On the Importance of Soft Skills in the U.S. Labor Market

Antonio Biondi Zacharias Kountoupis Joan Rabascall Costa Marco Solera

Barcelona Graduate School of Economics

June, 2020

Abstract

This paper explores the role of soft skills in the U.S. labour market. According to the previous literature, these skills – also called non-cognitive- are crucial as they allow firms to lower coordination costs by trading job tasks more efficiently. We look at both sides of the labour market. On the demand side, we collect 4,980 job ads from U.S. job portals through a web scraping technique, finding that larger firms require more job tasks and soft skills in their ads than the small and the medium ones. On the supply side, we match the skills from the O*NET dictionary with the Survey of Income and Program Participation (SIPP) of the United States from 2013 to 2016, estimating return to soft skills around 15% of hourly wage. Moreover, we find statistically significant soft skills wage premium in the big firms around 2.5%, up to 3.5% for highly educated workers. To the best of our knowledge, this is the first paper that finds a firm size wage premium for soft skills. These pieces of evidence suggest that larger enterprises are willing to pay more soft skills as they face higher coordination costs.

Keywords: United States, Non-cognitive, Skill, Ability, Wage Premium, SIPP.
JEL Classification: J22, J23, J24, J31

1 Introduction

An increasing literature is focusing on the role of "soft skills" as the critical driver for labour market success (Heckman et al. (2006); Weinberger (2014); Deming (2017)). A growing body of empirical evidence documents a reversal in demand for cognitive abilities and soft skills – also called "non-cognitive"- in the labour market. A large number of authors have shown substantial growth in the demand for occupations involving cognitive tasks starting from the '80s (Katz and Murphy (1992); Beaudry and Green (2005); Acemoglu (2002); Autor and Murnane (2003);Autor and Dorn (2013)). However, in recent years, some authors have found a stagnating or decreasing return to cognitive skills starting from the early 2000s in the U.S. (Castex and Dechter (2014); Beaudry and Sand (2016)). On the contrary, a secular increase in the return to non-cognitive skills has been observed from 1992 to 2013 (Edin, Fredriksson, Nybom, and Ockert (Edin et al.)). As pointed out by Autor (2015), a possible explanation is that soft-skills are associated with job tasks that are harder to replace by automation, as they are mainly composed of "tacit knowledge" that is hard to encode.

Empirical literature in the field faces several problems. First of all, defining "soft skills" could be problematic. Soft skills are described as the set of personality traits, motivations and preferences that are valued in the labour market (Sheikh (2015)): communication, cooperation, problem-solving, time management, critical thinking, leadership etc. Secondly, non-cognitive skills are typically hard to measure. Two approaches are generally followed in literature: on one side, the subjective measurement of the non-cognitive skills refer to the so-called "BigFive" (Conscientiousness, Openness, Emotional stability, Extraversion and Agreeableness), according to a personality survey collected by psychologists. On the other side, the objective measurement is derived from the classification of soft skills used by the Occupational Information Network (O*NET). In both cases, measurement errors arise when estimating personality traits. Moreover, the endogeneity problem could bias the estimation of non-cognitive skills on wage as the fact that people with higher soft skills tend to self-select into specific occupations.

Few papers estimate the impact of soft skills on wages. Dunifon (1998) evaluates the wage premium for personal efficacy, defined as the ability to reach long term goals through individual's intentional activities, estimated at around 14% in 1988-1992 in the U.S. Borghans et al. (2006) estimates a penalty around 5% for people skills in the U.S. and the United Kingdom, while Weinberger (2014) found a wage premium for leadership skills around 5.3% in the U.S. Similarly, there is no broad literature on the role of the non-cognitive skills in the labour market. From a theoretical point of view, soft skills allow workers in trading job tasks more efficiently and consequently reducing coordination costs within the firm (Becker (1992); Deming (2017)).

Our aim is to explore the importance of soft skills in the U.S. labour market and their impact on wages by evaluating both sides of the labour market. On the demand side, we create a dataset by collecting around five thousand job ads from the indeed.com job portal using a web scraping technique. We find that larger firms tend to require more job tasks and soft skills in their ads than small and medium ones. On the supply side, using data from the Survey of Income and Program Participation from 2013 to 2016, we estimate that soft skills account for around 15% of hourly wage in the U.S. Moreover, we find a statistically significant soft skills wage premium in the big firms around 2.5%, which increases to 3.5% for those highly educated. These results are in line with our assumption, according to which larger enterprises are willing to pay more soft skills as they face higher coordination costs.

The paper is organized as follows. Section II discusses related literature on cognitive and non-cognitive skills and labor market returns. Section III describes the datasets being used, and Section IV presents our empirical methodology. Section V summarizes the most relevant results and Section VI concludes.

2 The Economic Returns to Non-cognitive skills: Theory and Literature

The role of soft skills in the labor market and their impact on wage is still largely unknown. The definition of soft skills is quite broad, referring to non-cognitive skills as intangible, hard to measure and closely connected with individual attitudes, such as: problem-solving, critical thinking, cooperation, leadership and judgment, among others (Reber (1995)). They differ from the so called hard skills, which are easily observed or measured and are closely connected with an individual's knowledge. An increasing body of literature on non-cognitive skills has been growing recently. Borghans et al. (2006) showed a substantial growth in job tasks requiring soft skills from 1970 to 2002. Edin, Fredriksson, Nybom, and Ockert (Edin et al.) also documented a secular increase in returns to non-cognitive skills in Sweden, particularly pronounced in the private sector and the upper-end of the wage distribution. Bacolod et al. (2009) found a statistically significant correlation between requirements on cognitive and non-cognitive skills, supporting the idea that they complement each other. Measuring soft skills though implies dealing with problems such as endogeneity and measurement errors. For example, Heckman et al. (2006) recognize the potential presence of measurement errors when using score tests as a proxy for cognitive abilities, since cognitive skills are usually intertwined with non-cognitive skills.

To address this issue, Heckman et al. (2006) developed an hedonic model for wages that considers simultaneously cognitive and non-cognitive abilities, assuming independence between both sets of skills. Simultaneously, Sheikh (2015) and Heckman and Kautz (2012) suggested that reverse causality can also be present in such an approach. If the specification uses an "achievement test" as an explanatory variable for wages, this could generate a reverse causality problem because the estimates cannot distinguish whether higher skills cause higher wages or whether additional years of education cause both higher cognitive skills and higher wages. To tackle this inconvenience Maciente (2013) and Neves Jr et al. (2017) work under the premise of Spence's Signaling Model (Spence (Spence)), according to which, employees send signals of their skills to potential employers by their level of education, however, in this case, the signaling comes from the occupation each individual possesses. Matching this information with an objective measure of skills and abilities mitigates the impact of measurement errors.

2.1 Supply side literature

Regarding the impact of employees' soft skills on wages, a small empirical body of literature exists. On one side, soft skills are derived from big five subjective measures (Heckman and Kautz (2012)). This classification contains five types of personality traits: *neuroticism*, *extraversion*, *openness to experiences*, *agreeableness*, and *conscientiousness*. Borghans et al. (2006) estimated the return on wage for people skills, estimating a penalty (instead of a premium) around 5% in the United States labor market.

On the other side, soft skills are derived using an objective measure: job tasks and requirements from the Dictionary of Occupational Titles produced by the United States Department of Labor. Bacolod et al. (2009) found that the soft skills premium nearly doubled from 1968 to 1990. Moreover, they found a strong complementarity between cognitive and soft skills. Lindqvist and Vestman (2011) estimated the positive impact on wages of non-cognitive skills in the Swedish labor market.

