Skill Demand and Wages. Evidence from Online Job Postings in Austria

Lennart Ziegler*

University of Vienna

22 February 2019

PRELIMINARY DRAFT - PLEASE DO NOT CITE OR CIRCULATE WITHOUT PERMISSION

Abstract

What are the current skill requirements on the labor market and to what extent do these skills explain wages? While most studies analyze test score or survey data, this paper follows a different approach and infers skills from online job ads. Using more than 100,000 job postings published on Austria's major employment website, we identify the most common skill requirements mentioned in job descriptions. Because employers in Austria are legally required to state the minimum remuneration for advertised positions, it is possible to estimate the returns to these skill requirements using ad-level variation. Accounting for required education, firm and occupation fixed-effects, we find a robust association between the number of skill requirements and posted wages. In particular, job ads with many skill requirements pay substantially higher wages. While we estimate large effects for managerial and analytical skills, returns to most soft skills are small.

Keywords: Job postings, skills, wage differentials

JEL Classification Numbers: J23, J24, J31

*Address: University of Vienna - Department of Economics, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. E-mail: lennart.ziegler@univie.ac.at.

1 Introduction

Due to technological progress, labor markets around the world have been facing substantial changes in the demand for skills and their associated returns. A popular approach to measure skill demand is to refer to occupational dictionaries such as DOT or O*NET, which quantify the skill content of occupations and can readily be related to measures of pay or firm performance (see e.g. Autor et al., 2003). While these dictionaries provide a comprehensive and detailed summary of required skills, they only measure the average skill content of each occupation. In the analysis of skill returns, it is thus not feasible to account for unobserved differences between occupations. This paper follows an alternative approach and infers skill measures from online job ads which allow to gain valuable insights into the recent demand for skills and skill returns. Using more than 100,000 job posts published on the major employment website in Austria, we analyze job descriptions and identify the 14 most common skill requirements mentioned in ad texts. A useful feature of Austrian job ads is that employers are legally required to specify the minimum remuneration for each advertised vacancy. While employers are free to set higher wages, the posted wages cannot be below the industry- and occupation-specific collective bargaining wages. Exploiting ad-level variation in these wage posts, we can estimate associations between skill requirements and prospective pay. This paper is also a methodological contribution as we discuss the opportunities and limitations of job ads as a measure for skill demand in the labor market. More specifically, we examine what kind of information job postings comprise, how measured effects should be interpreted and to what extent measurement error can bias those estimates.

Numerous empirical studies have contributed to the literature on skill returns (Heckman et al., 2006; Lindqvist and Vestman, 2011; Autor and Handel, 2013; Hanushek et al., 2015). While most of these papers use common supply-side measures of cognitive and non-cognitive skills such as standardized test scores, we analyze skill requirements that are specified by the employer. In that sense, our analysis complements the existing literature by focussing on the demand side of skills. This study also relates to several previous studies on job vacancies. A rising number of papers exploits the new opportunities that online job boards offer to empirical researchers, ranging form studies on gender discrimination (Kuhn and Shen, 2012) to the effects of unemployment insurance programs (Marinescu, 2017). Two other recent studies have analyzed data on vacancies posted by the Austrian public employment services. Lalive et al. (2015) use the job listings to study market externalities of unemployment insurance programs, and Kettemann et al. (2018) examine the impact of vacancy duration on starting wages. Compared to the data that we analyze in our paper, the two studies focus on vacancies which cover an earlier period (1987-2014) and do not contain ad text information to infer detailed skill requirements.

Closely related to our paper is a recent study by Deming and Kahn (2018) who are among the first to analyze the impact of skill measures derived from keywords in US job ads. The empirical analysis shows strong correlations between skill requirements and average pay, and they also find evidence for complementarities between cognitive and non-cognitive skills. Because we follow a data driven approach to identify the most frequent skill requirements mentioned in job ads, the skill dimensions analyzed in our study differ somewhat from the ten general skills in Deming and Kahn (2018). Another key difference concerns the measure of earnings. Since potential earnings are usually not reported in vacancy posts in the United States, they supplement the sample of job ads with external earnings data and calculate average pay by occupations and metropolitan statistical areas. Using data on Austrian vacancies instead, we are able to exploit variation in posted wages between ads which allows to account for unobserved fixed-effects of firms and geographical regions.

In the first part of the analysis, we identify the most common keywords that describe the skill content of vacancies and group these into 14 different skill types. Despite a relatively short average text length of 180 words, we measure, on average, 1-2 skills per job ad. Vacancy posts for high educated workers report more than twice as many skills than those for low educated. Whereas soft skills such as reliability and teamwork competence are in high demand among employers who look for workers with a vocational degree or less, language and analytical skills are often frequented in posts for university graduates.

The second part of the analysis relates the skill measures to posted wages. Accounting for required education, and firm and occupation fixed-effects, we find that one additional skill increases posted wages by about one percent. Especially job ads with many skill requirements are associated with substantially higher wages. Furthermore, we estimate significant differences by individual skill types. Entrepreneurial and leadership skills show the largest impact, increasing wages by 8-10 percent. Whereas analytical skills and other hard skills also have positive returns, many soft skills such as stress-tolerance and reliableness are not associated with higher pay. Although the estimated returns to communication skills are relatively small, we find evidence for strong interaction effects with analytical skills. This is in line with recent empirical evidence for the US (Deming and Kahn, 2018; Weinberger, 2014), suggesting that the labor market is increasingly characterized by strong complementarities between social skills and cognitive skills. The large pool of job posts also allows to compare job ads between urban and rural districts in Austria. Consistent with the common notion that high productivity in metropolitan areas attracts higher skilled workers, the spatial analysis shows that some skill groups face a higher demand in more urban areas. For the same skills, we also estimate a higher wage return in these districts.

Because skill requirements are inferred from specific keywords, our estimates might be affected by measurement error. To analyze potential biases, we calculate potential effect sizes for different levels of under- and over-detection of skills and conclude that the underlying measurement error would have to be unrealistically large to explain the observed differences in skill returns. As an additional robustness check, we replicate our analysis using job ads of the largest private provider in Austria. Estimated wage effects are very similar to those obtained for our main sample, showing that the observed impact is not specific to the pool of ads on a specific job board.

The remainder of this paper proceeds as follows. In the next section, we provide details on wage setting and regulations for job postings in Austria. Section 3 outlines the setup of Austria's major online job board and describes the dataset. In particular, we explain how skills are inferred from job ads, and provide the corresponding statistics. The main estimation results are presented in Section 4. Section 5 analyzes spatial effects and provides additional robustness checks with regard to measurement error and representativeness. Finally, Section 6 concludes.

2 Institutional framework

Nearly 98 percent of workers in the Austrian labor market are covered by collective bargaining agreements (CBAs).[1] While most agreements are negotiated on the industry level, there also exist some firm-level CBAs which are often complementary to industry agreements and specify better conditions for workers. The negotiated pay scale differs by occupation and often increases with age and experience. Wages specified in the agreements are legally binding and cannot be undercut but firms are free to offer higher wages. In that sense, CBA wages should be interpreted as a lower bound for actual wages. As many firms point out, the degree of overpay depends on both qualification and experience. Next to the pay structure, CBAs regulate the extent of working hours and other working conditions. A national minimum wage does not exist in Austria.

