	0.371	0.046	0.140	0.663
SD	(0.278)	(0.222)	(0.437)	(0.344)	(0.291)	(0.287)	(0.146)	(0.483)	(0.210)	(0.141)	(0.473)
N	1668	2488	2031	1734	1613	1073	1942	2383	2658	2720	2720

Notes: Panel A presents the probability that a job posting requires a particular skill, conditional on the LFS respondent reporting that task-specific skill. Panel B presents the opposite conditional probability of Panel A; the probability that a LFS respondent is listing that task-specific skill conditional on the job posting requires the skill. The probabilities in Panel B are smaller than Panel A because most workers only report using one main skill. The probabilities in Panel B are, however, larger than the unconditional probabilities in Table 1, Column 2, Panel A, which confirms the positive correlation between individual skill measures in the JP and LFS data. The Estimation Sample, $n = 2750$, is used to calculate these probabilities — see Appendix Table A.1 for more details on the sample.

driving the issue: a frequent, non-categorized reported main skills is, for example, "cash register" (translated to English).[17]

Considering the matched JP-LFS sample in Columns 3 and 4 of Table 1, out of the workers reporting at least one main skill in the LFS, 72% of them report only one of the of nine skills we consider.[18] About 23% report skills that fall across two main skill categories, and 5% of workers report skills that fall across three or more skill categories. In job posts, employers tend to mention skills that fall across more skill categories. 94% of employees work in jobs for which three or more skills were listed in the corresponding job post. On average, employers demand skills that fall across six different skill categories. Because employees are free to list many main skills in the LFS, a possible interpretation of these facts is that workers are more specialized than what the skill indicators derived from job postings may suggest. In other words, the skill intensities within each of the employer-demanded skill categories may vary substantially. The job posting skills then capture the extensive margin of skills used on the job, and the LFS main skills capture the intensity of skills used on the job.

Next, we verify that if an individual lists a particular main skill in the LFS survey, then the job for which they are hired also requires that task-specific skill. Table 2, Panel A, presents the probability that a job posting requires a particular skill, conditional on the LFS respondent reporting that task-specific skill. For instance, in Column 2, we see that almost 85% of those LFS respondents who stated that one of their main skills was cognitive skills hold a job that advertised for cognitive skills. In Column 11 we see that 91.6% of jobs with one or more LFS skills have at least one of the skill represented with a JP skill. For 88.7% of the jobs with an LFS, the job has all the same JP skills (see Column 10). Regardless of the skill category, we can clearly see that for the vast majority of individuals, the main skill they use on the job corresponds, at the individual level, to a skill listed in the job posting to which they applied. This is our first piece of evidence that skill measures derived from job posts in fact capture skills used on the job.[19] Yet, as these conditional probabilities are not 1, there are employees reporting in the LFS that they use a main skill which their employer did not explicitly include in the job posting. Pooling across LFS main skill types, we find that this is true for 10% of workers.

Table 2, Panel B, presents the reverse conditional probability: the probability that an LFS respondent lists a task-specific skill conditional on the job post requiring that skill. For instance, about 8.5% of those who have a job that sought cognitive skills primarily use cognitive skills on the job. On the other hand, almost 40% of those who work in a job that advertised for customer service skills state that they mainly handle customer service. Column 10 shows that among those with a particular JP skill, on average 14% also have the corresponding LFS skill. Comparing these conditional probabilities to the unconditional probabilities in Table 1, Column 4, we see that the former is greater than the latter for all skills but character skills. This is our second piece of evidence that measures of skill demand derived from job postings reflect skills used on the job.

Table 3 continues this exercise by presenting the results of regressing an indicator of the LFS task-specific main skill measures on the skill indicators derived from the job postings data. If LFS respondents had been asked to report all of the skills used on the job (as opposed to just main skills), we would expect that coefficients along the diagonal would be close to 1, in the absence of measurement error and mismatch. However, given that the LFS captures an indicator for (only) the most important skills, this need not be the case.[20] For instance, we learn from Column 1, Table 3, that the probability that cognitive skills will be the main skill reported by a worker in the LFS increases by 4.5 percentage points (significantly different from 0) when cognitive skills are mentioned in the corresponding job posting, holding other extensive skill requirements constant. As all coefficients along the diagonal are positive, except for character skills, we take this as our

[17] The Danish word for "cash register" is "kasse" which can mean both "cash register" and "box", so this word cannot be categorized as a skill without context; not all skills are evident from individual keywords.

[18] Note that this is conditional on having at least one keyword, i.e. having responding to the survey.

[19] Another approach would be to calculate a "most important" skill from the JP and use it for comparison purposes. The issue with this approach is how exactly to construct such a variable. In the LFS data, the importance is implied in the question. In the JP, one might be tempted to assume that word frequency corresponds to importance; however, if this were true, the most important skill across occupations would be character — not necessarily because this skill is relatively more important, but because this skill category includes relatively more words. We explore such measures in Section 5.

[20] Table 3 is a correlation of the two skill measures that can give different results at the individual and occupational level. For instance, at the individual level, the correlation between skill measures would be low if few workers have a given skill as their main skill while the skill is relatively common in the JP. However, at the occupational level, the correlation between the same skill measures can be high if the occupational fractions of the few workers with the skill from the LFS data correlate across occupations with the higher occupational fraction of workers with the skill in the JP. One example of this could be if a main skill (from LFS) was used in two occupations with the probability of 0.01 and 0.02. If this same skill was used at the extensive margin (from the JP) in the same two occupations with probability 0.4 and 0.8, then the occupational correlation would be 1. However, the individual-level correlation could at most be 0.025, but it could also be zero or negative.

Table 3
Regression of LFS (main) skill on job posting skills.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Skill present on	*LFS (main) skill*								
job posting:	Cognitive	Social	Management	Financial	Computer, General	Computer, Specific	Writing/ Language	Customer service	Character
Cognitive	0.045***	0.026***	0.018	−0.004	0.034***	0.023**	0.004	−0.096***	−0.006
	(0.011)	(0.010)	(0.021)	(0.016)	(0.012)	(0.012)	(0.008)	(0.029)	(0.011)
Social	−0.009	0.024	−0.038	−0.052*	−0.025	0.001	−0.007	0.068	0.014
	(0.020)	(0.019)	(0.034)	(0.032)	(0.017)	(0.015)	(0.015)	(0.045)	(0.016)
Management	0.029**	0.015	0.079***	0.008	0.001	0.019*	−0.008	−0.019	0.019
	(0.012)	(0.013)	(0.023)	(0.020)	(0.012)	(0.011)	(0.009)	(0.034)	(0.018)
Financial	−0.001	−0.030***	0.010	0.131***	−0.010	−0.026**	−0.003	−0.059**	0.005
	(0.013)	(0.012)	(0.020)	(0.018)	(0.012)	(0.011)	(0.007)	(0.026)	(0.013)
Computer, General	0.018	−0.003	0.028	−0.055**	0.046***	0.025**	0.013*	−0.067**	0.026**
	(0.011)	(0.012)	(0.026)	(0.023)	(0.013)	(0.011)	(0.008)	(0.034)	(0.013)
Computer, Specific	0.017	−0.009	0.062***	−0.008	0.057***	0.050***	−0.015***	−0.060*	−0.015
	(0.020)	(0.010)	(0.023)	(0.019)	(0.015)	(0.014)	(0.006)	(0.033)	(0.010)
Writing/Language	0.013	−0.003	0.034	−0.016	−0.003	−0.006	0.015**	−0.056**	−0.007
	(0.009)	(0.011)	(0.021)	(0.018)	(0.011)	(0.012)	(0.007)	(0.028)	(0.012)
Customer service	−0.062***	−0.008	0.027	−0.034*	−0.014	−0.015	−0.023**	0.285***	0.005
	(0.023)	(0.015)	(0.027)	(0.020)	(0.016)	(0.016)	(0.011)	(0.031)	(0.014)
Character	−0.010	−0.042	−0.061	0.073*	−0.009	0.011	0.013	−0.018	-0.016
	(0.039)	(0.043)	(0.054)	(0.038)	(0.030)	(0.019)	(0.021)	(0.074)	(0.024)
Observations	2750	2750	2750	2750	2750	2750	2750	2750	2750
R-squared	0.029	0.007	0.031	0.039	0.039	0.028	0.009	0.072	0.006
Clusters	893	893	893	893	893	893	893	893	893

Notes: Columns 1 to 9 present the results of regressing each of the LFS task-specific main skill measures on job posting skills categories. The Estimation Sample is used to calculate these probabilities — see Appendix Table A.1 for more details on the sample. Note that the skill categories are not mutually exclusive; an individual can have more than one skill. Standard errors, in parentheses, clustered at the firm level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

third piece of evidence that skills advertised by employers are actually being used on the job.

Table 3 also illustrates how skill bundles sought by employers vary according to the main skills performed on the job. In Column 1, we see that workers who mainly use cognitive skills are more likely to hold a job that advertised for management skills, but less likely to be in a job that advertised for customer service skills. From Column 2, we see that workers who intensively use social skills are significantly more likely (about 2.6%) to be working for an employer who stated that their employees should possess cognitive skills and significantly less likely (3%) to be working for an employer seeking employees with financial skills.[21]

Next, we seek to better understand why 10.3% of employees report using a main skill not advertised by their employer in their corresponding job post.[22] We hypothesize that if there is skill mismatch, this misalignment would likely imply a lower match quality: such matches would have shorter durations and/or employees would receive lower wages.[23] On the other hand, if workers correctly report the skills they intensively use on the job, but employers implicitly, rather than explicitly, state that a skill is demanded in the job post, we expect no negative effects on match duration or wages.

Table 4 presents the results of regressing various measures of match quality on an indicator of skill misalignment: whether or not a worker is mainly using a skill type on the job that was not advertised by their employer. We look at the effect of this misalignment on the probability of the match lasting more than 1, 2 and 3 years in Columns 1, 2 and 3 respectively and find small, statistically insignificant effects. In Column 4, we regress the natural logarithm of average wage over the first year of a job spell on the same indicator of skill misalignment and again find no significant difference. We conclude from this exercise that there is no evidence of skill mismatch among main skills, but rather that skills extracted from job posts involve some measurement error.[24]

To better understand why some individuals report intensively using a skill that was not advertised in the job posting, we compare the relative frequencies of other JP skills required for that job. Appendix Table B.2 presents this comparison. Each row corresponds to a sample that is conditional on having a particular LFS skill, and further on whether or not the JP data match for a particular skill (match) or do not agree (not matched). For example, the first two rows correspond to the subsamples in which individuals have stated that cognitive is their main skill. The first row, "match", is also conditional on being in a job where cognitive skills were listed in the JP. The second row, "not matched", conditions on a worker stating that cognitive is their main skill in the LFS data, but also working in a job that did *not* advertise for a cognitive skill.

Columns 1–9 show the relative frequency of the JP skill indicated in the column header. Column 10 presents the average number of skill categories in each subsample. The most frequent JP skills in each subsample are shaded red (including ties), the second most frequent are shaded orange (including ties), and the third most frequent are shaded yellow (including ties).

When comparing the shaded cells of the "not matched" to the "match" for each skill, we generally observe that the ranking of the

[21] By simulating an upper bound of the regression coefficients if all LFS skills were a subset of JP skills, i.e., with no measurement error or mismatch, we have a comparison for the size of the regression coefficients. We provide these simulated correlations in Appendix Table B.1. Compared to the simulated upper bound, we see that the unconditional regression coefficients represent around 45% of the maximum unconditional correlation.

[22] We match individuals to a job post in the same firm-occupation cell if the job was posted in the month in which they started their new job or a maximum of four months prior. We check if the share of employees who report using a main skill not advertised by their employer depend on the length of this matching window. The share varies from 11.1% (3 months matching window), 10.3% (4 months matching window), to 9.6% (8 months matching window). See Section 4 for results from Table 3 with different matching windows.

[23] If the choice to report a skill as a main skill is driven only by the desire to be perceived in a particular way and is not correlated at all to an actual deprioritization of the true main skill, then we would not expect to see tenure or wage effects.

[24] We might be worried that attenuation due to measurement error is masking evidence of negative wage and tenure effects. But given the significant wage effects we find in Table 6 that are generally of the same magnitudes reported in the literature, we do not believe this is the case.

