FIRM HETEROGENEITY IN SKILL DEMANDS

David Deming
Lisa B. Kahn

WORKING PAPER 23328

FIRM HETEROGENEITY IN SKILL DEMANDS

David Deming
Lisa B. Kahn

Firm Heterogeneity in Skill Demands
David Deming and Lisa B. Kahn
NBER Working Paper No. 23328
April 2017
JEL No. F1,F16

ABSTRACT

We study variation in skill demands across firms and labor markets using job vacancy postings. We categorize a wide range of keywords found in job ads into ten general skills. We find a positive correlation between average wages and cognitive skill and social skill requirements that remains even after controlling for occupation, industry, and MSA fixed effects. We also find a positive correlation between these skill requirements and indicators of firm performance such as revenue per worker and whether the firm is publicly traded. These returns are concentrated among ads that specify both a cognitive and social skill requirement. Job skills have explanatory power for pay and firm performance, beyond what is available in widely-used labor market data.

David Deming
Harvard Graduate School of Education
Gutman 411
Appian Way
Cambridge, MA 02138
and Harvard Kennedy School
and also NBER
david_deming@gse.harvard.edu

Lisa B. Kahn
Yale University
School of Management
P. O. Box 208200
New Haven, CT 06520
and NBER
lisa.kahn@yale.edu

1   Introduction

Wage inequality in the U.S. has increased markedly over the last few decades. A large literature in economics links rising wage inequality in the U.S. and other developed countries to technological change, specifically the computerization of the labor market (e.g. Autor, Levy and Murnane 2003, Autor, Katz and Kearney 2008, Acemoglu and Autor 2011, Goos, Manning and Salomons 2014). This work focuses on changes in relative demand across occupations that differ in the content of the tasks that are typically performed on the job. For example, Autor, Levy and Murnane (2003) show that the falling price of computer capital led to employment and wage reductions in routine task-intensive occupations and increases in cognitive occupations.

One empirical limitation in the study of technological change is that existing data sources typically measure variation in workplace tasks *across* occupations, but not within them. This work typically proceeds by linking measures of occupation task content – obtained from data sources such as the Dictionary of Occupational Titles (DOT) or the Occupational Information Network (O*NET) - with employment data. Autor, Levy and Murnane (2003) show that cognitive-intensive occupations such as doctors and lawyers have become a larger share of the U.S. labor force since the 1970s. This work documents relative employment shifts across occupations, but says nothing about whether the tasks performed by doctors and lawyers have changed over time in response to technological change. Yet wage inequality has also increased sharply within occupations in recent decades (e.g., Firpo, Fortin and Lemieux 2009). Perhaps rising wage inequality across space and over time is linked to changes in task content, even within narrowly defined categories of occupations.

In this paper we study heterogeneity in employer skill demands using a dataset of job vacancies that encompasses nearly all jobs posted to an online source in the U.S. from 2010-2015. The data include detailed information about employers' skill requirements for individual jobs, including firm identifiers and other important observable characteristics such as occupation, industry, and education and experience requirements. The data also include a vast and detailed set of job requirements, ranging from general skills such as "problem-solving" or "negotiation" to specific capabilities such as proficiency with specialized software (e.g., Java and SQL). Importantly, these data allow us to move beyond the limitations of existing work and ask if there are labor market returns to skills even within narrowly-defined occupation categories.

We begin by winnowing down the large variety of keywords and phrases in job vacancy posting into ten commonly observed and recognizable job skills. These skill definitions are motivated by the task literature that ties shifts in relative demand across occupations to technological change. We first show that there is substantial heterogeneity in demand for these skills, even within occupation, industry and location and when controlling for other characteristics such as education and experience requirements. We next show that the prevalence

of these skills is correlated with proxies for job productivity. Differences in employer skill demands predict occupational wage premia across Metropolitan Statistical Areas (MSAs), even after controlling for a rich set of characteristics. We also show that variation in skill demands is positively correlated with measures of firm performance. Finally, we show that the ten job skills account for roughly 15 percent of the variation in our aggregated wage measure across firms. Overall, our results show the measures of advertised skill requirements found in job vacancy data are important predictors of wage patterns, beyond what is available in widely-used labor market data.