Regarding the role of soft skills in the labor market, Becker and Murphy (1992) developed a model in which the degree of specialization of the firms depends on the cost of coordinating specialized workers. In this context, workers with higher non-cognitive skills are crucial to lowering coordination costs by trading tasks with other workers more efficiently. Following this approach, Deming (2017) found an increasing return for social skills in the United States Labor Market. However, when focusing on particular skills, not all non-cognitive traits may have a positive impact on labor outcomes. For instance, while leadership abilities have a significant effect on economic and employability outcomes (Kuhn and Weinberger (2005); Lundin et al. (2019)), agreeableness is punished, especially for women (Nyhus and Pons (2005); Heineck (2007)).

2.2 Demand side literature

The research literature on labor markets and skills has historically devoted more attention to the supply side of the market, i.e. on the skills of job-seekers and job-incumbents, than to the demand side of the markets, which refers to the skills that employers require to their future or current workers. However, there exists literature on the evolution of trends in job skills demands during the last decades which suggests a raise in educational, cognitive and interpersonal skill requirements, while craft skills, physical demands and the frequency of repetitive physical tasks has declined (Handel (2012)). Furthermore, Cunningham and Villaseñor (2014) points out that employers and educators have different understandings of the types of skills valued in the labor market. Hence, they suggest that policymakers need to re-conceptualize education and training systems.

Our analysis, building on the assumption that large firms in the labour market face higher coordination costs, aims to display the importance of soft skills in the labor market and signify their relative difference depending on the firm size. Our areas of interest to indicate this are wage premiums and frequency of demanded soft-skills by firm.

To our knowledge, there is no research that connects soft skills and firm size to wages premiums. As far as we know, there is also no literature regarding the skill requirements depending on the firm-size. Hence, this would be our contribution on the relevant literature on both sides of the labour market.

3 Data

For our analysis, we aim to investigate the effects of and non-cognitive skills in both the supply and demand sides of the labor market. For that purpose, we are using multiple data sets that will act as representations of both sides of the U.S. labour market.

3.1 Supply side dataset

For the supply side analysis, we use the *Survey of Income and Program Participation* (SIPP), which is conducted by the United States Census Bureau. This longitudinal survey provides information on an individual level, where each individual is reporting information on their income, employment status, occupation, and industry in which they operate among many others, all necessary for our analysis. Several controls such as education level, sex, firm-size of the firm that employs the individual, marital status and age are also included in the dataset. The collection of data from the Census Bureau is conducted in quarterly waves in the span of approximately four years, hence the progression of each individual regarding their employment, income and residence -among many others- is observed and allows for a more thorough analysis. Regarding the part of the data that we are adopting for our analysis, only individuals who reported a salary were considered, with the four waves being used taken from 2013 till 2016, with each wave representing a year.

The overall sample that satisfies the aforementioned characteristics comprise 79,823 individuals. Tables 1, 2, and 3 provide descriptive statistics for wages, years of education, occupation, firm-size, and differences between male and female employees. These descriptive statistics are characterizing the supply-side dataset and are driving the results of our empirical model. It is important to note that individuals which reported being entrepreneurs were excluded since we are interested in knowing the impact on salaries. Moreover, we excluded public servants because the development in the private sector is driven by market forces to a greater extent than in the public sector.

In order to extract the occupation and industry information which are included in the SIPP database, as they are given in standardized codes, we use O*NET database, which contains hundreds of standardized and occupation-specific descriptors on almost 1,000 occupations covering the entire U.S. economy. Referring to the international vocabulary, as it is listed in O*NET, we then correspond each occupation to a list of skills. Then these extracted skills can be used for our empirical methodology. Some examples of the skills that are found in the O*NET database and characterize SIPP's occupations can be found in the Appendix's tables.

3.2 Demand side dataset

For the demand side analysis, we constructed a dataset from scratch. Using the *indeed.com* job portal, we extracted information from job ads using a web-scraping method. In order to construct a dataset which will allow us to compare the effects of skill premiums between the demand and supply side, the database was created by extracting jobs postings -in the U.S. job market- of the ten most frequently found occupation categories in the SIPP database, which can be seen in Table 14 of the Appendix. The constructed dataset consists of 4,980 observations. For each extracted job advertisement, information on the offered salary, firm's name, location, and job description are included. This information is necessary for our analysis, as the job description will be used to extract information about the skills, abilities, and knowledge demanded by firms. The description of each job advertisement is then analyzed using a Text Mining approach, where each keyword and phrase is being corresponded to specific job skills characterizing the labor market. To order abilities, skills, and knowledge we refer to the classification provided by the *Factor Analysis* conducted in the supply side via SIPP, as it will be displayed in the Empirical Methodology section of the report.

An additional variable to include is the size of each firm posting a job advertisement. Since job ads usually do not include the firm size, we use the Dun & Bradstreet (D&B) Corporation database as source of information. D&B contains more than 265 million business records about different characteristics of companies worldwide, such as employees in each firm, local branches, subsidiaries, and parent companies. Using the firm-size definition according to number of employees, we allocate a firm size for each observation. It was decided to allocate firm-size according to the size of the parent company and not of the local branch that has posted a job advertisement. We categorize the firm size according to the SME definition: small firms have less than 50 employees, medium firms have more than 50 and less than 200 employees, large firms have more than 200 employees.

A limitation that exists in the extracted dataset that is used for the demand side analysis is the lack of salary information for an important share of observations. Thus, we adopt a different methodology for the demand-side analysis, as it will be analyzed. A caveat that could potentially make our results slightly inaccurate is the way that each firm officially presents itself. By that we mean the following: For financial and tax-reduction purposes companies 'divide' themselves, presenting their branches as individual entities. This has a direct effect on the number of employees each firm presents on its business record, potentially leading to incorrect firm-size allocations in our analysis.

According to the O*NET definitions, cognitive abilities influence the acquisition and application of verbal information in problem-solving. Regarding the skills, process skills refer to procedures that contribute to the more rapid acquisition of knowledge and skill across a variety of domains. They stand for critical thinking and monitoring (i.e.: to control the performance of oneself and other individuals). Content skills turn to background structures needed to work with and acquire more specific skills in a variety of different domains. They incorporate reading comprehension, active listening, writing, speaking, mathematics, science. Complex problem-solving skills pertain to developed capacities used to solve novel, well defined problems in complex, real world settings. Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions. Resource management skills speak on developed capacities used to allocate resources efficiently. They encompass management of financial/material/personnel resources and time management. Social Skills refer to developed capacities used to work with people to achieve goals. System Skills allude to developed capacities used to understand, monitor, and improve socio-technical systems. They include judgment and decision making, system analysis, and system evaluation. Knowledge is divided by sector: business & management, manufacturing, engineering & technology, math & science, health services, education & training, arts & humanities, laws & public safety, communication, transportation. These abilities, skills, and knowledge (ASK) are elements that are specific to each one of the occupations of both the SIPP database, and of the sample that we created. Through these elements we identify the soft-skills, and conduct our analysis.