Table 4
Effects of skill discrepancy on match duration and wages.

	(1)	(2)	(3)	(4)
	Indicator for length of time in job:			Average wage in job
	1(years ≥ 1)	1(years ≥ 2)	1(years ≥ 3)	
	Panel A: Without individual controls			
Indicator for skill discrepancy	0.014	0.027	−0.002	0.010
	(0.033)	(0.037)	(0.039)	(0.031)
R-squared	0.062	0.079	0.076	0.205
	Panel B: With individual controls			
Indicator for skill discrepancy	0.016	0.027	−0.001	0.000
	(0.032)	(0.036)	(0.037)	(0.017)
R-squared	0.088	0.114	0.114	0.601
Observations	2733	2474	2148	2750
Year Indicators	YES	YES	YES	YES
Municipality Indicators	YES	YES	YES	YES
Individual controls	YES	YES	YES	YES
Clusters	887	809	710	893

Notes: Columns 1, 2 and 3 show the estimates of regressing an indicator of whether or not the worker–firm-occupation match last for at least 1, 2 and 3 years respectively on an indicator of skill discrepancy. The skill discrepancy indicator takes the value of 1 if a worker reports mainly using a skill category that the firm did not include in their job post. Panel A presents the estimates of these regressions when just year and municipality fixed effects are included in the regressions. Panel B presents the results if individual controls are included: age, age^2, experience, experience^2, years of education fixed effects, and a gender dummy. Note that for each of the regressions shown in Columns 1–3, we require that the individual started the job at least 1, 2 and 3 years prior to the end of our sample, respectively. This is why the sample sizes are smaller than the Estimation Sample, $n = 2750$. In Column 4, we regress the indicator of skill discrepancy on average wages during the job spell. We consider the natural logarithm of average wages within the first year of a job spell. All standard errors are clustered at the firm level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

most frequently used skills on the job is remarkably consistent, i.e., that the importance of skills (i.e., rank ordering) in jobs requiring a particular main skill are very similar, regardless of whether they are categorized as not matched or matched for the skill in question. Given that the relative frequencies of the skills in the skill bundles are so consistent, the omission of the JP skill corresponding to the intensely used skill appears to be indicative of measurement error.

Additionally, the last column shows that the not matched subsamples have, on average, a lower number of skill categories (and, not shown, a lower number of keywords), suggesting that the mismatch may be due to measurement error arising from the brevity of some employers, e.g., some employers may not list a general computer skill if they are listing specific computer skills as it may be considered implied.[25][26]

3.2. Skills and wages

Given that we find no evidence of mismatch, we move on to further understand skill premia or penalties by considering both the extensive and intensive skill margins. We first present results from estimating Eq. (2) for each skill separately, focusing on how the inclusion of the intensive skill measures from the LFS affects the associated returns to the extensive skill measures from the JP. Next, we present the results from estimating Eq. (3) in which all 9 skills are included together in order to allow for correlations between the skill measures. We consider this our main specification, from which we interpret the associated skill returns on the extensive and intensive margin.

The odd Columns 1, 3, …, 17 of Table 5 show the return to each of the 9 skills from the JP, when we restrict θ_k^1 to zero. The even Columns 2, 4, …, 18 show the resulting estimates for the unrestricted model, i.e., when we allow for nonzero effects of intensively using a skill. Table 5 includes controls for year and municipality fixed effects, while Appendix Table B.3 presents results that also include individual controls.

Column 1 of Table 5 shows the return to cognitive skills is 0.145 and significantly different from zero, such that having a job that requires any cognitive skill, on average, is associated with an increase in a worker's wage by 14.5% relative to jobs without cognitive skills. In Column 2, we further include an indicator for intensive use of cognitive skill, which is equal to one if the worker lists a cognitive skill as a main skill in the LFS. Adding an intensive margin of cognitive skill use only decreases returns at the extensive margin by a small amount, from 14.5% to 14.2%. The coefficient on the intensive use of cognitive skills is 0.068, which means that workers who use cognitive skills at both the intensive and extensive margins have an average of $14.2 + 6.8 = 21.0\%$ higher wages than workers not using cognitive skills.

The fact that the returns to using cognitive skills at the extensive margin is relatively unaffected by the inclusion of the intensive margin is not surprising given that there is a relatively low correlation between the use of cognitive skills at the extensive and intensive margins; only 8.5% of workers who use cognitive skills at the extensive margin also use them intensively (see Table 2). Moreover, the additional returns from intensively using cognitive skills are not exceptionally large. This, in conjunction with the fact that so few workers intensively use this skill, explains why the extensive margin return remains relatively unchanged, whether or not intensive skills measures are included.

We can compare the returns to cognitive skills in Columns 1 and 2 to the returns to management skills in Columns 3 and 4, as well as to the returns to financial skills in Columns 5 an 6. Both the management and financial skills categories have higher correlations between their extensive and intensive margins. The returns to using management skills at the extensive margin is 10.5% when we set $\theta_k = 0$ and decreases to 8.4% when we also include a term indicating intensive use of management skills. The returns on the extensive margin of management skills decreases when we include the intensive margin, both because the two margins are relatively highly correlated, but also because the additional returns to using management skills intensively is sizable at 23.0%. The overall wage increase for workers using management skills intensively is $8.4 + 23.0 = 31.4\%$ relative to workers who do not use any management skills. The returns to financial skills at the extensive margin do not change when we include a term indicating intensive use of financial skills because the additional returns to using financial skills

[25] Interestingly, the share of men is economically and statistically significantly larger in the not matched group.

[26] Job posts that have at least one match LFS-JP match have 52.6% more keywords (including noise words) than those that have at least one "not matched" skill (LFS skill not included in JP). This suggests measurement error simply because of shorter job posts.

Table 5
Wage regressions, separately for each skill.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Cognitive	0.145***	0.142***																
	(0.024)	(0.025)																
Cognitive as main skill		0.068*																
		(0.037)																
Social			0.034	0.034														
			(0.034)	(0.034)														
Social as main skill				0.104***														
				(0.034)														
Management					0.105***	0.084**												
					(0.035)	(0.035)												
Management as main skill						0.231***												
						(0.019)												
Financial							0.078***	0.078***										
							(0.027)	(0.028)										
Financial as main skill								0.003										
								(0.043)										
Computer, General									0.092**	0.081*								
									(0.043)	(0.043)								
Computer, General as main skill										0.167***								
										(0.042)								
Computer, Specific											0.101**	0.092**						
											(0.043)	(0.043)						
Computer, Specific as main skill												0.142***						
												(0.040)						
Writing/Language													0.069**	0.069**				
													(0.033)	(0.034)				
Writing/Language as main skill														−0.112*				
														(0.058)				
Customer service															−0.046	−0.011		
															(0.031)	(0.029)		
Customer service as main skill																−0.178***		
																(0.021)		
Character																	−0.061	−0.062
																	(0.047)	(0.045)
Character as main skill																		0.202***
																		(0.032)
Constant	5.053***	5.051***	5.111***	5.105***	5.065***	5.023***	5.093***	5.093***	5.087***	5.073***	5.105***	5.094***	5.095***	5.100***	5.179***	5.194***	5.197***	5.189***
	(0.035)	(0.035)	(0.037)	(0.037)	(0.034)	(0.034)	(0.033)	(0.033)	(0.034)	(0.033)	(0.033)	(0.032)	(0.034)	(0.034)	(0.039)	(0.037)	(0.047)	(0.044)
Observations	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750
R-squared	0.233	0.235	0.206	0.209	0.218	0.273	0.214	0.214	0.217	0.227	0.219	0.225	0.211	0.212	0.207	0.249	0.206	0.217
Year Indicators	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Municipality Indicators	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Individual controls	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Clusters	893	893	893	893	893	893	893	893	893	893	893	893	893	893	893	893	893	893

Notes: The dependent variable is the natural logarithm of average wages within the first year of a job spell. Odd Columns 1, 3... 17 present the results from estimating Eq. (2) when restricting θ_k to zero, and even Columns 2, 4...18 present the results from estimating Eq. (2) when allowing θ_k to vary. Individual controls include: age, age^2, experience, experience^2, years of education fixed effects, and a gender dummy. Note that the skill categories are not mutually exclusive; an individual can have more than one skill. Standard errors, in parentheses, clustered at the firm level. All regressions include year and municipality fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

intensively is zero (with a point estimate of 0.003). Finally, we consider character skills in Columns 17 and 18. Adding a term for intensive use of character skills has almost no effect on the returns at the extensive margin, even though the return to using character skills intensively are high at 20%. This result is driven by the use of character skills at extensive and intensive margins being insignificantly (and negatively) correlated (see Table 3).

Next, we present the results from estimating Eq. (3). Columns 1 and 3 of Table 6 show results when restricting θ_k to zero, whereas Columns 2 and 4 show estimates from unrestricted models. The first two columns present the results when year and municipality indicators, but not individual controls, are included whereas Columns 3 and 4 show the results when individual controls are also included.[27]

From Columns 1 and 3 in Table 6, which list the estimated wage skill premia at the extensive margin, we see that cognitive and management skills are associated with positive and significant wage returns whereas character skills are associated with significantly negative returns. Specifically, Column 1 shows that cognitive and management skills, on average, are associated with 11.4% and 5.5% higher wages when we do not control for individual characteristics. In contrast, having character skills are associated with a 12.8% lower wage. After controlling for individual characteristics in Column 3, the estimated returns decrease in absolute value (but remains statistically significant) demonstrating the importance of controlling for selection when estimating skill premia.[28] If we assume that firms advertise only the skills that will be used on the job, these coefficients are the weighted average of those who use these skills intensively and those who do not. In this sense, these estimates correspond to what is often reported in the literature.

In Columns 2 and 4, we include the main skill indicators from the LFS. Generally speaking, the inclusion of the main skill indicators has relatively little effect on the estimated coefficients on the JP skill measures, although the coefficients' absolute values decrease in all cases. In addition, the inclusion of the main skill indicators also greatly increases the explanatory power of the model when no individual controls are included. However, much of the variation in wages explained by the intensity of skill use is absorbed once individual controls are included in the regression. This suggests that extensive margin skill returns can be more precisely estimated if intensive skill measures and individual controls are included; something that is rarely done in the existing literature using skill measures from job posts.

Finally, we highlight three ways in which the returns to extensive and intensive skills vary according to the skill considered. First, we see additional positive wage returns to cognitive and management skills when these skills are used intensively. This is consistent with the idea that cognitive and management skills are specific forms of human capital, where higher levels of these skills are associated with greater returns. When individual controls are not included, employees with jobs that advertised for, and mainly use, cognitive skills earn wages that are 17% higher compared to employees who work in jobs that did not advertise for cognitive skills and who did not report using cognitive skills as a main skill. Employees who work in jobs that advertise for cognitive skills, but do not include cognitive skill as a main skill, earn wages that are only 10% higher than those who work in jobs that do not advertise for or use cognitive skills. When individual controls are added to the regression, these returns more than halve to about 7% and 4%, respectively. We also see that individuals whose main skill is management receive higher wages, a wage premium at around

[27] The individual controls include: age, age^2, experience, experience^2, years of education fixed effects, and a gender dummy.

[28] An interesting exercise to perform in this context would be to add occupation indicators. Jensen (2024) is able to further explore the effects of occupation and firm fixed effects on the returns to task-specific skills due to his larger estimation sample.

Table 6
Wage regressions.