A unique feature of the Austrian labor market is that as of March 2011, employers are required to specify the respective CBA starting gross wage in job postings. Since August 2013, a similar regulation also holds for the few sectors that are not covered by collective bargaining agreements. Here, the posted wage should be at least the lower bound of the typical pay for an offered vacancy. Posted wages should exclude bonus or other extra payments, and the unit of time has to be reported (hour, month or year). Employers who do not comply with these rules may be fined for violation by local authorities. Whereas the posted wage cannot be below the CBA figure, employers can post higher wages. If firms want to attract qualified applications and are able to overpay, it can be advantageous to post a higher wage. It is thus probable that for some ads observed wages exceed the wage of the respective collective bargaining agreement.When we control for firm and occupation fixed-effects in the empirical analysis, most of the variation in wages should result from these ads.

Skill effects on posted wages can differ from those on actual pay. In fact, it is likely that candidates who better fit the specified job profile should be able to negotiate a higher final wage. For this reason, our estimates can be seen as lower bound for the impact on final wages.

Little empirical evidence exists on the wedge between collective bargaining

[1]See report by OECD (2012).

wages and actual wages. While final wages are observed in the Austrian social security records, it is in many cases hard to figure out which bargaining agreement applies to a specific spell. A policy report by Leoni and Pollan (2011) estimates a gap of around 20 percent for the industrial sector. A few job ads in our sample contain some information on expected differences. When employers directly enter a vacancy post on the job board, they have the option to report separately the CBA wage, the final wage or both. A small subset of firms reports both figures ($N = 964$). While this sample is too small for the main analysis, it allows to obtain an estimate for the difference between final pay and collective bargaining wage. Using all job ads in which both wages are reported, we calculate an average log wage difference of 0.153 ($SD = 0.136$). It is not clear what motivates firms to (not) reveal both wage figures. Reporting employers might find it difficult to fill vacancies and use the comparison to attract more applications, which makes more sense when the wedge is large. If there is selection in reporting with respect to pay differences, one would thus expect that the average difference across all ads is smaller. To examine whether selection based on observables matters, we use all characteristics for which we have sufficient observations in the subsample (extent of work, required education and state of workplace) to predict log wage differences for all job ads. The predicted mean (0.154) is very comparable to the unadjusted mean.

3 Data

3.1 Job board

The *AMS e-Jobroom* is the online job board of the public employment administration (AMS) in Austria. Having about $60,000$ active postings at a time, this platform offers by far the biggest pool of vacancies in Austria. A comparison to the largest private competitor is provided in Section 5.3. According to a representative quarterly survey among establishments conducted by Statistics Austria, the AMS has covered about 50-60 percent of all open vacancies in recent years.

Companies can either directly enter job posts on the website or inform the public employment office about open vacancies. In the latter case, the AMS processes the provided information and updates the corresponding job posts daily. We observe that only about 10 percent of vacancies are directly posted by companies. The daily checks by the AMS greatly reduce the number of inactive vacancies, which are common on many private job boards (Cheron and Decreuse, 2016). Contrary to many private competitors, the employment office does not charge companies for job ads. To be eligible for postings on the job board, the offered jobs have to be located in Austria.

Job seekers can either use the open search mask or filter ads by several characteristics such as location or occupation. It is also possible to register for free and set up an application profile, which enables firms to get in contact with registered job searchers. Job posts on the e-Jobroom contain both structured and unstructured information. Structured information are characteristics which are reported in separate columns for all vacancies and mostly correspond to the filters in the search mask. These data include firm name, location (on post code level), occupation, required education and extent of work. Ads are grouped using the AMS occupation classification which distinguishes 450 different occupations. Its partition level is thus similar to the five-digit Standard Occupational Classification (SOC) of 461 occupation used in the United States. In total, there exist five categories of education requirements, ranging from compulsory schooling to university education. Unstructured information are attributes described in the open text section and need to be extracted using text pattern matching. A typical ad text is about 100-300 words long and often contains a short description of the company and the advertised job followed by a characterization of the profile of a suitable applicant. Characteristics that we obtain from the ad text are all skill requirements and the posted wage.[2]

All available vacancies were scraped from the AMS e-Jobroom in two to four weeks' intervals between November 2018 and January 2019. To prevent double-counting of vacancies, we drop ads that are equal in all observable characteristics.[3]

[2]If hourly or annual wages are listed, we convert wages to monthly pay. For a few job ads which have been posted directly by the employer, wages are reported in a separate column. Unreasonably low values (below 500 euros) and high values (above $10,000$ euros) are dropped from the sample to minimize measurement error.

[3]Because of the daily checks for inactivity by the AMS, we expect that most duplicate observations are vacancies that have not been filled yet. Another possibility is that some repeated observations are reposts of vacancies with the exact same profile.

Table 1: Job ad characteristics

	Mean	Std. Dev.	Min	Max
Monthly salary	2,052.23	585.33	500	10,000
# words	181.87	86.85	10	857
Urban area	0.26	0.44	0	1
Extent of work:				
- Full-time	0.74	0.44	0	1
- Part-time	0.17	0.38	0	1
- Full- or part-time	0.09	0.29	0	1
Education:				
- Compulsory schooling	0.33	0.47	0	1
- Vocational training	0.51	0.50	0	1
- Higher voc.-techn. schools/gymnasium	0.10	0.30	0	1
- Applied university	0.02	0.13	0	1
- University	0.04	0.19	0	1

# ads per firm	1	2-5	6-10	11-100	>100
Firms (N=22,547)	9,906	9,878	1,551	1,112	100

NOTE: N=100,872. A vacancy is classified as *urban* if the job is located in a district of the five largest Austrian cities (Vienna, Graz, Linz, Salzburg, Innsbruck and Klagenfurt).

For the empirical analysis, we make a few sample restrictions. A small share of posts on the job board do not contain any ad text. Because skill requirements cannot be inferred for these vacancies, we restrict our estimation sample to ad texts with at least 10 words, which corresponds to 95 percent of all scraped ads. Another five percent of observations do not report the salary or minimum required level of education for the vacancy. Finally, we drop vacancies for which the firm name is not reported (6-7 percent). The final estimation sample contains $100,872$ job ads, or 84 percent of the raw dataset.

Table 1 summarizes the main characteristics of the job ads. We observe that the length of an average ad text is 180 words. Firms post a gross wage of, on average, $2,050$ euros per month. While most job ads are for full-time positions, we find substantial heterogeneity in the level of required education. Contrary to many other online job boards, the e-Jobroom lists many vacancies for lower educated workers with basic school education or vocational training. About 16 percent require a higher secondary degree or university education. According to the Austrian labor force survey in 2017, approximately 30 percent of workers fall into the latter category. Although the reported levels are *minimum* requirements and should therefore understate the average education of successful applicants, the large difference to the composition of the current workforce suggests that jobs for lower educated workers are overrepresented on the job board. Finally, the last two rows of Table 1 report the distribution of job ads across firms. To include firm fixed-effects in the analysis, we need to observe multiple job ads per firm. The distribution shows that most firms on the job board post more than one job ad during the period of observation.