In principle, measures of skill requirements extracted from job vacancy data can be used to study the returns to a variety of skills across occupations, labor markets and other contexts. Here we focus on just two - cognitive skills and social skills. A large literature shows that the returns to cognitive skill have increased in recent decades, perhaps because computers complement workers performing cognitively complex, abstract tasks (Autor, Levy and Murnane 2003, Autor, Katz, and Kearney 2008). More recent work suggests that the return to cognitive skills may be evolving, with important consequences for income inequality. Beaudry, Green, and Sand (2016) show that cognitive jobs as a whole saw little wage or employment growth in the last decade, while researchers instead point to a rising complementarity between cognitive and social skills (Borghans et al. 2014, Weinberger 2014, Deming 2016). Indeed social skills may become increasingly important as computers substitute for a wider range of non-interactive tasks (Autor 2014, Lu 2015).[1] We highlight cognitive and social skills because of their prominence in this literature linking technological change to wage inequality, and because of the sense that they are generally important in a wide range of jobs. This focus provides a clear illustration of what can be learned from the skill requirements found in job postings, relative to existing data.

To understand whether stated preferences for cognitive and social skills are indeed indicative of higher skilled and more productive jobs, we investigate the relationship between this stated demand and proxies for productivity of workers in the advertised jobs. Because job ads tend not to specify wages (Kuhn and Shen 2013, Marinescu and Wolthoff 2016), we link our data to two external sources measuring wages and performance at aggregated levels. First, we obtain average wages at the location-occupation – MSA by 6-digit SOC occupation – level from the Occupational Employment Statistics (OES) program. Second, we obtain two measures of firm performance: whether the firm posting the add is publicly traded (as proxied by whether we can link the name to a firm in Compustat) and revenue per worker for publicly traded firms.

We find that cognitive skill and social skill requirements in job ads positively predict occupational wage differences across local labor markets, even after controlling for education

---

[1]For example, advances in automation of some types of cancer diagnoses and treatment, as well as in legal e-discovery, make possible the automation of some aspects of even doctors' and lawyers' jobs (Brynjolfsson and McAfee 2014; Demng 2016).

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 average frequency of firm vacancy postings within locations and occupations. These returns are largely concentrated among ads that post both cognitive and social skill requirements, which is consistent with the complementarity between cognitive skills and social skills found in Weinberger (2014) and Deming (2016).

Our results suggest that variation in pay and performance across MSAs and firms may partly reflect differences in skill demands. We estimate that demand for cognitive and social skills accounts for 5 percent of the variation in firm pay, within our detailed set of controls (occupation, industry, education and experience requirements, and the other eight general skill requirements), and similar fractions of variation in firm performance. Our pay measure masks considerable variation across firms since it is only available at an aggregated MSA-occupation level. We might expect this percentage to be even larger for worker-level pay.

Our results are thus consistent with wage inequality across firms being driven in part by differences in skill utilization, perhaps as some firms take better advantage of modern production technologies. Naturally, we do not observe production technology in our data and cannot rule out that an omitted factor, even within our detailed controls, drives both stated skill demand and outcomes. However, these results are interesting in light of recent research that points to widening pay differentials across firms and across U.S. cities as important contributors to the overall rise in inequality.[2] We use our short panel to link variation in pay to variation in demand for tasks that are particularly associated with rising inequality over time, even though we cannot speak to the time series on rising pay and performance differentials across firms and cities. We therefore 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 (Weinberger 2014, Beaudry, Green and Sand 2016, Deming 2016).

Research on the task content of occupations has furthered our understanding of broad labor market trends in inequality (e.g., Autor, Levy, and Murnane 2003, Autor and Dorn 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 advertised skill requirements in explaining heterogeneity in wages across firms and labor markets, even *within* narrowly defined occupations. 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 Pren-

---

[2]For research on firms, see other papers in this volume, including Barth, Bryson, Davis and Freeman (2016), Card Heining and Kline (2016), and Song et al. (2015). For research on geographic sorting and agglomeration economies, see Glaeser and Gottlieb (2009), Moretti (2013), Diamond (2016).

nushi (1997)). We also 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 (Hershbein and Kahn 2016), firm preferences for discrimination (Kuhn and Shen 2013) and the general equilibrium impacts of unemployment insurance (Marinescu 2017).