4 Empirical Methodology

4.1 Supply Analysis

To verify and estimate the impact of the firm size on soft skill wage premium, we need to consider different identification issues. Firstly, the endogeneity problem must be taken into account: unobserved and time-invariant characteristics could affect the wage. Since we cannot use a *Fixed Effects* (FE) model as the skills would be absorbed by the individual fixed effects, we will rely on a *Pooled OLS* model and contextually we use cluster standard errors at an individual level. Moreover, there could be a presence of spatial sorting related to unobserved characteristics: the solution is provided by including location fixed effects (i.e. using the state of worker's residence) and a control for metropolitan areas. In addition, we control for occupation fixed effects to consider the job sorting. The empirical specification follows a Mincerian specification:

$$
\begin{aligned}
log\left(wage_{i,t}\right) = \alpha + \delta Skill_{i(k)} + \sum_{j=1}^{J} \beta_j Abil_{i(k)} + \sum_{l=1}^{L} \pi_l Know_{i(k)} + \\
+ \rho med_i + \tau big_i + \gamma Skill \cdot med_i + \sigma Skill \cdot big_i + \theta X_{i,t} + \eta_k + \eta_s + \eta_t + \varepsilon_{i,t} \quad (1)
\end{aligned}
$$

Where i is the worker, t is the year, k is the occupation and s is the state worker's residence. The dependent variable is the logarithm of hourly wage in U.S. dollars. We include several controls: gender, age, marital status, experience, education, race, being part of a union, living in metropolitan areas, interactions between marital status and female gender, and interactions between skills and abilities. Moreover, we create interactions between skills and dummy variables for firm size.

One issue of concern is the potential collinearity among abilities, skills, and knowledge of the information obtained from O*NET. To address this problem, we will group them together using a *Factor Analysis* approach. The results can be seen in table 4. We use *Factor Analysis* for each category separately. In the first two columns, we get the results for abilities: Factor 1 explains 77% of the variance of abilities, which are mainly driven by psycho motor, physical, and sensory abilities. Factor 2 is driven by cognitive and sensory abilities. Although Factor 2 has an Eigenvalue below 1 (the threshold), we include it into our analysis as it clearly explains something different compared to Factor 1. In the case of skills, Factor 1 explains the 90% of the variance, led by resource management, content, social, process, problem-solving, and system skills. On the other side, we discard Factor 2 as it explains the variance poorly and it is headed uniquely by technical skills. Regarding knowledge, Factor 1 embodies mainly humanities such as: business & management, health services, education & training, arts & humanities, laws, communication, and transportation.

Factor 2 is mainly driven by "hard" knowledge, such as: manufacturing, engineering, math & science. After predicting the variables according to the *Factor Analysis*, we obtain: two variables for abilities (that we call "physical" and "cognitive"), one variable for skills (we call it "soft skills"), and two for knowledge ("soft knowledge" the first one, "hard knowledge" the second one). We use them as the regressors on wages.

Finally, cognitive and soft skills could still be noisy measures of the same underlying ability. To further address this issue, we include interactions between cognitive ability and soft skills.

4.2 Demand Analysis

Regarding the empirical strategy adopted for the demand-side analysis of the labour market, the lack of information on salary does not allow us to make an analysis that estimates the non-cognitive skill premiums. For that reason, our demand-side analysis focuses on two different aspects, both important in understanding each firm's behavior and consideration of skills in the U.S. labor market, depending on its size.

One part that we examine is the frequency of the skills that are found in the demand-side dataset. Skills and abilities-such as Computer and Electronics knowledge, Management skills, and writing among many others- that each firm requires for its offered position are extracted from the Job description in each advertisement. Key phrases that correspond to each one of the skills being examined are used to determine whether they are demanded by them. By this, we will determine which primary skills are demanded by firms depending on its size. The second characteristic of the demand-side that is being examined is what firms are demanding by their potential employees compared with the standard skills required for each occupation-according to the O*NET specification. Hence we try to estimate whether firms tend to demand more than what they need.

5 Results

5.1 Supply Analysis

After predicting the variables through the *Factor Analysis*, we run the *Pooled OLS* on log hourly wage. Results can be seen in *Table 5*. Columns (1), (2) and (3) show the estimates for naïve regressions by including abilities, skills, and knowledge, firm size dummy variables and the interactions of our interest between soft skills and firm size. We also include an interaction between cognitive abilities and soft skills. Column (4) displays the estimates after controlling for a wide set of controls. Columns (5) and (6) include occupation and state fixed effects, respectively. Column (7) shows the results after including all the controls and adding occupation, state, and year fixed effects.

We found a statistically significant return to soft skills: an increase in one standard deviation in soft skills is associated with an increase in hourly wages of around 15%, while one standard deviation in cognitive abilities is associated with an increase on hourly wages of around 4%. Regarding our coefficient of interest, we found a statistically significant soft skills wage premium in big firms, around 2.5%, consistent with our prediction based on Becker (1992) model. Since soft skills tend to lower coordination cost, they are more valuable in larger firms.

To understand better what drives our results, we run other regressions by gender and education. Results can be seen in *Table 6*. Males do not show a soft skill premium in larger firms, no matter what their level of education is. On the contrary, we find a positive and large coefficient for females.

The soft skill wage premium is around 2.9% in larger firms, and it levels up to 4.5% when considering females with more than 12 years of education. There results suggest that soft skills trade job task more efficiently when increasing level of education.

Then, we run regressions on hourly wage by using each single soft skill as the regressor instead of the variable derived from the *Factor Analysis*. Regarding resource management skills, we find a wage premium for males. The premium is not significant for females, regardless their level of education. On the contrary, we find a statistically significant wage premium in large firms around 3.5% for system skills. The coefficient is not significant when considering only males; conversely it is significant and around 3.7% when considering only females; 5.6% when considering only females with more than 12 years of schooling. Regarding process skills, problem-solving and content skills we found similar results. As it can be observed in *Table 7, 8, 9, 10, 11 and 12*, the coefficients for the cognitive skills and soft skills are negative, implying that they are associated with lower wages. On the contrary, the interactions between these two variables are positive and statistically significant. This implies that labour market tends to value cognitive and non-cognitive skills jointly, whereas the occupations that require only one of them are associated with lower wages. These findings are consistent with the empirical literature.

We found a statistically significant wage premium in larger firms for: problem-solving skills (4% overall, 5.4% for female and 6% for female with more than 12 years of schooling); content skills (3.2% overall, 7.3% for female with more than 12 years of schooling); process skills (3.2% overall, 4.7% for female with more than 12 years of schooling); resource management skills (3.8% overall, 4.5% for male and 5.2% for male with more than 12 years of schooling); system skills (3.5% overall, 5.6% for female with more than 12 years of schooling).

5.2 Demand Analysis

By evaluating the importance of soft-skills from the Demand-side of the labor market, we intend to reinforce our findings displaying the wage premiums for non-cognitive skills for large firms, and display their higher needs for soft skills to cover their coordination costs. Without being able to extract the wage premiums for non-cognitive skills as they are offered by firms, we are extracting the amount of skills that are being demanded, to determine their relative importance by firm size. After extracting the relevant non-cognitive skills from each job advertisement for each firm size we derived the following results: Large firms demand on average 3.72 soft skills in each advertisement, while small and medium firms demand 3.35 and 3.31 soft skills respectively. Results can be seen in *Table 17*. Tests have also been conducted to verify that the difference in the demanded skills is statistically significant for large firms in comparison with medium and small, as it can be seen in *Table 20, 21, and 22.* This result aligns with the assumption that large firms require more soft-skills from their employees as they have to face greater co-ordination costs. Through this 'channel', the wage premiums for non-cognitive skills become higher, as it has been displayed.

A caveat with this approach has to do with the job description itself, as each job ad depends on the firm size. We found statistical evidence that job ads for big firms are larger than small and medium firms job ads. It is possible that large firms, having a greater capital to exploit, post bigger job descriptions, hence 'demanding' more soft skills as with our approach more words which correspond to soft skills are found in the description. This could be one reason for which more soft skills are demanded by large firms, aside from the higher coordination costs.

A further result derived from the demand-side analysis concerns the excess demand of skills observed by all types of firms. On average 6.3% of the firms demand from potential employees more skills than the ones that are necessary for an occupation's needs, according to the O*NET specifications. Our sample consists of job advertisements that were active during May, 2020. This result could potentially be tied with the Covid-19 pandemic and its economic implications. It is likely that firms demand more skills from their potential employees since the impeding high unemployment rates would need employees to do more for their firms, than what normally their occupation requires.

