Firm Heterogeneity in Skill Demands

David Deming
Harvard University and NBER

Lisa B. Kahn
Yale University and NBER

June 3, 2016

Abstract

We study variation in skill demands across firms and labor markets using job vacancy postings. We categorize a wide range of open text fields into ten general skills. There is substantial heterogeneity in demand for these skills, which provide explanatory power in pay and firm performance regressions, even within a detailed set of controls. We provide an illustration of the importance of cognitive and social skill demands in explaining wage heterogeneity across labor markets and firms. Finally, job skills account for a substantial fraction of firm variation in these measures of pay and performance, even after conditioning on controls.

*Correspondence: David Deming, Harvard Graduate School of Education, 6 Appian Way, Gutman 411, Cambridge, MA 02138. Email: david_deming@gse.harvard.edu. Lisa B. Kahn, Yale School of Management, 165 Whitney Ave, PO Box 208200, New Haven, CT 06511. Email: lisa.kahn@yale.edu. We are grateful to seminar participants at NBER and the Trans Pacific Labor Seminar 2016. We are especially indebted to Brad Hershbein, Dan Restuccia, Jake Sherman, and Bledi Taska for help and provision of the Burning Glass data.

1 Introduction

A large literature documents increased wage dispersion in the U.S. over the last few decades. Rising wage inequality has been attributed to a variety of factors, including skill-biased technological change, declining unionization and the shrinking real value of the minimum wage (e.g. Katz and Murphy 1992, Dinardo, Fortin and Lemieux 1996, Katz and Autor 1999, Lee 1999, Card and DiNardo 2002, Autor, Katz and Kearney 2008). A number of authors have also highlighted the important contribution to wage inequality of geographic sorting and agglomeration economies in cities (e.g. Glaeser and Gottlieb 2009, Moretti 2013, Diamond 2016). More recently, several papers in this volume have further highlighted the important role of high-paying firms in widening wage inequality (Barth, Bryson, Davis and Freeman 2016, Card Heining and Kline 2016, Song et al 2015).

An important body of work studies the impact of technological change on the relative demand for different types of occupations and job tasks. Autor, Levy and Murnane (2003) argue that growth in computing power has lowered the relative price of routine tasks and led to machines substituting for workers in occupations that are relatively routine task-intensive. Subsequent work in the U.S. and in a number of other developed countries has shown that routine, middle-skill occupations have experienced relative employment and wage declines over the last two decades, a pattern referred to as "polarization" (e.g. Autor, Katz and Kearney 2008, Acemoglu and Autor 2011, Goos, Manning and Salomons 2014). Largely because of data limitations, most work in labor economics focuses on relative shifts in employment across occupations rather than changes in skill demands *within* occupations - both across space and over time. This is an important limitation, because wage inequality varies widely across labor markets, and has increased sharply within occupations as well as across them (e.g. Firpo, Fortin and Lemieux 2009). Perhaps computer programmers, for example, perform a different set of tasks in some firms and in some labor markets than in others. Can variation in these tasks account for variation in pay across labor markets and firms, even within narrowly defined occupations?

In this paper we study heterogeneity in employer skill demands using a dataset of job vacancies that encompasses the near-universe of jobs posted to an online source across all major U.S. labor markets. We first categorize a set of ten job skills based on the prevalence of key words and phrases extracted from the job vacancies. We next show that the propensity to request these job skills varies widely across other observable characteristics of the ads. In addition, the ten job skills provide important explanatory power in external measures of job pay and firm performance. Finally, we show that they account for substantial fractions of firm-level variation in these pay and performance measures. The job vacancy data therefore provide useful explanatory power, above and beyond information that is currently available in other labor market data, including matched employer-employee records.

Our data come from Burning Glass Technologies, a firm that collects and codes data from the near-universe of online job vacancies from 2010-2015. The data contain thousands of key words and phrases, coded and regularized from the text of the job vacancy. We cull the most prevalent of these into ten categories of job skills that could be useful in a wide range of jobs. To understand the information content of these job skills, we link our data to two external sources yielding proxies for productivity of the job ad. First, we obtain msa-occupation average wages from the Occupational Employment Statistics program. Second, we proxy for whether the firm posting the add is publicly traded with whether we can link the name to a firm in Compustat. We also use revenue per worker as an outcome for publicly traded firms.

In principle, job vacancy data can be used to study heterogeneity in skill demands within a wide variety of occupations, industries and labor markets. Here we illustrate one particular application - the correlation between relative wages and measures of cognitive skill and social skill requirements of job vacancies. Cognitive occupations have been favored by machine technologies that replace routine workers and a rise in the return to cognitive ability was likely an important contributor of inequality in the 1990s (e.g., Acemoglu and Autor 2011). However, the nature of this return changed in the 2000s where, as a whole, cognitive occupations were less rewarded, while occupations that required both cognitive and social tasks grew (Beaudry, Green, and Sand (2016), Deming 2016).[1] These skills thus feature prominently in the evolution of technological change and inequality. With our data, we ask two questions regarding the return to cognitive and social skill requirements. First, does the return to cognitive and social skills hold *within* occupations? Second, given the importance of firms in explaining the rise in U.S. wage inequality, can differential demand for cognitive and social skills account for any variation in firm pay and performance?

We find a positive correlation between both types of skill requirements and MSA-occupation average wages after controlling for education and experience requirements, MSA fixed effects, and industry and occupation (6 digit) fixed effects. We also find a positive correlation between demand for both skills and firm performance, even after controlling for other skill requirements and the locations and occupations firms tend to post in. This suggests that variation in pay and performance across MSAs and firms may partly reflect differences in skill demands. We then estimate that demand for cognitive and social skills accounts for 5 percent of the residual variation in firm pay and similar fractions of residual variation in firm performance. Our results are thus consistent with wage inequality across firms being driven in part by differences in skill demands, perhaps as some firms take better advantage of modern production technologies.

Research on the task content of occupations has furthered our understanding of broad labor market trends in inequality (e.g., Auter, Levy, and Murnane 2003, Autor and Dorn

[1]Beaudry, Green, and Sand document a "Great Reversal" in demand for cognitive skills in the 2000s, showing that employment and wages stagnated in cognitive occupations. Also, Castex and Dechter (2014) find evidence of a secular decline in the wage return to cognitive ability. Deming (2016) finds evidence of increasing relative demand for occupations that have both high cognitive skill and social skill requirements.

2013), as well as human capital and worker mobility (Gathmann and Schonberg (2010), Poletaev and Robinson (2008), Speer (2016)). Our results show the usefulness of job vacancy data in explaining heterogeneity in wages across firms and labor markets, even *within* narrowly defined occupations. We also provide a specific illustration of how our data can speak to the recent debate in the inequality literature based on the changing demand for cognitive and social skills, and the increasing complementarity between them (Beaudry, Green and Sand (2016), Deming (2016), Weinberger (2014)). We thus join a growing set of papers that use job vacancy data to understand a variety of labor market issues, for example how and when firms respond to technological change (Kahn and Hershbein 2016), firm preferences for discrimination (Kuhn and Shen 2013) and the general equilibrium impacts of unemployment insurance (Marinescu 2015).

That performance persistently differs across seemingly similar firms is a well-known puzzle in the literature (Gibbons and Henderson 2012). Crucially, firm identifiers in our dataset allow us to relate internal firm practices (i.e., stated preferences) to performance, yielding additional explanatory power. In that spirit we are related to the important literature measuring and relating management practices to firm performance (Bloom and Van Reenen (2007), Ichniowski, Shaw and Prennushi (1997)).

Overall, our results show the usefulness of job vacancy data in explaining heterogeneity in wages across firms and labor markets. We find consistent evidence that job skills collected from vacancy data add explanatory power beyond information - such as occupation and industry - that is available in more traditional data sources. We also provide a specific illustration of the importance of cognitive skill and social skil demands in explaining wage heterogeneity across labor markets and firms.