	(1)	(2)	(3)	(4)
Job posting skill indicators:				
Cognitive	0.114***	0.100***	0.041**	0.043***
	(0.026)	(0.023)	(0.017)	(0.017)
Social	−0.006	0.005	0.007	0.008
	(0.031)	(0.028)	(0.025)	(0.025)
Management	0.055*	0.03	0.032*	0.022
	(0.028)	(0.027)	(0.018)	(0.018)
Financial	0.028	0.025	0.011	0.012
	(0.024)	(0.023)	(0.017)	(0.017)
Computer, General	0.022	−0.001	0.006	−0.001
	(0.035)	(0.031)	(0.018)	(0.018)
Computer, Specific	0.055*	0.031	0.023	0.018
	(0.032)	(0.028)	(0.016)	(0.016)
Writing/Language	0.007	−0.001	−0.009	−0.009
	(0.025)	(0.023)	(0.016)	(0.015)
Customer service	−0.100***	−0.067***	−0.008	−0.01
	(0.028)	(0.025)	(0.020)	(0.020)
Character	−0.128**	−0.111**	−0.101***	−0.093***
	(0.051)	(0.045)	(0.037)	(0.034)
LFS (main) skill indicators:				
Cognitive		0.071**		0.026
		(0.030)		(0.022)
Social		0.067**		0.038
		(0.030)		(0.023)
Management		0.213***		0.095***
		(0.019)		(0.014)
Financial		0.005		0.001
		(0.039)		(0.023)
Computer, General		0.098***		0.027
		(0.035)		(0.022)
Computer, Specific		0.111***		0.041*
		(0.033)		(0.021)
Writing/Language		−0.104*		−0.077**
		(0.055)		(0.032)
Customer service		−0.123***		−0.024*
		(0.020)		(0.013)
Character		0.134***		0.091***
		(0.030)		(0.022)
Observations	2750	2750	2750	2750
Clusters	893	893	893	893
R-squared	0.252	0.347	0.608	0.623
F-statistic	–	21.99***	–	8.57***
Individual controls	NO	NO	YES	YES

Notes: The dependent variable is the natural logarithm of average wages within the first year of a job spell. Columns 1 and 3 presents the results from estimating Eq. (2) when restricting θ_k to zero, and Columns 2 and 4 presents the results from estimating Eq. (2) when allowing θ_k to vary. The F-statistic shown is the result of a joint hypothesis test with that null that all of the θ_k are 0. See Appendix Table B.I for details on the sample. Individual controls include: age, age^2, experience, experience^2, years of education fixed effects, and a gender dummy. Note that the skill categories are not mutually exclusive; an individual can have more than one skill. Standard errors, in parentheses, clustered at the firm level. All regressions include year and municipality fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

12%. In contrast, we find no positive association between wages and management skills for employees who work in jobs that advertise for management skills, but who do not use those skills intensely. Recall that managers, as defined from occupation codes, are not included in this analysis due to poor representation in the LFS, suggesting that there are high returns to the use of management skills prior to entering a management occupation.

Second, we find that those who work in a job that advertised for character skills, but who do not mainly use character skills, face wage penalties, but those in jobs that both advertise for character skills and who use character skills as a main skill face no such penalties. These results reflect the fact that, uniquely, character skills are negatively correlated in the LFS and JP (see Table 3). One way to interpret this result is that jobs requiring some character skills in the JP are also lower paying jobs, e.g., a lower paying job may require "good mood", but this may be implicitly understood (or not demanded) in a higher paying jobs.

Third, those working in jobs in which writing/language and customer service skills are used intensively receive lower wages. However, employees holding jobs that advertised for these skills, but who do not report them as main skills, see no negative effects on wages. After controlling for individual characteristics, mainly using writing/language skills is associated with 7.7% lower wages, and workers who mainly use customer service have 2.4% lower wages on average. Since language skills are some of the core skills in humanities (see e.g., De Dijn et al., 2023), the negative returns to intensively using writing/language skills align well with the education literature, which finds that humanities have the lowest returns in terms of earnings out of all fields of education (see e.g., Kirkeboen et al., 2016; Daly et al., 2022).

We next compare our estimated returns to skill use at the extensive margin to Deming and Kahn (2018), though we note that they use data aggregated at the MSA-occupation-level, and our wage Eq. (3) is estimated the individual level.[29] Deming and Kahn (2018) report estimates of returns to cognitive and social skills with and without controls. Without controls, Deming and Kahn (2018) report an estimated coefficient on cognitive skills equal to 0.113, which is very close to the 0.114 estimated coefficient we find on cognitive skills. However, they find positive and even larger returns to social skills with an estimated coefficient of 0.429, whereas we find an insignificant estimate of −0.006. There are a variety of reasons why the estimates may not be the same across US and Denmark, both because of institutional settings, but also because our estimates are at the individual level, and thus, exclude any potential positive externalities of, for example, social skills on team members. Jensen (2024) discusses this at length in the Danish setting and shows that aggregated estimates are similar to those of Deming and Kahn (2018) despite individual-level estimates not indicating positive returns to the interaction between social and cognitive skills.[30]

In sum, our wage results suggest that considering the intensive use of skills flexibly in a Mincerian-style wage equation can increase the precision on the extensive margin skill returns while also allowing for a better understanding of whether or not a skill can be interpreted as a human capital, where increasing skill levels yield a higher wage. Deming and Noray (2020) analyze the effect of job skill change on the return to experience, and Braxton and Taska (2023) analyze how technological change affect earnings after a job loss. Both use job postings data to identify skills and find large earnings losses associated with skill or technological change. Perhaps not surprisingly, our results suggests that these earnings losses will be heterogeneous, such that the losses will be even larger for workers who would use the skills that become obsolete intensively and smaller for workers who do not.

Since our measure of intensive skill usage from the LFS is only available for small samples, our results suggest that any measure of the intensity of skill use may be of value and help increase our understanding of returns to task-specific skills. Alternatively (Alekseeva et al., 2021) and Jensen (2024) include continuous measures of JP skills — one of several approaches that we will further consider in Section 5.

4. Robustness

In this section, we verify that our conclusions are robust to changing the construction of our estimation sample and our skill measures. We consider different ways of matching job spells to job posts, alternative keyword-to-skill categorizations, and different sample selection criteria. Throughout, we focus on the robustness of our key results from Tables 3 and 6. Although all robustness checks are discussed here, results are mainly reported in Appendix C.

Time from job post to job start: For our main analyses, we match individuals to a job post in the same firm-occupation cell if the job was posted in the month in which they started their new job or a maximum of four months prior. Next, we derive skill measures from the relevant job posts. For some jobs, a four months matching window may be too restrictive; for others, matches may form sooner. To examine how our results depend on the length of the matching window between job posts and new job spells, we allow the length to vary from 3 to 8 months where 4 months is our default. We plot the results equivalent to the diagonal in Table 3 in Fig. 1. Our default estimates, considering job posts maximum four months prior to job start, are emphasized by the horizontal dashed line. We see that the correlation between cognitive skills in the JP and LFS start declining after the 4-months point, and the correlation between management skills stabilizes at the 4-months point. For the remaining skills, the correlations are stable across the various matching windows. This indicates that considering job posts maximum four months prior to job start yields the skill measures most informative about skill use.

Next, we consider if our results from Table 6 change depending on the matching window. These results are reported in Figure C.1. In line with our results from Fig. 1, Figure C.1 shows that the correlation between cognitive skills and wages is reduced after the 4-months point, and the correlation between management skills and wage stabilizes at the 4-months point. Overall, we interpret our results as robust to varying the time between job post and job start, but that the 4-months matching window gives the most informative skill measure.

Matching to one job post only: We may observe multiple job posts when considering a firm-occupation cell in the same month as a new job spell starts and four months prior. An individual may be hired to either of the posted jobs within this matching window, but we cannot observe which one. Therefore, we aggregate the skill requirements of all the observed job posts within the matching window. 53% of our estimation sample match with two job posts or fewer; 37% match to a single job post. To check if our results are affected by aggregation of job posts for the remainder of the sample, we rerun our analyses focusing specifically on individuals who match with a single job post. We report the results of this exercise in Figures C.2 and C.3. In Figure C.2, we see that correlations between LFS and JP skills are generally similar when considering those who only match with one job post. Exceptions are financial skills and both measures of computer skills that have stronger, but still similar, correlations. In contrast, the correlation with customer service skills is weaker.

Job spells starting at new firms: In our main analyses, we consider new job spells both of individuals who change occupations within a firm, but also those who start a job in a new firm. To make sure that our results are not driven by within-firm transitions (e.g., due to firm-specific human capital accumulation/returns to tenure), we separately consider the 79.7% of our estimation sample who start a job at a new firm. Results equivalent to Tables 3 and 6 are reported in Figures C.4 and C.5. In line with Fig. 1 and C.1, we allow the matching window between job posts and new job spells to vary. Our estimates focusing on individuals starting jobs at new firms are very similar to those from the full estimation sample.

Keywords and skill categorization: When comparing two different sources of text; that is self-reported task-specific skills in the LFS, and job posts written by firms, misalignment in skill measures could arise from linguistic differences between the text sources. Employees may choose different words to describe their skills compared to, e.g., HR-departments. To assess potential distortion due to linguistic differences between LFS and JP, we undertake a robustness check where we only categorize JP keywords if they also appear in the LFS text. As such, we fix the lexicon across the two sources of text. The results from this exercise are reported in Figures C.6 and C.7. With this alternative categorization of JP keywords, we find slightly larger correlations between JP and LFS skills when considering social and specific computer skills, but smaller correlations when considering general computer skills in Figure C.6. Considering wage regressions in Figure C.6 results are very similar to those in Fig. 1 with the exception of the coefficient on JP character skills generally being statistically insignificant in Figure C.6.

We observe similar results with this alternative skill category because it highly correlates with the original skills category. On average, for the original JP skill category, 88.7% of jobs with an LFS skill also

[29] This means that the interpretation of our coefficients is different than theirs: their estimated extensive margin return is the effect on average wages at the MSA-occupation-cell that results when 0 workers in an MSA-occupation-cell using a skill switch to *all* workers in the MSA-occupation-cell using the skill.

[30] One potential explanation of this may be because *"workers supplying both social and cognitive skills may be paid their individual marginal product, but not necessarily be rewarded for their positive externalities on co-workers' productivity if employers fail to internalize these"* (Jensen, 2024, p. 21).

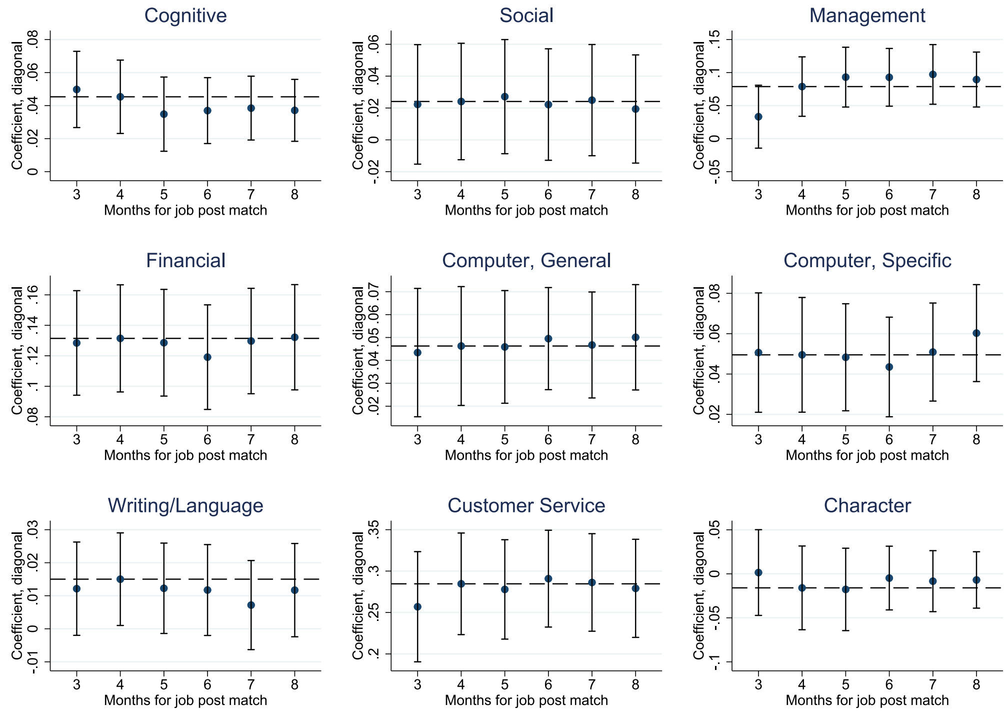

Fig. 1. Regression of LFS (main) skill on job posting skills: Estimation sample. *Notes*: This figure reproduces results from Table 3, varying the lead time required for a job post to match with a new job spell. Our default lead time is 4 months before the job start; estimates using this default are indicated by a horizontal dashed line. The reported coefficients are estimated by regressing each of the LFS task-specific main skill measures on the 9 job posting skills categories. We report only the coefficient on the job posting skill that correspond to the relevant LFS skill; these are the diagonal elements of Table 3. Standard errors clustered at the firm level; 95%-confidence interval indicated.

had the same JP skill (see Table 2). When we define the JP skills categories to only include LFS words, 87.5% of jobs with an LFS skill also had the same JP skill. The small decline in this conditional average is due to the fact that the JP skills are indicators taking the value 1 if any of the keywords in a category are listed in the job post. Conditional on requiring an HBS skill, jobs have at the median, 5 keywords that belong to that skill. When we condition on the keywords belonging to the LFS keywords, the median remains the same. As such, only 88.7%-87.5% = 1.2% of jobs with an LFS skill change their extensive use of the JP skill when we condition JP skills to only include the LFS keywords. The difference is relatively small, as the most frequent keywords appear in both the LFS and JP data.

