Signaling Soft Skills

by

Alexandra Edquist

(Under the Direction of Meghan Skira)

Abstract

This paper finds evidence consistent with the notion that employers use education level to screen for soft skills in the hiring process. Soft skills or interpersonal skills include traits like service orientation and persuasion. This paper uses skills data from O*Net to measure the importance of soft skills relative to other skills in an occupation. If employers use education level, especially a bachelor's degree, to screen for soft skills, we would expect to see in occupations where soft skills are important that employers pay a higher wage premium for more education, that a greater share of workers are overqualified for their occupations, and that there is be more wage variability within the occupation. This paper finds strong evidence supporting this for overqualification and wage premiums, and some evidence for this for wage variability when separated by skill quantile.

Index Words: Labor economics, soft skills, overqualification, wage premium, wage variability

Signaling Soft Skills

by

Alexandra Edquist

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements of the Degree

Master of Arts

Athens, Georgia

2016

Signaling Soft Skills

by

Alexandra Edquist

Major Professor:	Meghan Skira
Committee:	Ian Schmutte
	Christopher Cornwell

Electronic Version Approved:

Suzanne Barbour

Dean of the Graduate School

The University of Georgia

August 2016

Acknowledgements

I thank my thesis committee, especially my chair, Dr. Meghan Skira, who has been an incredible source of help and support over this year. I also thank Dr. Chris Cornwell and Dr. Ian Schmutte. Dr. Cornwell was one of the reasons I chose to attend the University of Georgia, to study economics, and to pursue this particular program.

I would also like to thank my family and friends for supporting me through this process, especially the current and previous economics graduate students who have helped me so much throughout this program. Eilidh Geddes, Kayla Wilding, Andrew Kane, Kevin Sun, Meghan Esson, Caitlin Thomas, Kim Groover, and everyone else: I couldn't have been successful without you.

I also thank the UGA Honors Program, Foundation Fellowship, and the Terry College of Business for their incredibly generous financial, professional, and personal support over my four years at the University of Georgia.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Several of the puzzles in labor economics relate to educational mismatch between workers and jobs (Brynin & Longhi (2009)) and the earnings disparity between workers in the same occupation (Wheeler (2005) and Bowles, Weel, & Weinberg (2001)). An additional gap in the literature is the root of the "sheepskin effect" from completing an educational credential that many researchers have found boosts an individual's wages by more than the improvement in human capital caused by the education itself (Hungerford & Solon (1987), Jaeger & Page (1996), and Murnane, Tyler, & Willett (2000)).

The literature has evaluated many possible causes for these phenomena, but one hypothesis that has not yet been tested in the literature is whether employers screen potential workers for soft skills using educational attainment. Workers with more education are more likely to be productive, and therefore, workers with less education as likely to be less productive in occupations where soft skills are relatively important. We would expect to see higher wage premiums for additional education, more workers overqualified for their job, and more wage variability or inequality within the occupation if soft skills are important and employers indeed use education to screen for soft skills.

Rising wage inequality, especially by level of educational attainment, has been well documented in the empirical literature. Most of the literature explores the effect of technology on wage inequality. Autor, Katz, & Kearney (2008) find that growing wage inequality in the United States since the 1960s can be explained by growing demand for skilled workers. Wheeler (2005) also finds evidence that growing wage inequality is driven by increasing use of technology on the job and thus, the demand for better-educated workers.

However, the above hypothesis could be another explanation for growing wage inequality driven by differences in educational attainment. If employers use higher education as a proxy for soft skills, and soft skills are becoming more important in the economy, then we would see growing wage inequality because employers are paying wage premiums to bachelor's degree holders more often. Borghans, Weel, & Weinberg (2006) find that soft skills are indeed becoming more important in the economy. They estimate that "people tasks" became 3.6 times as important in 2002 as they were in 1971. We also know that wage inequality is growing. Autor, Katz, & Kearney (2008) document that overall wage inequality rose sharply in the 1980s and continued to rise more slowly in the 1990s and early 2000s.

This paper also examines whether there is evidence that employers use higher education as a proxy for soft skills by analyzing overqualification and wage premiums. If higher education signals soft skills, then we would expect occupations requiring more soft skills to have more workers with more education than is strictly required for the job and for employers to pay those workers a higher wage premium for having that additional education. If employers screen for soft skills using education, this means that the "soft skill signal" of education would have a large enough benefit to warrant the cost of additional education (see Brynin & Longhi (2009) for other situations in which overqualification is rational).

Section 2 provides a review of the literature. Section 3 provides an overview of the data used in this paper. Section 4 outlines the empirical approach. Section 5 presents results. Section 6 presents robustness checks. Section 7 concludes.

Chapter 2

Literature Review

The paper upon which the job skills literature is based is Becker (1962). Becker's model separated workers into "high skill" and "low skill" roles that were affected differently by changes in technology. In his model, technology is seen as a substitute for low-skill roles but as a complement for high-skill roles. Therefore, improvements in technology would favor high-skill workers, hence the phrase "skill-biased technological change" (SBTC). However, while SBTC explained rising wage inequality in the 1970s and 1980s, it did not adequately explain the lack of rising wage inequality in the 1990s even though use of technology, especially computers, continued to grow, as described in Card & DiNardo (2002).

Autor, Levy, & Murnane (2003) offer another model, defining worker roles as "task-based." They separate tasks into four categories: manual routine, manual non-routine, cognitive routine, and cognitive non-routine. They find that technology is a substitute for routine tasks. As detailed in Autor & Dorn (2013), this model helps to explain the well-documented polarization (see Autor, Katz, & Kearney (2008)) in the job market in which middle-wage jobs generally focused on routine tasks are losing share relative to low-wage and high-wage jobs generally focused on non-routine tasks. It also explains the rise of the service industry, which is mostly composed of low-skill but non-routine jobs, better than SBTC.

The task-based model extends naturally into analysis of skills. Because it is difficult to observe skills directly, researchers often create measures for skills using job tasks. For example, Krueger & Schkade (2007) use data from the Texas DRM in which workers gave a diary of their working day. They use the proportion of time the workers spent interacting with other people as a measure of how important social skills were on the job. Weinberger (2014), Borghans, Weel, & Weinberg (2006), and Bacolod & Blum (2010) are three papers that construct measures of job skills from data on job tasks from the Dictionary of Occupational Titles (DOT), and Black & Spitz-Oener (2007) construct similar measures from job task data from West Germany. The switch in the United States from the Dictionary of Occupational Titles (DOT) to the Occupational Information Network (O*Net) in 1998 has made analyzing skills easier because its focus is on job skills rather than job tasks (Mariani (1999)). Hirsch (2005) and Deming (2015) are examples of papers that use O*Net's skill measures directly to estimate skills.

There have been many papers detailing various aspects of job skills; however, the economic literature on soft skills is rather small. This is mostly because soft skills are difficult to define and measure, as detailed in Balcar (2014). Much of the literature examines psychological or personality traits. Heckman & Kautz (2012) use the "Big Five" personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism as a measure of soft skills. Urzua (2006) uses attitudes as measured in the Rosenberg Self-Esteem and Rotter Locus of Control scales to measure soft or non-cognitive skills. In addition, as discussed above, researchers have used task and skill data from the DOT and O*Net to create measures of soft skills. Deming (2015) averages the four job skills from O*Net of social perceptiveness, coordination, negotiation, and persuasion, for example to define an occupation's social skill intensity.