This paper proceeds as follows. In section 2 we decribe the datasets and explain how we classify ten job skills from the thousands of open text fields available in our data. In section 3 we show that these skills have explanatory power in labor market pay and firm performance regressions, even within a detailed set of job controls. In section 4 we decompose variation in lfirm abor market pay and performance into components attributable to subsets of skills measures, as well as controls. Section 5 concludes with a brief discussion of how job vacancy data could be used in the furture.

2 Data

2.1 Overview

Our primary data source is a database of employment vacancies provided by Burning Glass Technologies (hereafter BG), an employment analytics and labor market information firm. BG examines nearly 40,000 online job boards and company web sites, parses and deduplicates them into a systematic, machine-readable form. They use the resulting database to create labor market analytic products. BG claims their database covers the near-universe of job vacancies posted online during the time period of measurement. The BG microdata were first used by Hershbein and Kahn (2016) to study whether recessions accelerate routine-biased technological change.

The BG data include education and experience requirements, detailed industry and occupation codes (6-digit SOC), the location and firm identifiers (where available) for each job vacancy. BG also parses the actual text of each job vacancy and codes key words and phrases as additional job requriements. In our sample, roughly 95 percent of ads have at least one such requirement, and, conditional on having any requirement, post an average of 7. In the next subsection, we describe how we distill over ten thousand unique text fields into a subset of general "job skills" that could be useful across a wide range of jobs.

We make two main sample restrictions. First, we restrict the sample to "professional" occupations.[2] This category includes almost all occupations employing college-educated workers, and it has the most complete coverage in the BG data (e.g. Carnevale, Jayasundera and Repnikov 2014, Beaudry, Green and Sand 2016).

Second we restrict the sample to the approximately 60 percent of ads with non-missing firms, and we further focus our attention on firms that have posted at least ten ads total and ads in at least two different professional occupations and at least two different Metropolitan Statistical Areas (MSAs) over our sample period.[3] These restrictions help clean out some noise in the firm-level variation in skill requirements that we explore.

Our resulting dataset contains nearly 23 million ads, annually, for the years 2010-2015, across 92,000 firms. Only 16 percent of these ads post the wage that is offered, which prevents us from studying the relationship between skill demands and wages for individual job vacancies.[4] Instead, we use two sources of external data that allow us to relate average skill demands to wages across labor markets and performance across firms.

We obtain data on average wages by occupation across MSAs from the Occupational Employment Statistics (OES) program produced by the Bureau of Labor Statistics (BLS). The OES is a large survey of non-farm establishments, especially designed to produce data at sub-state levels. We compute mean wages by MSA and occupation by taking the unweighted average across the 2010-2015 releases. Four percent of job vacancies are posted in MSA-

[2] Specifically, we restrict to major SOC categories 11-29, which include management, business and financial operations, computer and mathematical, architecture and engineering, the sciences, community and social services, legal, education, arts and entertainment, and healthcare practitioners and technical occupations.

[3] Ads that do not contain a firm identifier are typically obtained from recruiter websites where the poster does not wish to reveal the information. Our results that do not require firm identifiers are qualitatively similar when performed on the full sample. These are available upon request. The restriction on firm coverage excludes 6 percent of ads with a non-missing firm and at least ten posts (firms with a very small number of posts likely represent data errors, containing fragments of text from the ad that do not correspond to firm name). We also exclude Micropolitan Statistical Areas from our analysis (5.6 percent of ads) since external data on wages and location characteristics is either unavailable or much less precise.

[4] It is well-known that vacancies rarely post wage offers. See for example Kuhn and Shen (2012) and Marinescu and Wolthoff (2016).

occupations that cannot be matched to OES data, likely because they are too small, and these cells are removed from this analysis.

We also obtain MSA demographic characteristics to use as control variables from American Community Survey (ACS) data, averaging 2010-2014 annual data.[5] Here, again, we can match all but a small fraction of MSAs (making up 5 percent of professional ads) to the ACS, excluding some small cities.[6]

Finally, we link firms in the BG sample to Compustat North America by Standard & Poors.[7] All publicly traded companies are required to track accounting and balance sheet data, making Compustat the most complete database of this information for U.S. firms. We are able to match about 30 percent of job vacancies to a publicly traded firm in Compustat.[8] We use the link to Compustat to measure firm performance. Our main performance measures are whether or not the firm is publicly traded, as measured by whether or not the firm can be linked to Compustat, and conditional on the link being possible, the firm's revenue per worker, taking an unweighted average across years 2010-2015.

Appendix table A1 provides summary statistics for three main sample, each weighted by the number of ads in a given cell.[9] Panel A summarizes data at the msa-occupation (six digit SOC) level for cells that can be matched to OES data. The average cell contains 412 ads (unweighted) and there is a wide range. The average mean wage is $42/hour among professional occupations.

Panel B provides provides summary statistics for the firm-level sample. The more than 90,000 firms post an average of 250 ads (unweighted) each, again with quite a range.[10] Given our restrictions and weighting by ads, the average cell is attributed to a firm that posts in 100 6-digit professional occupations (out of 352) and 100 MSAs (out of 371). Panel C restricts the firm-level sample to those that can be matched to Compustat. Firms in this sample are bigger, posting in an average of 1762 ads and across a larger number of MSAs and occupations. Revenue per worker averages roughly half a million dollars using our weighting.

[5] The 2015 data had yet to be released at the time we conducted our analysis.

[6] Demographic controls from the ACS include MSA-level share female, black, Hispanic, asian, married, and moved in the last year. We also control for education (high school dropouts, exactly high school, some college, exactly BA) and age (less than 18, 19-29, 30-39, 40-49, 50-64) distributions. We set all controls to zero if the MSA did not match and include a dummy for whether the MSA matched to ACS data.

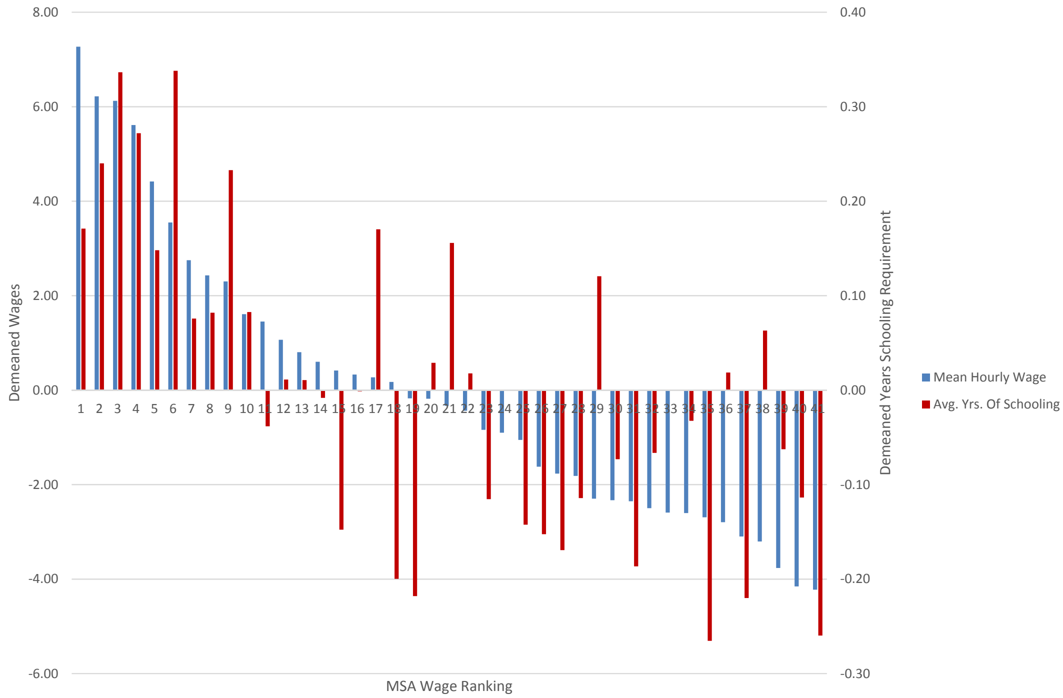
[7] We obtain these data via Wharton Research Data Services.