Conditioning on an LFS skill: In our main analyses, we condition on survey participants providing some answer to the question on task-specific skills used on the job, i.e., not leaving a blank reply. Nevertheless, some participants may still provide very brief or non-sensible answers to the relevant question. To make sure our results are not driven by including such participants in our analyses, we undertake a robustness check considering the 72% of our sample with at least one LFS keyword indicative of a task-specific skill. We report results in Figures C.8 and C.9. As one would expect, we find stronger correlations for some skills, namely cognitive, social, financial and computer skills in Figure C.8. When considering correlations between wages and skills in Figure C.9, coefficient on JP skills are generally similar to those for the full estimation sample in C.1. In Panel (b), we see relatively larger coefficients on cognitive and management skills from the LFS. However, those coefficients are very similar between in Panels (d) of Figure C.9 and C.1 where we control for selection by education and experience.

5. Extensions

Given that most researchers using JP data will not have access to intensity measures from additional sources such as the LFS, a natural question arises: can measures from the JP data be derived to capture some of the variation explained by the LFS skill importance/intensity indicators? To explore this possibility, we develop several measures of skill interactions and skill intensity from the JP data, and next, assess their ability to explain variation in wages and the extent to which they absorb some of the explanatory power of the intensive LFS skill measure. In this section, we present results without individual controls as most researchers will not have access to such individual level data, though all results with added controls are included in Appendix D.

First, we investigate whether the LFS skill intensity measure captures interactive effects among JP skills, e.g., that some JP skills increase the returns to other skills. For instance, a worker may consider social skills important because they are used in conjunction with other skills such as cognitive or management skills. If the combination of social skills with other skills makes these skills more valuable, we would expect interaction terms involving social skills to absorb part of the effects of the LFS variables.

Second, the importance of a skill as captured by the LFS measure may partially reflect the complexity of a job. For example, an employee working in a job that requires 8 skills may be less likely to state a particular skill as a main skill relative to an employee working in a job that demands only 2 skills. Therefore, we test how incorporating a measure of job complexity – specifically, the number of different skill types required by an employer – affects the explanatory power of the set of LFS indicator variables as well as how it interacts with the main skill used on the job.

Third, we evaluate several candidate measures of skill intensity derived directly from the JP data. We assess their explanatory power with respect to wages, both independently and in conjunction with the extensive margin JP indicators of skill, and analyze how the model changes when we include the LFS skill intensity measures.

JP skill interactions: Fig. 2 presents the results of extending Eq. (3) to include interaction terms among the JP extensive margin skill measures:

$$wage_i = \gamma_0 + \sum_{k=1}^{9} \gamma_k s_{ki}^{JP} + \sum_{k=1}^{8} \sum_{j=k+1}^{9} \alpha_{kj} s_{ki}^{JP} s_{ji}^{JP} + x_i \delta + \epsilon_i \tag{4}$$

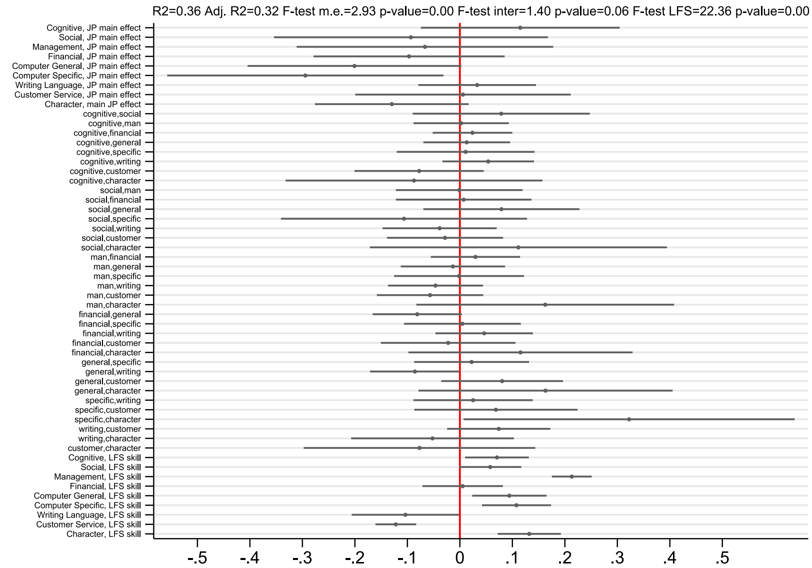

Fig. 2. Wage regressions, two-way interactions. *Notes*: This figure plots the estimated effects from regressing the natural logarithm of average wages within the first year of a job spell on JP indicators, 2-way interactions among these indicators, LFS indicators along with indicators for municipality and year. Standard errors clustered at the firm level; 95%-confidence interval indicated. In addition to the R^2 and Adjusted R^2, the figure displays the F-statistic on the joint significance of the main effects (interaction effects), F-test m.e. (F-test inter) followed by its p-value and the F-test on the joint significance of the LFS skill indicators followed by its p-value.

where the γ_k reflect the main effects of the JP skill indicators and the α_{kj} reflects the two-way interaction effects between the skill indicators. As in the main specification, x_i contains year and municipality indicators.[31] The inclusion of interaction terms in the model leads to a marginally significant increase in explanatory power, and the p-value on the F-test on these interaction terms is 0.06. Due to the relatively small sample size and the high correlation between skills, the coefficients on the interaction terms are estimated with too little precision to draw conclusions about the association between wages specific two-way skill combinations. Only one interaction term is statistically significant, well within the expected number of Type 1 errors given the number of tests conducted. The coefficients on the LFS skill indicators remain nearly identical to those in the main results (see Table 6, Column 2), as does their joint significance, indicating that the LFS variables are not capturing the effects of these two-way interactions.

We next attempt to absorb some of the explanatory power of the LFS intensity measure by including a full set of JP skill bundle indicators (i.e., indicators for each unique combination formed by concatenating the 9 JP skill indicators) instead of the main and two-way interaction effects in Eq. (4). Fig. 3 displays the resulting coefficient estimates for the LFS skill indicators.[32] Once again, the coefficients on the LFS skill indicators are virtually identical to those presented in Table 6, Column 2, as are their individual and joint significance levels.

Finally, we explore a data-driven approach to identify JP skill clusters using k-means clustering. We present results for $k = 10$ clusters here to give an alternative to the use of single skill indicators with similar degrees of freedom.[33] Wages are then regressed on indicators for each of these k-means clusters, along with the LFS skill indicators, year and municipality fixed effects. Fig. 4 presents these results, by first plotting the estimated effects of the k-means cluster indicators with labels 1 though 10, then the estimated coefficients on the LFS skill indicators. It is immediately clear that the clustering does not add explanatory power relative to the one-way JP skill indicators (e.g., none of the JP clusters are significantly different than 0 as opposed to three of the JP skill indicators in the main specification). As expected given these results, the estimated coefficients on the LFS skill indicators remain unaffected. Taken together, the results in this subsection suggest that the type of intensity or importance of skill use picked up by the LFS indicators is not highly correlated with interaction effects among variables.

Job complexity: Next, we examine how job complexity, measured as the number of different types of JP skills demanded in a job posting, affects wages and interacts with the main skill used in the job, as indicated by the LFS skill indicators. As a starting point, Fig. 5 displays the results of regressing wages on indicators of the number of skill categories a particular job posting covers, along with municipality and year fixed effects.[34] Compared to jobs where all skill categories are required (the omitted condition), jobs requiring fewer skills tend to be less well-compensated. Interestingly, the returns to complexity exhibit a U-shaped pattern: jobs that cannot be classified as either specialized (i.e., those requiring 1 or 2 skills) or generalist (i.e., those requiring 6 to 9 skills) are the most penalized.[35]

To determine if returns to complexity capture some of the explanatory power of the LFS variables, we next examine how the inclusion of the LFS skill variables affects the estimated effects of the number of skill category indicators. We also compare the resulting estimates of the LFS skill indicators to those found in the main results of Table 6. Fig. 6 presents these results.[36] We observe that the estimated effects of the LFS indicator variables are effectively unchanged by the inclusion of the complexity measure, and vice versa, suggesting that these two sets

[31] Appendix Figure D.1 presents results when individual controls are included: age, age^2, experience, experience^2, years of education indicators, and a gender dummy variable.

[32] The estimated effects when individual controls are included are shown in Appendix Figure D.2. The regression coefficients for the skill bundles are also reported in Appendix Figures D.3 and D.4, without and with controls, respectively.

[33] We report results for $k = 20$ and $k = 30$ with and without individual controls in Appendix Figures D.7–D.10, corresponding to models with similar degrees of freedom as the model with two-way interactions, as well as the estimates for $k = 10$ with individual control variables in Appendix Figure D.6. Heatmaps displaying the JP skill concentrations in each cluster are also provided in Appendix Figure D.5.

[34] Note that a full set of JP skill indicators is perfectly collinear with indicators of the sum of JP skills required on the job.

[35] See Appendix Figure D.11 for results when individual controls are included in the regression.

[36] See Appendix Figure D.12 for results when individual controls are included in the regression.

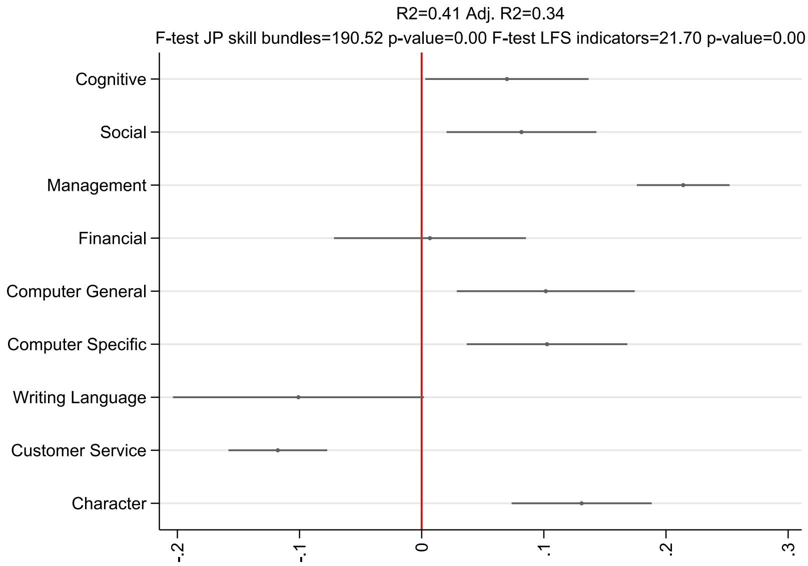

Fig. 3. Wage regressions, skill bundles. *Notes*: This figure plots the estimated effects from regressing the natural logarithm of average wages within the first year of a job spell on mutually exclusive JP skill bundles (i.e., the concatenation of the 9 binary skill indicators), LFS skill indicators and municipality and year fixed effects. Standard errors clustered at the firm level; 95%-confidence interval indicated. In addition to the R^2 and Adjusted R^2, the figure displays the F-statistic on the joint significance of the JP skill bundles followed by its p-value and the F-test on the joint significance of the LFS skill indicators followed by its p-value.