Soft skills are undoubtedly important in job-market success. Mueller & Plug (2006) show that personality traits taken together have an effect on wages that is comparable to that of cognitive abilities (IQ, for example). Heckman & Kautz (2012) show how cognitive ability alone cannot explain differences in economic outcomes. They further show that personality traits cause differences in outcomes. Finally, they explain the success of preschool programs on participants' outcomes, even though such programs have no effect on cognitive ability, through improvements in soft skills that participants get from such programs. Bowles, Gintis, & Osborne (2001) also show that personality traits and behaviors (for example, acting aggressively and getting arrested) are needed to explain wage differences in individuals with similar education levels and cognitive abilities. Kuhn & Weinberger (2005) use high-school leadership as a proxy for soft skills, and find that high-school leaders have higher wages as adults, controlling for demographics and cognitive ability. Weinberger (2014) finds that soft skills are increasingly valuable when paired with cognitive skills and that there has been a shift in the last twenty years from single-skilled individuals (those using either cognitive or soft skills) to multi-skilled individuals (those using both cognitive and social skills), with an associated wage premium. Deming (2015) creates a model in which soft skills allow workers to act as a team and utilize their comparative advantages in tasks to become more productive and finds empirical evidence that the labor market rewards soft skills.

Soft skills can also be used to explain inequality. For example, Borghans, Weel, & Weinberg (2006) find that differences in interpersonal style (which they define as a caring-directness spectrum) can explain part of the gender wage gap. Urzua (2006) finds that differences in noncognitive abilities explains part of the black-white gap in incarceration rates (although he did not find that it explained a significant part of the wage gap). Bacolod & Blum

(2010) find that rising returns to people skills relative to motor skills can help explain the narrowing gender wage gap in the 1970s and 1980s. Soft skills can also be used to explain instances of reduced inequality, most notably the narrowing of the gender gap, as Black & Spitz-Oener (2007) describe.

Soft skills can potentially explain other puzzles in the labor literature. Overqualification, for example, seems at first like an economically irrational decision. Brynin & Longhi (2009) demonstrate that overqualification resulting in an additional certificate can sometimes be a rational decision because it results in a wage premium. Working a job requiring lots of soft skills could be a situation in which it is rational to be overqualified if it results in a higher wage premium. As detailed in Buchel (2001), while there is an extensive literature measuring overqualification and its effects, the decision of the employer to hire an overqualified worker "almost never features" in the empirical literature. Buchel also discusses the lack of soft skill analysis in the overqualification research, and he suggests that soft skills and "professional expertise" could be substitutes – meaning that an employer would be willing to hire either a worker perceived as having greater soft skills, such as a recent college graduate, than a worker with fewer soft skills but more experience.

Chapter 3

Data

This paper uses two major datasets. The first is the Current Population Survey (CPS) Merged Outgoing Rotation Groups (MORG) data from the National Bureau of Economic Research. This rich dataset includes demographic information, wage information, and industry and occupation information for each outgoing (in their 4th or 8th month of interviewing) member of the 50,000-60,000 households the CPS surveys every month. The second comes from O*Net, which is the successor to the Dictionary of Occupational Titles (DOT) and is the United States' primary source of occupational data. It includes skills and education required, for each occupation by Standard Occupational Classification (SOC) code. I used the skill and jobzone (education required) datasets from O*Net.

3.1 O*Net data

The O*Net data includes measures of importance for 35 skills in seven categories of skills for 953 occupations. It measures the importance of each skill using two different methods

Table 3.1: Skills by Category

Content	Reading Comprehension	Active Listening
	Speaking	Writing
	Mathematics	Science
Process	Critical Thinking	Active Learning
	Learning Strategies	Monitoring
Social	Social Perceptiveness	Coordination
	Persuasion	Negotiation
	Instructing	Service Orientation
Problem-Solving	Complex Problem Solving	
Technical	Operations Analysis	Technology Design
	Equipment Selection	Installation
	Programming	Operational Monitoring
	Operation and Control	Equipment Maintenance
	Troubleshooting	Repairing
	Quality Control Analysis	
Systems	Judgment and Decision Making	Systems Analysis
	Systems Evaluation	
Resource Management	Time Management	Mgmt of Financial Resources
	Mgmt of Personnel Resources	Mgmt of Material Resources

("importance" and "level"), for a total of 66,710 data points. I use the "importance" measure, as that is the measure O*Net uses to define if a skill is important or not in its database. I define "soft skills" in this paper as the six skills that O*Net classifies as "social skills." There are skills that O*Net classifies as another type of skill that employers may consider to be a soft skill, such as "active listening" which is classified as a content skill. However, because there is no strict definition of a soft skill in the literature, I use the O*Net classification. Deming (2015) also uses O*Net's defined social skills to create a measure of "social skill intensity", although he used the 1998 version, which did not include "Instructing" or "Service Orientation" as skills (it did include "Service Orientation" as a task, which Deming included in his measure of "service task intensity"). The skills are not time-varying over the period I use; this version of O*Net has been in use since 2011. Table 3.1 lists the seven categories of skills as described by O*Net.

Because I am interested in the relative importance of "soft" skills (defined by O*Net as "social" skills), I calculated the ratio of social skills to total skills needed for the job for each occupation. I define this as an occupation's skill ratio. The literature uses ratios to measure an occupation's task intensity (for example, Black & Oener-Spitz (2007) use task ratios identical to these skill ratios for different categories of tasks, and Autor & Dorn (2013) create measures of routine task intensity by taking the natural logarithm of the ratio of routine tasks to non-routine tasks). A high skill ratio (closer to 1) signifies an occupation that requires relatively more social skills than other skills. A low skill ratio (closer to 0) signifies an occupation that requires relatively fewer social skills. I classify occupations requiring zero total skills as having a skill ratio of 0. In robustness analysis, I modify how zero total skill occupations are classified, and it does not greatly affect the results.

O*Net classifies a skill as "important" to an occupation if it ranks at least a 3 on a 1 to

5 scale from the surveys. I counted the number of "important" soft skills and total skills when calculating these ratios. I did not weight the skills by importance. For example, O*Net classifies nineteen total skills and six soft skills as "important" for food service managers, so the skill ratio for food service managers would be $\frac{6}{19} = 0.316$.

The skill ratios ranged from 0 to 1, with a mean and median value of 0.19046 and 0.19048, respectively. Ten of the 516 occupations required no skills (for example, laborers and mail carriers). O*Net classifies these types of occupations as needing "knowledge" like public safety and "abilities" like manual dexterity rather than skills to carry out their occupational tasks. When I remove those occupations from the dataset, the mean skill ratio is 0.193 and the median is 0.199. Figure 3.1 shows a histogram for the distribution of skill ratios including the zero-skill occupations. Table 3.2 displays a sampling of occupations with high, medium, and low skill ratios.

The distribution of the histogram has a large collection (92 of the 517 occupations in the data) of occupations for which zero social skills are important, most occupations falling somewhere between 0.1 and 0.35, and two occupations for which only social skills are important. Some of the occupations requiring zero soft skills are in the left-hand column of Table 3.2. The middle column has occupations near the median (middle ten percent) of the skill ratio distribution. The right column has occupations at or near the maximum (upper ten percent) of the skill ratio distribution.