[8] After cleaning firm names in both BG and Compustat to remove words like "Incorporated" and its associated abbreviations, as well as all punctuation, we match based on exact name (80 percent of matched firms). We then use a fuzzy match alogrithm to link firms with at least a 95 percent chance of being the same (5 percent of matched firms). Finally we match based on a regularized subset of words. Hershbein and Kahn (2016) also use this match. They point out that total employment in Compustat was 50 percent of US employment in these sample years, though employment is not collected in a standardized way in Compustat and includes foreign affiliates.

[9] We take an unweighted average across years 2010-2015 to obtain these cells. Thus the number of ads in a given cell is an unweighted annual average.

[10] Because we weight cells by number of ads, the means vary from Panel A only because Panel B includes a small number of firms that post in MSA-occupations that cannot be matched to OES data.

Figure 1: Wages and Education Requirements by City Wage Rank



Notes: Blue(red) bars show MSA average wages (education requirements) ordered by MSA wage rank for the top 41 employment MSAs, with demeaned values plotted on the left (right) axis. Wage (education) data are OES (Burning Glass) 2010-2015 averages.

2.2 Job Skills in the Burning Glass Data

The primary contribution of this paper is to distill and analyze the key words and phrases coded from the open text of ads in the BG data. But BG also codes the more standard skill measures, education and experience. From table A1, 65 percent of professional ads specify an education requirement that averages 15.7 years of school (fitting modal years to degree requirements). 62 percent of ads specify a requirement for experience in the field that averages 4 years.

Figure 1 shows demeaned hourly wages (blue bars) from the OES ordered by their city wage rank for the 41 largest MSAs (cities employing at least 250,000 workers in professional occupations according to the OES). Wage bars are followed by the corresponding demeaned years of schooling required for job vacancies posted in that same MSA. As Figure 1 shows clearly, average wages (from the OES data) and average years of schooling required (from the BG data) are strongly correlated (about 0.6 when weighted by total employment). This correlation is reassuring that the BG data are broadly representative and generally preserve the ranking of labor markets by skill.

We next explain how we make sense of the key words and phrases coded by BG. Using the more than ten thousand unique fields as our starting point, we create ten categories of job skills that could be useful across a wide range of jobs. Table 1 lists the ten skills and provides the corresponding words and phrases that fall in each category. We code an ad as having a particular job skill requirement if it has at least one of the key words or phrases listed,

Table 1: Description of Job Skills

Job Skills	Key words and phrases
Cognitive	Problem Solving, Research, Analytical, Critical Thinking, Math, Statistics
Social	Communication, Teamwork, Collaboration, Negotiation, Presentation
Character	Organized, Detail-oriented, Multi-tasking, Time Management, Meeting Deadlines, Energetic
Writing	Writing
Customer Service	Customer, Sales, Client, Patient
Project Management	Project Management
People Management	Supervisory, Leadership, Management (not project), Mentoring, Staff
Financial	Budgeting, Accounting, Finance, Cost
Computer (general)	Computer, Spreadsheets, Common Software (e.g. Microsoft Excel, Powerpoint)
Software (specific)	Programming language or specialized software (e.g. Java, SQL, Python, etc.)

Notes: Authors categorization of open text fields in Burning Glass data.

though it may have many. The skills are mutually exclusive but not collectively exhaustive - indeed, there are many other categories of job skills that one could study.[11]

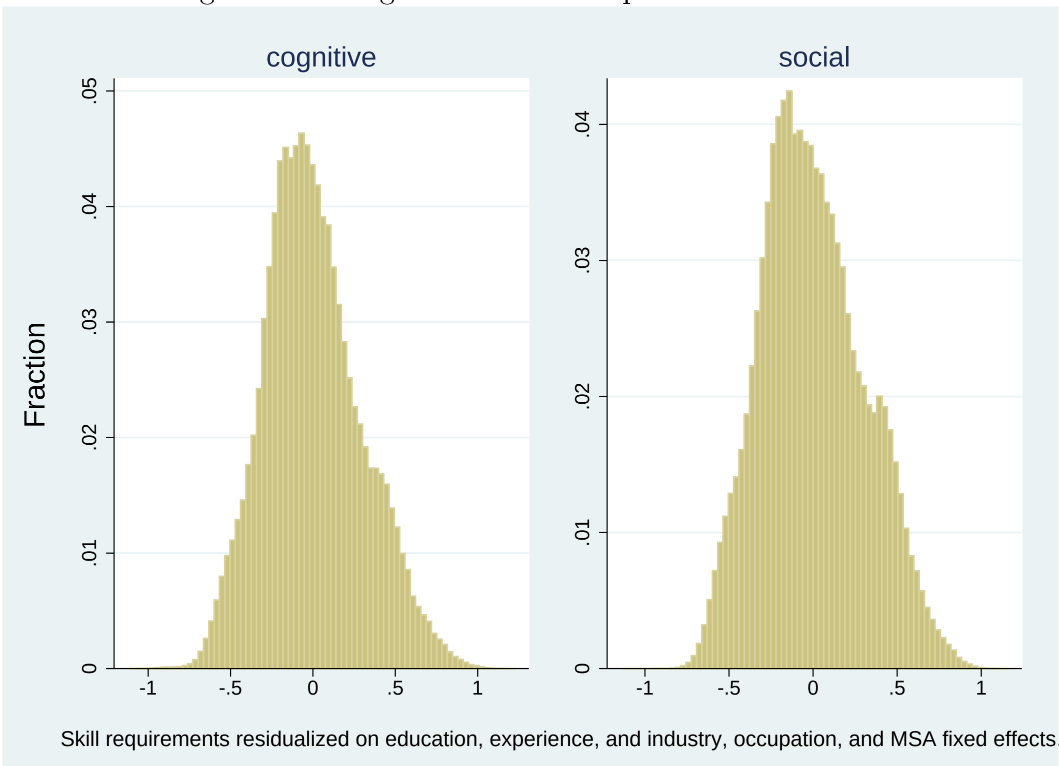
The first two skills listed in Table 1 are “cognitive” and “social”. Job vacancies that require cognitive skills ask for key words and phrases such as “problem solving”, “research”, and “analytical”. We chose these skills deliberately to match the description of the “nonroutine analytical” job tasks used in Autor, Levy and Murnane (2003) and other related work. We group key words such as communication, teamwork, and collaboration under the heading of “social skills”, following closely the definition used in Deming (2016). Beaudry, Green, and Sand (2014, 2016) show a decline in demand for cognitive occupations in recent years, while Deming (2016) shows relative growth in employment and wages of occupations that require social skills and especially *both* cognitive and social skills, as measured by O*NET. We are thus particularly interested in understanding whether there is variation *within* occupations and other typical controls in demand for these skills, and whether they can account for differences in pay across firms.

We categorize eight additional skill groups and show that these are important in explaining pay differentials across labor markets and performance differentials across firms. The third skill, “character”, is an umbrella term for key words and phrases such as “organized”, “detail-oriented”, and “time management”. Here we follow the large literature on non-cognitive or “soft” skills, which discusses the labor market returns to personality traits such as conscientiousness and agreeableness as well as personal attributes such as self-control and positive affect (e.g. Heckman and Kautz 2012).

Our criteria for the other seven job skills is that they be commonly listed and generally applicable to a wide range of jobs. For example, writing, customer service, sales and project management are among the top fifteen most commonly listed text fields in BG data from 2014. We also create the skill categories “people management” and “financial” from a range of related key words and phrases. Finally, we include categories for general and specific computer skills. The former encompasses the generic phrase “computer skills” as well as key

[11]Examples range from skills that are particularly useful in certain jobs (e.g. plumbing) to business processes (e.g. six sigma) to very general attributes that are hard to categorize (e.g. quick learner).

Figure 2: Histograms of Skill Requirement Residuals



phrases for common software such as Microsoft Excel, while the latter includes specialized software such as SQL, Java, or C++.

Table A1 provides summary statistics for these measures. We find 42 percent of professional jobs specify a social skill requirement and the same fraction specify a cognitive requirement. A quarter of ads specify both requirements. The other skill measures appear frequently ranging from being required in an eighth of ads (project management) to a third (character).