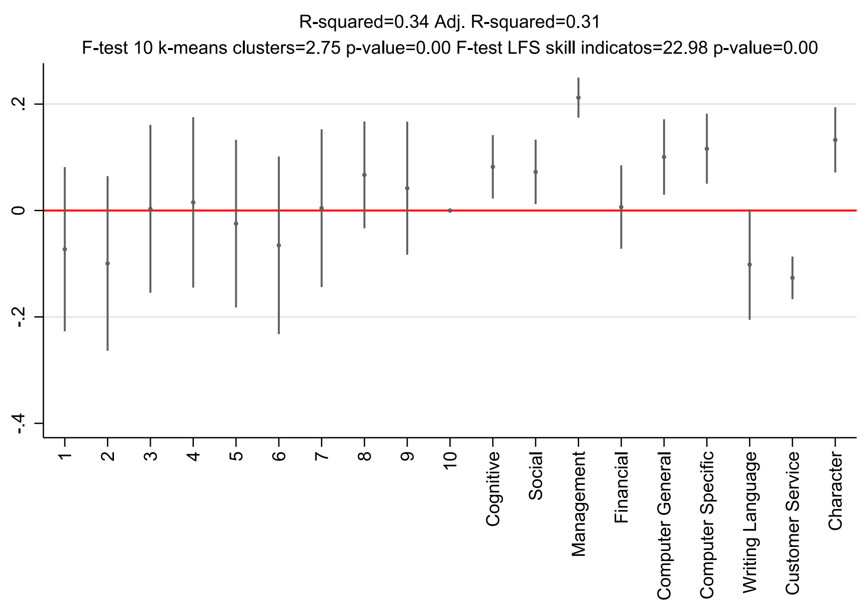

Fig. 4. Wage regressions, 10 K-means clusters with LFS indicators. *Notes*: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on indicators of 10 k-means clusters, the LFS skill indicators and time and municipality fixed effects. The omitted condition corresponds to cluster 10. Standard errors are clustered at the firm level, with 95% confidence intervals indicated.

of variables have little covariance.[37] This indicates that the importance or intensity being captured by the main skills as identified by workers is not strongly correlated with the complexity of the job as measured by the number of skill categories listed in the job ad.

Of course, the representation of skills by the number of skills present in the ad may not be random. To investigate this, Fig. 7 plots the relative frequency of each JP skill conditional on the number of JP skill demanded. The red line indicates the fraction of skills we should expect if the skills were randomly allocated. The figures reveals that the more specialized jobs are more likely to be, for instance, those requiring customer service jobs or character.

To better understand the importance of complexity and the return to skills, we next regress wages on the interaction between indicators for the number of skill categories in the job posting and JP skill indicators. Fig. 8 plots the interaction estimates.[38] For instance, the top left subgraph of the figure displays the estimated effects of the JP cognitive

[37] To make this comparison properly, without the inclusion of JP skill indicators, please see Column 9 of Table 7.

[38] See Appendix Figure D.13 for results when individual controls are included in the regression.

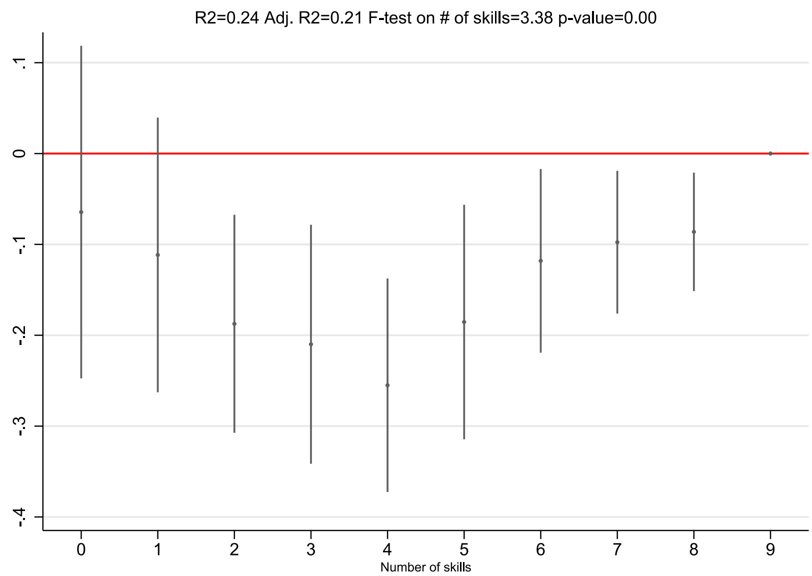

Fig. 5. Wages regressions, number of JP skill categories. *Notes*: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on indicators of the number of skill categories demanded, along with time and municipality fixed effects. The omitted condition is for workers employed in jobs that demanded all skill categories (i.e., 9 skills). Standard errors clustered at the firm level; 95%-confidence interval indicated.

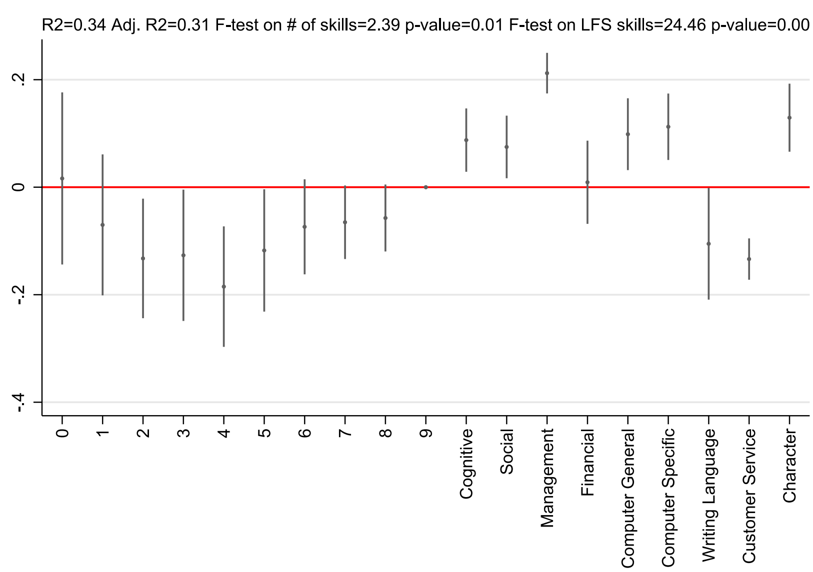

Fig. 6. Wage regressions, number of skill categories and LFS skill indicators. *Notes*: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on indicators of the number of skill categories demanded on job postings, LFS skill indicators along with time and municipality fixed effects. The omitted condition is for workers employed in jobs that demanded all skill categories (i.e., 9 skills). Standard errors clustered at the firm level; 95%-confidence interval indicated.

skill when it is the only demanded skill (1 on x-axis), one of two skill demanded (2 on x-axis), etc. The omitted condition is 9 skills. Relative to the case of working in a job that requires all skills, Fig. 8 suggests that cognitive skills are least rewarded when combined with 3 or 4 other skills, but more highly rewarded when used in specialized jobs. No strong evidence is otherwise detected between the other skill variables and the complexity of the job. As is clear from Fig. 7, the representation of skills in each bundle is not random, and hence the JP skill returns depicted in the figure could be confounded with interaction effects with the other skills with which they are bundled. In terms of explaining variation in wages, this model does not perform better than the main specification which just includes the JP indicators and no measure of job complexity (both models have an adjusted $R^2 = 0.22$).

Fig. 9 presents the same exercise, except it uses LFS skill indicators instead of JP skill indicators and interacts them with indicators for the number of skill categories in the job posting.[39] Generally, job complexity does not appear to significantly affect the estimated effects of the LFS skill variables, with exception of management and customer service skills. The top rightmost sub-graph presents the effect of reporting management as a main skill interacted with an indicator for the number of skills demanded on the job. The figure displays no clear relationship, suggesting no strong correlation between complexity and the returns to the main skill used on the job.

Additional JP skill intensity measures: We next consider several candidate measures of skill intensity derived directly from the JP data and compare their estimated effects and combined explanatory value with those of the LFS skill intensity measures. The first eight columns

[39] See Appendix Figure D.14 for results when individual controls are included in the regression.

Table 7
Comparison of alternative skill measures.

Variables	(1) JP Indicator	(2) JP Indicator	(3) JP Fraction Add	(4) JP Fraction Add	(5) JP Fraction Max	(6) JP Fraction Max	(7) JP Standardized	(8) JP Standardized	(9) LFS Indicator	(10) LFS Indicator
Cognitive	0.114***	0.095***	0.008***	0.010***	−0.014	0.002	−0.008	0.001	0.098***	0.087***
	(0.026)	(0.022)	(0.002)	(0.002)	(0.047)	(0.039)	(0.026)	(0.022)	(0.028)	(0.030)
Social	−0.006	−0.013	−0.007	−0.003	0.159***	0.084**	0.196***	0.103**	0.075**	0.074**
	(0.031)	(0.031)	(0.004)	(0.003)	(0.038)	(0.042)	(0.047)	(0.052)	(0.030)	(0.030)
Management	0.055*	0.084***	0.009*	0.012***	0.079*	0.083*	0.090*	0.094*	0.222***	0.211***
	(0.028)	(0.023)	(0.005)	(0.004)	(0.047)	(0.045)	(0.053)	(0.051)	(0.020)	(0.020)
Financial	0.028	0.034	−0.001	−0.004	0.030	0.008	0.020	0.006	0.015	0.011
	(0.024)	(0.023)	(0.007)	(0.006)	(0.025)	(0.026)	(0.017)	(0.017)	(0.039)	(0.040)
Computer, General	0.022	0.064***	−0.003	−0.002	−0.048**	−0.061***	−0.034**	−0.043***	0.120***	0.102***
	(0.035)	(0.022)	(0.007)	(0.007)	(0.022)	(0.017)	(0.016)	(0.012)	(0.031)	(0.035)
Computer, Specific	0.055*	0.100***	0.028**	0.024**	0.075***	0.043***	0.069***	0.039***	0.128***	0.114***
	(0.032)	(0.024)	(0.011)	(0.011)	(0.014)	(0.016)	(0.013)	(0.015)	(0.028)	(0.032)
Writing/Language	0.007	0.039*	−0.006	−0.009*	−0.023	−0.037**	−0.019	−0.030**	−0.107**	−0.102*
	(0.025)	(0.023)	(0.005)	(0.005)	(0.016)	(0.015)	(0.013)	(0.012)	(0.052)	(0.052)
Customer service	−0.100***	−0.067**	−0.041***	−0.032***	−0.224***	−0.126**	−0.285***	−0.161**	−0.142***	−0.135***
	(0.028)	(0.026)	(0.008)	(0.004)	(0.053)	(0.055)	(0.067)	(0.070)	(0.018)	(0.020)
Character	−0.128**	−0.146***	−0.033***	−0.032***	−0.082***	−0.104***	−0.107***	−0.134***	0.129***	0.129***
	(0.051)	(0.052)	(0.005)	(0.006)	(0.027)	(0.026)	(0.035)	(0.034)	(0.032)	(0.032)
Standardized skills in add		−0.349***		−0.133**						
		(0.053)		(0.054)						
Standardized skills in add, quadratic		0.087***		0.025						
		(0.021)		(0.021)						
Standardized skills in add, cubic		−0.007***		−0.002						
		(0.002)		(0.002)						
Number of skill categories in add						−0.135***		−0.135***		−0.101**
						(0.043)		(0.043)		(0.042)
Number of skill categories in add, quadratic						0.027***		0.027***		0.020**
						(0.009)		(0.009)		(0.010)
Number of skill categories in add, cubic						−0.001*		−0.001*		−0.001
						(0.001)		(0.001)		(0.001)
Observations	2750	2750	2750	2750	2750	2750	2750	2750	2750	2750
Clusters	893	893	893	893	893	893	893	893	893	893
R-squared	0.252	0.327	0.322	0.343	0.271	0.320	0.271	0.320	0.324	0.337
Adjusted R2	0.220	0.300	0.290	0.320	0.240	0.290	0.240	0.290	0.300	0.310
F-test on skill measures	10.70	23.30	25.98	26.97	98.49	108.8	98.49	108.8	38.17	23.95
p-value for skill measures	0	0	0	0	0	0	0	0	0	0