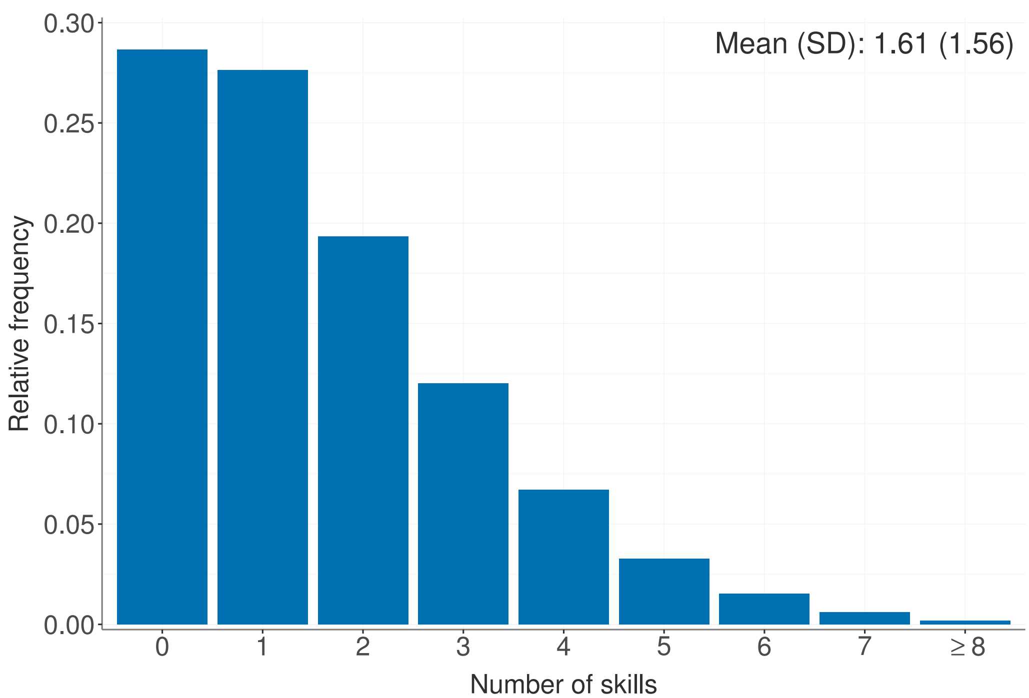
Table 2: Classification of skills

Skill group	**Specific skill**	**Examples (in German)**
Analytical skills	Analytical skills	Problemlösungskompetenz, analytisches Denken
Communication skills	Communication skills	Kommunikationsfähigkeit, kommunikativ
Managerial skills	Entrepreneurial skills	Unternehmerisches Denken/Geist
	Leadership skills	Führungsstärke, Führung
Other hard skills	Programming	Programmierer, Programming, Python, SQL
	MS-Office skills	MS Office, MS-Office, Microsoft Office
	Foreign language	Englisch, Französisch, Italenisch
Other soft skills	Teamwork	Teamwork, gerne im Team arbeiten, Teamplayer
	Organizational skills	Organisationstalent, Organisationsfähigkeit
	Self-reliance	Eigeninitiative, eigenverantwortlich
	Assertiveness	Durchsetzungsvermögen
	Creativity	Kreativität, kreativ
	Stress tolerance	Belastbarkeit, Stressresistenz, Stress
	Reliability	Zuverlässigkeit, Verlässlichkeit

3.2 Skill requirements

To identify the most common skill requirements, we split up the ad texts into words and rank them according to their overall frequency. Next, we filter out all words that describe skill requirements.[4] These can either be general skills such as being communicative or specific skills like a programming language. Finally, we group the terms into skill categories. Table 2 provides an overview of this classification procedure. Column two and three list the 14 identified skills along with an excerpt of used keywords for each skill. To reduce the dimensionality of the skill set, we further group these skills into five major skill groups. From the list of skills in Table 2, it becomes apparent that job posts mainly describe skill requirements for higher skilled occupations. These vacancies are often characterized by a higher complexity of tasks, which can explain a higher skill request. Also, the skill set of university graduates might be more diverse than that of workers with a vocational degree.[5]

Figure 1: Distribution of skills per job ad



[4]In some cases, we have to rely on multiword expressions to avoid over-detection caused by expressions that do not necessarily refer to the skill profile.

[5]Note that Austria has a well-developed apprenticeship system with national standards and centralized examination, which helps employers to better assess the skills of applicants with vocational training.

It is clear that job post cannot provide a full characterization of a worker's profile. Compared to occupational dictionaries, ad texts provide a shorter, more superficial description of the required skill set. Furthermore, it is more difficult to describe the relative importance of skills. For this reason, firms might put more emphasis on major skill requirements, while minor skills are more likely to be omitted. In contrast to continuous skill measures, estimated effects should thus be interpreted as the impact of prioritized skills within a given occupation. It is also possible that skill requirements are only observed with some degree of measurement error because of over- or under-detection of described skill attributes. Potential sources of measurement error and its consequences will be discussed in Section 5.2.

Table 3: Skill shares by education

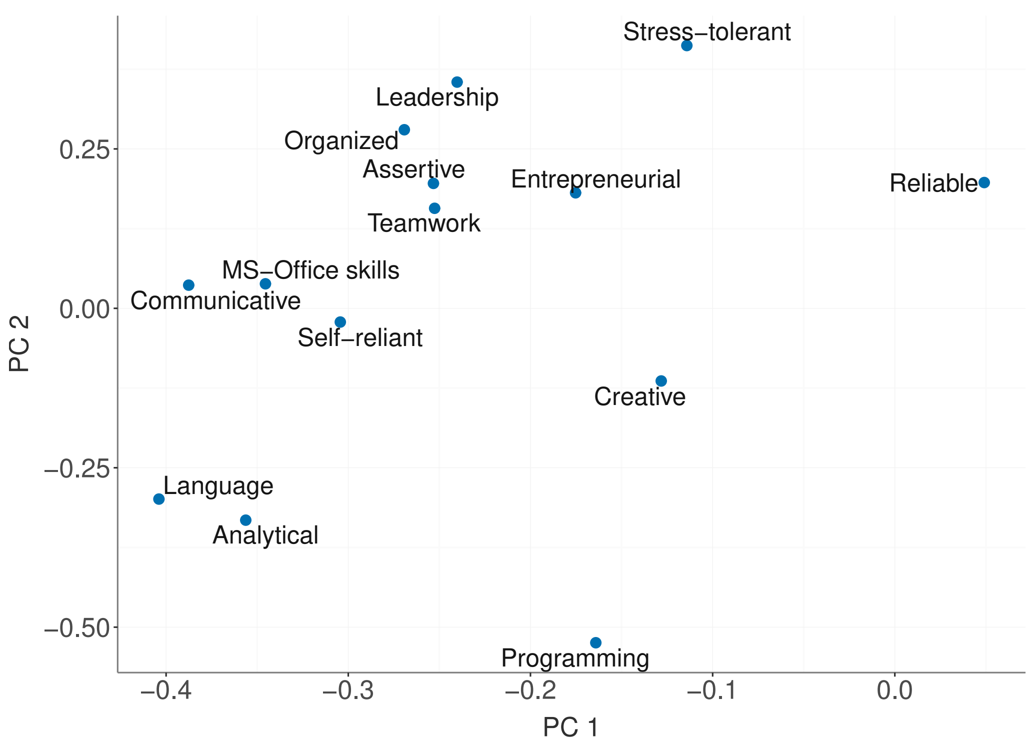
Low educated		**High educated**	
Skill	Share	Skill	Share
Reliable	0.30	Language	0.45
Teamwork	0.26	Communicative	0.43
Communicative	0.16	Teamwork	0.39
Stress-tolerant	0.15	Self-reliant	0.30
Self-reliant	0.13	Analytical	0.28
Leadership	0.07	MS-Office skills	0.22
Language	0.06	Reliable	0.19
MS-Office skills	0.06	Programming	0.19
Creative	0.04	Leadership	0.15
Analytical	0.04	Stress-tolerant	0.12
Organized	0.03	Creative	0.09
Programming	0.02	Organized	0.09
Assertive	0.01	Assertive	0.06
Entrepreneurial	0.01	Entrepreneurial	0.04
Mean skills: 1.35		Mean skills: 3.00	
N=84,941		N=15,931	