There is also substantial variation in skill requirements across ads. From table A1, the standard deviations for most variables range from a tenth to a quarter. Much of this variation holds after residualizing on a rich set of controls. Figure 2 plots histograms for the two main skills variables, cognitive and social, after residualizing on education and experience requirements, and occupation, industry and MSA fixed effects.[12] Even within these controls, there is still substantial variation in skill requirements. Historgrams for the remaining eight skills are shown in appendix figure A1. Most of these retain substantial variation, though some, like finance, are fairly uniform once we control for occupation, industry, and location.

To provide more intuition for these job skills, appendix table A2 shows the share of job vacancies that require each skill for selected occupations. The first thing to note is that job skills line up well with occupation titles and perceived job tasks. 81 percent of

[12]Education controls include both the share of ads with any education requirement and the average number of years required if there is a requirement (otherwise zero), and experience controls are analogous. Occupation is at the six-digit SOC level and industry is at the two-digit NAICS level.

vacancies for accountants and auditors require financial skills, compared to an average of only 16 percent across all professional occupations. Similarly, the share of sales manager occupations requiring customer service is particularly high, as is the share of computer programmer occupations requiring skill in a specific software program.

The second important point from table A2 is that while cognitive, social and character skills are required across a broad range of occupations, there is a lot of variation in skill requirements even within detailed occupation categories. For example, about one third of manager vacancies have some sort of financial skill requirement. 39 percent of computer programmer vacancies require teamwork and/or collaboration. Finally, job skills appear to be listed more frequently for occupations that are less well-defined. Registered nurse and teacher vacancies almost universally require professional certification, and are probably more standardized than an occupation like management analyst in terms of job function.

3 Skill Demands and Worker and Firm Outcomes

The goal of this paper is to understand whether pay differentials across markets and firms can be attributed to differences in skill demands. In particular, are firms and markets that pay more also more likely to require employees to have cognitive skills, social skills or both? This is interesting because differences in skill demand and associated returns to skills may reflect variation in production technology, especially within narrowly defined occupations. Since we do not measure productivity or wages directly in our dataset, we use three proxies. The first is wages in the MSA-occupation cell, obtained from OES data, which should be correlated with wages paid by firms posting ads in those same cells. The second is whether or not the firm is publicly traded, measured by whether we can link the firm name to Compustat. Publicly traded firms are generally larger, higher-paying, more successful firms and we can ask whether this firm-performance measure is correlated with skill demand. The third is log revenue per worker, conditional on being matched to Compustat. For this subset of firms, all of whom are at least successful enough to be publicly traded, we can ask whether differences in skill demand correlate with differences in their bottom line.

Table 2 shows simple bivariate correlations between each skill measure and our three outcome variables.[13] The first thing to note is that - echoing Figure 1 - wages are strongly correlated with education and experience requirements (0.570 and 0.621 respectively).[14] Second, most of the job skills are positively correlated with wages. Cognitive skills and social skills are more highly correlated with wages (0.30 and 0.298, respectively) than most of the other skills. Though, project management and people management skills are the most

[13]The unit of observation for wages is an MSA-occupation, while the latter two columns are measured at the firm level. All the correlations are weighted by the total number of job vacancies in each cell. The results are all very similar when we weight the data by total employment in each MSA-occupation cell.

[14]When there is no education or experience requirement, the years variable is set to 0. Regressions include an indicator for no requirement.

Table 2: Correlations between Skills and Job Outcomes

	Correlation with :		
	Log Hourly Wages	Publicly Traded	Log Revenue per Worker
Years of Schooling Required	0.570	-0.009	0.212
Years of Experience Required	0.621	0.222	0.283
Cognitive	0.305	0.249	0.339
Social	0.298	0.271	0.237
Cognitive and Social	0.358	0.274	0.241
Character	0.121	0.244	0.138
Customer Service	-0.071	0.092	-0.137
Writing	0.160	0.146	0.184
Project Management	0.352	0.219	0.253
People Management	0.371	0.146	-0.056
Financial	0.251	0.162	0.169
Computers (General)	0.012	0.221	0.193
Specific Software	0.168	0.122	0.151

Notes: Table shows the bivariate correlations between the skill variables and 3 outcome measures: MSA-occupation wages, whether the firm is publicly traded (an indicator equal to one if the firm can be matched to Compustat), and log revenue per worker of the firm. Wages are estimate at the MSA-occupation level while the other outcomes are firm level, each weighted by number of ads posted. Years schooling and experience equal zero if there is no requirement. See table 1 for skills definitions.

highly correlated with wages (0.352 and 0.371 respectively). Third, the patterns are similar for the firm outcomes. While firm-level demand for most job skills is positively correlated with whether or not the firm is publicly traded and log revenue per worker, the correlations are among the largest for cognitive and social skills, ranging from 0.22 to 0.34.

In thinking about the relationship between job skills and wages, a few examples may be instructive. 1) 40 percent of computer programmer vacancies in Washington, DC require social skills, compared to only 26 percent in Manchester-Nashua, NH, and hourly wages are about 25 percent higher. At the same time, though firms posting for computer programmers in Washington, DC have a similar probability of being publicly traded to those in Manchester, the former have nearly 10 percent higher revenue per worker. 2) 85 percent of management analyst vacancies in San Jose, CA require cognitive skills, compared to only 71 percent in San Diego, CA, and hourly wages are about 23 percent higher in the former compared to the latter. Firms posting for management analysts in San Jose are also about 2 percentage points more likely to be publicly traded, compared to those in San Diego.

Our hypothesis is that, while programmers typically require cognitive skills, the programmers in DC that also require social skills are performing more complex functions, perhaps strategizing with clients or overseeing coworkers, and are therefore more productive. Similarly, while all management analysts likely require social skills to interact with clients, those in San Jose also perform more complex analyses than perhaps a more plug-and-chug analyst in San Diego, and this makes them more productive. Naturally, we cannot observe production technology in our data. But we find the subsequent evidence on stated preferences of firms compelling and consistent with this hypothesis.

Of course, wages and job skill requirements may be correlated for a variety of reasons. Washington, DC and San Jose are more expensive cities with more educated workers and different amenities than their counterparts. We want to understand whether the skill requirements reflect differences in how workers produce or simply pick up city- or occupation-wide differences. To that end, we estimate regressions of the following general form:

$$log(wage)_{om} = \alpha + \overline{Skill_{om}}\beta' + X_{om}\gamma' + \delta_o + \epsilon_{om} \tag{1}$$

For a 6-digit SOC occupation, o, and MSA, m, we regress the mean log wage ($log(wage)$) as measured by the OES on a vector of average skill requirements ($\overline{Skill_{om}}$) of ads posted in the MSA-occupation in the BG data. We also add controls for the average years of education and experience required and the share of ads that have any education and experience requirements, industry composition within MSA-occupation cells as measured by two-digit (NAICS) codes, occupation fixed effects (δ_o) and MSA characteristics from the ACS or fixed effects. All regressions are weighted by the number of ads in each MSA-occupation cell, although results are very similar when we instead weight by employment.

Table 3 presents results from estimates of equation (1). Column 1 summarizes a sparse model that includes the 10 skill measures and the education and experience variables. The coefficient of 0.136 on cognitive, significant at the 1 percent level, implies that a 10 percentage point increase in the share of job vacancies requiring cognitive skills increases wages by about 1.4 percent. Alternatively, since the ad-weighted standard deviation of the cognitive skill measure is 0.17 in this sample, a one standard deviation increase in cognitive skill requirements increases wages by 2.3 percent. The coefficient on social skills is much larger in magnitude and also significant at the 1 percent level. A standard deviation increase (0.13) in demand for social skill is associated with a 4 percent wage increase.

Column 2 adds the share of vacancies in each MSA-occupation cell with *both* a cognitive skill and social skill requirement. Weinberger (2014) and Deming (2016) shows that cognitive skills and social skills are complements in an earnings regression. Similarly, we find strong evidence of complementarity between cognitive skills and social skills. Each of the positive correlations between cognitive or social skill requirements and wages is explained by vacancies that ask for *both* types of skills. The coefficient implies that a one standard deviation increase (0.11) in the share of vacancies requiring cognitive and social skills increases wages by 10 percent.