Note: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on various measures derived from the JP data with exception to the last 2 columns. Columns 1 and 2 use indicator variables corresponding to whether or not a skill is present in a job posting. Columns 3 and 4 use the fraction of keywords in the ad that belong to specific skill categories, labeled as *Fraction Ad*. Columns 5 and 6 use the number of keywords in a particular category divided by the maximum number of keywords in that category across all ads, labeled as *Fraction Max*. Columns 7 and 8 use a standardized measure of keywords, labeled as *Standardized*, within a category: the number of keywords in a category in an ad minus the mean number of keywords in that category across all ads, divided by the standard deviation of the number of keywords in that category. Columns 9 and 10 use indicator variables corresponding to the LFS skill intensity measure. The regressions include fixed effects for municipality and year. Standard errors are clustered at the firm level, with 95% confidence intervals indicated. Robust standard errors are shown in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

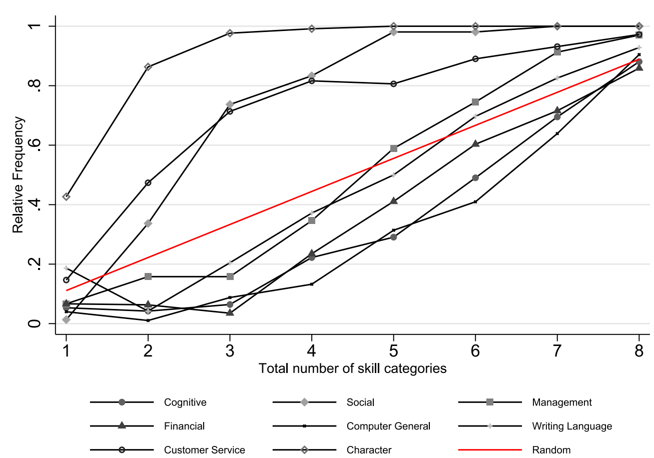

Fig. 7. JP skill relative frequency by number of JP skill categories. *Notes*: The figure shows the relative frequencies of workers employed by firms that demand specific skills, conditional on the number of skills required by the employer (as indicated on the x-axis). The red line represents the expected frequency if skills were randomly distributed in bundles of the size indicated on the *x*-axis.

of Table 7 present the results of regressing wages on various skill measures derived from the JP data. Columns 9 and 10 show results when wages are regressed on the LFS indicators (note that JP indicators are not included in these regressions). All regressions include year and municipality fixed effects; results with individual controls are available in Appendix Table D.1. Even-numbered columns also include measures of job complexity. Column 1 reproduces Column 1 of Table 6, a regression of wages on JP skill indicators. Column 2 performs the same regression, now including measures of job complexity. Columns 3 and 4 use the fraction of keywords in the ad that belong to specific skill categories as an intensive measure of skill use, labeled as *Fraction Ad*. Columns 5 and 6 instead use the number of keywords in a particular skill category divided by the maximum number of keywords in that category across all ads, labeled as *Fraction Max*. Columns 7 and 8 use a standardized measure of keywords within a category, labeled as *Standardized*.[40] Columns 1(2) and 3(4) both deliver similar signs and significance, though the explanatory power of the models using the fraction of skill in the ad is significantly higher than those using JP indicators, especially when no measure of complexity is included (Adjusted $R^2 = 0.29$ in Column 3 rather than an Adjusted $R^2 = 0.22$ in Column 1) demonstrating the important role that skill intensity plays.

Columns 5–8 present results for JP skill intensity measures defined relative to the number of skills in all job postings, as a fraction of the maximum number of keywords across all ads in Columns 5 and 6 or as a standardized value of the number of keywords in a category where the standardization is taken across all ads in Columns 7 and 8. The model performance is the same, regardless of whether one chooses an intensity measure derived as the fraction of the maximum, or whether one chooses a standardized variable. Choosing between the two specifications comes down to the ease with which one can interpret the resulting estimates. Here we see notable differences in the coefficient estimates relative to those attained when using the extensive margins in Columns 1 and 2. For instance, in Column 2, we see that the extensive margin return to social skill is effectively zero. In Column 8, we see that there is a large and significant increase of about 10% in wages for an increase of 1 standard deviation in the number of keywords identified in the ad as being social.

The last two columns of the table present the estimated effects associated with the LFS skill indicators. As is clear from comparing these two columns to the corresponding columns for the other candidate intensity skill measures, the LFS skill measure seems to be picking up

40 The number of keywords in a category in an ad minus the mean number of keywords in that category across all ads, divided by the standard deviation of the number of keywords in that category.

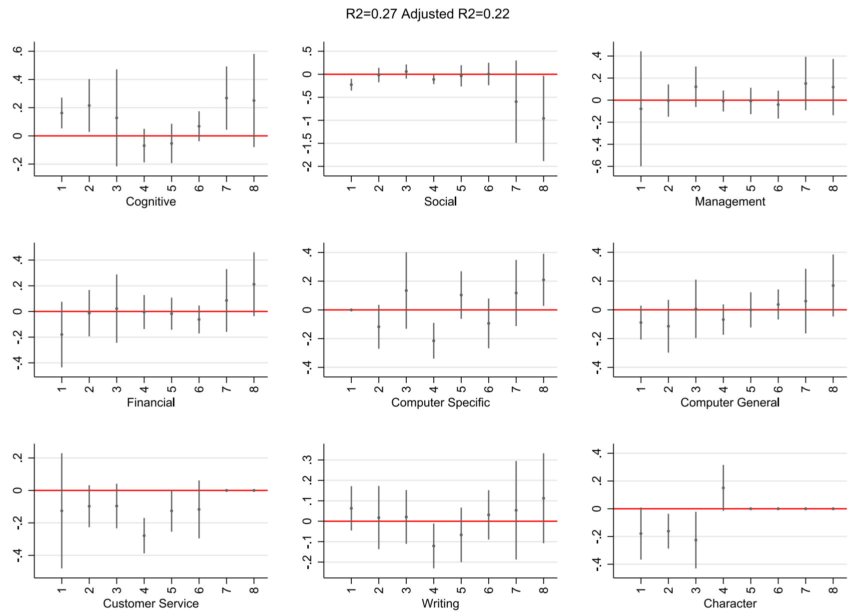

Fig. 8. Wage regression, effect of JP skills moderated by job complexity. *Notes*: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on interactions between the number of skill categories demanded and JP skill indicators, along with time and municipality fixed effects. The omitted condition is for workers employed in jobs that demanded all skill categories (i.e., 9 skills). Standard errors clustered at the firm level; 95%-confidence interval indicated.

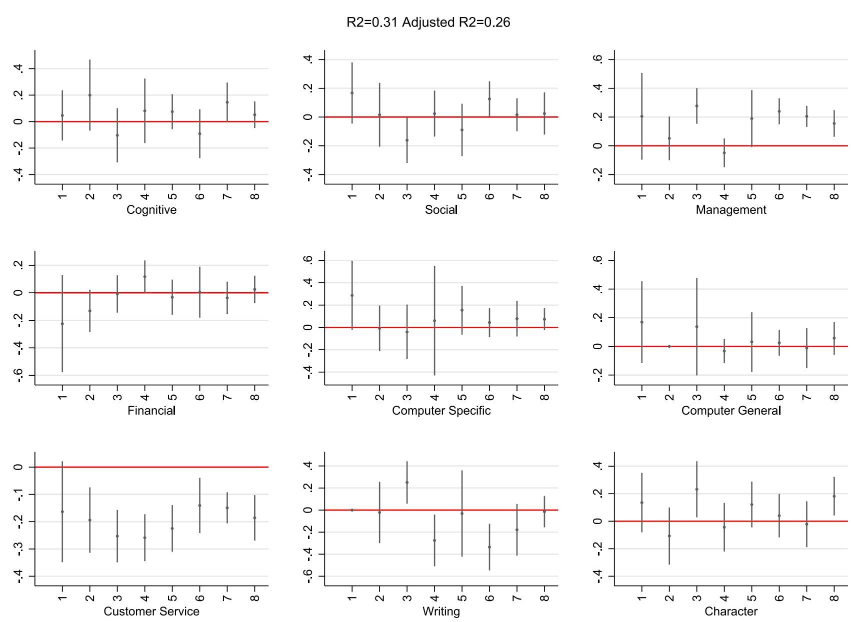

Fig. 9. Wage regression, effect of LFS skills moderated by job complexity. *Notes*: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on interactions between the number of skill categories demanded and LFS skill indicators, along with time and municipality fixed effects. The omitted condition is for workers employed in jobs that demanded all skill categories (i.e., 9 skills). Standard errors clustered at the firm level; 95%-confidence interval indicated.

a different source of variation. To test this, we next embed the set of JP indicators, a JP intensity measure, and the LFS intensity indicators into the same model. We perform this exercise using the standardized JP intensity measure.

Column 1 of Table 8 presents a model based on both extensive and intensive skill measures that can be derived from the JP data. Column 2 presents the main results of the paper, i.e., JP extensive margin indicators and LFS intensive margin indicators. Column 3 omits the JP extensive margin estimates but includes both the standardized JP intensive skill measure and the LFS intensive skill measure. Finally, Column 4 presents a model with the JP extensive margin skill indicators, JP intensive margin skill measures, and LFS intensive margin skill indicators.[41]

Comparing Column 1 to Column 4, we see that the explanatory power of the model significantly increases when the LFS skill indicators are included: the adjusted R^2 increases from 0.30 to 0.372. Yet, the

[41] See Appendix Table D.2 for the results of a model that includes individual controls.

Table 8
Model comparison.

	(1)	(2)	(3)	(4)
Cognitive, JP extensive	0.096***	0.100***		0.087***
	(0.022)	(0.023)		(0.021)
Social, JP extensive	−0.014	0.005		−0.006
	(0.031)	(0.028)		(0.029)
Management, JP extensive	0.070***	0.030		0.046**
	(0.024)	(0.027)		(0.022)
Financial, JP extensive	0.031	0.025		0.027
	(0.023)	(0.023)		(0.022)
Computer, General, JP extensive	0.058**	−0.001		0.034
	(0.024)	(0.031)		(0.021)
Computer, Specific, JP extensive	0.076***	0.031		0.060***
	(0.024)	(0.028)		(0.023)
Writing/Language, JP extensive	0.038*	−0.001		0.025
	(0.023)	(0.023)		(0.021)
Customer service, JP extensive	−0.074***	−0.067***		−0.055**
	(0.026)	(0.025)		(0.025)
Character, JP extensive	−0.145***	−0.111**		−0.124***
	(0.052)	(0.045)		(0.046)
Cognitive, JP intensive	−0.002		−0.003	0.000
	(0.021)		(0.020)	(0.017)
Social, JP intensive	0.115**		0.145***	0.098**
	(0.051)		(0.043)	(0.047)
Management, JP intensive	0.092*		0.087**	0.091**
	(0.051)		(0.039)	(0.039)
Financial, JP intensive	0.009		0.026	0.016
	(0.017)		(0.018)	(0.018)
Computer, General, JP intensive	−0.044***		−0.036**	−0.043***
	(0.012)		(0.014)	(0.011)
Computer, Specific, JP intensive	0.037***		0.042***	0.022*
	(0.014)		(0.012)	(0.013)
Writing/Language, JP intensive	−0.024**		−0.011	−0.016
	(0.012)		(0.012)	(0.011)
Customer service, JP intensive	−0.174**		−0.207***	−0.143**
	(0.070)		(0.054)	(0.058)
Character, job posting, JP intensive	−0.128***		−0.107***	−0.124***
	(0.033)		(0.031)	(0.031)
Cognitive, LFS		0.071**	0.084***	0.055**
		(0.030)	(0.023)	(0.024)
Social, LFS		0.067**	0.057**	0.045*
		(0.030)	(0.027)	(0.027)
Management, LFS		0.213***	0.216***	0.197***
		(0.019)	(0.019)	(0.019)
Financial, LFS		0.005	−0.021	−0.028
		(0.039)	(0.038)	(0.036)
Computer, General, LFS		0.098***	0.090***	0.058**
		(0.035)	(0.025)	(0.024)
Computer, Specific, LFS		0.111***	0.107***	0.078***
		(0.033)	(0.027)	(0.027)
Writing/Language, LFS		−0.104*	−0.117**	−0.117**
		(0.055)	(0.053)	(0.057)
Customer service, LFS		−0.123***	−0.125***	−0.100***
		(0.020)	(0.016)	(0.016)
Character, LFS		0.134***	0.133***	0.135***
		(0.030)	(0.032)	(0.029)
Observations	2750	2750	2750	2750
Clusters	893	893	893	893
R-squared	0.330	0.347	0.368	0.401
Adjusted R2	0.300	0.318	0.340	0.372
F-test on JP indicators	18.70	6.122	–	11.55
p-value for JP indicators	0.00	0.00	–	0.00
F-test on JP percents	88.61	–	80.35	62.94
p-value for JP percents	0.00	–	0.00	0.00
F-test on LFS indicators	–	21.99	35.21	25.99
p-value for LFS indicators	–	–	0.00	0.00

Note: The reported coefficients are estimated by regressing the natural logarithm of average wages within the first year of a job spell on various skill measures. The regressions include fixed effects for municipality and year. Standard errors are clustered at the firm level, with 95% confidence intervals indicated. Robust standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

frequency with which a skill is mentioned in the job post does correlate with the importance that individuals assign to that skill, as can be seen by comparing the estimates of the LFS measures in Column 2 to Column 4; for instance, the estimated effect of LFS cognitive drops by about 20 percent whereas the estimated effect of LFS social drops by almost a third when JP intensity measures are introduced.