3.2 CPS MORG data

I used the CPS MORG data from 2011 to 2014, a time period during which the same set of occupational codes were used (the codes were edited between 2010 and 2011). I include only observations for working-age people 18-65 years old who reported an industry,

Figure 3.1: Skills Histogram

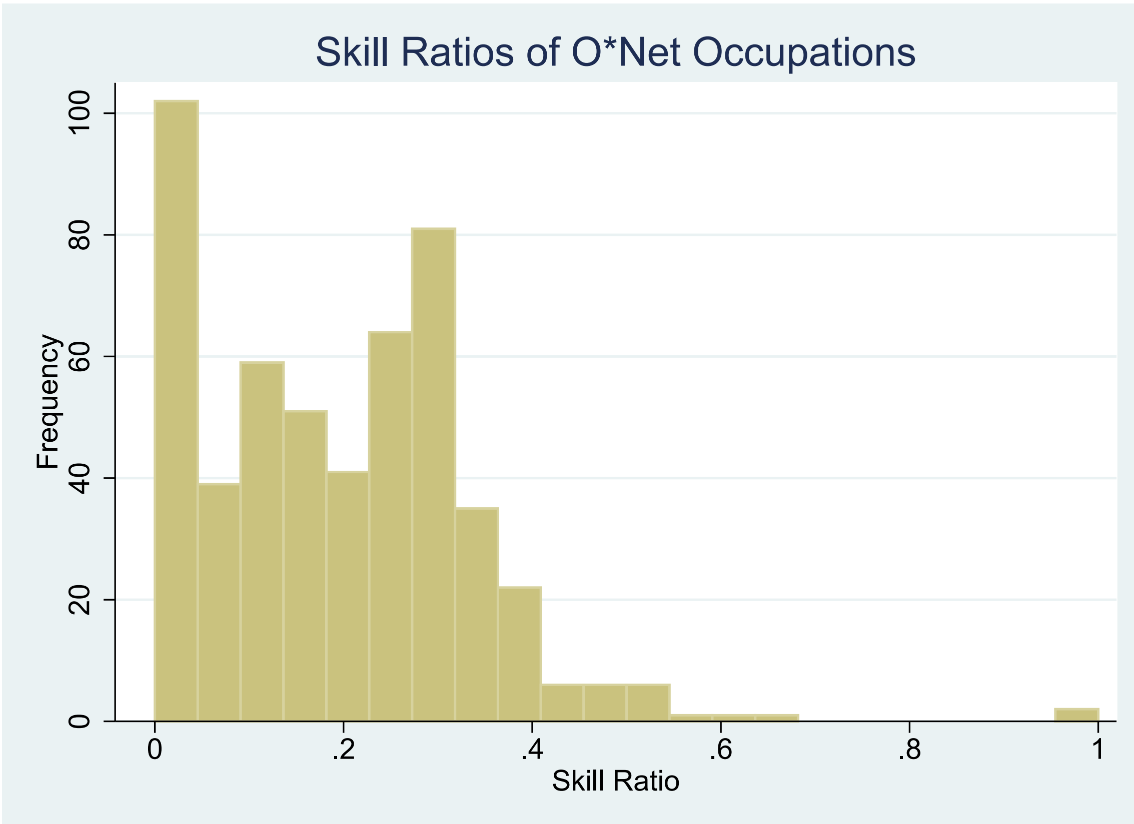

Table 3.2: Examples of Occupations with Low, Medium, and High Soft Skill Ratios

Minimum Skills Ratio	Median Skills Ratio	Maximum Skills Ratio
Riggers	Sociologists	Dining Room Attendants
Shoe Repairers	Animal Caretakers	Maids and Housekeeping Cleaners
Mail Carriers	Brokerage Clerks	Telemarketers
Typists	Tax Preparers	Tour and Travel Guides
Boiler Operators	Aircraft Pilots	Hosts and Hostesses
Electric Tool Repairers	Computer Operators	Special Education Teachers
Fishers	Marketing Specialists	Retail Salespersons
Furnace Operators	Actors	Nursing Aides
Computer Programmers	Paralegals	Models
Biological Technicians	Fire Inspectors	Pharmacy Aides

an occupation, and earnings and who were not full-time students. I adjusted earnings to reflect 2000 dollars, multiplied the top-coded earnings by 1.4 as is standard in the literature, and took the log of earnings. For people who reported hours and weekly earnings but not hourly earnings, I approximated hourly earnings by dividing weekly earnings by hours per week. 144,997 of the 715,420 observations that reported an occupation had to be dropped because I could not calculate hourly earnings for them (132,038 did not report any earnings, and 12,959 reported only weekly earnings but neither hourly earnings nor hours per week). Because I was interested in hourly wages rather than weekly earnings, I did not exclude part-time workers, following Card & DiNardo (2002). For a sub-analysis without the effects of gender and race, I limit the data to include only white men. For a second sub-analysis, I limit the data to include only people with five years' or less of potential experience.

I do not weight earnings by hours. I do this because my research question relates to the hourly wage paid to an individual, similar to papers like Acemoglu & Autor (2011) which also do not weight earnings by hours.

3.3 Merging the O*Net and CPS data

To merge these two datasets, I used Census occupation and industry crosswalks from the Bureau of Labor Statistics to match the Census occupation codes in the CPS data to the SOC codes in the O*Net data. There were many observations in the CPS data where the occupation was less detailed (4- or 5-digit SOC code) than in the O*Net data (6-digit SOC code). For example, an occupation in the CPS data might be "elementary and middle school teachers" while O*Net will break that category down into "elementary school teachers," "middle school teachers," and "career/technical education teachers, middle school." To merge these observations, I paired the 6-digit codes with their corresponding 4- or 5-

digit codes (for example, 123456 would pair with 123450 and 123400). If there was only one more detailed O*Net occupation in the broader CPS occupation, then I simply appended the O*Net data to the CPS observation. However, there were 117 CPS occupations that contained multiple detailed O*Net occupations within them.

When an occupation in the CPS had more than one possible occupation in the O*Net data, I follow Autor, Levy, & Murnane (2003) and Acemoglu & Autor (2011) and used data from the Occupational Employment Statistics (OES) Survey, which provided a crosswalk on how many people worked in each detailed occupation. I used this data to create a weight for each detailed occupation (for example, if there were 500,000 elementary school teachers, 400,000 middle school teachers, and 100,000 career/technical education teachers, the weights would be 0.5, 0.4, and 0.1 respectively). I then took the weighted average of data from each detailed occupation to create the measures for the broader occupations. For example, if a person surveyed in the CPS was a "marketing and sales manager" (SOC code 112020), then to find that person's occupational skill ratio, I took the average of the skills ratio for "marketing manager" (SOC code 112021) which was 0.286, weighted by 0.34 (which is the share of marketing managers in 112020), and the average of the skills ratio for "sales manager" (SOC code 112022) which was 0.273, weighted by 0.66, to give an occupational skill ratio of 0.277.

The exception to this weighting strategy was required education for an occupation. Because I used required education to measure overqualification, I took the maximum value to avoid overstating overqualification. For example, the CPS occupation "financial managers" includes both O*Net occupations "treasurers and controllers" which require a graduate degree and "financial managers, branch or department" which require only a bachelor's degree, so I classify "financial managers" as requiring a graduate degree. Forty-one of the 517 CPS occupations for the 2011-2014 time period had different levels of required education for different O*Net occupations within the CPS occupation.

There was one broader occupation (computer support specialists) that did not have a breakdown for the number of workers in the more detailed occupations (computer user support specialists and computer network support specialists) in the 2011 OES Survey, so I assigned those 2011 occupations their 2012 weights. This only affected 585 individual observations in the CPS data.

Chapter 4

Empirical Strategy

To evaluate the hypothesis that employers screen for soft skills on the basis of education, I consider three different dependent variables: overqualification, intra-occupation wage variance, and wage premiums to employees who have at least a bachelor's degree but whose occupation does not require it.