We also note F-statistics on the cognitive and social variables (including the requirement for both in even columns) and for the full set of skills. In both columns we can strongly reject that these groups of coefficients are equal to zero ($p = 0.000$).

Columns 3 and 4 of Table 3 add our main set of controls: four-digit SOC occupation fixed effects, MSA characteristics and the distribution of ads across two-digit NAICS industries.

Table 3: Average Wages and Skill Requirements

	Dependent variable: Log(mean wages) in MSA-Occ					
	(1)	(2)	(3)	(4)	(5)	(6)
Cognitive	0.136***	-0.257***	0.173***	0.165***	0.0538***	0.0341***
	(0.00902)	(0.0158)	(0.00761)	(0.0127)	(0.00720)	(0.0102)
Social	0.307***	-0.0429**	0.149***	0.142***	0.0360***	0.0188*
	(0.0130)	(0.0173)	(0.0101)	(0.0136)	(0.00761)	(0.00989)
Both required		0.882***		0.0178		0.0425***
		(0.0291)		(0.0223)		(0.0156)
Years of Education	0.124***	0.122***	0.0556***	0.0557***	0.00338***	0.00341***
	(0.000816)	(0.000811)	(0.000853)	(0.000854)	(0.000826)	(0.000826)
Years Experience	0.163***	0.165***	0.0615***	0.0615***	0.0194***	0.0195***
	(0.00117)	(0.00116)	(0.00122)	(0.00122)	(0.000979)	(0.000979)
Base Controls			X	X		
Detailed Controls					X	X
F-stat (cog and social)	472.6	626.0	440.3	293.7	49.17	35.28
F-stat (all 10 skills)	1937	2055	393.2	357.5	28.59	26.67
MSA-Occ Cells	53,510	53,510	53,510	53,510	53,510	53,510
R-squared	0.687	0.693	0.851	0.851	0.941	0.941

*** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions control for the share of ads with each of the 8 other job skill, education and experience requirements. Years of education and experience equal 0 if the msa-occupation cell has no ads that specify requirements. Dependent variable is the log of median hourly earnings in the MSA-occupation, obtained from OES data. Observations are weighted by number of ads in the MSA-occupation cell. Base controls include four-digit SOC occupation fixed effects, MSA characteristics from ACS data, and the share of ads in the MSA-occupation that are in each of the two-digit NAICS industries. Detailed controls include industry shares and six-digit SOC occupation and MSA fixed effects. See table 1 for skills defintions.

The patterns from Columns 1 and 2 primarily hold, although the coefficients become smaller in magnitude. The return to cognitive and social skill requirements also load primarily on each term individually, rather than on the joint requirement. We can strongly reject that the coefficients on cognitive and social skills are jointly equal to zero, and likewise for all ten skills together ($p = 0.000$ in both cases).

Columns 5 and 6 fully saturate the model, controlling for six-digit occupation fixed effects and MSA fixed effects, in addition to industry shares. MSA fixed effects control for any pay differences across labor markets that are due to common factors such as cost of living or local amenities. Even in this highly controlled specification, we can strongly reject that the coefficients on cognitive and social skills - and all skills jointly - are equal to zero ($p = 0.000$ in both cases). This shows clearly that data on the average skill requirements of jobs in an MSA have explanatory power beyond what is available in conventional data sources. Interestingly, we find positive, significant coefficients on cognitive skills, social skills and both included together. It is also worth noting that the controls for the other job skills suggest that the returns to cognitive and social skills are not driven by their relative concentration in particular job types, such as managerial or finance jobs, that earn high returns.

Overall, the results in Table 3 suggest that a substantial share of occupation wage premia across labor markets can be explained by differences in employer skill demands, even within very narrowly defined occupation categories. In particular, we find that labor markets paying relatively high wages to a particular occupation are likely to have relatively higher cognitive and social skill requirements. In contrast, we do not find this same pattern for "character" skills, or for other specific skill requirements such as people management, finance or customer service.[15]

Next we ask whether larger and more productive firms post systematically different skill requirements for the same occupations. We estimate the correlation between firm performance and skill requirements in a regression framework:

$$firm\_perf_f = \alpha_0 + \overline{Skill_f}\beta' + \bar{I}_f^o + \bar{X}_f\gamma' + \bar{I}_f^m + \theta_n + \epsilon_f \tag{2}$$

$\overline{Skill_f}$ is a vector of firm average shares of job vacancies that require each skill, $\bar{I}_f^o$ is the share of each firm's postings belonging to each occupation code, $\bar{X}_f$ is a vector of average education and experience requirements, $\bar{I}_f^m$ is the share of each firm's postings belonging to each MSA, or the ad-weighted average MSA characteristics of the firm, and θ_n is industry fixed effects. We weight by the share of ads posted by the firm.

Table 4 presents estimates of equation (2) for both firm performance outcomes for our primary set of controls (four-digit SOC occupation distribution, average MSA characteristics and industry fixed effects) and for the full set of controls (distributions of six-digit occupations

[15] These results are available upon request.

and MSAs, and industry fixed effects). Columns 1-4 present results where the outcome is an indicator variable for whether the firm is publicly traded. In column 1, the coefficients imply that a one standard deviation increase in the share of vacancies with a skill requirement (about 0.2 for both skills) increases the probability that a company is publicly traded by 1.7 and 9 percentage points for cognitive skills and social skills, respectively. This is non-trivial, relative to the baseline 30 percent probability that an ad is posted to a publicly traded firm. Column 2 adds the share of vacancies that require both types of skills. We see that the bulk of the return to cognitive or social loads on both being required. A standard deviation increase in requiring both (0.16) increases the probability of being publicly traded by 5 percentage points. Note also that the social and cognitive coefficients are jointly significant, as is the full set of skill requirements. These skill measures thus add explanatory power above and beyond the typically available control variables. Furthermore, results are quite robust to the more detailed set of controls included in columns 3 and 4.

Table 4: Firm Outcomes and Average Skill Requirements

	Publicly Traded				Log(Revenue per Worker)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive	0.0852***	-0.0783***	0.0725***	-0.0838***	0.559***	0.896***	0.534***	-0.0838
	(0.0113)	(0.0166)	(0.0119)	(0.0173)	(0.122)	(0.197)	(0.139)	(0.224)
Social	0.115***	-0.0107	0.0994***	-0.0182	0.338***	0.621***	0.312**	-0.174
	(0.0105)	(0.0141)	(0.0106)	(0.0142)	(0.108)	(0.169)	(0.126)	(0.187)
Both required		0.323***		0.302***		-0.581**		1.083***
		(0.0242)		(0.0242)		(0.267)		(0.307)
Years of Education	-0.00413***	-0.00354***	-0.000451	-8.28e-05	0.0336	0.0326	0.0245	0.0278
	(0.00121)	(0.00121)	(0.00121)	(0.00121)	(0.0232)	(0.0232)	(0.0271)	(0.0270)
Years Experience	0.0245***	0.0244***	0.0147***	0.0149***	0.0911***	0.0891***	0.0551***	0.0554***
	(0.00139)	(0.00139)	(0.00146)	(0.00146)	(0.0149)	(0.0149)	(0.0182)	(0.0182)
Base Controls	X	X			X	X		
Detailed Controls			X	X			X	X
F-stat (cog and social)	110.3	133.3	76.27	102.9	18.89	14.19	11.53	11.87
F-stat (all 10 skills)	165.9	167.3	129.8	132.4	9.850	9.395	4.177	4.943
# Firms	92,349	92,349	92,349	92,349	3,690	3,690	3,690	3,690
R-squared	0.284	0.285	0.344	0.345	0.512	0.513	0.737	0.738

*** p<0.01, ** p<0.05, * p<0.1

Notes: All regressions control for the share of ads with each of the 8 other job skill, education and experience requirements. Years of education and experience equal 0 if the msa-occupation cell has no ads that specify requirements. In columns 1-4 the dependent variable is an indicator equal to one if the firm can be matched to Compustat; in columns 5-8 it is equal to the log of revenue per worker, conditional on being matched to Compustat. Observations are weighted by number of ads posted by the firm. Base controls include the distribution of ads across four-digit SOC occupations, MSA characteristics averaged across all ads to a given firm, and two-digit NAICS industry fixed effects. Detailed controls include industry fixed effects and six-digit occupation and MSA distributions across ads. See table 1 for skills defintions.