6 Conclusions

We explored the role of soft skills in U.S. Labor Market and their impact on wages. Relying on Becker (1992) according to which the non-cognitive skills lower the coordination costs, we test the link between soft skills and firm size. In our framework, soft skills are more valuable when increasing the size of firms as they are supposed to face higher coordination costs compare to small firms.

We proceeded to analyze both, the supply and demand, for soft skills in the U.S. labor market. In the case of the supply side we found a statistically significant return to soft skills: around 15% on hourly wages, while cognitive abilities account approximately 4% on hourly wages. In addition, we found a statistically significant soft skills wage premium in larger firms, around 2.5%, 3.5% for those highly educated. To the best of our knowledge, this is the first paper that finds a firm size wage premium for soft skills. This finding is consistent with our prediction based on Becker's model. Results are partly driven by education and females; whose premium is around 3% and increases to 4,5% when considering only those with more than 12 years of schooling.

Results suggest that soft skills trade job tasks more efficiently when increasing the level of education. Moreover, the promotion of gender equality in larger firms could encourage females to ask for higher wage (thus, facilitating those with higher soft skills to negotiate higher salaries), explaining why we observe this gender wage premium.

The demand side analysis supports our results. We found a statistically significant difference between the average number of soft skills demanded by big firms, compared the small and medium ones. In addition, we also found that 6.30% of firms tend to demand more skills than they need from their potential employees.

References

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature, 40*(1), 7–72.

Autor, D. H. (2015). Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives 29*, 3–30.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–1597.

Autor, D., F. L. and R. Murnane (2003). The skill content of recent technological change: An empirical exploration. *Quarterly journal of economics 118*(4), 1279–1333.

Bacolod, M., B. S. Blum, and W. C. Strange (2009). Skills in the city. *Journal of Urban Economics 65*(2), 136–153.

Beaudry, Paul, D. A. G. and B. Sand (2016). The great revearsal in the demand for skill and cognitive tasks. *Journal of Labor Economics 34*(1), 199–247.

Beaudry, P. and D. A. Green (2005). Changes in u.s. wages, 1976-2000: Ongoing skill bias or major technological change? *Journal of Labor Eocnomics 23*(3), 1976–2000.

Becker, G. S. and K. M. Murphy (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics 107*(4), 1137–1160.

Becker, M. K. (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics 107*(4), 1137–1160.

Borghans, L., B. Ter Weel, and B. A. Weinberg (2006). People people: Social capital and the labor-market outcomes of underrepresented groups. Technical report, National Bureau of Economic Research.

Castex, G. and E. K. Dechter (2014). The changing roles of education and ability in wage determination. *Journal of Labor Economics 32*(4), 685–710.

Cunningham, W. and P. Villaseñor (2014). *Employer voices, employer demands, and implications for public skills development policy*. The World Bank.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics 132*(4), 1593–1640.

Dunifon, R., D. G. J. (1998). Long-run effects of motivation on labor- market success. *Social Psychology Quarterly 61*(1), 33–48.

Edin, P.-A., P. Fredriksson, M. Nybom, and B. Ockert. The rising return to non-cognitive skill. *Working Paper Series*.

Handel, M. J. (2012). Trends in job skill demands in oecd countries.

Heckman, J. J. and T. Kautz (2012). Hard evidence on soft skills. *Labour economics 19*(4), 451–464.

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

Heineck, G. (2007). Does it pay to be nice? personality and earnings in the uk. *LASER Discussion Paper 3*.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics 107*, 35–78.

Kuhn, P. and C. Weinberger (2005). Leadership skills and wages. *Journal of Labor Economics 23*(3), 395–436.

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

Lundin, M., O. N. Skans, and P. Zetterberg (2019). Leadership experiences, labor market entry, and early career trajectories. *Journal of Human Resources*, 0617–8866R3.

Maciente, A. (2013). *The determinants of agglomeration in Brazil: input-output, labor and knowledge externalities*. Ph. D. thesis, University of Illinois at Urbana-Champaign.

Neves Jr, E. C., C. R. Azzoni, A. Chagas, et al. (2017). Skill wage premium and city size. Technical report, University of São Paulo (FEA-USP).

Nyhus, E. K. and E. Pons (2005). The effects of personality on earnings. *Journal of Economic Psychology 26*(3), 363–384.

Reber, A. S. (1995). *The Penguin dictionary of psychology*. Penguin Press.

Sheikh, S. B. (2015). Non-cognitive skills and labor market success. *North Hall 3047*.

Spence, M. Job market signaling. *The Quarterly Journal of Economics 87*.

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

7 Annex

Table 1: Descriptive Statistics, All Sample

Sample of 79,823 individuals	Small firm		Medium firm		Large firm	
	Percent.	Wages	Percent.	Wages	Percent.	Wages
By group						
Female	47.9	18.2	46.9	19.7	47.8	20.2
Male	52.1	22.6	53.1	26.0	52.2	25.8
Age (mean)	41.1	20.5	41.7	22.8	40.8	22.6
Years of Education (mean)	13.0	20.5	13.4	22.8	13.3	22.6
Married	49.9	22.5	51.5	26.0	48.1	27.0
Metropolitan area	77.4	20.6	77.4	23.4	79	23.7
White	83.0	20.4	81.9	23.5	78.8	23.5
By occupation						
Administrative Support	13.3	18.0	12.7	16.8	12.8	17.4
Architect & Engineering	1.3	33.4	2.3	36.7	2.2	38.3
Arts, Entertainment, & Media	1.8	28.9	2.1	35.2	1.7	26.7
Cleaning and Maintenance	4.9	13.8	3.4	15.9	4.2	13.5
Community and Social Service	1.7	17.6	2.1	20.1	1.4	22.7
Computer & Mathematics	1.6	38.5	2.4	44.5	2.5	41.4
Construction and Extraction	7.2	17.3	4.7	22.9	4.9	21.7
Education	2.4	17.5	2.9	24.4	3.2	23.9
Farming, Fishing, and Forestry	2.0	18.3	1.0	12.0	1.0	11.7
Financial Operations	3.6	28.8	4.7	30.7	4.3	33.5
Food	8.9	11.4	7.3	12.3	8.0	13.0
Healthcare Pratitioners and Technician	4.9	33.3	4.7	35.4	5.8	32.0
Healthcare Support	3.0	14.7	3.6	14.4	2.8	16.1
Installation, Maintenance, & Repair	4.2	29.3	3.4	21.4	3.7	21.7
Legal	1.3	30.8	1.0	39.3	0.7	35.8
Life, Physical, and Social Science	0.2	25.6	0.7	38.1	0.8	36.2
Management	8.7	28.9	12.4	32.7	8.8	37.4
Material Moving	2.3	12.5	3.7	14.3	3.8	14.8
Personal Care	3.2	14.5	1.4	12.9	2.4	13.4
Production	4.9	16.2	6.1	15.1	7.7	18.1
Protective Service	1.0	14.5	1.4	21.9	1.1	15.6
Sales	12.7	20.1	10.9	20.0	11.7	19.0
Transportation	4.1	17.7	4.4	19.1	3.7	21.4
Observations	5,395		4,978		69,450	