I use the control variables whose summary statistics are presented at the individual level in the first column of Table 4.1. I also include industry dummies in one analysis, but because there are 271 industries in the data, I do not include them in the summary statistics table. I include as many control variables as I can that have been shown to be significant to explaining wages in the literature. Because the CPS data is from Census, these are mostly individual demographic controls. These control variables (quartic for potential experience and dummy variables for the following: race, sex, union membership, metropolitan area, marital status, children present in household, foreign-born, job sector, level of education completed, region, and industry) are consistent with the soft skills literature (Balcar (2014), Hirsch (2005)). All variables except experience can be found in the CPS data. I define potential experience as age minus years of schooling minus six, as is standard in the literature. If potential experience for an individual using this formula is a negative number, I change it to zero. I then make it a quartic, following Lemieux (2003).

For each dependent variable, I estimate three regressions. The first includes only white men, following Wheeler (2005), Weinberger (2014), and others that do this to eliminate race and gender effects. Summary statistics for the sample of white men are presented in the second column of Table 4.1. The second regression includes all observations. The third regression includes only individuals who have five years or less of potential experience, as Belman & Heywood (1997) show that sheepskin effects from signaling are strongest for individuals with the least work experience. If employers are using education to screen for soft skills, then we would expect the soft skill ratio of an occupation to have a larger effect for the less experienced individuals. Summary statistics for this sample are presented in the third column of Table 4.1.

4.1 Overqualification

I defined an individual in the CPS data as "overqualified" if he or she had more education than is required for his or her occupation according to O*Net's required education. I estimate a linear regression because about twenty percent of individuals in the CPS data are overqualified, which is far enough away from zero, although the results are similar if a probit regression is used instead. For this regression, I removed from the dataset all individuals that could not be overqualified (who were in occupations requiring the highest possible job zone or who had the lowest possible level of education: high school dropouts). The results are robust to this adjustment.

Table 4.1: **Summary Statistics**

Variable	**All Observations**	**White Men**	**Inexperienced People**
Hourly Wages	2.599	2.759	2.264
Skills Ratio	0.247	0.205	0.269
OQ	0.202	0.189	0.267
White	0.703		0.690
Black	0.096		0.089
Hispanic	0.119		0.130
Female	0.493		0.508
Experience	21.587	22.031	2.751
Union Member	0.121	0.130	0.065
Metropolitan	0.807	0.772	0.825
Married	0.576	0.615	0.177
Children	0.367	0.343	0.127
Immigrant	0.157	0.047	0.098
Public Sector	0.172	0.149	0.111
Nonprofit	0.076	0.050	0.073
HS Dropout	0.071	0.043	0.047
HS Grad	0.457	0.474	0.475
Associate's	0.114	0.110	0.082
Bachelor's	0.233	0.247	0.285
Advanced Degree	0.125	0.125	0.111
New England	0.109	0.129	0.102
Mid-Atlantic	0.094	0.092	0.093
East North Central	0.116	0.132	0.113
West North Central	0.124	0.149	0.137
South Atlantic	0.179	0.162	0.178
East South Central	0.046	0.050	0.042
West South Central	0.087	0.070	0.089
Mountain West	0.104	0.112	0.105
Pacific West	0.142	0.102	0.141
N	536,657	192,164	65,894

Source: CPS MORG data, 2011-2014

I estimate:

$$OQ_{i,j,t} = \alpha + SkillsRatio_j\gamma + Z_{i,t}\zeta + \delta_t + \epsilon_{i,j,t} \tag{4.1}$$

where $OQ_{i,j,t}$ is the overqualification status of individual i in occupation j at time t, δ_t is a vector of time dummies, $SkillsRatio_j$ is the soft skills ratio of occupation j, $Z_{i,t}$ is a vector of individual characteristics, and $\epsilon_{i,j,t}$ is a mean-zero error term.

In $Z_{i,t}$, I use a variety of demographic dummies and a measure of potential experience. I include dummies for gender and race, union membership, geographic location (metropolitan/non-metropolitan area and regional indicators), marital status and presence of children in household, job sector, and immigration status. The regression including only white men does not have the race and gender dummies.

Note that the educational dummies are not included in these control variables. This is because I define overqualification as having an educational credential greater than required for the job, so including the educational dummies creates collinearity problems. However, I do control for the occupation's required education as a robustness analysis.

I use robust standard errors for all regressions in this paper.

4.2 Wage Premium

I consider all individuals in the CPS data who had at least a bachelor's degree and who worked in occupations that required less than a bachelor's degree. I then calculated the difference between their logged hourly earnings and the average logged hourly earnings of their occupation, then I undid the logarithm and subtracted one to make this number a percentage difference between the hourly earnings and the average hourly earnings. I call this the "wage premium," $WP_{i,t}$ (note: even though I call this a premium, some values are negative because people are earning less than the occupation average). I then estimate the following linear regression:

$$WP_{i,j,t} = \alpha + SkillsRatio_j \gamma + \delta_t + Z_{i,t}\zeta + \epsilon_{i,j,t} \tag{4.2}$$

The individual characteristics in $Z_{i,t}$ include the same characteristics as those in Equation **??**, but with the addition of a dummy for an advanced degree (I can include education because my wage premium measure does not have the same education colinearity problem as my overqualification measure, but because I include only individuals with at least a bachelor's degree, the advanced degree dummy is the only one I can include).

I again estimate regressions for the full sample, only white men, and only people with five years' or less possible experience.

4.3 Wage Variance

Then, I consider intra-occupation wage variance. If employers are screening for soft skills on the basis of education, then we would expect to see greater wage variance in occupations in which soft skills are relatively more important because workers with more education are expected to be relatively more productive and workers with less education are expected to be relatively less productive than their counterparts in occupations where soft skills are relatively less important.

Total wage variance is calculated as:

$$V_t = \frac{1}{N_t}\sum_{j=1}^{J_t}\sum_{i=1}^{N_{j,t}}(w_{j,i,t} - \bar{w_{j,t}})^2 + \frac{1}{N_t}\sum_{j=1}^{J_t}\sum_{i=1}^{N_{j,t}}(\bar{w_{j,t}} - \bar{w}_t)^2 \tag{4.3}$$

Table 4.2: Occupations with High, Medium, and Low Levels of Intra-Occupation Wage Variability

Minimum Wage Variability	Median Wage Variability	Maximum Wage Variability
Textile Machine Operators	Security Guards	Computer Research Scientists
Fish and Game Wardens	Fishers	Advertising Sales Agents
Glass-forming Machine Operators	Truck Drivers	Pilots
Etchers and Engravers	Ticket Agents and Travel Clerks	Petroleum Technicians
Dishwashers	Brick, Block, and Stone Masons	Authors
Tool Sharpeners	Construction Workers	Benefits Managers
Roustabouts	Supervisors of Production Workers	Actors
Veterinary Assistants	Biomedical Engineers	Procurement Clerks
Explosives Workers	Occupational Therapy Assistants	Optometrists
Pressers	Urban Planners	Judicial Law Clerks

where t is year, i is worker, j is occupation, $N_{j,t}$ is the number of workers in occupation j in year t, $w_{j,i,t}$ is an individual's hourly wage, $\bar{w_{j,t}}$ is the occupation's average wage, and $\bar{w}_t$ is the year's average wage. The first term represents intra-occupation wage inequality (e.g. one doctor getting paid more than another doctor) and is the term of interest. The second term represents inter-occupation wage inequality (e.g. doctors getting paid more than nurses).