Columns 5-8 report results for log revenue per worker, among the 3,690 firms in our data that are publicly traded. Column 5 shows that firms with higher revenue per worker also have significantly higher cognitive skill and social skill requirements. The coefficients imply that a one standard deviation increase in the share of vacancies with each type of skill requirement (about 0.2) increases revenue per worker by about 11 percent and 7 percent for cognitive and social, respectively. As with log wages, in our base specification returns load primarily on the main effects of cognitive and social and not on the requirement for both (column 6). However with the full set of controls, the return is entirely concentrated in ads that have a joint requirement. Again, in all specifications, we can reject the hypotheses that skills are jointly equal to zero. Overall, the results in Table 4 show that skill requirements are strongly correlated with proxies for firm productivity.

In this section, we have shown that job skills have significant explanatory power for wages and firm performance, above and beyond variation explained by typical controls of detailed occupation, industry and location. Furthermore, we find some evidence of a complementarity between stated preferences for cognitive and social skills in that they are positively related to pay and firm performance, especially with our strictest set of controls. Our preferred explanation is that the stated preferences indicate that these firms produce differently, taking advantage of modern technologies (which may require cognitive skills, social skills, or both) and resulting in higher productivity. While we cannot rule out that these correlations are driven by unobserved aspects of labor markets and firms, we show they hold up to a rich set of controls.

4 Heterogeneity across Firms and Skill Demand

We have shown that jobs requiring more cognitive and social skills pay more and that firms requiring more of these skills perform better. Furthermore, the full set of job skills explains a significant amount of residual variation in worker pay and firm performance. Variation in firm pay, controlling for worker quality, accounts for an important fraction of the rise in income inequality over the last 30 years. How much heterogeneity is there in skill demands? Based on the link between skill demands and pay/performance, can changing skill demand account for much of the rise in firm pay inequality?

Table 5 gives a general sense of firm variability in outcomes and skill demands. We take firm averages of each variable and show the standard deviations, all weighted by the number of postings. The wage measure masks substantial heterogeneity since it simply maps MSA-occupation OES wages onto firms by taking a weighted average across markets a firm posts in. However, in row 1 column 1, we still find substantial variation across firms; a firm that pays a standard deviation higher wages pays about 21 percent more. In column 2 we first residualize on our baseline set of controls (average MSA characteristics, the distribution of ads across four-digit occupations and industry fixed effects). The residual standard deviation

Table 5: Standard Deviations of Firm Effects in Outcomes and Skills

	No Controls	Full Controls
Log Hourly Wages	0.191	0.056
Publicly Traded	0.458	0.397
Log Revenue per Worker	0.829	0.594
Cognitive	0.204	0.153
Social	0.203	0.173
Cognitive and Social	0.163	0.135
Character	0.190	0.156
Customer Service	0.180	0.134
Writing	0.155	0.133
Project Management	0.107	0.072
People Management	0.125	0.109
Financial	0.141	0.089
Computers (General)	0.188	0.152
Specific Software	0.246	0.101

Notes: Table shows standard deviation in firm-level averages of each variable, weighted by the number of postings to each firm. Column 2 first residualized each variable on the average distribution across six-digit occupations, the average MSA characteristics and as well as industry fixed effects. See table 1 for skills definitions.

in firm effects implies that moving up a standard deviation results in a 6 percent pay increase.

There is also substantial variation across firms in performance. Even after residualizing on our full set of controls, a one standard deviation higher firm is 40 percentage points more likely to be publicly traded and producing about 80 percent more revenue per worker (when restricting to the Compustat sample).

The table also shows standard deviations in firm average propensity to specify each job skill, and again we find substantial variation across firms. A firm one standard deviation more likely to specify a cognitive skill requirement has a 20 percentage point higher probability, with a similar number for social skills. These standard deviations fall by about a quarter when we first residualize on controls.

Given the substantial returns to posting these jobs skills documented in the previous section, and the substantial variation across firms in their propensity to post, we ask whether different demand for these job skills can account for differences in firm pay and performance outcomes. To that end, we present a simple decomposition.[16]

Average log wages can be expressed as follows for an occupation, o, MSA, m, and firm f. β^f are firm fixed effects and $controls$ are typical variables found in other datasets (our base controls).

$$log(wages)_{omf} = \beta^f + controls + \epsilon_{omf} \tag{3}$$

[16] This decomposition follows the spirit of Altonji, Kahn, and Speer (2014) who decompose variation in the returns to college major over time into changes in the returns to tasks preformed in jobs typically held by workers with a given major.

Firm fixed effects incorporate unobserved systematic differences in earnings across firms. We express these as a function of a vector of skill demands ($s\bar{k}ill^f$) and a residual (ν^f) in equation 4. δ is the coefficient on the residual and α is a vector of coefficients on these skills variables.

$$\beta^f = s\bar{k}ill^f \alpha' + \delta \nu^f \tag{4}$$

Plugging equation 4 into 3 yields

$$log(wages)_{omf} = s\bar{k}ill^f \alpha' + \delta \nu^f + controls + \epsilon_{omf} \tag{5}$$

This model provides a way to decompose the variance of log wages. In particular, the variance in the firm effects, $Var(\beta^f)$, is equal to $Var(\overline{skill_f}\alpha' + \delta\nu_f)$, and we can calculate the variance in β^f that is attributable to the skills variables, or to a subset of the skills variables, by setting the other coefficients (components of the vector, α) to 0 and calculating the variance.

We estimate equations 3-5 on a disaggregated msa-occupation-firm-level dataset. In practice, we do not have variation in wages across firms within occupation-MSA cells. Instead β^f represents the ad-weighted average wage across labor markets the firm posts in. Results are presented in the first three columns of table 6. Column 1 shows a simple decomposition which includes no controls and only the social, cognitive, and combined skill measures. Column 2 adds the remaining 8 job skills, as well as education and experience requirements. Column 3 adds our base controls (excluding industry fixed effects since those vary only at the firm level).

Table 6: Decomposing Firm Effects in Outcomes on Skill Demands

	Log (Wages)			Publicly Traded			Log Revenue per Worker		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Total Standard Deviation of Firm Effect	0.188	0.188	0.083	0.458	0.458	0.452	0.829	0.829	0.820
Share attributed to skills:									
Total	11.8%	32.6%	20.9%	8.4%	14.5%	10.1%	14.9%	20.1%	15.6%
Social and Cognitive Skills	11.8%	5.5%	4.5%	8.4%	2.5%	1.8%	14.9%	11.1%	7.4%
Other Skills		10.8%	7.3%		3.6%	3.5%		1.7%	2.4%
Education and Experience		16.3%	9.1%		8.4%	4.8%		7.4%	5.9%
Residual	88.2%	67.4%	79.1%	91.6%	85.5%	89.9%	85.1%	79.9%	84.4%
Additional Skills		X	X		X	X		X	X
Controls			X			X			X
# Firms		92,727			92,349			3,690	

Notes: When included, control variables are MSA characteristics and the four-digit occupation. Social and cognitive skills include requirements for each and the share of ads specifying both. Other skills include the 8 additional job skills listed in table 1, as well as the education and experience requirements. Decomposition is estimated on the msa-occupation-firm sample. For log wages, we obtain firm fixed effects from a regression (that includes controls if indicated) weighted by number of ads in the cell. We next regress these firm fixed effects on firm-level averages for the skills measures and obtain the residuals. We finally regress firm-level skills variables and the residuals on log wages (and controls if applicable). We use coefficients to fit the overall variance in the firm effect and the share attributed to various components (by fitting variances with the other coefficients set to 0). Each component also includes terms for correlations with the other skills (if included in the regression).