Table 2: Descriptive Statistics, subsample of Females

Sample of 43,784 female	Small firm		Medium firm		Large firm	
	Percent.	Wages	Percent.	Wages	Percent.	Wages
By group						
Age (mean)	41.2	17.8	41.8	19.2	41.1	19.5
<= 12 Years of Education	11.2	12.6	11.3	13.8	11.2	13.7
>12 Years of Education	14.7	21.1	14.8	22.5	14.8	22.8
Married	47.1	19.9	47.8	21.9	44.9	22.5
Metropolitan area	77.8	18.3	77.7	19.8	79.0	20.1
White	82.5	17.9	79.5	20.1	77.3	19.9
By occupation						
Administrative Support	22.3	15.3	19.0	17.1	18.8	16.5
Architect & Engineering	0.1	23.7	0.8	30.9	0.6	31.3
Arts, Entertainment, & Media	1.8	16.8	1.5	36.0	1.7	25.6
Cleaning and Maintenance	4.5	15.6	3.5	10.8	4.0	12.8
Community and Social Service	2.1	16.8	3.0	19.0	1.8	22.1
Computer & Mathematics	0.8	24.7	1.0	39.3	1.3	37.9
Construction and Extraction	0.3	9.4	0.2	21.8	0.3	16.3
Education	3.7	15.9	4.8	21.2	4.7	21.7
Farming, Fishing, and Forestry	1.2	34.5	0.8	11.3	0.6	9.4
Financial Operations	4.2	24.2	6.4	24.5	5.2	26.6
Food	9.8	10.2	9.2	12.1	9.2	13.8
Healthcare Pratitioners and Technician	7.3	28.7	7.5	30.7	9.6	30.0
Healthcare Support	5.8	14.6	6.7	14.5	5.3	16.1
Installation, Maintenance, & Repair	0.3	18.4	0.1	18.5	0.3	17.8
Legal	1.8	24.3	1.4	32.9	0.9	32.5
Life, Physical, and Social Science	0.2	19.0	0.5	31.0	0.6	30.0
Management	8.7	24.7	11.0	26.1	7.9	30.0
Material Moving	0.7	9.9	1.2	14.1	1.8	12.2
Personal Care	5.3	14.7	2.2	13.6	3.8	13.4
Production	3.2	13.3	4.3	12.3	5.2	14.3
Protective Service	0.7	13.7	1.0	13.2	0.6	13.7
Sales	14.1	18.4	12.4	15.1	14.2	14.8
Transportation	0.5	16.1	0.7	13.2	0.8	18.2
Observations	2,934		2,758		38,092	

Table 3: Descriptive Statistics, subsample of Female with more than 12 Years of Schooling

Sample of 22,947 female	Small firm		Medium firm		Large firm	
	Percent.	Wages	Percent.	Wages	Percent.	Wages
By occupation >12 Years of Education						
Administrative Support	22.4	16.1	18.8	17.3	18.9	17.4
Architect & Engineering	0.2	27.9	1.0	36.4	0.8	32.9
Arts, Entertainment, & Media	2.2	18.7	2.0	38.9	2.2	27.2
Cleaning and Maintenance	2.3	20.9	1.0	11.2	1.5	13.4
Community and Social Service	2.9	17.2	4.1	19.5	2.5	22.4
Computer & Mathematics	1.3	24.7	1.5	39.3	1.9	38.4
Construction and Extraction	0.2	9.3	0.2	16.9	0.2	16.6
Education	4.7	17.6	6.7	22.2	6.5	22.9
Farming, Fishing, and Forestry	0.6	81.0	0.2	14.7	0.2	11.0
Financial Operations	5.5	23.8	8.6	24.7	6.9	27.7
Food	6.2	11.9	6.0	13.8	6.2	17.3
Healthcare Pratitioners and Technician	10.6	30.6	10.1	32.6	13.7	30.8
Healthcare Support	5.2	15.3	6.5	16.3	4.7	17.8
Installation, Maintenance, & Repair	0.1	33.5	0.1	17.6	0.3	18.7
Legal	2.8	25.1	1.9	36.1	1.2	34.5
Life, Physical, and Social Science	0.4	19.0	0.5	37.0	1.0	30.9
Management	11.8	26.5	13.2	28.9	10.2	32.3
Material Moving	0.1	9.1	0.5	23.1	0.8	14.5
Personal Care	4.2	19.6	2.3	14.2	3.1	14.0
Production	2.4	14.8	1.9	13.8	2.7	15.8
Protective Service	0.6	15.5	0.5	21.5	0.5	15.6
Sales	12.3	24.3	11.7	15.1	12.3	17.0
Transportation	0.1	14.8	0.4	14.5	0.7	17.1
Observations	1,446		1,416		20,085	

Table 4: Factor Analysis

	Variable	Factor1	Factor2	Factor1	Factor2	Factor1	Factor2
Abilities	Cognitive	-0.40	0.71				
	Psycomotor	0.95	0.04				
	Physical	0.88	-0.18				
	Sensory	0.67	0.59				
Skills	Resource Management			0.81	0.08		
	Content			0.94	-0.00		
	Process			0.97	-0.07		
	Social			0.89	-0.32		
	Problem-Solving			0.95	0.16		
	Technical			0.04	0.61		
	Systems			0.97	0.11		
Knowledge	Business & Management					0.63	0.03
	Manufacturing					0.16	0.50
	Engineering & Technology					0.04	0.41
	Math & Science					0.41	0.68
	Health Services					0.87	0.02
	Education & Training					0.60	-0.50
	Arts & Humanities					0.86	-0.10
	Laws & Public Safety					0.78	-0.33
	Communication					0.72	0.23
	Transportation					0.74	0.09
Eigenvalue		2.31	0.88	5.17	0.53	4.16	1.33
Cumulative		0.77	1.06	0.90	1	0.68	0.90

Table 5: *Pooled OLS*, All Sample

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Physical abilities	-0.0618***	-0.0619***	-0.0619***	-0.0175***	-0.1127***	-0.1067***	-0.1064***
	(0.0052)	(0.0052)	(0.0052)	(0.0050)	(0.0080)	(0.0080)	(0.0080)
Cognitive abilities	0.1360***	0.1358***	0.1358***	0.0824***	0.0395***	0.0403***	0.0397***
	(0.0073)	(0.0073)	(0.0073)	(0.0076)	(0.0097)	(0.0096)	(0.0096)
Soft Skills	0.2172***	0.2163***	0.1938***	0.1652***	0.1482***	0.1476***	0.1480***
	(0.0085)	(0.0084)	(0.0143)	(0.0134)	(0.0142)	(0.0141)	(0.0140)
Soft Knowledge	0.0080	0.0085	0.0083	-0.0140**	-0.0253***	-0.0265***	-0.0267***
	(0.0060)	(0.0060)	(0.0060)	(0.0057)	(0.0069)	(0.0069)	(0.0069)
Hard Knowledge	0.0746***	0.0746***	0.0746***	0.0513***	0.0450***	0.0450***	0.0448***
	(0.0047)	(0.0047)	(0.0047)	(0.0048)	(0.0079)	(0.0078)	(0.0078)
Medium Size		0.1158***	0.1187***	0.0871***	0.0833***	0.0812***	0.0811***
		(0.0158)	(0.0160)	(0.0153)	(0.0149)	(0.0148)	(0.0147)
Big Size		0.1126***	0.1154***	0.0873***	0.0831***	0.0826***	0.0826***
		(0.0119)	(0.0123)	(0.0118)	(0.0115)	(0.0114)	(0.0114)
Soft Skills*Medium Size			0.0138	0.0162	0.0140	0.0146	0.0156
			(0.0161)	(0.0148)	(0.0144)	(0.0143)	(0.0142)
Soft Skills*Big Size			0.0249**	0.0289**	0.0242**	0.0246**	0.0247**
			(0.0125)	(0.0116)	(0.0112)	(0.0111)	(0.0111)
Soft Skills*Cognitive				0.0548***	0.0310***	0.0305***	0.0305***
				(0.0035)	(0.0042)	(0.0042)	(0.0042)
Constant	2.8001***	2.6948***	2.6920***	-0.4169***	-0.1072	-0.2086*	-0.2236**
	(0.0036)	(0.0117)	(0.0120)	(0.1052)	(0.1038)	(0.1064)	(0.1064)
Controls	No	No	No	Yes	Yes	Yes	Yes
Occupation FE	No	No	No	No	Yes	Yes	Yes
State FE	No	No	No	No	No	Yes	Yes
Year FE	No	No	No	No	No	No	Yes
Observations	75,276	75,276	75,276	75,276	75,274	75,274	75,274
R^2	0.2018	0.2031	0.2032	0.2868	0.3120	0.3208	0.3219

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Abilities, soft skills and knowledge are predicted from Factor Analysis. Small, medium and big firm size are dummy variables that take value 1 if employees are < 50, < 200 and > 200 respectively. Controls are: gender, age, marital status, experience, education, race, being part of the unions, living in metropolitan areas, interactions between marital status and female and also interactions between skills and abilities.