NOTE: **Low education**: Compulsory schooling and vocational training. **High education**: Higher voc.-techn. schools/gymnasium and (applied) university.

As illustrated in Figure 1, we observe on average 1.6 skills per job ad. Whereas 29 percent of posts do not list any of the discussed skill types, around six percent request more than four skills. Table 3 shows the relative frequency of all skill types separately for low and high educated workers. As expected, most skills are more often mentioned in job posts for vacancies that require higher education. As reported at the table bottom, job ads for these workers specify, on average, more than twice as many skill requirements. Whereas social skills like communication and teamwork are frequent in both groups, cognitive skills such as analytical skills and programming skills are rarely asked of lower educated workers.

When examining the joint mentioning of skills within job ads, we find evidence for substantial correlations between skill requirements (see Table A.1 in the appendix). To better understand these associations, we conduct a principal component analysis.

Figure 2: Principal components of skills



By estimating orthogonal linear combinations of the skill indicators that maximize the variance in the data, principal components allow to visualize the comovement of skills in job postings.[6] Figure 2 illustrates the first two principal components of all 14 skill types. While we observe similar principal components for most soft skills, those of hard skills are more scattered. This suggests that soft skills more often refer to a similar skill profile. Programming knowledge and reliableness stick out as these skills are often mentioned in isolation. Entrepreneurial skills are very centrally located in the graph of Figure 2. This is consistent with the view that entrepreneurs need to possess a variety of skills (see Lazear, 2004).

4 Skill returns

4.1 Estimation strategy

To analyze the impact of skills on wages, we estimate the following regression equation

$$log(wage_{ijk}) = \alpha_j + \beta_k + S'_{ijk}\gamma + X'_{ijk}\delta + u_{ijk}$$

where S_{ij} denotes the skill measure of ad i posted by firm j for occupation k. As skill measures, we use (i) the number of skills ($\# \, skills$), and (ii) a vector of skill type indicators. Vector X_{ijk} contains other job characteristics reported in the ad, including required level of education and a variable that indicates whether the ad was entered directly by the employer or via the public employment administration. Finally, α_j and β_k denote firm and occupation fixed-effects, respectively. The identification assumption for the marginal returns to skills (γ) is that conditional on education, location, occupation and firm fixed-effects, stated skill requirements are not correlated with unobserved factors that explain wage differences.

A standard wage determination model predicts that wages are determined by both productivity and bargaining power.[7] In this framework, we think of skill returns as returns to productivity. It is thus important to account for differences

[6] To account for large differences in skill shares, we normalize the standard deviation of each skill indicator to one.

[7] See e.g. Cahuc et al. (2006) for a theoretical framework.

in bargaining power. As mentioned in the previous section, most collective bargaining agreements in Austria are industry-wide, with a few firm-level exceptions. Controlling for firm fixed-effects allows to account for any confounding effects on firm- or industry-level which are constant between listings.[8] Furthermore, this specification allows to take out differences in the wage policy of firms, which might be correlated with skill requirements. Empirical studies that estimate firm-specific wage premiums often find evidence for substantial firm heterogeneity (e.g. Abowd et al. 1999; Card et al. 2013).

Although job descriptions in ad postings cannot provide a full characterization of the desired applicant profile, the occupation itself contains information on skill requirements. It should be known to applicants which skills are necessary to master all duties in a specific occupation. Including occupation indicators allows a within-occupation comparison of skill types which reduces the threat of biased estimates due to unreported 'self-explanatory' skills requirements that are correlated with our skill measures. A drawback of this specification is that we can only identify returns to skills which vary within occupations. We do not expect that many skills fall into this category as this information should already be communicated by the occupation itself. For completeness, all main regression results are reported with and without occupation fixed-effects.

4.2 Wage regressions

Following the estimation equation outlined above, we next examine the impact of skills on posted wages. Estimation results for five different specifications are given in Table 4. The upper row reports point estimates for the impact of the number of skills. In all regressions, we obtain small but robust coefficients, indicating a one percent increase in wages per additional skill. Neither occupation nor firm or region fixed-effects appear to influence the association between number of skills and prospective pay. As shown in Section 3.2, there exists substantial heterogeneity in the number of skill requirements across job ads. It is possible that also marginal effects change with increasing skill requirements. Using the richest specification

[8] Here, we assume that firms operate in one industry, which applies to the vast majority of Austrian firms. Even though individual bargaining power also influences final wages, this channel cannot affect the skill returns to posted wages.

Table 4: Wage regressions

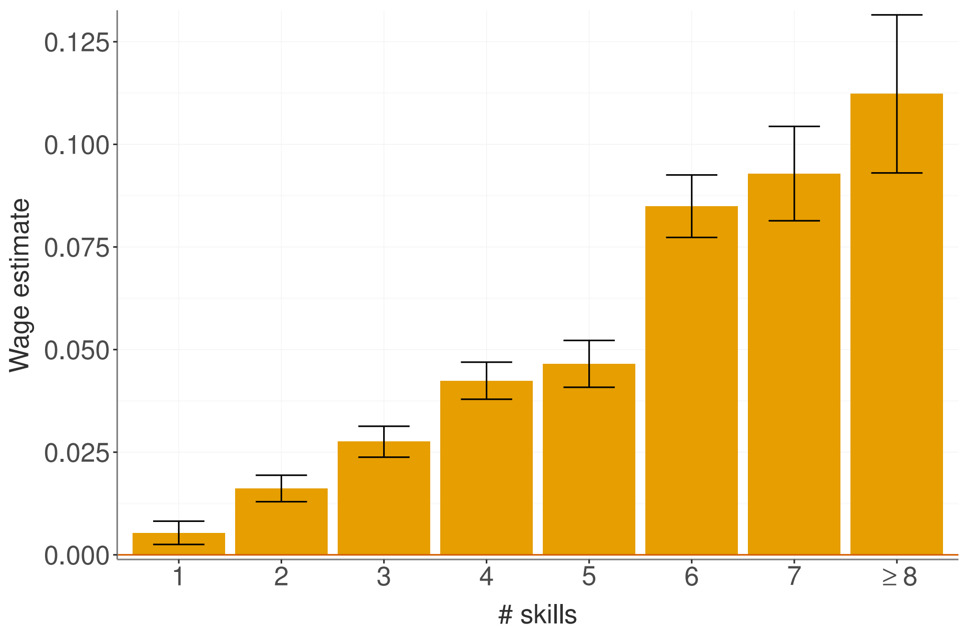
	(1)	(2)	(3)	(4)	(5)
# skills	**0.009**	**0.010**	**0.011**	**0.012**	**0.012**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Analytical skills	**0.085**	**0.039**	**0.050**	**0.030**	**0.030**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Communication skills	**-0.024**	**0.009**	-0.003	**0.008**	**0.008**
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Managerial skills	**0.051**	**0.065**	**0.084**	**0.079**	**0.079**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Other hard skills	**0.034**	**0.026**	**0.017**	**0.019**	**0.019**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Other soft skills	-0.002	**-0.003**	-0.003	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Occupation FE		✓		✓	✓
Firm FE			✓	✓	✓
Region FE					✓

NOTE: N=100,872. **Bold** coefficients indicate significance at the 1%-level. All regressions control for education category and whether the ad was placed directly by the employer. See Table 2 for skill group classification.