This paper proceeds as follows. In section 2 we describe the datasets and explain how we classify ten job skills from the thousands of key words found in the job ads. 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 firm labor 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 future.

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, removes duplicates and parses the ads into a systematic, machine-readable form. They use the resulting database to create labor market analytic products. BG claims their database encompasses nearly all jobs posted to an online source during the time period of measurement. The BG microdata were first used by Hershbein and Kahn (2016) to study whether the Great Recession accelerated the adoption of routine-labor saving technologies.

We restrict our attention to ads posted in "professional" occupations (roughly 60 percent of all ads).[3] This category includes almost all occupations employing college-educated workers, and it has the most representative coverage in the BG data (Carnevale, Jayasundera and Repnikov 2014, Hershbein and Kahn 2016).[4] Furthermore, professional occupations are the focus of the primary debate to which we would like to speak. Focusing on professional occupations, Beaudry, Green, and Sand (2016) show that occupations requiring cognitive skills saw no employment or wage growth in the 2000s. Deming (2016) instead finds growth in occupations that require both cognitive and social skills, providing evidence for an increasing complementarity precisely across higher paying occupations. The Burning Glass data provide a unique view of the cross-sectional variation in skill requirements *within* detailed

---

[3]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.

[4]Naturally, data on job ads posted to an online source will not be perfectly representative of all job ads or of employment, even within professional occupations. See Hershbein and Kahn (2016), especially the data appendix, for more detail on sample representativeness, how the BG sample changes over time, and how skill requiements relate to the skills of employed workers.

occupations. The large sample sizes allow us to explore heterogeneity in skill demands across cities and firms, with a particular focus on professional occupations.

Our resulting dataset contains nearly 45 million ads, for the years 2010-2015.[5] The BG data include education and experience requirements, detailed industry and occupation codes (6-digit SOC), the location, the date the ad was posted 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 requirements. In our sample, 93 percent of ads have at least one such requirement, and, conditional on having any requirement, post an average of 9. In the next subsection, we describe how we distill over ten thousand unique keywords and phrases identified by BG into a subset of general "job skills" that could be useful across a wide range of jobs.[6]

We focus on two samples: (1) the full set of professional ads (hereafter, the full sample) and (2) the 63 percent of ads with a non-missing firm.[7] This latter sample allows us to estimate heterogeneity in skill requirements across firms. For the firm sample, we further focus our attention on the remaining 96 percent of ads from firms that have posted at least ten ads total and ads in at least two different professional occupations and at least two different MSAs over our sample period.[8] These restrictions help clean out some noise in the firm-level variation in skill requirements that we explore. The resulting database, hereafter, the firm sample, contains nearly 86,000 unique firms.

Only 13 percent of professional ads post the wage that is offered, which prevents us from studying the relationship between skill demands and wages for individual job vacancies.[9] 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 for MSA-occupation (6-digit SOC) cells from the 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. Six percent of job vacancies are posted in MSA-occupations that cannot be matched to OES data, or have suppressed wage data, likely because they are too small, and these cells are removed from this analysis.

We obtain firm performance data from Compustat North America by Standard & Poors.[10]

---

[5]We exclude Micropolitan Statistical Areas from our analysis (roughly 5 percent of ads) since external data on wages and location characteristics is either unavailable or much less precise.

[6]BG cleans and codes the text of job ads into a taxonomy of thousands of unique, but standardized requirements. Beginning with a set of pre-defined possible skills, BG using machine learning algorithms to search text in an ad for an indication that the skill is required. For example, for team work, they search for the key words "team work" but also look for variations such as "ability to work as a team."

[7]Ads that do not contain a firm identifier are typically obtained from recruiter websites where the poster does not wish to reveal the information. We explore robustness to this sample restriction, below.

[8]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.

[9]It is well-known that vacancies rarely post wage offers. See for example Kuhn and Shen (2013) and Marinescu and Wolthoff (2016).

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

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 ads in the firm sample to a publicly traded firm in Compustat.[11] Our main firm 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.