From column 1, the total standard deviation of the firm wage effect without controls is 0.188. We find that 12 percent of this variation can be accounted for by variation in the propensity to list cognitive and social skill requirements and the positive wage return of these skills. The remaining 88 percent goes unexplained. Column 2 shows that about half of the 12 percent is because the cognitive and social requirements are correlated with other skill measures. Education and experience requirements account for the largest fraction, 16 percent, while the other eight job skills account for 11 percent. In total, we can account for a third of the variation in raw earnings gaps across firms with these skill measures.

Once we include our full set of controls, the firm wage gap narrows by more than half. Also, we find that the full set of skills measures accounts for 21 percent of this residual variation. All of the skill measures are important in explaining for this variation, with nearly half accounted for by education and experience requirements, a third by the other eight skill measures, and a fifth by the cognitive and social requirements.

The table also presents a form of the decomposition for the firm performance outcomes. These measures vary only at the firm level thus equation 3 has no residual. Also, we cannot separetely identify firm effects and controls. Instead, in columns 6 and 9, we first residualize firm performance and each of the skill measures on our base set of controls and estimate equations 4 and 5 on firm averages of the residuals.

We find that the skill measures account for a modest fraction of the variation in firm performance. They account for 10 percent of the residual probability of being publicly traded (column 6), after taking into accont variation in occupations, industries, and locations. As with wages, about half of the explained component is driven by education and experience requirements, with still important roles for the job skills.

Finally, job skills account for 15.6 percent of the firm-level variation in log revenue per worker. Here social and cognitive skills are the most important predictor, accounting for about half the explained variation in firm performance residuals.

This exercise is only illustrative since we do not have ad-level variation in wages or firm outcomes. However, we still find that firms vary widely in both wages (as measured by average wages in the markets they post ads in) and performance outcomes. Also, importantly, we find that firm-level variation in demand for the job skills explains substantial fractions of firm-level variation in these outcomes. This suggests the job skill measures we define in this paper provide useful information about how firms produce and why they have diverse outcomes.

5 Conclusion

In this paper we use data from the near-universe of online job vacancies in the U.S. between 2010 and 2015 to study heterogeneity in employer skill demands. We use key words and phrases from the actual text of job vacancies to form measures of a variety of distinct "job skills" such as cognitive skills, social skills, management and finance. We show that the prevalence of job skills is positively correlated with relative wages across labor markets and performance differences across firms, even after controlling for education and experience requirements and detailed occupation and industry codes. This suggests that our measures of job skills add explanatory power beyond what is available in typical labor market data. Using a simple decomposition, we show that variation in skill requirements can explain about 20 percent of the residual variance in wages across labor markets and between 10 and 15 percent of the variance in firm productivity in our sample.

Turning to specific job skills, we find consistent evidence that higher paying labor markets and firms demand higher levels of cognitive skills and social skills from their employees. Morever, we find particularly large correlations between wages and whether a job vacancy requires *both* types of skill. This is consistent with recent theory and evidence from Weinberger (2014) and Deming (2016), who focus on the growing importance of social skills and growing complementarity between cognitive skills and social skills in the labor market. We also find that more productive firms - as measured by log revenue per worker as well as whether the firm is publicly traded - have greater demand for cognitive skills and social skills even after controlling for a detailed set of job characteristics such as industry and occupation. Our decomposition shows that cognitive and social skill requirements, in particular, account for 5 percent of residual variation in wages and between 2 and 7 percent of variation in firm productivity.

Rising variation in pay across firms, holding constant the quality of workers they employ, as been shown to be a primary contributor of rising wage inequality. Our results show that some of this variation can be accounted for by variation across firms in demand for cognitive and social skills. We cannot measure production technologies directly in the BG data. However, our results are consistent with at least some of the rise in inequality across firms being driven by a differential ability to adopt better production technologies that are more complementary with machines.

More generally, this paper demonstrates the usefulness of job vacancy data for studying employer skill demands and relative wages across labor markets and occupations. Future work could pair job vacancy data with detailed information about firm characteristics and production technology. This would enable researchers to study changes in the returns to specific skill requirements with a better understanding of how heterogeneity in skill demands translates to firm production technology.

References

[1] Acemoglu, D. and Autor, D. (2011), "Skills, tasks and technologies: Implications for employment and earnings," *Handbook of Labor Economics,* 4, 1043-1171.

[2] Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer (2014), "Trends in Earnings Differentials across College Majors and the Changing Task Composition of Jobs," *American Economic Review Papers and Proceedings*, 104(5): 387–393.

[3] Autor, David H. and David Dorn (2013), "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5): 1553-1597.

[4] Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2008), "Trends in US Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90(2): 300–323..

[5] Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman (2016), "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States." *Journal of Labor Economics*, 34(2): S67- S97.

[6] Beaudry, Paul , David A. Green, and Benjamin M. Sand (2014), "The Declining Fortunes of the Young since 2000," *American Economic Review*, 104(5): 381–386

[7] Beaudry, Paul , David A. Green, and Benjamin M. Sand (2016), "The Great Reversal in the Demand for Skill and Cognitive Tasks," *Journal of Labor Economics*, 34(S1 Part 2): S199–S247.

[8] Bloom, Nicholas and John Van Reenen (2007), "Measuring and Explaining Management Practices across Firms and Countries," *Quarterly Journal of Economics*, 122(4): 1351–1408.

[9] Card, David and John E. DiNardo (2002, "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4): 733-783.

[10] Card, David, Jörg Heining, and Patrick Kline (2013), "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 128(3): 967–1015.

[11] Carnevale, Anthony P., Tamara Jayasundera, and Dmitri Repnikov, (2014), *Understanding Online Jobs Data: A Technical Report*, Washington D.C.: Georgetown University Center on Education and the Workforce.

[12] Castex, Gonzalo, and Evgenia K. Dechter, (2014), "The Changing Role of Education and Ability in Wage Determination," *Journal of Labor Economics*, 32(4): 685–710.

[13] Deming, David J. (2015), "The Growing Importance of Social Skills in the Labor Market," NBER Working Paper #21473.

[14] Diamond, Rebecca (2016), "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106(3): 479-524.

[15] Firpo, S., N. Fortin, T. Lemieux (2009), "Unconditional Quantile Regressions," *Econometrica*, 77(May): 953-973.

[16] Gathmann, Christina and Uta Schonberg (2010), "How General is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 28(1): 1-49.

[17] Gibbons, Robert and Rebecca Henderson (2012), "What Do Managers Do?" *Handbook of Organizational Economics*, Princeton University Press.

[18] Glaeser, Edward and Joshua D. Gottlieb (2009), "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature.*

[19] Goos, M., A. Manning, and A. Salomons (2014), "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *The American Economic Review* 104(8): 2509-2526.

[20] Heckman, J. J. and T. Klautz (2012), "Hard Evidence on Soft Skills," *Labour Economics*, 19(4), 451-464.

[21] Herbein, Brad and Lisa B. Kahn (2016), "Do Recessions Accelerate Routine-Biased Technological Change," mimeo.

[22] Ichniowski, Casey, Kathryn Shaw and Giovanna Prennushi (1997), "The Effects of Human Resource Management Practices on Productivity," *American Economic Review* 87(3): 291-313.

[23] Kuhn, Peter, and Kailing Shen. (2013), "Gender Discrimination in Job Ads: Evidence from China," *Quarterly Journal of Economics*, 128(1): 287–336.

[24] Marinescu, Ioana and Ronald Wolthoff (2016), "Opening the Black Box of the Matching Function: the Power of Words," mimeo.

[25] Moretti, Enrico (2013), "Real Wage Inequality," *American Economic Journal: Applied Economics*, 5(1).

[26] Poletaev, Maxim and Chris Robinson (2008), "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000," *Journal of Labor Economics*, 26(3): 387–420.

[27] Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality", NBER working paper #21199.

[28] Speer, Jamin (2016), "Pre-Market Skills, Occupaitonal Choice, and Career Progression," *Journal of Human Resources*, forthcoming.