I take the first term of Equation 4.3 and represent it as :

$$Var_{j,t} = \frac{1}{N_{j,t}} \sum_{i=1}^{N_{j,t}} (w_{j,i,t} - \bar{w_{j,t}})^2 \tag{4.4}$$

which is the same as the first term of Equation 4.3, but without summing over all occupations. Table 4.2 gives a sampling of occupations with the lowest (smallest ten percent), median (middle ten percent), and highest (top ten percent) wage variability.

I then estimate the following regression:

$$Var_{j,t} = \alpha + \delta_t + X_{j,t}\beta + \epsilon_{j,t} \tag{4.5}$$

where δ_t is a vector of time dummies, $X_{j,t}$ is a vector of occupation characteristics, and $\epsilon_{j,t}$ is a mean-zero error term. Note that because the skills ratio of the occupation does not vary across time, I cannot include occupation fixed effects. Instead, I control for a large set of occupational characteristics.

In $X_{j,t}$ I attempt to include as many occupational controls as the CPS dataset will allow me to. Because the CPS dataset mostly contains individual characteristics, as my occupational controls I use a variety of demographic variables that represent the decimal of each demographic group represented in the data in each occupation, with the exception of potential experience, which is given as the average of the potential experience of the individuals in the occupation. I include measures for gender and race, union membership, geographic location (metropolitan/non-metropolitan area and regional indicators), marital status and presence of children in household, job sector, immigration status, and educational attainment. Note that these are the same variables as in the overqualification and wage premium regressions, but with the addition of a full set of educational attainment controls, and at the occupational rather than individual level. The regression including only white men does not have the race and gender controls.

Chapter 5

Results

5.1 Overqualification

Table 5.1 shows the results of Equation 4.1. In the table, I leave out the regional dummies and the year dummies. While some of the regional dummy coefficients were statistically significant (probably because of the very large sample size), their coefficients were very small. The year dummies were not statistically significant.

Here, there is strong evidence that in occupations where soft skills are relatively more important there is more overqualification. In all three regressions, the estimated skills ratio coefficient is positive and statistically significant. In addition, its coefficient is the largest of any of the tested variables. Further, the coefficient on the skills ratio in the sample of individuals with five years' or less experience is larger than in the full sample. These are results I would expect if employers were indeed screening for soft skills using education.

The difference between an occupation with zero soft skills and only soft skills (ratio of 0 and 1) is 18.6 percent, 24.2 percent, and 32.5 percent higher overqualification for the entire sample, inexperienced people, and white men, respectively. This is a rather large effect especially because only 20 percent of the entire sample is overqualified. However, the R^2 is very small, indicating that there are many other factors affecting overqualification.

5.2 Wage Premium

Table 5.2 shows the results of Equation 4.2. Again, I leave out the regional dummies (some of which were statistically significant but had very small coefficients) and the year dummies (none of which were statistically significant).

The estimated skill ratio coefficient is positive and statistically significant. It is again the largest coefficient out of any of the tested variables except in the case of the inexperienced individuals, for which the advanced degree dummy is slightly larger. I construct the wage premium measure so that the coefficient is a percentage change. This estimation indicates that an individual who has a bachelor's degree but whose occupation does not require one will see a 20 percent higher wage premium in an occupation with a skill ratio of one (requiring only soft skills) than in an occupation with a skill ratio of zero (requiring no soft skills).

The advanced degree dummy coefficient is also relatively large, positive, and statistically significant, which makes sense because I would expect that advanced degree holders would command a larger wage premium than those holding only a bachelor's degree. The advanced degree effect from soft skills is much larger – twenty percent higher – for inexperienced workers. The wage premium regressions also offer evidence that employers do indeed screen for soft skills on the basis of education.

Table 5.1: Effect of Soft Skills on an Individual's Overqualification

	(1) White Men Only	(2) All Observations	(3) 0-5 Yrs Exp
Skills Ratio	0.325***	0.186***	0.242***
	(41.35)	(40.51)	(18.74)
Experience	-0.00218	-0.00312**	-0.0319
	(-1.37)	(-3.18)	(-1.71)
Exp Squared	-0.000368**	-0.000255**	0.0356*
	(-2.68)	(-3.00)	(2.07)
Exp Cubed	0.0000205***	0.0000148***	-0.0126*
	(4.62)	(5.38)	(-2.35)
Exp Fourth	-0.000000302***	-0.000000232***	0.00131*
	(-6.37)	(-7.86)	(2.48)
Union Member	0.000441	0.0313***	0.0334***
	(0.14)	(14.22)	(3.89)
Metropolitan	0.0430***	0.0471***	0.0582***
	(17.74)	(28.33)	(11.25)
Married	0.0442***	0.0368***	0.105***
	(17.09)	(23.95)	(17.13)
Immigrant	0.117***	0.0969***	0.0833***
	(19.77)	(39.89)	(10.89)
Working in Public Sector	0.0884***	0.0857***	0.0230**
	(26.24)	(42.65)	(3.23)
Working in Nonprofits	0.0939***	0.0425***	-0.0118
	(15.14)	(14.24)	(-1.33)
Children in HH	-0.0264***	-0.0326***	-0.0624***
	(-9.55)	(-19.45)	(-9.66)
White		-0.0268***	-0.0113
		(-9.10)	(-1.46)
Hispanic		-0.128***	-0.103***
		(-39.19)	(-11.90)
Black		-0.0689***	-0.0808***
		(-19.53)	(-8.28)
Female		0.0340***	0.0618***
		(23.59)	(14.58)
N	156257	416766	54497
R^2	0.038	0.041	0.037

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are regional dummies and year dummies

Table 5.2: Effect of Soft Skills on an Individual's Wage Premium

	(1) White Men Only	(2) All Observations	(3) 0-5 Yrs Exp
Skills Ratio	0.189***	0.216***	0.315***
	(6.54)	(13.00)	(5.57)
Experience	0.0369***	0.0379***	-0.0107
	(6.49)	(11.83)	(-0.15)
Exp Squared	-0.00107	-0.00153***	0.0342
	(-1.88)	(-4.68)	(0.50)
Exp Cubed	0.0000114	0.0000287*	-0.00490
	(0.55)	(2.37)	(-0.23)
Exp Fourth	-7.84e-08	-0.000000249	0.0000806
	(-0.32)	(-1.70)	(0.04)
Union Member	0.0380**	0.0383***	0.0258
	(3.22)	(4.99)	(0.77)
Metropolitan	0.0990***	0.108***	0.0661**
	(8.97)	(15.46)	(2.74)
Married	0.0497***	0.0304***	0.0373
	(4.75)	(5.36)	(1.63)
Immigrant	-0.0855***	-0.0738***	-0.0127
	(-5.00)	(-9.10)	(-0.32)
Working in Public Sector	-0.0470***	-0.0199**	-0.0243
	(-4.19)	(-3.03)	(-0.80)
Working in Nonprofits	-0.0530**	-0.0324***	-0.0150
	(-2.98)	(-3.66)	(-0.41)
Children in HH	0.0581***	0.0196**	-0.0341
	(5.13)	(3.21)	(-1.12)
Advanced Degree	0.0702***	0.0749***	0.267***
	(5.50)	(10.20)	(7.30)
White		0.00102	-0.0408
		(0.10)	(-1.07)
Hispanic		-0.0261*	-0.0993*
		(-2.23)	(-2.47)
Black		-0.0367**	-0.139**
		(-3.00)	(-3.19)
Female		-0.127***	-0.131***
		(-25.09)	(-7.72)
N	17028	48263	8146
R^2	0.068	0.062	0.046

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are regional dummies and year dummies

5.3 Wage Variance

Table 5.3 shows the results of Equation 4.5. Although the regression included regional controls, a marriage control, a foreign-born control, a children-present control, job sector controls, and year dummies, I suppress the coefficients on those variables in Table 5.3 because none of them were statistically significant.