(c.f. column (5) of Table 4), we estimate wage differences for each skill count relative to wages in job ads that do not mention any of the described skills. The plot of Figure 3 shows that the estimated impact sharply increases with the number of listed skills. Compared to job posts with just one mentioned skill, wages in posts with more than seven skill requirements are higher by a margin of more than 10 percent. The graph also demonstrates that the overall effect is not driven by changes at the extensive margin. In fact, estimated wages are very similar in job ads with zero and one skill listing.

The lower panel of Table 4 reports results for separate skill groups. In the first rows, we focus on analytical skills and communication skills, which are among the most frequent skills asked of higher educated workers. Furthermore, they serve as a good proxy for cognitive and non-cognitive capabilities. Controlling for education, we find a positive impact of 8.5 percent for analytical skills. The point estimate declines by more than 50 percent when we additionally account for

Figure 3: Wage effect by number of skills

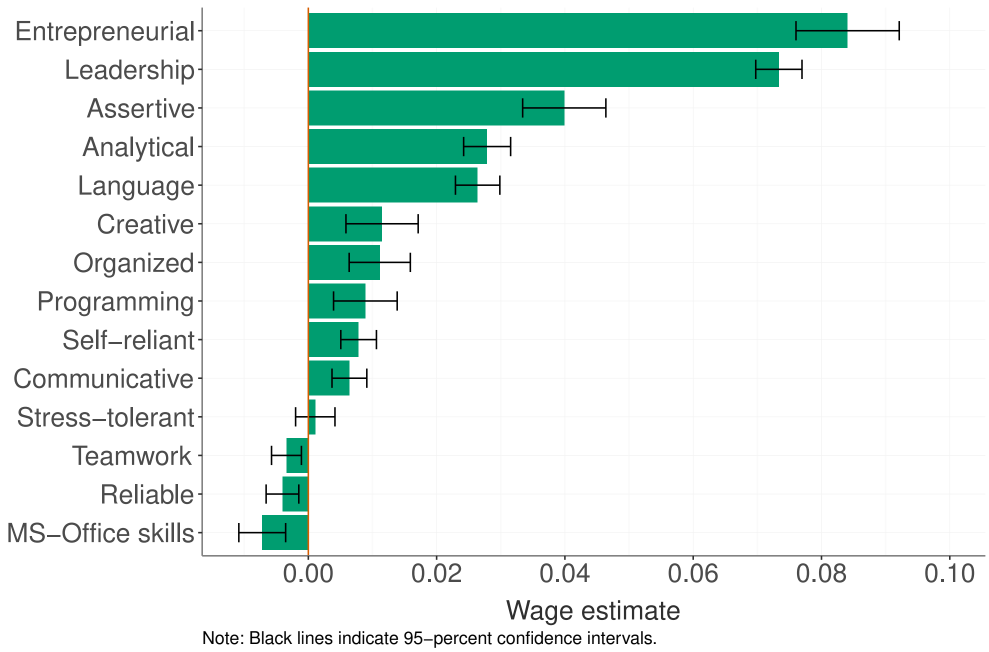


Note: Black lines indicate 95−percent confidence intervals.

occupation fixed-effects. In the last specification, we estimate a wage difference of three percent. On the contrary, communication skills are associated with lower wages in the first specification. Controlling for differences between occupations removes this negative effect. Instead, we measure a small impact of about one percent, which is robust to firm and region fixed-effects. These large changes in effect size compared to the simple regression show that communication skills are more frequented in lower paying occupations while analytical skills are more common in higher paying occupations. By far, the strongest wage effect can be observed for managerial skills. When firm differences are taken into account, employers who mention these requirements post, on average, eight percent higher wages. The last two rows Table 4 report estimates for other hard or soft skills. While the remaining soft skills have no joint effect, we find that ads which require additional hard skills pay somewhat higher wages.

To analyze the impact of specific skills, we next estimate wage regressions for all 14 skill types. Again, we account for education, firm, occupation and region fixed-effects. As shown in Figure 4, there exist large differences in the impact

Figure 4: Wage effect by skill type



of hard skills. For language skills, we estimate a relatively large effect similar in size to the coefficient on analytical skills. The impact of programming skills is only half of that, and MS-Office skills are even associated with lower wages. This might partly be due to the fact that returns to profession-specific skills are absorbed by the occupation fixed-effect. Indeed, we measure that the occupation of a programmer pays above average wages (also conditional on education). When we exclude occupation fixed-effects from the regression, the point estimate on programming skills becomes 4-5 times larger. As for hard skills, we also observe some heterogeneity in wage estimates of soft skill requirements. Whereas assertiveness is associated with a relatively strong return of four percent, stress-tolerance, teamwork capabilities and reliability show no positive effects.

To uncover potential complementarities between hard and soft skills, we estimate additional wage regressions, which include interaction terms of analytical and communication skills. As reported in Table A.2 in the appendix, we find in all specifications evidence for strong interaction effects. Whereas the point estimates of the separate skills are relatively small, requiring both skills jointly increases posted wages by around two percent.