Table 6: *Pooled OLS* by Gender and Education

Variables		Years of Schooling		Male			Female		
	All sample	<= 12	> 12	Overall	<= 12 Years	> 12 Years	Overall	<= 12 Years	> 12
Physical abilities	-0.1064***	-0.0626***	-0.1185***	-0.1097***	-0.0368**	-0.1237***	-0.0914***	-0.0584***	-0.1072***
	(0.0080)	(0.0121)	(0.0104)	(0.0115)	(0.0187)	(0.0148)	(0.0112)	(0.0168)	(0.0147)
Cognitive abilities	0.0397***	0.0333**	0.0150	0.0428***	0.0374*	0.0226	0.0376***	0.0092	0.0085
	(0.0096)	(0.0151)	(0.0130)	(0.0141)	(0.0217)	(0.0197)	(0.0131)	(0.0211)	(0.0171)
Soft Skills	0.1480***	0.1525***	0.1681***	0.1581***	0.1551***	0.1946***	0.1319***	0.1591***	0.1300***
	(0.0140)	(0.0203)	(0.0202)	(0.0194)	(0.0255)	(0.0291)	(0.0207)	(0.0350)	(0.0276)
Soft Knowledge	-0.0267***	-0.0471***	-0.0312***	-0.0506***	-0.0470***	-0.0625***	0.0048	-0.0444***	0.0084
	(0.0069)	(0.0101)	(0.0092)	(0.0094)	(0.0136)	(0.0130)	(0.0106)	(0.0170)	(0.0133)
Hard Knowledge	0.0448***	0.0603***	0.0665***	0.0514***	0.0535**	0.0777***	0.0361***	0.0723***	0.0478***
	(0.0078)	(0.0139)	(0.0094)	(0.0121)	(0.0213)	(0.0148)	(0.0104)	(0.0196)	(0.0121)
Medium Size	0.0811***	0.0841***	0.0786***	0.0949***	0.0893***	0.0978***	0.0655***	0.0798**	0.0605**
	(0.0147)	(0.0223)	(0.0215)	(0.0215)	(0.0307)	(0.0329)	(0.0196)	(0.0316)	(0.0275)
Big Size	0.0826***	0.0520***	0.0875***	0.0936***	0.0579***	0.1075***	0.0712***	0.0496**	0.0693***
	(0.0114)	(0.0162)	(0.0167)	(0.0171)	(0.0214)	(0.0276)	(0.0150)	(0.0245)	(0.0197)
Soft Skills*Medium Size	0.0156	-0.0006	0.0219	0.0087	-0.0104	0.0107	0.0207	0.0122	0.0235
	(0.0142)	(0.0212)	(0.0205)	(0.0193)	(0.0277)	(0.0280)	(0.0204)	(0.0328)	(0.0287)
Soft Skills*Big Size	0.0247**	-0.0185	0.0362**	0.0198	-0.0218	0.0242	0.0291*	-0.0137	0.0456**
	(0.0111)	(0.0168)	(0.0159)	(0.0153)	(0.0203)	(0.0227)	(0.0158)	(0.0287)	(0.0208)
Soft Skills*Cognitive	0.0305***	0.0108*	0.0248***	0.0220***	0.0119	0.0196**	0.0543***	0.0156	0.0405***
	(0.0042)	(0.0065)	(0.0059)	(0.0055)	(0.0085)	(0.0077)	(0.0067)	(0.0109)	(0.0092)
Constant	-0.2236**	0.3135	-1.4283***	0.0178	0.5315**	-1.0875***	-0.3228**	0.0961	-1.6083***
	(0.1064)	(0.1908)	(0.1642)	(0.1447)	(0.2397)	(0.2358)	(0.1509)	(0.3126)	(0.2249)
Observations	75,274	30,732	44,542	39,319	17,689	21,630	35,955	13,043	22,912
R^2	0.3219	0.1896	0.3297	0.3279	0.1950	0.3349	0.2971	0.1412	0.3000

Table 7: *Pooled OLS* of Content Skill on Wage

Variables	All Sample	Male		Female	
		Overall	> 12 Years	Overall	> 12 Years
Physical Abilities	-0.0613***	-0.0722***	-0.0926***	-0.0357***	-0.0588***
	(0.0083)	(0.0114)	(0.0147)	(0.0118)	(0.0155)
Cognitive Abilities	-0.2020***	-0.0661	0.1359*	-0.4328***	-0.1784**
	(0.0390)	(0.0547)	(0.0773)	(0.0584)	(0.0796)
Content Skills	-0.1427***	-0.1544**	0.0362	-0.2679***	-0.0600
	(0.0465)	(0.0673)	(0.0863)	(0.0654)	(0.0872)
Medium Size	0.0286	0.0473	0.0412	0.0110	-0.0166
	(0.0574)	(0.0738)	(0.1182)	(0.0884)	(0.1372)
Big Size	-0.0058	0.0119	0.0401	-0.0292	-0.1349
	(0.0437)	(0.0557)	(0.0934)	(0.0687)	(0.0974)
Content*Medium Size	0.0197	0.0192	0.0214	0.0195	0.0285
	(0.0212)	(0.0277)	(0.0392)	(0.0323)	(0.0463)
Content*Big Size	0.0319*	0.0307	0.0253	0.0357	0.0734**
	(0.0163)	(0.0216)	(0.0312)	(0.0250)	(0.0334)
Content*Cognitive	0.1287***	0.1081***	0.0458**	0.1900***	0.0990***
	(0.0130)	(0.0176)	(0.0218)	(0.0193)	(0.0259)
Constant	-0.1784	-0.1139	-1.6809***	0.1984	-1.5348***
	(0.1419)	(0.1864)	(0.3089)	(0.2125)	(0.3146)
Observations	75,274	39,319	21,630	35,955	22,912
R^2	0.3220	0.3264	0.3322	0.2994	0.3006

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Each regression includes all the controls and all the fixed effects.