Although the JP skill intensity measures are highly significant, the LFS skill intensity measures also remain highly significant, implying that the way workers deem a skill to be the main skill is at least in part driven by something other than how frequently the skill was mentioned in the job ad. It is not unreasonable to imagine several scenarios in which this might be true: A job ad that reads, "Applicants absolutely must have exceptionally strong cognitive skills. In addition, professional, competent people who enjoy working with others are preferred". The current text-to-skill classification is not capable of picking up the implied intensity in statements such as "absolutely must have", and as such, the resulting measurement error embedded within JP intensive skill measures can be somewhat offset by the LFS intensity skill measures.

6. Conclusion

The aim of this paper is to better understand both how well skill measures derived from job postings data capture skills used on the job, and next, the extent to which the availability of an intensive measure of skill use, as opposed to just an extensive measure of skills, improves the estimates of returns to task-specific skills. We empirically test the assumption that skills demanded in job posts reflect skill use on the job by comparing the demand of task-specific skills from job posts to individual self-reported main skill use extracted from the Danish Labour Force Survey (LFS) for private employees over the 10-year period beginning in 2007. To our knowledge, we are the first to compare skills demanded in job posts with self-reported skills use at such a granular level.

We explore the degree to which measures of task-specific skills derived from job postings correspond with skills used on the job as reported by workers in the LFS. We find that workers who report a particular main skill are extremely likely to be in a job for which that particular skill was demanded in the corresponding job post. Moreover, we generally find positive and significant correlations between skills derived from job posts and self-reported main skills from the LFS. We move on to investigate the relative employer-employee match quality and find no evidence of substantial mismatch between skill supply and demand, allowing us to propose a framework in which to understand the two sets of skill measures: we interpret skill demand as captured by job posts and main skills reported on the LFS as extensive and intensive measures of skill use, respectively.

Our wage regression results suggest that extensive margin skill returns can be more precisely estimated if intensive skill measures and individual controls are included. We find several large differences between the extensive margin returns to advertised skills and the returns to skills mainly used on the job as captured by the LFS. For example, returns to cognitive and management skills increase with the intensity of their use, whereas writing/language and customer service skills are associated with lower wages when these skills are used more intensively. Our finding of consistently positive returns to cognitive skills is in line with the existing literature, e.g., Spitz-Oener (2006), Black and Spitz-Oener (2010), Beaudry et al. (2016), Deming and Kahn (2018) and Atalay et al. (2020).[42]

We conduct a number of robustness exercises and implement various extensions to our basic model of wage returns to skills and find that our results are robust to various matching strategies between the LFS and JP data (i.e., different matching windows between the dates of job posting and job start, matching to only one job post, and to only include firm switchers as new job spells). We also show that our results are robust to different skill categorizations (i.e., harmonizing lexicons between LFS and JP data, and conditioning on at least one LFS skill).

Given that most researchers using JP data will not have access to intensive skill use measures from additional sources such as the LFS, we try to understand if measures derived from the JP data can capture some of the variation explained by the LFS skill indicators. We develop several measures of skill interactions, skill intensity, and job complexity from the JP data and assess their ability to explain variation in wages and the extent to which they absorb some of the explanatory power of the intensive LFS skill measure. Based on this analysis, we recommend researchers who only have JP data available to consider both intensive and extensive measures of skills along with a job complexity measure — in our setting, such a model yields the most precise estimates. In addition to this, if feasible, researchers should also consider augmenting JP data with survey data capturing skill intensity or importance when developing such models.

In summary, our findings suggest that the skills measures derived from job postings data typically used in the literature capture main skills used on the job, and thus, using job postings data to analyze skill use on the job is generally a valid empirical strategy. However, a rich dimension of heterogeneity in skill returns may be missed if only the extensive margin of skill demand is considered. Our data allow us to study how skills measures from job posts and self-reported skills covary at the individual level, and thus, we avoid issues of interpretation that arise when studying correlations between more aggregated variables. While much of the existing literature has focused on the returns to skill demand on the extensive margin, we believe our ability to study both the intensive (derived from both the JP data and also from skill use data) and extensive margins of task-specific skills at the job level is new to the literature.

An interesting avenue for future research to consider is even more nuanced measures of skill importance and intensity from job postings, e.g., by using language models that can recognize cognitive skills as particularly important if highlighted in statements such as "strong cognitive skills are an essential requirement".

CRediT authorship contribution statement

Moira Daly: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Fane Groes:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mathias Fjællegaard Jensen:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Novo Nordisk Foundation grant number 16OC0021056, and the Kraks Fond project number 301061. Jensen also thanks the Independent Research Fund Denmark for financial support through grant number 1058-00011B.

[42] Some of these papers consider routine and non-routine cognitive skills separately. Beaudry et al. (2016) describe a decline in the demand for cognitive skills after 2000, but still find that cognitive skills are associated with higher wages.

Table A.1
Comparison of matched samples.

	(1)		(2)		(3)		(4)	
	Population		JP-Population Matched		LFS		JP-LFS Matched *Estimation Sample*	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Female	0.519		0.523		0.497		0.494	
Age	35.790	(11.621)	34.420	(11.288)	36.390	(11.501)	35.325	(10.890)
Immigrant	0.078		0.083		0.061		0.063	
Married	0.361		0.327		0.406		0.390	
Number of children under 18	0.677	(0.975)	0.639	(0.954)	0.740	(0.997)	0.729	(0.975)
Registered experience	11.731	(10.848)	10.704	(10.405)	12.060	(10.905)	11.132	(10.237)
Potential experience (years since ended education)	12.470	(15.867)	11.344	(15.091)	11.974	(14.247)	10.699	(10.584)
Years of education (monthly)	14.314	(2.270)	14.180	(2.290)	14.738	(2.334)	14.766	(2.277)
Student at any point in month	0.168		0.171		0.101		0.096	
Home region indicators:								
Northern Jutland	0.083		0.083		0.092		0.089	
Mid-Jutland	0.215		0.222		0.232		0.241	
Southern Denmark	0.174		0.170		0.184		0.167	
Capital Region	0.405		0.401		0.373		0.387	
Zealand	0.124		0.124		0.119		0.117	
Work region indicators:								
Northern Jutland	0.078		0.077		0.084		0.078	
Mid-Jutland	0.208		0.216		0.228		0.233	
Southern Denmark	0.167		0.162		0.179		0.169	
Capital Region	0.454		0.459		0.425		0.446	
Zealand	0.091		0.086		0.084		0.075	
Person-year observations	2,821,996		499,645		13,138		2750	
Share of new jobs matched to job post			*17.71%*				*20.93%*	

Notes: "Population" refers to the full population of new individual-level job spells in the private sector starting in January 2008 to July 2017 in ISCO-groups 2 to 5. "JP-Population Matched" refers to the subsample of new job spells that can be linked to corresponding job post(s). "LFS" refers to the subsample of new job spells that can be linked to an observation in the LFS within the first year of commencing the job spell. "JP-LFS Matched / Estimation Sample" refers to the subsample of new job spells that can be linked to both corresponding job post(s) and an observation in the LFS within the first year of commencing the job spell.

Appendix A. Data

A.1. Job postings data

The JP are supplied by HBS Economics (HBS). The data are provided after an initial cleaning procedure has been performed. HBS asserts that their data contain the near universe of publicly accessible Danish online job posts from 2007 to 2017. Duplicates are removed and the data cleaned before machine reading the job posts. HBS extracts the date on which a given job post was posted online, a firm ID, and an occupation code. If the firm identifier is not listed directly in the job post, HBS imputes it from publicly accessible registers using the firm name listed in the job post. Importantly, HBS also extracts keywords from the raw text in the job post. In many ways, the resulting data are similar to the US job postings data supplied by Burning Glass Technologies.

We group individual keywords from the job posts into 9 different skills categories, using the categories in (2018). We do this by manually assigning the most frequently occurring keywords (around 2000 terms) to a skill category or noise tag. The remaining words are categorized using synonyms or machine-learning methods based on each word's dictionary definition (see Jensen, 2024). The most frequent keywords for each skill category are reported in Table A.2

A.2. Danish labour force survey

Similar to the job postings data, we also group individual keywords from the self-reported skills use into 9 different skills categories. Free-text answers to the task/skill use question are cleaned by removing stop words (e.g., "and", "or") and are spell checked. Next, the same mapping of keywords to skill groups used with the JP are used for the remaining LFS words, categorizing the majority of them. We then manually categorize approximately 800 of the most frequent additional keywords from the LFS text data, such that slightly more than 75% of all keyword observations from the LFS are categorized. As mentioned above, we interpret a reported task as the utilization of a task-specific skill, and therefore prefer to refer to the measures extracted from the LFS as the main skills used on the job. The most frequent keywords for each skill category are reported in Table A.3.

A.3. Data match

To understand the representativeness of our samples, we compare the population of professionals, technicians, and associate professionals, clerical workers, and service and sales workers who are in the first year of a new private-sector job (either in a new occupation or in a new firm) as captured by Danish register data, shown in Column 1 of Appendix Table A.1, to the JP-Population Matched sample, shown in Column 2. Compared to the population as a whole, individuals in the JP-Population Matched sample are slightly younger and less experienced. This stems from the fact that entry-level jobs are more often posted on an online platform relative to more senior jobs, which are often filled either internally without advertisement or via established networks, explaining the relatively lower levels of experience shown in Column 2.

Next, we look at the subset of individuals from the register data who have answered the LFS in order to compare their self-reported main skills to the skills advertised in the JP. Column 3 in Appendix Table A.1 presents the results from merging the LFS with register data that capture first-year individual-level employment spells in the private sector so that firm identifiers can be appended. In the LFS sample, workers are less likely to be students and more likely to be older and live outside of Copenhagen. In our regression specifications, we include controls for age, education, experience, and municipality in order to account for this finding. Column 4 presents the subsample for whom information is available on both the job posting skills and self-reported

Table A.2
Keywords accounting for top 50% of character keyword observations, JP.