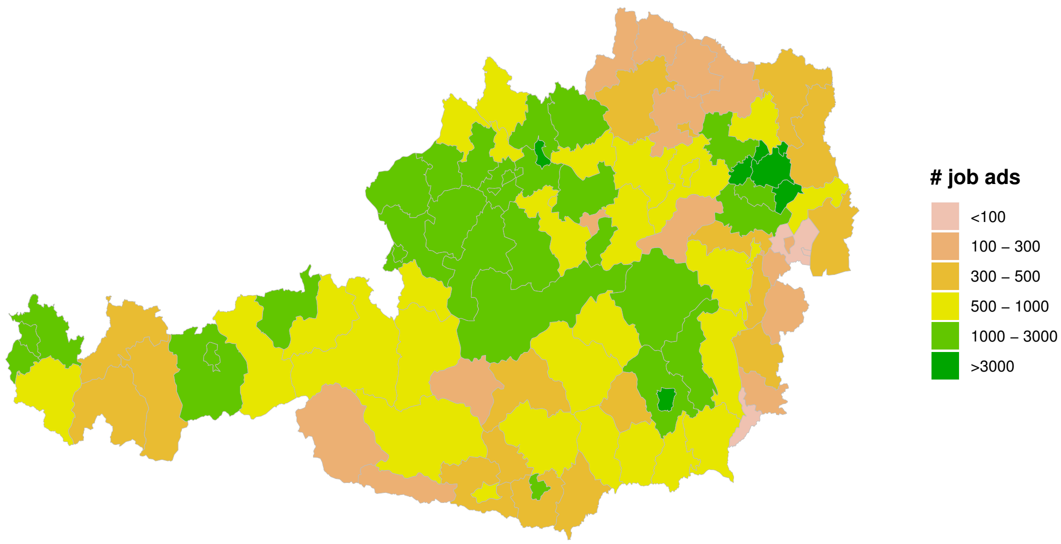
4.3 Interpretation of findings

Compared to the findings of Deming and Kahn (2018), who estimate the effect of skills on wages using US job ads, our estimates are overall smaller. This might be explained by differences in the measure of pay. While they use average wages by region and occupation as outcome, we focus on posted wages, which can be lower than final wages. As discussed in Section 2, we expect the impact on final wages to be larger because better suitable candidates should be able to negotiate a higher wage. Because our estimates suggest that the difference between posted and final wages is rather small (15 percent), it is possible that other factors such as a different estimation strategy explain lower effects, too. Using ad-level variation in wage postings allows to take out firm and region fixed-effects and, thus, accounts for potential biases. Indeed, our estimates in Table 4 show that most coefficients decrease when we control for these differences.

Institutional differences between the United States and Austria might also play a role. The Austrian labor market is more homogenous both in terms of workforce characteristics and pay. First, wages in Austria are less dispersed than in the United States, which suggests that skill returns are lower, too. Second, it is possible that the skill profile of Austrian workers is less heterogeneous. One reason could be differences in education. While US institutions of higher education differ substantially in the quality of education (e.g. measured in terms of school resources), differences among universities and colleges in Austria are much less pronounced. The uniform apprenticeship system further guarantees homogeneous skill profiles among lower educated workers. This might make it less important to filter applicants by skill requirements in job postings. If less firms explicitly refer to specific skills, the explanatory power in wage regressions should be lower.

Consistent with previous studies is our finding of strong interaction effects between analytical and communications skills in all regressions. As pointed out by Weinberger (2014), skill-biased technical change in the labor market induces a rising demand for workers with both cognitive and non-cognitive skills. This is also line with our finding that entrepreneurial and leadership skills have the

Figure 5: Job ads by district



largest wage effects. These skills require cognitive and non-cognitive capabilities, and are often mentioned jointly with other hard and soft skills as shown in the previous section.

5 Discussion

5.1 Spatial differences

Because of differences in amenities, infrastructure and industry spillovers, it is likely that skill demand and skill returns differ across regions (Glaeser and Mare, 2001; Moretti, 2004). To examine whether such differences exist, we construct two spatial measures that exploit variation between the 94 administrative districts in Austria. Figure 5 illustrates the frequency of job ads by district. Although the number of postings greatly varies, our sample is large enough to also contain a substantial number of job ads in most rural districts. To measure spatial differences, we classify a job location as urban if the respective location belongs to the districts of the six largest Austrian cities (Vienna, Graz, Linz, Salzburg, Innsbruck and Klagenfurt). As shown in Table 1, this is true for about 26 percent of the sample. To obtain an alternative spatial measure, we merge data on district-level population density to our sample of job ads. This should serve as a continuous proxy for urbanity. For the estimation, we use the logarithm of population per squared kilometer and standardize it to have a standard deviation of one.

The corresponding regression results are given in Table 5. Column (1) and (2) report spatial differences in skill shares for both measures. Accounting for differences in occupations and education, the spatial wedge in skill demand is relatively modest. Employers in urban districts, list 0.05 more skill requirements. Analytical, communication and other hard skills are mentioned 0.5-1 percentage points more often in urban areas. For the remaining two skill groups, we do not observe significant differences. When using the logarithm of population density as explaining variable, we find again similar differences although the difference in analytical skills is not present here. A one standard deviation increase in population density is associated with 0.8 percentage point higher demand for communication skills and other hard skills.

The remaining two columns report estimated coefficients on the interaction terms of spatial and skill measures in wage regressions. Compared to the overall return of one percent per additional skill requirement, we find that this impact is around 20 percent larger in the cities. We further estimate significant differences in skill returns for analytical and communication skills. The point estimates indicate approximately one percent higher returns in urban areas.

These findings are consistent with a spatial equilibrium model of the labor market with heterogeneous workers and imperfect labor mobility (see Enrico, 2011). Firms in metropolitan areas tend to be more productive and face a higher demand for skills. To attract the most talented workers, they pay higher returns than in more rural areas.

5.2 Measurement error

Since we use text pattern matching to infer the skill content of job postings, it is possible that the skill measures suffer from some degree of measurement error. More specifically, our keywords may fail to identify skills (*under-detection*) or wrongly attribute skills (*over-detection*) in a few job ads. Some employers might

Table 5: Spatial differences

	Difference in shares		Diff. impact on log(wage)	
	D(Urban)	Log(pop dens.)	D(Urban)	Log(pop dens.)
# skills	**0.049**	**0.025**	**0.003**	**0.001**
	(0.011)	(0.006)	(0.001)	(0.000)
Analytical skills	**0.005**	0.002	**0.013**	**0.006**
	(0.002)	(0.001)	(0.004)	(0.002)
Communication skills	**0.007**	**0.008**	**0.007**	**0.007**
	(0.003)	(0.002)	(0.003)	(0.001)
Managerial skills	0.002	0.000	-0.001	0.001
	(0.003)	(0.001)	(0.003)	(0.002)
Other hard skills	**0.008**	**0.009**	0.001	-0.001
	(0.003)	(0.002)	(0.003)	(0.001)
Other soft skills	0.005	-0.001	0.001	-0.002
	(0.004)	(0.002)	(0.002)	(0.001)

NOTE: N=100,872. **Bold** coefficients indicate significance at the 1%-level. Log population density is standardized to have mean zero and standard deviation one. All regressions control for education category, firm and district fixed-effects, and whether the ad was placed directly by the employer. See Table 2 for skill group classification.

paraphrase required skills in the job description instead of naming them directly. Typing errors or the use of infrequent terms that we have not identified as common keywords can be other sources of mis-measurement. Conversely, it is also possible that we wrongly attribute keywords to the applicant's profile. One example are keywords that describe the workplace rather than the applicant. Although both errors are rather unlikely, it is informative to analyze whether minor inaccuracies can lead to significant changes.