Table 8: *Pooled OLS* of Problem-Solving Skill on Wages

Variables	All Sample	Male		Female	
		Overall	> 12 Years	Overall	> 12 Years
Physical Abilities	-0.0774***	-0.0708***	-0.0725***	-0.0714***	-0.0693***
	(0.0078)	(0.0109)	(0.0115)	(0.0111)	(0.0115)
Cognitive Abilities	-0.1824***	-0.1282**	-0.0311	-0.3329***	-0.2524***
	(0.0385)	(0.0506)	(0.0549)	(0.0603)	(0.0656)
Problem Solving Skills	-0.1598***	-0.0582	0.0329	-0.3679***	-0.2782***
	(0.0456)	(0.0620)	(0.0662)	(0.0675)	(0.0712)
Medium Size	0.0007	0.0712	0.0879	-0.0528	-0.0424
	(0.0659)	(0.0907)	(0.0956)	(0.0923)	(0.0999)
Big Size	-0.0370	0.0188	0.0149	-0.0867	-0.1017
	(0.0509)	(0.0705)	(0.0755)	(0.0719)	(0.0768)
Solving*Medium Size	0.0284	0.0097	0.0057	0.0410	0.0388
	(0.0225)	(0.0307)	(0.0320)	(0.0318)	(0.0337)
Solving*Big Size	0.0410**	0.0258	0.0276	0.0547**	0.0600**
	(0.0176)	(0.0244)	(0.0258)	(0.0249)	(0.0263)
Solving*Cognitive	0.1193***	0.0953***	0.0618***	0.1780***	0.1458***
	(0.0134)	(0.0177)	(0.0188)	(0.0206)	(0.0219)
Constant	-0.1207	-0.1463	-0.8370***	0.2981	-0.3522
	(0.1520)	(0.2025)	(0.2393)	(0.2268)	(0.2551)
Observations	75,274	39,319	34,849	35,955	32,969
R^2	0.3230	0.3287	0.3251	0.2984	0.2919

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Each regression includes all the controls and all the fixed effects.

Table 9: *Pooled OLS* of Process Skill on Wages

Variables	All Sample	Male		Female	
		Overall	> 12 Years	Overall	> 12 Years
Physical Abilities	-0.0920***	-0.0887***	-0.1128***	-0.0791***	-0.1106***
	(0.0079)	(0.0112)	(0.0144)	(0.0112)	(0.0105)
Cognitive Abilities	-0.0838*	-0.0142	0.2677***	-0.2976***	0.1689**
	(0.0433)	(0.0610)	(0.0926)	(0.0653)	(0.0661)
Process Skills	-0.3563***	-0.2925***	-0.1509*	-0.5344***	-0.1900***
	(0.0449)	(0.0628)	(0.0833)	(0.0637)	(0.0593)
Medium Size	0.0117	0.0576	0.0565	-0.0272	-0.0229
	(0.0649)	(0.0879)	(0.1437)	(0.0936)	(0.1039)
Big Size	-0.0173	0.0264	0.0381	-0.0629	-0.0567
	(0.0498)	(0.0678)	(0.1150)	(0.0724)	(0.0787)
Process*Medium Size	0.0227	0.0132	0.0152	0.0296	0.0334
	(0.0214)	(0.0294)	(0.0440)	(0.0304)	(0.0315)
Process*Big Size	0.0320*	0.0221	0.0239	0.0428*	0.0471*
	(0.0166)	(0.0234)	(0.0356)	(0.0234)	(0.0242)
Process*Cognitive	0.1418***	0.1207***	0.0573**	0.2048***	0.0721***
	(0.0135)	(0.0187)	(0.0248)	(0.0199)	(0.0183)
Constant	-0.0036	0.0169	-1.6553***	0.4718**	-1.7522***
	(0.1547)	(0.2055)	(0.3423)	(0.2317)	(0.2456)
Observations	75,274	39,319	21,630	35,955	44,542
R^2	0.3205	0.3258	0.3314	0.2968	0.3270

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Each regression includes all the controls and all the fixed effects.

Table 10: *Pooled OLS* of Resources Management Skills on Wages

Variables	All Sample	Male		Female	
		Overall	> 12 Years	Overall	> 12 Years
Physical Abilities	-0.1039***	-0.0986***	-0.1160***	-0.0921***	-0.1054***
	(0.0077)	(0.0107)	(0.0140)	(0.0112)	(0.0148)
Cognitive Abilities	0.2387***	0.2772***	0.4732***	0.1112***	0.2512***
	(0.0261)	(0.0338)	(0.0504)	(0.0423)	(0.0600)
RES Skills	-0.3677***	-0.3508***	-0.1949**	-0.4727***	-0.3435***
	(0.0423)	(0.0572)	(0.0771)	(0.0641)	(0.0854)
Medium Size	0.0658	0.0586	0.0705	0.1070*	0.1060
	(0.0402)	(0.0553)	(0.0889)	(0.0580)	(0.0816)
Big Size	0.0008	-0.0020	-0.0050	0.0285	0.0007
	(0.0310)	(0.0432)	(0.0740)	(0.0436)	(0.0553)
RES* Medium Size	0.0078	0.0191	0.0165	-0.0209	-0.0193
	(0.0186)	(0.0249)	(0.0363)	(0.0277)	(0.0365)
RES*Big Size	0.0385***	0.0449**	0.0522*	0.0201	0.0360
	(0.0146)	(0.0204)	(0.0310)	(0.0204)	(0.0246)
RES*Cognitive	0.1080***	0.0956***	0.0360*	0.1582***	0.1028***
	(0.0128)	(0.0165)	(0.0215)	(0.0204)	(0.0272)
Constant	-0.6181***	-0.4458***	-1.9715***	-0.4561**	-2.1027***
	(0.1301)	(0.1705)	(0.2799)	(0.1970)	(0.2831)
Observations	75,274	39,319	21,630	35,955	22,912
R^2	0.3198	0.3254	0.3317	0.2952	0.2977

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Each regression includes all the controls and all the fixed effects.

Table 11: *Pooled OLS* of Social Skills on Wages

Variables	All Sample	Male		Female	
		Overall	> 12 Years	Overall	> 12 Years
Physical Abilities	-0.1112***	-0.1016***	-0.1222***	-0.1047***	-0.1158***
	(0.0079)	(0.0113)	(0.0144)	(0.0111)	(0.0147)
Cognitive Abilities	0.0295	0.0496	0.3083***	-0.1092*	0.0823
	(0.0409)	(0.0567)	(0.0882)	(0.0625)	(0.0905)
Social Skills	-0.5374***	-0.5604***	-0.4328***	-0.6320***	-0.4566***
	(0.0452)	(0.0657)	(0.0955)	(0.0631)	(0.0885)
Medium Size	-0.0145	-0.0274	-0.0642	0.0132	0.0546
	(0.0667)	(0.0901)	(0.1459)	(0.0976)	(0.1541)
Big Size	0.0015	-0.0021	-0.0440	0.0032	-0.0313
	(0.0509)	(0.0697)	(0.1194)	(0.0745)	(0.1108)
Social Skills* Medium Size	0.0329	0.0444	0.0559	0.0168	0.0029
	(0.0234)	(0.0329)	(0.0496)	(0.0330)	(0.0489)
Social Skills*Big Size	0.0274	0.0340	0.0529	0.0224	0.0354
	(0.0180)	(0.0261)	(0.0411)	(0.0250)	(0.0356)
Social Skills*Cognitive	0.1637***	0.1642***	0.0994***	0.2027***	0.1355***
	(0.0146)	(0.0207)	(0.0296)	(0.0212)	(0.0295)
Constant	0.0657	0.2930	-1.2399***	0.3000	-1.4679***
	(0.1573)	(0.2081)	(0.3564)	(0.2377)	(0.3540)
Observations	75,274	39,319	21,630	35,955	22,912
R^2	0.3207	0.3264	0.3325	0.2963	0.2980

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Each regression includes all the controls and all the fixed effects.