Keywords - English	Keywords - Danish	Skill	Frequency
Committed	Engageret	Character	309 701
Responsibility	Ansvar	Character	299 171
Self Employed	Selvstændig	Character	298 694
Professional	Faglig	Character	191 637
Friendly	Venlig	Character	186 248
Active	Aktiv	Character	184 816
Flexible	Fleksibel	Character	181 899
Honest	Ærlig	Character	132 007
Nature	Natur	Character	128 892
Dynamic	Dynamisk	Character	121 374
Positive	Positiv	Character	118 939
Open	Åben	Character	108 771
Personal	Personlig	Character	99 536
Joy	Glæde	Character	98 588
Professional	Professionel	Character	97 590
Structured	Struktureret	Character	96 414
Good Mood	Godt Humør	Character	91 386
Humor	Humor	Character	80 531
Targeted	Målrettet	Character	79 881
Burner	Brænder	Character	79 728
Drive	Drive	Character	78 783
Informal	Uformel	Character	75 214
Overview	Overblik	Character	71 501
Busy	Travl	Character	68 029
Order	Orden	Character	67 111
Stable	Stabil	Character	66 880
Values	Værdier	Character	65 210
Respect	Respekt	Character	62 315
Solution	Løsning	Cognitive	117 829
Logical	Logisk	Cognitive	75 394
Research	Forskning	Cognitive	59 794
Optimization	Optimering	Cognitive	48 250
Analysis	Analyse	Cognitive	29 163
Issues	Problemstillinger	Cognitive	24 744
Technical	Teknisk	Computer, General	132 799
SUPPRESS[a]	SUPPRESS[a]	Computer, General	105 486
System	System	Computer, General	58 856
Data	Data	Computer, General	40 914
IDENT[a]	Ident	Computer, Specific	65 442
Program	Program	Computer, Specific	28 246
Padding	Padding	Computer, Specific	27 177
Platform	Platform	Computer, Specific	21 438
E Security	E Security	Computer, Specific	7048
Server	Server	Computer, Specific	7031
Hardware	Hardware	Computer, Specific	6279
Databases	Databaser	Computer, Specific	6043
Service	Service	Customer service	237 279
Sell	Sælge	Customer service	199 260
Customer	Kunde	Customer service	198 198
Orders	Ordrer	Customer service	73 799
Guide	Vejlede	Customer service	37 558
Serves	Betjener	Customer service	34 490
Economy	Økonomi	Financial	60 573
Budget	Budget	Financial	42 324
Financial Accounting	Regnskab	Financial	40 103
Purchase	Indkøb	Financial	36 767
Resources	Ressourcer	Financial	31 373
Turnover	Omsætning	Financial	31 139
Margin	Margin	Financial	24 407
Accounting	Bogføring	Financial	20 284
Import	Import	Financial	19 540
Bookkeeping	Bogholderi	Financial	18 373
Reconciliation	Afstemning	Financial	17 999
Balance	Balance	Financial	17 929
Management	Ledelse	Management	105 495
Plan	Planlægge	Management	96 861
Operation	Drift	Management	92 595
Implement	Implementere	Management	66 631
Administration	Administration	Management	64 596
Coordinate	Koordinere	Management	63 942
Supervision	Supervision	Management	45 611
Control	Styring	Management	33 495
Management	Forvaltning	Management	32 727
Organize	Organisere	Management	29 302
Cooperation	Samarbejde	Social	440 544
Team	Team	Social	363 728
Communication	Kommunikation	Social	149 524
Extroverted	Udadvendt	Social	126 718
Social	Social	Social	98 160
Dialog	Dialog	Social	70 567
Danish	Dansk	Writing/Language	160 608
English	Engelsk	Writing/Language	109 639
Write	Skrive	Writing/Language	105 023
Language	Sprog	Writing/Language	68 069

Notes:
[a] IDENT refers to words that have been anonymized as they could otherwise potentially identify a firm. SUPPRESS refers to the group of words of sufficiently low frequency. Keeping the actual low frequency words would make it possible to identify individual observations from the raw data.

Table A.3
Keywords accounting for top 50% of character keyword observations, LFS.

Keywords - English	Keywords - Danish	Skill	Frequency
Responsibility	Ansvar	Character	5003
Personal	Personlig	Character	3471
Research	Forskning	Cognitive	3578
Analysis	Analyze	Cognitive	2330
Mathematics	Matematik	Cognitive	1690
Technical	Teknisk	Computer, General	3265
Software	Software	Computer, General	2023
System	System	Computer, General	1919
Program	Program	Computer, Specific	1608
Server	Server	Computer, Specific	1045
Graphic	Grafisk	Computer, Specific	920
Edb	Edb	Computer, Specific	845
System Developer	Systemudvikler	Computer, Specific	555
Sell	Sælge	Customer service	21 094
Customer	Kunde	Customer service	18 319
Serves	Betjener	Customer service	6846
Service	Service	Customer service	4838
Expediting	Ekspederer	Customer service	2613
Financial Accounting	Regnskab	Financial	7995
Bookkeeping	Bogholderi	Financial	5833
Purchase	Indkøb	Financial	3996
Economy	Økonomi	Financial	3293
Accounting	Bogføring	Financial	2811
Management	Ledelse	Management	23 071
Administration	Administration	Management	6560
Manager	Manager	Management	3816
Plan	Planlægge	Management	3340
Manager	Direktør	Management	3048
Operation	Drift	Management	2389
Department Manager	Afdelingsleder	Management	2376
Social	Social	Social	12 957
People	Folk	Social	3459
Dansk	Dansk	Writing/Language	2993
Write	Skrive	Writing/Language	2728
English	Engelsk	Writing/Language	1337

main skills (i.e., the result of merging the JP and LFS samples), and as expected, the resulting sample looks quite similar to the JP-Population Matched sample shown in Column 2. The takeaway from this exercise is that the matched JP-LFS (the estimation sample) is representative of the JP-Population Matched sample. We refer to Jensen (2024) for results on a larger sample of workers with only JP skills and wage regressions including occupation and firm fixed effects.

A.4. Keywords

In Table A.2 and Table A.3, we show the keywords of each skill category for the LFS and JP skills. As a robustness check, we have performed the wage regressions where the skill categories from the JP skills only include key words that are also in the top 50 percent of the LFS skills. The coefficients in wage regression do not significantly change and we therefore conclude that differences across skill measures in the individual keywords included in the skill are not driving the results.

Appendices B to D. Additional results, robustness, and extensions

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.labeco.2024.102661.

Data availability

The data that has been used is confidential.

References

Adams, A., Balgova, M., Qian, M., 2020. Flexible work arrangements in low wage jobs: Evidence from job vacancy data. In: CEPR Discussion Paper Series, (DP15263).

Alekseeva, L., Azar, J., Giné, M., Samila, S., Taska, B., 2021. The demand for AI skills in the labor market. Labour Econ. 71 (May), 102002.

Arntz, M., Gregory, T., Zierahn, U., 2016. The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis, Vol. 189. OECD Social, Employment and Migration Working Papers, oecd.

Atalay, E., Phongthiengtham, P., Sotelo, S., Tannenbaum, D., 2020. The evolution of work in the United States. Am. Econ. J.: Appl. Econ. 12 (2), 1–36.

Autor, D.H., Levy, F., Murnane, R.J., 2003. Skill demand, inequality, and computerization: Connecting the dots. In: Ginther, D.K., Zavodny, M. (Eds.), Technology, Growth, and the Labor Market. Kluwer Academic Publishers, Netherlands, pp. 107–129, number December, chapter 6.

Azar, J., Marinescu, I., Steinbaum, M., Taska, B., 2020. Concentration in US labor markets: Evidence from online vacancy data. Labour Econ. 66 (101886).

Bagger, J., Fontaine, F., Galenianos, M., Trapeznikova, I., 2022. Vacancies, employment outcomes and firm growth: Evidence from Denmark. Labour Econ. 75, 102103.

Beaudry, P., Green, D.A., Sand, B.M., 2016. The great reversal in the demand for skill and cognitive tasks. J. Labor Econ. 34 (1), S199–S247.

Black, S.E., Spitz-Oener, A., 2010. Explaining women' s success : Technological change and the skill content of women' s work. Rev. Econ. Stat. 92 (120), 187–194.

Blair, P.Q., Deming, D.J., 2020. Structural increases in skill demand after the great recession. AEA Pap. Proc. 110 (May), 362–365.

Botero, J.C., Djankov, S., La Porta, R., Lopez-De-Silanes, F., Shleifer, A., 2004. The regulation of labor. Q. J. Econ. 119 (4), 1339–1382.

Braxton, J.C., Taska, B., 2023. Technological change and the consequences of job loss. Amer. Econ. Rev. 113 (2), 279–316.

Clemens, J., Kahn, L.B., Meer, J., 2020. Dropouts need not apply? The minimum wage and skill upgrading. J. Labor Econ. 39 (S1), S107–S149.

Daly, M., Jensen, M.F., le Maire, D., 2022. University admission and the similarity of fields of study: Effects on earnings and skill usage. Labour Econ. 75, 102118.

De Dijn, M., Jacobs, C., Zenner, E., Ihalainen, L., Palander-Collin, M., Peterson, E., Arens, S., De Baar, M., Touwen, J., Heyvaert, L., 2023. Skills as stepping stones for employability: Perception research into the skills of Humanities students. Arts Humanit. High. Educ. 22 (2), 194–210.

Deming, D.J., 2022. Four facts about human capital. J. Econ. Perspect. 36 (3), 75–102.

Deming, D.J., Kahn, L.B., 2018. Skill requirements across firms and labor markets: Evidence from job postings for professionals. J. Labor Econ. 36 (S1), S337–S369.

Deming, D.J., Noray, K., 2020. Earnings dynamics, changing job skills, and STEM careers. Q. J. Econ. 135 (4), 1965–2005.

Fluchtmann, J., Glenny, A.M., Harmon, N., Maibom, J., 2022. Unemployed Job Search Across People and Over Time: Evidence from Applied-For Jobs. Working Paper.

Forsythe, E., Kahn, L.B., Lange, F., Wiczer, D., 2020. Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims. J. Public Econ. 189.

Goos, M., Manning, A., Salomons, A., 2014. Explaining job polarization: Routine-biased technological change and offshoring. Amer. Econ. Rev. 104 (8), 2509–2526.

Grinis, I., 2019. The STEM requirements of "Non-STEM" jobs: Evidence from UK online vacancy postings. Econ. Educ. Rev. 70, 144–158.

Groes, F., Kircher, P., Manovskii, I., 2015. The U-shapes of occupational mobility. Rev. Econ. Stud. 82 (2), 659–692.

Heckman, J., Landersø, R., 2022. Lessons for Americans from Denmark about inequality and social mobility. Labour Econ. 77 (August), 101999.

Hershbein, B., Kahn, L.B., 2018. Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. Amer. Econ. Rev. 108 (7), 1737–1772.

Javorcik, B., Stapleton, K., Kett, B., O'kane, L., 2020. Unravelling Deep Integration: Local Labour Market Effects of the Brexit Vote. Oxford Discussion Paper, pp. 5–6, (November).

Jensen, M.F., 2024. Gender Differences in Returns to Skills: Evidence from Matched Vacancy-Employer-Employee Data. Working Paper.

Kettemann, A., Mueller, A.I., Zweimüller, J., 2018. Vacancy durations and entry wages: Evidence from linked vacancy-employer- employee data. In: IZA Discussion Paper Series, (11852).

Kircher, P., 2022. Job search in the 21st century. J. Eur. Econom. Assoc. 20 (6), 2317–2352.

Kirkeboen, L.J., Leuven, E., Mogstad, M., 2016. Field of study, earnings, and self-selection. Q. J. Econ. 131 (3), 1057–1111.

Kreiner, C.T., Svarer, M., 2022. Danish flexicurity: Rights and duties. J. Econ. Perspect. 36 (4), 81–102.

Levine, R., Rubinstein, Y., 2017. Smart and illicit: who becomes an entrepreneur and do they earn more? Q. J. Econ. 132 (2), 963–1018.

Marinescu, I., Rathelot, R., 2018. Mismatch unemployment and the geography of job search. Am. Econ. J.: Macroecon. 10 (3), 42–70.

Mincer, J., 1974. Schooling, experience, and earnings. Human behavior & social institutions no. 2.

Modestino, A.S., Shoag, D., Ballance, J., 2016. Downskilling: changes in employer skill requirements over the business cycle. Labour Econ. 41, 333–347.

Modestino, A.S., Shoag, D., Ballance, J., 2020. Upskilling: Do employers demand greater skill when workers are plentiful? Rev. Econ. Stat. 102 (4), 793–805.

Papageorge, N.W., Ronda, V., Zheng, Y., 2019. The Economic Value of Breaking Bad: Misbehavior, Schooling and the Labor Market. Technical Report, National Bureau of Economic Research.

Spence, M., 1978. Job market signaling. In: Uncertainty in Economics. Elsevier, pp. 281–306.

Spitz-Oener, A., 2006. Technical change, job tasks, and rising educational demands: Looking outside the wage structure. J. Labor Econ. 24 (2), 235–270.