For none of the three regressions was the estimated soft skills ratio coefficient statistically significant. Only the coefficients on education dummies and the dummy for urban area were statistically significant. All four of the dummies included had a negative coefficient, suggesting that advanced degrees would have had a positive coefficient had I included that dummy and left out a different education dummy.

Generally, these results do not offer evidence supporting the hypothesis for wage variability being accounted for by soft skills required by an occupation. The lack of evidence could be either economic or methodological. If employers are indeed screening for soft skills on the basis of education, then rather than adjusting the wages they pay in a way that would increase wage variability, it is possible that they simply would not hire a potential worker with less education (this would be consistent with the results in the test for overqualification). However, it is equally possible that occupations where soft skills are important indeed have more wage variability, but the sample size (less than 2000) was too small to show this conclusively. We see that in the sample of all individuals, the coefficient on the soft skills ratio is positive, and its t-value is 1.45, which is almost significant at the 90% level, indicating that this might be the case.

Table 5.3: Effect of Soft Skills on an Occupation's Wage Variability

	(1) White Men Only	(2) All Observations	(3) 0-5 Yrs Exp
Skills Ratio	-0.127	0.0982	0.0764
	(-0.67)	(1.45)	(0.84)
Experience	-0.206	-0.0745	-0.299
	(-1.18)	(-0.62)	(-1.10)
Exp Squared	0.0174	0.00640	0.334
	(1.20)	(0.62)	(1.40)
Exp Cubed	-0.000543	-0.000187	-0.118
	(-1.17)	(-0.57)	(-1.54)
Exp Fourth	0.00000561	0.00000173	0.0126
	(1.13)	(0.50)	(1.61)
Avg. Union Membership	-0.181*	-0.0526	-0.0155
	(-2.25)	(-1.35)	(-0.24)
Percent Metropolitan	0.0695	0.160**	0.0814*
	(0.92)	(2.74)	(2.00)
Percent HS Dropouts	-0.566	-0.187	-0.353
	(-1.71)	(-1.18)	(-1.95)
Percent HS Grads	-0.274	-0.211	-0.284
	(-1.00)	(-1.71)	(-1.74)
Percent with Associate's	-0.487*	-0.366***	-0.239
	(-2.18)	(-3.67)	(-1.35)
Percent with Bachelor's	-0.217	-0.130	-0.212
	(-0.60)	(-0.73)	(-1.00)
Percent White		0.220	0.0993
		(1.77)	(1.23)
Percent Black		-0.0437	-0.152
		(-0.28)	(-1.67)
Percent Hispanic		-0.00000880	0.119
		(-0.00)	(1.14)
Percent Female		-0.00999	-0.0262
		(-0.35)	(-0.88)
N	1893	1914	1758
R^2	0.045	0.083	0.027

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are regional controls, a marriage control, a foreign-born control, a children-present control, job sector controls, and year dummies

Chapter 6

Robustness Checks

I then performed a series of robustness checks on my results. For all robustness checks, I used all observations rather than either the white men or the inexperienced individuals subsets.

6.1 Required Education

The first robustness check I perform is estimating a regression including an occupation's required education as a control variable. I add the set of dummy variables to each of the three main regressions using wage variability, overqualification, and wage premiums as dependent variables. Note that each dependent variable uses a different set of dummy variables for required education. Overqualification only uses three of the four dummies because I drop all individuals whose job requires an advanced degree and therefore cannot be overqualified. Wage premium only uses two of the four dummies because I keep only individuals who have at least a bachelor's degree but whose job does not require one, so I do not use the two year or four year degree dummies. The results can be seen in Table 6.1.

In the overqualification regression, adding required education reduced the skills ratio variable coefficient, but it was still statistically significant. This is not a surprise because I define overqualification as having more education than is required, so the higher the required education, the fewer individuals can possibly be overqualified. The coefficient on the dummy for an occupation not requiring a high school education (the lowest possible required education) is enormous, but again this is not a surprise because I drop all individuals who do not have a high school degree because they cannot be overqualified, so all individuals in the sample are overqualified for these occupations.

In the wage premium regression, adding required education dummies actually increases the skills ratio coefficient and decreases the R^2, suggesting that these dummies are not necessarily appropriate for this dependent variable. However, as expected, wage premiums are larger the less education the occupation requires, which makes sense because all individuals in this test have at least bachelor's degrees, so the less education the occupation requires, the "more overqualified" they are.

Adding required education had little effect on the coefficient of the skills ratio in the wage variance regression. However, the greater the education required by the occupation, the greater than wage variability is, and for occupations requiring bachelor's degrees, this effect is highly significant. This is likely because the greater an occupation's required education, the higher its wages generally are, which allows for more wage variability compared to low-wage occupations.

Overall, controlling for required education did not have a very large effect on the results.

Table 6.1: Controlling for Required Education

	(1) Indiv. Overqualification	(2) Indiv. Wage Premiums	(3) Occ. Wage Variability
Skills Ratio	0.0909***	0.466***	0.0979
	(23.20)	(15.82)	(1.46)
No HS Req	0.821***	0.106***	-0.0523
	(581.54)	(4.47)	(-0.82)
HS Req	0.0866***	0.0726***	-0.00775
	(54.60)	(8.21)	(-0.17)
2 Yr Deg Req	0.0235***		0.0276
	(14.22)		(0.74)
4 Yr Deg Req			0.0710**
			(2.92)
N	416766	48263	1914
R^2	0.166	0.057	0.089

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are quartic experience, year dummies, and dummies for the overqualification and wage premium tests and controls for the wage variability test for: race, gender, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, and educational controls for the wage premium and wage variability tests

6.2 Controlling for Zero-Skill Occupations

I define an occupation's soft skill ratio as the number of soft skills required divided by the total number of skills required. However, there were ten occupations for which no skills were required. In my main regressions, I classify these occupations as having a skill ratio of zero. To examine whether the zero-skill occupations have a large effect on my results, I create a dummy for occupations having zero skills and re-estimate my main three regressions. I then remove the zero-skill occupations from my dataset and re-estimate the three regressions without those occupations. The results can be seen in Table 6.2.

There is very little effect for the overqualification and wage variability regressions. However, controlling for zero-skill occupations more than doubles the effect of soft skills on wage premiums. This is probably because the individuals in this sample (who have at least bachelor's degrees and who are working in jobs that do not require them) work in zero-skill occupations at half of the rate as all individuals working in jobs that do not require bachelor's degrees (there are no occupations that require both zero skills and more than a high school diploma). Table 6.3 shows the numbers and percentages of workers with and without bachelor's degrees in zero-skill occupations.

I also included the results for the dummy for female variable in Table 6.2 because including the zero-skill dummy had a very large effect on the coefficient of female dummy in the wage premium regression, even though the proportion of women with a bachelor's degree was very similar to the overall proportion in zero-skill occupations, as can be seen in Table 6.3. Because there are few women in zero-skill occupations and women earn a lower wage premium for having a bachelor's when they do not require one, including the zero-skill dummy led to the coefficient for female becoming even more negative, and the coefficient for the skills ratio to become even more positive.