Table 12: *Pooled OLS* of System Skills on Wages

Variables	All Sample	Male		Female	
		Overall	> 12 Years	Overall	> 12 Years
Physical Abilities	-0.0681***	-0.0616***	-0.0794***	-0.0581***	-0.0743***
	(0.0081)	(0.0113)	(0.0147)	(0.0117)	(0.0157)
Cognitive Abilities	-0.1337***	-0.1323***	0.0177	-0.2215***	-0.1002
	(0.0365)	(0.0499)	(0.0768)	(0.0554)	(0.0800)
System Skills	-0.1816***	-0.1347***	-0.0072	-0.3173***	-0.1565**
	(0.0396)	(0.0515)	(0.0695)	(0.0615)	(0.0778)
Medium Size	0.0388	0.0816	0.0741	0.0067	0.0026
	(0.0461)	(0.0614)	(0.1003)	(0.0671)	(0.0996)
Big Size	-0.0077	0.0176	0.0172	-0.0250	-0.0750
	(0.0354)	(0.0466)	(0.0801)	(0.0523)	(0.0722)
System Skills* Medium Size	0.0167	0.0069	0.0118	0.0223	0.0227
	(0.0177)	(0.0235)	(0.0338)	(0.0260)	(0.0349)
System Skills*Big Size	0.0346**	0.0297	0.0349	0.0372*	0.0564**
	(0.0139)	(0.0186)	(0.0275)	(0.0203)	(0.0260)
System Skills*Cognitive	0.1142***	0.1042***	0.0593***	0.1555***	0.0961***
	(0.0109)	(0.0142)	(0.0188)	(0.0173)	(0.0226)
Constant	-0.0878	0.0720	-1.3574***	0.0867	-1.5008***
	(0.1321)	(0.1737)	(0.2939)	(0.1994)	(0.2959)
Observations	75,274	39,319	21,630	35,955	22,912
R^2	0.3218	0.3275	0.3331	0.2974	0.2991

Notes: Cluster standard errors at individual level in parentheses: ***p<0.01, **p<0.05, *p<0.1. Each regression includes all the controls and all the fixed effects.

Table 13: Overall size of scraped dataset with salary information

Firm Size	**Observations**	**Observations with salary**	**Percentage**
Small	1,427	298	20.9%
Medium	475	103	21.7%
Large	3,078	698	22.7%
Total	**4,980**	**1,099**	**22.1%**

Table 14: Scraped dataset by occupation

Occupations	**Observations**
Administrative Support	739
Cleaning	445
Construction	445
Education	500
Financial Operations	432
Food	445
Healthcare	300
Management	500
Production	445
Sales	729
Total	**4,980**

Table 15: Most frequent demanded skills and abilities in Administrative Support Job Ads

Small Firms		
Skill, Ability, or Knowledge	**Frequency**	**%**
Administration and Management	78	6.27%
Written Comprehension	70	5.63%
Writing	70	5.63%
Written Expression	70	5.63%
Computers and Electronics	62	4.98%
Education and Training	56	4.50%
Customer and Personal Service	49	3.94%
Information Ordering	47	3.78%
Reading Comprehension	43	3.46%
Complex Problem-Solving	42	3.38%

Medium Firms		
Skill, Ability, or Knowledge	**Frequency**	**%**
Administration and Management	24	7.45%
Customer and Personal Service	20	6.21%
Reading Comprehension	16	4.97%
Information Ordering	16	4.97%
Computers and Electronics	14	4.35%
Education and Training	14	4.35%
Clerical	13	4.04%
Written Comprehension	11	3.42%
Writing	11	3.42%
Problem Sensitivity	11	3.42%

Large Firms		
Skill, Ability, or Knowledge	**Frequency**	**%**
Administration and Management	186	5.92%
Education and Training	174	5.54%
Computers and Electronics	142	4.52%
Written Expression	126	4.01%
Writing	126	4.01%
Customer and Personal Service	126	4.01%
Written Comprehension	126	4.01%
Information Ordering	116	3.69%
Clerical	115	3.66%
Reading Comprehension	95	3.02%

Table 16: Soft Skills demand in relation to overall skills

Firm Size	ASK	Soft Skills	Percentage
Small	15,665	4,783	30.5%
Medium	4,999	1,573	31.5%
Large	37,701	11,462	30.4%
Total	**58,365**	**17,818**	**30.5%**

Table 17: Soft skills demand in relation to firm size

Firm Size	Number	Soft Skills	Average
Small	1,427	4,783	3.35
Medium	475	1,573	3.31
Large	3,078	11,462	3.72
Total	**4,980**	**17,818**	**3.58**

Table 18: Overall skills demand in relation to firm size

Firm Size	Number	ASK	Average ASK
Small	1,427	15,665	11.0
Medium	475	4,999	10.5
Large	3,078	37,701	12.2
Total	**4,980**	**58,365**	**11.7**

Table 19: Excess soft skill demand compared to O*NET specification by fir size

Firm Size	Excess, % of firms
Small	6.3%
Medium	7.9%
Large	6.1%
Total	**6.3%**

Table 20: *T-test* for mean difference between medium and large firms soft skill demand

Group	**Obs**	**Mean**	**Std. Err.**	**Std. Dev.**	**[95% Conf.**	**Interval]**
Medium	475	3.31	0.12	2.63	3.07	3.55
Large	3,078	3.72	0.05	2.76	3.63	3.82
Combined	3,553	3.67	0.05	2.74	3.58	3.76
Difference		-0.41	0.14		-0.68	-0.15

Ha : Difference $\neq 0$
$\Pr(|T| > |t|) = 0.0023$

Table 21: *T-Test* for mean difference between small and large firms soft skill demand

Group	**Obs**	**Mean**	**Std. Err.**	**Std. Dev.**	**[95% Conf.**	**Interval]**
Small	1,427	3.35	0.08	2.87	3.20	3.50
Large	3,078	3.72	0.05	2.76	3.63	3.82
Combined	4,505	3.61	0.04	2.80	3.52	3.69
Difference		-0.37	0.09		-0.55	-0.20

Ha : Difference $\neq 0$
$\Pr(|T| > |t|) = 0.0000$

Table 22: *T-Test* for mean difference between small and medium firms soft skill demand

Group	**Obs**	**Mean**	**Std. Err.**	**Std. Dev.**	**[95% Conf.**	**Interval]**
Small	1,427	3.35	0.08	2.87	3.20	3.50
Medium	475	3.31	0.12	2.63	3.07	3.55
Combined	1,902	3.34	0.06	2.81	3.22	3.47
Difference		0.04	0.15		-0.25	0.33

Ha : Difference $\neq 0$
$\Pr(|T| > |t|) = 0.7870$

Table 23: *T-Test* for mean difference in the number of characters in job ads between medium and large

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
Medium	475	2,760.41	87.78	1,913.08	2,587.93	2,932.89
Large	3,078	3,695.18	49.94	2,770.81	3,597.26	3,793.11
Combined	3,553	3,570.21	45.14	2,690.79	3,481.71	3,658.72
Difference		-934.78	131.73		-1,193.06	-676.49

Ha : Difference $\neq 0$
$\Pr(|T| > |t|) = 0.0000$

Table 24: *T-Test* for mean difference in the number of characters in job ads between small and large

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
Small	1,427	3,039.01	69.31	2,618.20	2,903.05	3,174.97
Large	3,078	3,695.18	49.94	2,770.81	3,597.26	3,793.11
Combined	4,505	3,487.33	40.83	2,740.17	3,407.30	3,567.37
Difference		-656.18	87.22		-827.17	-485.18

Ha : Difference $\neq 0$
$\Pr(|T| > |t|) = 0.0000$

Table 25: *T-Test* for mean difference in the number of characters in job ads between small and medium firms

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
Small	1,427	3,039.01	69.31	2,618.20	2,903.05	3,174.97
Medium	475	2,760.41	87.78	1,913.08	2,587.93	2,932.89
Combined	1,902	2,969.43	56.49	2,463.58	2,858.64	3,080.22
Difference		278.60	130.38		22.90	534.30

Ha : Difference $\neq 0$
$\Pr(|T| > |t|) = 0.0327$