We also obtain MSA demographic characteristics to use as control variables from American Community Survey (ACS) data. Here, again, we can match all but a small fraction of our sample (making up 3.5 percent of ads) to the ACS, excluding some small cities.[12]

Both wages and revenue per worker are available at an annual frequency. However, because the time series in the BG data is fairly short and because job ads likely represent intentions for the flow of new workers, while wages and firm performance variables reflect the stock of existing workers, we aggregate all data sources, taking an unweighted average across the years 2010-2015. We thus restrict our attention to cross-sectional relationships between average skill requirements at the MSA-occupation or firm-level and wages or firm performance. This aggregation has the advantage of smoothing out time variation in skill requirements driven by factors outside the scope of this analysis, such as labor supply shocks or labor market conditions.[13]

Appendix table A1 provides summary statistics for ads in the full sample (Panel A), the firm sample (Panel B) and the Compustat-matched sample (Panel C). The full sample has coverage in 56,611 MSA-occupation (6-digit SOC) cells, containing an average of nearly 800 ads, though there is a wide range. The ad-weighted average wage is $42/hour in our sample of professional occupations.

The firm sample (Panel B) contains nearly 86,000 firms, posting in an average of 301 ads, each, again with a wide range. The average ad is from a firm that posts in roughly 100 occupations (out of 352) and roughly 100 MSAs (out of 371). This sample has coverage across nearly as many MSA-occupation cells as the full sample (55,241), though with typically fewer ads per cell.

---

[11]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 (75 percent of matched firms). We then use a fuzzy match algorithm to link firms with at least a 95 percent chance of being the same (3 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.

[12]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.

[13]Hershbein and Kahn (2016) show a persistent shift in skill requirements in MSAs hit harder by the Great Recession that is consistent with an increased adoption of routine-labor saving technologies. Sasser Modestino, Shoag and Ballance (2016) show that skill requirements are responsive to a labor supply shock: returning U.S. troops from Iraq and Afghanistan.

Panel C shows that the 3,622 firms that can be matched to Compustat are bigger. They post in an average of 2,138 ads across a larger number of MSAs and occupations. They still have coverage in over 45,000 MSA-occupation cells, with an average of 173 ads per cell. The average ad is posted to a firm with roughly half a million dollars in revenue per worker and they pay slightly more for professional workers ($44/hour).

2.2   Job Skills in the Burning Glass Data

One contribution of this paper is to distill and analyze the keywords 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, 59 percent of professional ads specify an education requirement that averages 15.7 years of school (fitting modal years to degree requirements). 58 percent of ads specify a requirement for experience in the field that averages 4 years. The propensity to specify these requirements is higher in the firm sample and even higher in the Compustat-matched sample. Interestingly, Hershbein and Kahn (2016) show that average education requirements in BG align well with the education levels of employed workers at the occupation and MSA levels. This shows that, for the most part, firms continue to specify skill requirements, even for high-skilled jobs where an education requirement might be assumed.[14]

Figure 1 provides some intuition for the relationship between education requirements and earnings among professional occupations. We plot demeaned hourly wages (blue bars) from the OES ordered by their city wage rank for the 50 largest MSAs in terms of employment 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 in the full sample. 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 0.84 when weighted by MSA employment). This correlation is reassuring that the BG data 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, 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.[15]

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

---

[14]For example, most ads posted on the Job Openings in Economics website specify a Ph.D. requirement.