Table 6.2: Controlling for Zero-Skill Occupations

	(1) OQ, Dummy	(2) OQ, Omitted	(3) WP, Dummy	(4) WP, Omitted	(5) Var, Dummy	(6) Var, Omitted
Skills	0.182***	0.186***	0.494***	0.501***	0.0984	0.0957
Ratio	(39.33)	(39.66)	(16.48)	(16.49)	(1.44)	(1.37)
Zero	-0.0181***		-0.0583**		0.00249	
Skill	(-5.40)		(-2.88)		(0.14)	
Female	0.0337***	0.0331***	-0.223***	-0.226***	-0.00996	-0.0104
	(23.41)	(22.56)	(-24.44)	(-24.33)	(-0.35)	(-0.36)
N	416766	401182	48263	46866	1914	1874
R^2	0.041	0.042	0.055	0.056	0.083	0.082

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are quartic experience, year dummies, and dummies for the overqualification and wage premium tests and controls for the wage variability test for: race, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, and educational controls for the wage premium and wage variability tests

Table 6.3: Workers In Zero-Skill Occupations

	(1) All Workers	(2) With Bachelor's	(3) Without Bachelor's
Number in Zero-Skill Jobs	18,574	1,397	17,177
Percent in Zero-Skill Jobs	5.70%	2.89%	6.19%
N	325,740	48,341	277,399
Percent Female	27.20%	28.92%	27.07%

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.3 Controlling for Industry

Many studies control for industry in occupation-related regressions. I do the same here by creating dummies for each of the 271 industries in the CPS data and re-estimating the three regressions. The results can be seen in Table 6.4. I did not include the industry coefficients in the tables because of space constraints.

In the overqualification regression, controlling for industry led to a smaller coefficient on the skills ratio variable. While the coefficient was still large, positive, and statistically significant, it was no longer the largest coefficient (the dummy for being Hispanic became the largest coefficient when this regression was estimated). Most of the industry dummies were statistically significant, which is to be expected because of the very large number of observations. However, several of the industry dummies had very large coefficients.

In the wage premium regression, controlling for industry actually led to a larger coefficient on the skills ratio variable. Similarly to the overqualification regression, most of the industry dummies were statistically significant, and several had large values.

In the wage variability regression, a handful of the industry dummies were statistically significant, but the majority were not. Including the industry dummies had a large negative effect on the skill ratio coefficient, although the skill ratio variable was still not statistically significant. Controlling for industry results in the dummies for education becoming more statistically significant and in the dummy for union membership becoming slightly statistically significant.

Table 6.4: Controlling for Industry

	(1) Indiv. Overqualification	(2) Indiv. Wage Premiums	(3) Occ. Wage Variability
Skills Ratio	0.101***	0.439***	-0.0747
	(20.73)	(13.60)	(-0.62)
N	416766	48263	1914
R^2	0.082	0.075	0.231

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are 270 industry dummies, quartic experience, year dummies, and dummies for the overqualification and wage premium tests and controls for the wage variability test for: race, gender, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, and educational controls for the wage variability test

6.4 Controlling for Individual Soft Skills

I next estimated the same three regressions but used the six individual soft skills (the social skills defined by O*Net: social perceptiveness, persuasion, instructing, coordination, negotiation, and service orientation) instead of the skills ratio measure. I created six dummies for whether O*Net defined each skill as important to the occupation. The results can be seen in Table 6.5.

Interestingly, two of the individual soft skills were statistically significant for the wage variability regression. Social perceptiveness had a positive coefficient and was highly significant, and instructing had a negative coefficient and was significant. This suggests that certain individual soft skills have effects on wage variability, but taken together they have little effect.

For the overqualification regression, all of the skills except for negotiation were highly statistically significant, and of those, social perceptiveness and service orientation were positive, and instructing, coordination, and persuasion were negative. Service orientation had the largest coefficient by far, suggesting that for overqualification, service orientation is the driving skill behind the skill ratio effect.

For the wage premium regression, all skills except for persuasion were statistically significant. As for overqualification, social perceptiveness and service orientation had positive coefficients and coordination and instructing had negative coefficients, but here, negotiation was positive rather than negative. All of the coefficients were similar in size, so for the case of wage premiums, it appears that no single skill is driving the overall effect, and it is the ratio of skills together that has a larger effect.

It is interesting that "coordination" was a relatively unimportant soft skill considering the very strong results in Deming (2015), which develops a model in which soft skills are important because it allows workers to coordinate trading tasks and therefore be more productive as a team.

Overall, this analysis suggests that disaggregating soft skills is important, and that each may have different effects.

6.5 Separating by Skill Level

I then divide the data into three parts by skill level. Each quantile has a roughly equal number of occupations, but not necessarily an equal number of individuals. I divide the data by skill level in two ways. First, I divide by the total number of skills required by an occupation. "Low skill" occupations require 10 skills or fewer. "Medium skill" occupations require more than 10 skills but fewer than 17 skills. "High skill" occupations require 17 skills or greater. Second, I separate by the average wage, following Autor & Dorn (2013). 1216 occupation-year observations were in the same category (low, medium, or high) for

Table 6.5: Effect of Specific Soft Skills

	(1) Indiv. Overqualification	(2) Indiv. Wage Premiums	(3) Occ. Wage Variability
Social Perceptiveness	0.0192***	0.0629***	0.105**
	(8.17)	(5.02)	(2.60)
Persuasion	-0.0380***	0.0319	0.0123
	(-12.54)	(1.54)	(0.42)
Instructing	-0.0245***	-0.0844***	-0.0528*
	(-12.47)	(-6.14)	(-2.24)
Coordination	-0.0274***	-0.0318**	-0.00954
	(-15.32)	(-2.88)	(-0.42)
Negotiation	-0.00595*	0.0916***	0.0402
	(-1.96)	(4.22)	(1.14)
Service Orientation	0.0868***	0.0765***	-0.0476
	(40.15)	(5.98)	(-1.33)
N	416766	48263	1914
R^2	0.044	0.055	0.095

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are quartic experience, year dummies, and dummies for the overqualification and wage premium tests and controls for the wage variability test for: race, gender, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, and educational controls for the wage variability test both number of skills and wages, and 698 occupation-year observations were in different categories for the two measures.

The results for overqualification can be found in Table 6.6. Separating by wages, the middle wage jobs clearly have the highest effect from the skills ratio. This could be because in low-wage occupations, the premium that employers would have to pay for overqualified workers is greater than the value of the soft skills signal because these occupations are unattractive for overqualified workers, and in the high-wage occupations, the premium is greater than the value of the signal because the wage is already relatively high. Separating by number of skills required, the more total skills are required, the greater the effect of the skills ratio on overqualification. This is likely because higher skill jobs with high skills ratios require a greater number of different soft skills, whereas a low skill job with a high skill ratio might just require one or two soft skills, so the signal is more valuable.

The results for wage premiums can be found in Table 6.7 and tell a similar story to those of overqualification. Again, the medium-skill occupations have the largest effect of the skills ratio on wage premiums, which we would expect because the skills ratio also has the largest effect on overqualification in those occupations. The exception is that the effect of the skills ratio on wage premiums is slightly higher for medium skill occupations than for high skill occupations separating by number of total skills.

The results for wage variability can be found in Table 6.8. For both measures of skill, the effect of the soft skills ratio on wage variability was larger for the higher-skilled occupations. This finding is consistent with Weinberger (2014) and Deming (2015), which finds that soft skills and cognitive skills (which are usually associated with high-skill, high-wage occupations) are complementary. Because soft skills and cognitive skills are complementary, then we would expect that workers in high-skill occupations requiring soft skills to be relatively more productive if they possess soft skills and relatively less productive if they do not, which would result in higher wage variance than in low-skill occupations.