To quantify the impact of over- and under-detection of skills, we assume in the following that the error (similar to a classical measurement error) is not correlated with the error term u_{ij} in our wage regressions. In other words, posted wages conditional on all observables should not differ by the degree of measurement error in job ads. Under this assumption, we can back out actual wage effects for given rates of under- and over-detection. Let indicator variables $\tilde{S}_i$ and S_i denote observed and actual occurrence of a specific skill in job ad i. Over- and under-detection rates, $p_o = P(S_i = 0|\tilde{S}_i = 1)$ and $p_u = P(S_i = 1|\tilde{S}_i = 0)$, are defined as the probabilities that we attribute the skill to ads which do not contain it, and vice versa. The true skill effect on outcome y_i is given by

$$\beta = \underbrace{E(y_i|S_i = 1)}_{\mu_1} - \underbrace{E(y_i|S_i = 0)}_{\mu_0}$$

Under over- and under-detection of skills, we instead observe

$$\begin{aligned} b &= E(y_i|\tilde{S}_i = 1) - E(y_i|\tilde{S}_i = 0) \\ &= (1 - p_o)\mu_1 + p_o\mu_0 - [(1 - p_u)\mu_0 + p_u\mu_1] \\ &= (1 - p_o - p_u)\beta \end{aligned}$$

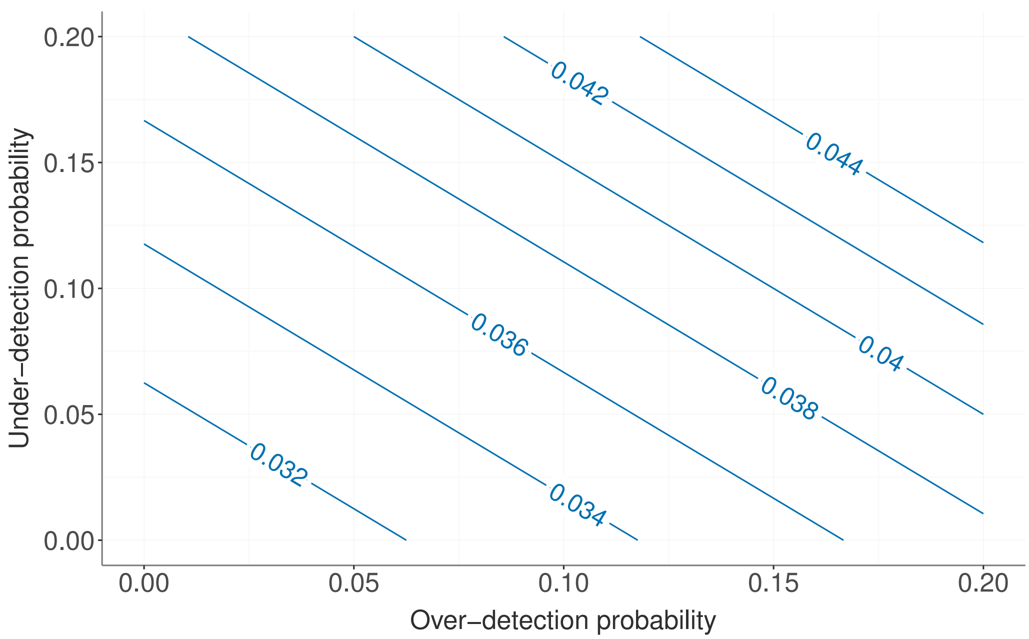
Most likely, p_o and p_u are lower than 0.5, which would mean purely random assignment of $\tilde{S}_i$.[9] As a result, mis-measurement of skills leads to a downward bias of the true impact towards zero. Because the bias increases exponentially in error rates, small levels of measurement error should not have a considerable impact on our estimates. The relation between true and estimated effect sizes is illustrated in Figure 6. Using the estimated impact of analytical skills as an example ($b = 0.03$), we plot the combinations of p_o and p_u that are associated with different levels of actual skill returns. The contour plot shows that even under substantial over- and under-detection rates of around 20 percent, the underlying effect would not be much larger than the estimated effect.

5.3 Alternative data source

Job ads posted on the AMS job board might not be representative for labor demand in the entire labor market. As discussed above, the public employment service only covers about 50-60 percent of all vacancies, and there might be selection with respect to education, profession and other unobservables. Next to the AMS, several private providers operate job posting websites in Austria. To test the robustness of our estimates, we redo the previous analysis of skill requirements and wage effects using job ads of the biggest competitor on the Austrian

[9] Fore p_o and p_u larger than 0.5, assignment of skills would be reversed, leading to effect size measures between zero and $-\beta$.

Figure 6: Actual effect sizes under measurement error ($b = 0.03$)



labor market. Compared to our main data source, the number of job ads is much smaller on this website. Whereas around $60,000$ job ads are available on the AMS job board at a time , this provider only lists about $20,000$ posts. Furthermore, the classification of job ads on this website is much less detailed. Job ads are clustered into 20 occupation groups, and there is no classification of required education. This makes it more difficult to appropriately take out confounding effects due to differences in occupation and education requirements. Despite these limitations, it is informative to examine as a robustness check whether skill shares and wages estimates are comparable to those found in the main analysis.

All available job posts were scraped from the website every two to four weeks between June and October 2018. Identification of skills and wages follows again the procedure outlined in Section 3.2. Restricting the sample to observations with non-missing information on skills, location and salary, the final sample consists of $41,374$ ads. Because job posts are not classified by education, we infer from the ad text whether any higher education is required. Measured skill shares and other descriptives statistics are provided in Table A.3 in the appendix. Compared to our main sample, we observe clear differences in most characteristics. Job ads on the website of the private competitor are, on average, longer, offer higher wages and are more often located in the six urban districts. Also, the measured skill requirements are substantially higher. This observation is consistent with the notion that private employment websites overrepresent job posts for professionals (see e.g. Deming and Kahn, 2018). In fact, the estimated skill shares are comparable in size to those of ads on the AMS job board for high educated workers as reported in Table 3. The comparison of ad characteristics between the two samples suggests that both sources complement each other. Whereas the private website clearly lacks the majority of job posts for low-skilled workers, the public employment service lists fewer job postings for high skilled workers.

Table A.4 in the appendix reports results of the corresponding wage regressions. The five specifications closely mimic the regression analysis of our main sample. Yet, due to the more superficial classification of ads, the included controls for occupation and education are less detailed. Estimated effects for both the number of skills and the skill types are largely in line with the main results presented in Section 4. One additional skill leads to about one percent higher wages, which is robust to the inclusion of occupation, firm and district fixed-effects. For the separate skill groups, we find initially again large point estimates that tend to decrease in richer specifications. Managerial skills are associated with the highest wage gains, followed by analytical skills. In this sample, the impact of other soft skills remains even in the last specification significantly negative. Overall, coefficients are somewhat larger than in the main analysis. This is consistent with our finding that effect sizes are less pronounced in specifications with many controls. If we could use the exact same set of controls, regression estimates for both samples might be even closer aligned.

This shows that our findings are robust to the use of a different job board as datasource. Although the private website covers very different vacancies with respect to most observable characteristics, estimated wage associations remain similar.

6 Conclusion

Given the vast amount of publicly available data on the Internet, many economists have started to exploit its potential and provide new evidence for a variety of research questions. This paper argues that online job ads contain useful information for the analysis of skill demand on the labor market. Due to regulations in Austria which require employers to report a lower bound for wages in job posts, the analysis of Austrian ads additionally allows to use variation between job posts in the estimation of skill differentials.

Analyzing a large number of job ads posted on the major online job board in Austria, we identify the 14 most common skill requirements. Even when we control for education, firm and occupation fixed-effects, we estimate a small but robust return to the number of mentioned skills. While managerial and analytical skills are associated with relatively high returns, most soft skills have small effects on posted wages. Consistent with predictions of a standard spatial model, the analysis also provides evidence for higher skill demand and skill returns in more urban districts of Austria. To investigate potential biases due to measurement error, we calculate potential skill effects assuming different degrees of over- or under-detection of skill requirements. This exercise shows that even under considerable measurement error, actual effect sizes would not be very different from obtained estimates.