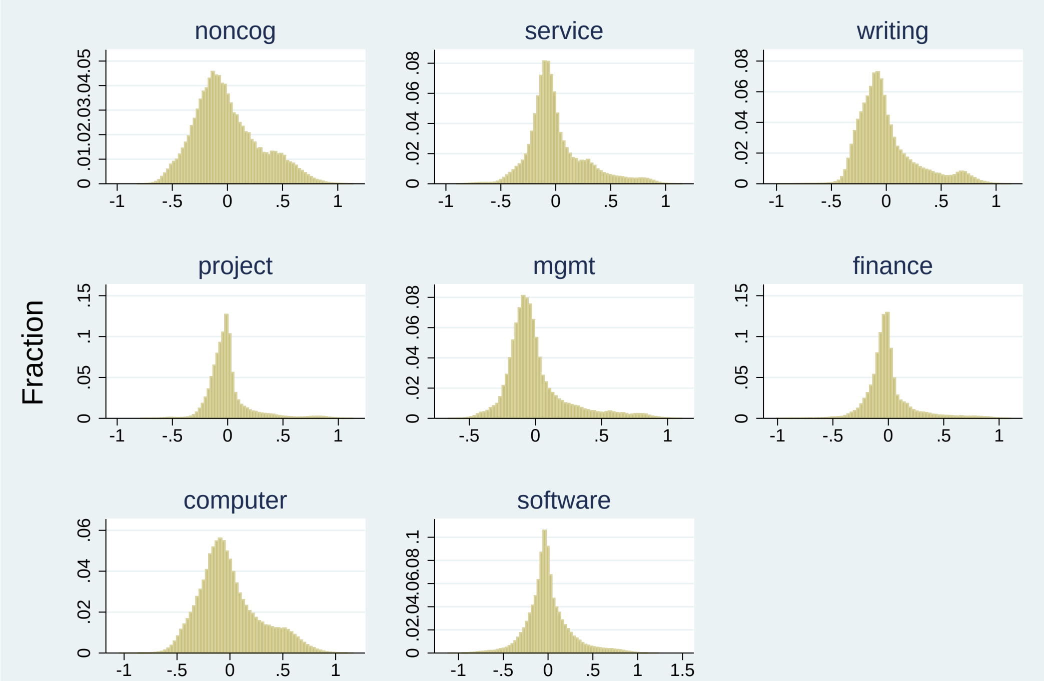
[29] Weinberger, C. J. (2014), "The Increasing Complementarity between Cognitive and Social Skills," Review of Economics and Statistics, 96(4): 849-861.

Table A1: Summary Statistics

	Panel A: MSA-Occ Sample				Panle B: Firm Sample				Panel C: In Compustat Sample			
	Mean	St Dev	Min	Max	Mean	St Dev	Min	Max	Mean	St Dev	Min	Max
Share of Ads with Skill Requirement:												
Cognitive	0.420	0.174	0	1	0.416	0.204	0	1	0.496	0.168	0	1
Social	0.425	0.126	0	1	0.420	0.203	0	1	0.505	0.169	0	1
Cognitve and Social	0.250	0.112	0	1	0.246	0.163	0	1	0.316	0.150	0	1
Character	0.345	0.132	0	1	0.340	0.190	0	1	0.413	0.153	0	1
Writing	0.231	0.103	0	1	0.228	0.155	0	1	0.264	0.132	0	1
Customer Service	0.237	0.163	0	1	0.230	0.180	0	1	0.254	0.149	0	1
Project Management	0.132	0.125	0	1	0.128	0.107	0	1	0.165	0.093	0	0.88
People Management	0.185	0.109	0	1	0.184	0.125	0	1	0.212	0.107	0	1
Finanical	0.167	0.185	0	1	0.163	0.141	0	1	0.199	0.118	0	1
Computer (general)	0.317	0.138	0	1	0.311	0.188	0	1	0.376	0.154	0	1
Software (specific)	0.318	0.273	0	1	0.309	0.246	0	1	0.356	0.199	0	1
Education	0.653	0.153	0	1	0.652	0.216	0	1	0.714	0.168	0	1
Experience	0.622	0.129	0	1	0.611	0.204	0	1	0.672	0.161	0	1
Years School, conditional	15.69	1.26	12	21	15.68	0.94	12	21	15.61	0.75	12	21
Years Experience, conditional	4.14	1.40	0.08	15	4.13	1.48	0.08	15	4.63	1.36	0.31	13.43
Missing ACS link	0.05	0.22	0	1	0.05	0.12	0.00	1	0.04	0.07	0.00	1.00
MSA-Occ median wage	42.15	15.44	8.15	135.00	41.94	7.21	10.53	123.52	43.76	6.04	17.74	88.39
Publicly Traded	0.308	0.144	0	1	0.300	0.458	0	1	1.000	1.000	1.000	1.000
Revenue per Worker ($ millions)	0.474	0.696	0	122.125	0.485	4.974	0.000	1122	0.485	4.975	0.000	1122
# Occupations Posted in	97.64	18.70	1.91	287	100.35	64.28	2	287	137.79	55.50	2	241
# MSAs Posted in	97.51	33.04	1.07	365	99.70	98.75	2	370	164.27	100.98	2	354
# Ads per cell (unweighted)	412	2495	1	154187	247	2032	2	230817	1762	6984	2	147576
# Cells		53,510				92,349				3,690		
# Ads, total		22,022,211				22,752,731				6,501,640		

Notes: Sample includes BG ads from 2010-2015, restricted to ads with non-missing firms that post at least 10 ads in the sample as a whole and ads in at least 2 MSAs and 2 occupations. In panel A, an observation is an MSA-occupation (6-digit SOC) cell, restricted to cells that can be matched to OES data. In panels B and C, an observation is a firm cell and panel C restricts to firms that can be matched to Compustat. Data are weighted by number of ads posted in the cell, unless otherwise noted. Skill requirements come from BG data; see table 1 for description. Median wage comes from OES data and is the average in the MSA-occupation from 2010-2015. Revenue pe Worker comes from Compustat and is the firm-level average from 2010-2015.

Figure A1: Histograms of Skill Requirement Residuals, Other Skills



Skill requirements residualized on education, experience, and industry, occupation and MSA fixed effects.

Table A2: Skill Requirements for Selected Occupations

Occupation (SOC, 6 digit)	Job Skills									
	Cognitive	Social	Character	Writing	Cust Serv	Proj Mgmt	Ppl Mgmt	Financial	Computer	Software
Accountants and Auditors	0.54	0.40	0.42	0.29	0.11	0.09	0.19	0.81	0.48	0.33
Computer Programmers	0.47	0.39	0.26	0.25	0.15	0.14	0.14	0.05	0.29	0.80
Computer User Support Specialists	0.43	0.43	0.35	0.23	0.38	0.08	0.12	0.05	0.58	0.46
Computer and Info. Systems Analysts	0.50	0.53	0.45	0.25	0.21	0.37	0.38	0.29	0.32	0.56
Elementary School Teachers	0.20	0.27	0.19	0.11	0.02	0.01	0.14	0.01	0.08	0.02
Financial Analysts	0.88	0.51	0.50	0.26	0.11	0.10	0.10	0.83	0.61	0.38
Financial Managers	0.55	0.48	0.43	0.26	0.30	0.11	0.39	0.62	0.38	0.20
Managers (General, Operations)	0.42	0.46	0.43	0.24	0.23	0.12	0.42	0.33	0.33	0.13
Lawyers	0.41	0.44	0.34	0.32	0.10	0.05	0.14	0.12	0.18	0.07
Loan Officers	0.37	0.49	0.39	0.39	0.51	0.01	0.13	0.31	0.44	0.07
Management Analysts	0.80	0.49	0.44	0.31	0.16	0.27	0.18	0.25	0.40	0.39
Registered Nurses	0.23	0.31	0.19	0.10	0.46	0.01	0.14	0.02	0.15	0.04
Sales Managers	0.36	0.53	0.43	0.20	0.69	0.08	0.28	0.19	0.33	0.11
Postings-Weighted Overall Average	0.42	0.42	0.34	0.23	0.23	0.13	0.18	0.16	0.31	0.31

Notes: Based on Burning Glass data. See table 1 for job skills definitions.