Table 6.6: Effect of Soft Skills on an Individual's Overqualification

	(1) Low Skill	(2) Medium Skill	(3) High Skill	(4) Low Wage	(5) Medium Wage	(6) High Wage
Skills Ratio	0.187***	0.241***	0.261***	0.160***	0.272***	0.0113
	(30.50)	(25.39)	(22.86)	(27.66)	(27.11)	(0.95)
N	156748	137457	122561	183246	120944	112576
R^2	0.047	0.042	0.071	0.031	0.046	0.094

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are quartic experience and dummies for race, gender, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, and year

Table 6.7: Effect of Soft Skills on an Individual's Wage Premium

	(1) Low Skill	(2) Medium Skill	(3) High Skill	(4) Low Wage	(5) Medium Wage	(6) High Wage
Skills Ratio	0.257***	1.055***	0.879***	0.438***	0.658***	0.159
	(7.64)	(13.98)	(10.50)	(12.06)	(12.53)	(1.40)
N	19240	17630	11393	23222	21216	3825
R^2	0.046	0.066	0.072	0.056	0.072	0.044

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are quartic experience and dummies for race, gender, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, advanced degree, and year

Table 6.8: Effect of Soft Skills on an Occupation's Wage Variability

	(1) Low Skill	(2) Medium Skill	(3) High Skill	(4) Low Wage	(5) Medium Wage	(6) High Wage
Skills Ratio	0.0662*	0.122	0.297*	-0.121	0.281***	0.267*
	(2.20)	(0.71)	(2.31)	(-1.53)	(3.88)	(2.27)
N	671	626	617	639	639	636
R^2	0.133	0.152	0.115	0.524	0.171	0.118

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Other variables included in the regression but not in the table are quartic experience, race, gender, union membership, a metropolitan indicator, regional indicators, marital status, children in household, immigration status, job sector, and educational attainment, and dummies for year

Chapter 7

Conclusion

There is some evidence that employers do screen for soft skills using education. I find that an occupation's soft skill ratio is correlated with overqualification and higher wage premiums. The effects are even larger for individuals who have five years or less of potential job experience, for whom employers are most likely to rely upon signals from education rather than work history when making hiring decisions. I do not find that wage variability can be explained by an occupation's soft skill ratio. However, this does not necessarily disprove the hypothesis of screening for soft skills on the basis of education because the mechanism by which screening would lead to more wage variability is the most difficult to explain.

Chapter 8

Works Cited

Acemoglu, Daron and David Autor. “Skills: Tasks and Technologies: Implications for Employment and Earnings.” Handbook of Labor Economics 4b (2011): 1043:1166. Web.

Autor, David H. and David Dorn. “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market.” American Economic Review 103.5 (2013): 1553-1597. Web.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. “Trends in U.S. Wage Inequality: Revising the Revisionists.” Review of Economics and Statistics 90.2 (2008): 300-323. Web.

Autor, David H., Frank Levy, and Richard J. Murnane. “The Skill Content of Recent Technological Change: An Empirical Exploration.” The Quarterly Journal of Economics 118.4 (2003): 1279-1333. Web.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. "Trends in U.S. Wage Inequality: Revising the Revisionists." Review of Economics and Statistics 90.2 (2008): 300-23. Web.

Bacolod, Marigee and Bernando Blum. "Two Sides of the Same Coin: U.S. 'Residual' Inequality and the Gender Gap." The Journal of Human Resources 45.1 (2010): 197-242. Web.

Balcar, Jiri. "Soft Skills and their Wage Returns: An Overview of the Empirical Literature." Review of Economic Perspectives 14.1 (2014): 3-15. Web.

Becker, Gary S. "Investment in Human Capital: A Theoretical Analysis." The Journal of Political Economy 70.5 (1962): 9-49. Web.

Belman, Dale and John S. Heywood. "Sheepskin Effects by Cohort: Implications of Job Matching in a Signaling Model." Oxford Economic Papers 49.4 (1994): 623-637. Web.

Black, Sandra E. and Alexandra Spitz-Oener. "Explaining Women's Success: Technological Change and the Skill Content of Women's Work." National Bureau of Economic Research (2007): n. page. Web.

Borghans, Lex, Bas Ter Weel, and Bruce Weinberg. "People People: Social Capital and the Labor-Market Outcomes of Underrepresented Groups." National Bureau of Economic Research (2006): n. pag. Web.

Bowles, Samuel, Herbert Gintis, and Melissa Osbourne. "The Determinants of Earnings: A Behavioral Approach." Journal of Economic Literature 39.4 (2001): 1137-1176. Web.

Brynin, Malcolm, and Simonetta Longhi. "Overqualification: Major or minor mismatch?" Economics of Education Review 28 (2009): 114-21. Web.

Buchel, Felix. "Overqualification: reasons, measurement issues, and typological affinity to unemployment." In *Training in Europe: Second report on vocational research in Europe 2000*, edited by Pascaline Descy and Manfred Tessaring, 453-560. Luxembourg: Office for Official Publications of the European Communities, 2001.

Card, David and John E. DiNardo. "Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." National Bureau of Economic Research (2002): n. pag. Web.

Deming, David J. "The Growing Importance of Social Skills in the Labor Market." National Bureau of Economic Research (2015): n. pag. Web.

Heckman, James J. and Tim D. Kautz. "Hard Evidence on Soft Skills." National Bureau of Economic Research (2012): n. pag. Web.

Hirsch, Barry T. "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." Industrial and Labor Relations Review 58.4 (2005): 525-551. Web.

Hungerford, Thomas and Gary Solon. "Sheepskin Effects in the Returns to Education." The Review of Economics and Statistics 69.1 (1987): 175-177. Web.

Jaeger, David A. and Marianne E. Page. "New Evidence on Sheepskin Effects in the Returns to Education." The Review of Economics and Statistics 78.4 (1996): 733-740. Web.

Krueger, Alan B. and David Schkade. "Sorting in the Labor Market: Do Gregarious Workers Flock to Interactive Jobs?" National Bureau of Economic Research (2007): n pag. Web.

Kuhn, Peter and Catherine Weinberger. "Leadership Skills and Wages." Journal of Labor Economics 23.3 (2005): 395-436. Web.

Lemieux, Thomas. "The "Mincer Equation" Thirty Years After *Schooling, Experience, and Earnings*." Center for Labor Economics, Working Paper No. 62 (2003): n. pag. Web.

Mariani, Matthew. "Replace with a database: O*Net replaces the Dictionary of Occupational Titles." Occupational Outlook Quarterly (1999): 1-9. Web.

Mueller, Gerrit and Erik J.S. Plug. "Estimating the Effect of Personality on Male and Female Earnings." Industrial & Labor Relations Review 60.1 (2006): 1:22. Web.

Murnane, Richard J., John H. Tyler, and John B. Willett. "Estimating the Labor Market Signaling Value of the GED." The Quarterly Journal of Economics 115.2 (2000): 431-468. Web.

Urzua, Sergio. "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices." The Journal of Human Resources 43.4 (2008): 919-971. Web.

Weinberger, Catherine J. "The Increasing Complementarity between Cognitive and Social Skills." The Review of Economics and Statistics 96.5 (2014): 849-861. Web.

Wheeler, Christopher H. "Evidence on Wage Inequality, Worker Education and Technology." Federal Reserve Bank of St. Louis Review 87.3 (2005): 375-93. Web.