Our analysis of job board data comes with two shortcomings. First, observed skill requirements can only serve as a rough proxy for a job's actual skill content. Because employers need to depict the job profile in just a few sentences, job ads cannot comprise the same level of detail as occupational dictionaries. Furthermore, it is more difficult to describe the relative importance of mentioned skills. Another limitation is that we estimate returns in terms of posted wages, which not always equal final wages. Although we observe that final wages are only about 15 percent higher, it is possible that the difference is correlated with skill requirements. Highly skilled applicants that meet all stated requirements should have a better bargaining position and negotiate higher wages. As a result, skill differentials will be higher for final wages, suggesting that the estimated effects should be interpreted as a lower bound for the skill return in terms of actual pay.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Autor, D. H. and Handel, M. J. (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, 31(S1):S59–S96.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Cahuc, P., Postel-Vinay, F., and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):323–364.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics*, 128(3):967–1015.

Cheron, A. and Decreuse, B. (2016). Matching with phantoms. *Review of Economic Studies*, 84(3):1041–1070.

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

Enrico, M. (2011). Local labor markets. In *Handbook of Labor Economics*, volume 4, pages 1237–1313. Elsevier.

Glaeser, E. L. and Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics*, 19(2):316–342.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., and Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, 73:103–130.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Kettemann, A., Mueller, A. I., and Zweimüller, J. (2018). Vacancy durations and entry wages: Evidence from linked vacancy-employer-employee data. Technical report, National Bureau of Economic Research.

Kuhn, P. and Shen, K. (2012). Gender discrimination in job ads: Evidence from China. *Quarterly Journal of Economics*, 128(1):287–336.

Lalive, R., Landais, C., and Zweimüller, J. (2015). Market externalities of large unemployment insurance extension programs. *American Economic Review*, 105(12):3564–96.

Lazear, E. P. (2004). Balanced skills and entrepreneurship. *American Economic Review*, 94(2):208–211.

Leoni, T. and Pollan, W. (2011). Lohnentwicklung und Lohnunterschiede in der Industrie seit 2000. *WIFO Monatsberichte (Oktober)*.

Lindqvist, E. and Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 3(1):101–28.

Marinescu, I. (2017). The general equilibrium impacts of unemployment insurance: Evidence from a large online job board. *Journal of Public Economics*, 150:14–29.

Moretti, E. (2004). Human capital externalities in cities. In *Handbook of Regional and Urban Economics*, volume 4, pages 2243–2291. Elsevier.

OECD (2012). Labour losing to capital: What explains the declining labour share? Technical report, OECD Employment Outlook 2012.

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, 96(4):849–861.

Appendix

Table A.1: Skill correlation matrix

	Ana.	Comm.	Entr.	Lead.	Progr.	MS-Off.	Lang.	Team.	Org.	Self-rel.	Asser.	Creat.	Str.-tol.
Comm.	**0.18**												
Entr.	**0.07**	**0.09**											
Lead.	**0.05**	**0.13**	**0.13**										
Progr.	**0.20**	**0.06**	**-0.01**	**-0.03**									
MS-Off.	**0.15**	**0.17**	**0.05**	**0.08**	**-0.02**								
Lang.	**0.27**	**0.23**	**0.06**	**0.06**	**0.18**	**0.27**							
Team.	**0.09**	**0.16**	**0.02**	**0.08**	**0.05**	**0.10**	**0.10**						
Org.	**0.07**	**0.13**	**0.09**	**0.14**	**-0.02**	**0.16**	**0.09**	**0.07**					
Self-rel.	**0.13**	**0.16**	**0.05**	**0.10**	**0.08**	**0.14**	**0.14**	**0.12**	**0.10**				
Asser.	**0.12**	**0.12**	**0.09**	**0.14**	**-0.01**	**0.12**	**0.11**	**0.05**	**0.11**	**0.06**			
Creat.	**0.06**	**0.05**	**0.02**	**0.02**	**0.06**	**0.02**	**0.07**	**0.05**	**0.06**	**0.08**	**0.01**		
Str.-tol.	**-0.01**	**0.08**	0.00	**0.09**	**-0.05**	**0.03**	-0.00	**0.14**	**0.08**	**0.05**	**0.05**	-0.01	
Rel.	**-0.03**	**-0.02**	**-0.02**	**-0.05**	**-0.04**	-0.01	**-0.05**	**0.08**	-0.00	**-0.02**	**-0.02**	**-0.02**	**0.06**

NOTE: N=100,872. **Bold** coefficients indicate significance at the 1%-level.

Table A.2: Wage regressions - Complementarities

	(1)	(2)	(3)	(4)	(5)
Analytical skills	**0.072**	**0.032**	**0.037**	**0.021**	**0.021**
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Communication skills	**-0.028**	**0.007**	**-0.007**	**0.004**	**0.004**
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Analyt. × comm. skills	**0.031**	**0.016**	**0.029**	**0.021**	**0.021**
	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)
Firm FE			✓	✓	✓
Occupation FE		✓		✓	✓
Region FE					✓

NOTE: N=100,872. **Bold** coefficients indicate significance at the 1%-level. All regressions control for education category, whether the ad was placed directly by the employer and remaining skill types. See Table 2 for skill group classification.

Table A.3: Job ad characteristics (alternative job board)

	Mean	Std. Dev.	Min	Max
Monthly salary	2701.85	866.08	500	10000
# words	266.57	95.55	32	1101
Urban area	0.63	0.48	0	1
University required	0.24	0.43	0	1
For job beginners	0.04	0.19	0	1
# days online	5.26	3.80	0	15
# skills	3.29	1.75	0	12

Skill	Share	Skill	Share
Communicative	0.48	Leadership	0.18
Language	0.48	Stress-tolerant	0.15
Teamwork	0.40	Programming	0.14
Self-reliant	0.34	Organized	0.10
Analytical	0.32	Creative	0.09
MS-Office skills	0.28	Assertive	0.07
Reliable	0.20	Entrepreneurial	0.05

NOTE: N=41,374. A vacancy is classified as *urban* if the job is located in a district of the five largest Austrian cities (Vienna, Graz, Linz, Salzburg, Innsbruck and Klagenfurt).

Table A.4: Wage regressions (alternative job board)

	(1)	(2)	(3)	(4)	(5)
# skills	**0.012**	**0.013**	**0.013**	**0.014**	**0.014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Analytical skills	**0.089**	**0.056**	**0.059**	**0.045**	**0.044**
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Communication skills	**0.013**	**0.020**	**0.015**	**0.016**	**0.016**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Managerial skills	**0.122**	**0.121**	**0.128**	**0.117**	**0.118**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Other hard skills	**0.054**	**0.048**	**0.014**	**0.021**	**0.020**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Other soft skills	**-0.066**	**-0.046**	**-0.028**	**-0.019**	**-0.018**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Occupation group FE		✓		✓	✓
Firm FE			✓	✓	✓
District FE					✓

NOTE: N=41,374. **Bold** coefficients indicate significance at the 1%-level. All regressions control for the number of days that an ad is online and indicators for university education and whether an ad is specifically for job beginners. See Table 2 for skill group classification.