[15]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 1: Wages and Education Requirements by City Wage Rank
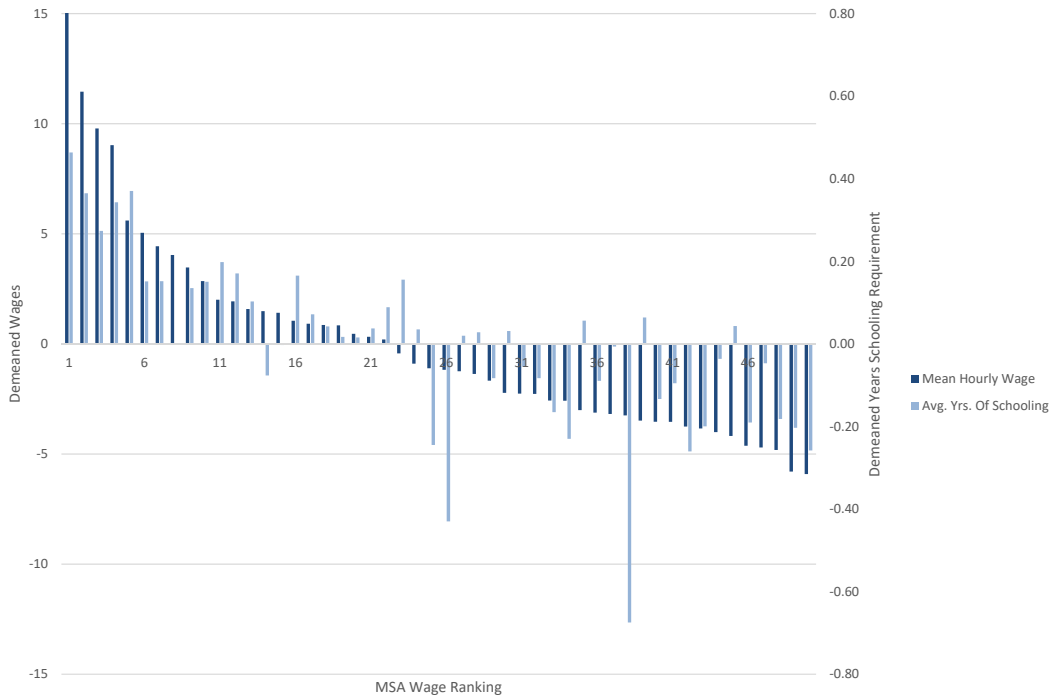
Notes: Dark bars show MSA average wages for professional occupations, obtained from the OES 2010-2015. Light bars show MSA average years schooling required in Burning Glass job ads 2010-2015, conditional on having any stated requirement. Chart includes top 50 MSAs (in terms of employment in professional occupations) and is ordered by MSA wage rank.

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.

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). For both cognitive and social skills, we check each key phrase to ensure that its prevalence corresponds with occupations that the previous literature tends to classify as cognitive or social.[16]

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. BG codes these categories themselves. The former encompasses the generic phrase "computer skills" as well as key phrases for common software such as Microsoft Excel, while the latter includes specialized software such as SQL, Java, or C++.

To provide intuition for these job skills, appendix table A2 shows the share of job vacancies that require each skill for selected occupations in the full sample. As can be seen, this skill demand is fairly high for cognitive, social and character skills, which are required across a broad range of jobs. Also, job skills line up well with occupation titles and perceived job tasks. 84 percent of vacancies for accountants and auditors require financial skills, compared to an average of only 16 percent across all professional occupations. Similarly, the share of computer systems analysts requiring project management skills is particularly high, as is the share of computer programmers requiring skill in a specific software program. Ads for registered nurses are more likely to specify a customer service requirement. Finally, there is a lot of variation in skill requirements even within detailed occupation categories. For example, among managers, financial manager vacancies are much more likely to require financial skills, than sales and general managers. In contrast, customer service skills are much more likely to be requested for sales managers than for the other types.

Our exercise falls very much in the spirit of the task literature, which catalogs task requirements at the occupation level using the DOT or O*NET and explores variation over time or over a career (Autor, Levy and Murnane 2003, Poletaev and Robinson 2008, Gath-

---

[16]Specifically, at the occupation level, we find that the share of ads containing a given key word is positively related to an index score for how cognitive or social the occupation is. Index scores are based on O*NET 18.1 and stem from measures used in Deming (2016) based on an earlier time period. Hershbein and Kahn (2016) also follow this approach.

mann and Schonberg 2010). In what follows, we primarily focus on cognitive skills and social skills because they are important predictors of productivity and wages in a wide variety of occupations and because of their prominence in the literature linking technological change to wage inequality (e.g. Weinberger 2014, Deming 2016). While we can show that the remaining skill measures have explanatory power in pay and performance regressions, we refrain from interpreting (or showing) the signs or magnitudes of the coefficients on the other eight job skill measures. This is because we do not have a general framework for analyzing them. Some skills may be valuable, the higher the requirement, while some may reflect horizontal differentiation in production or selection of workers, and others may substitute or complement each other. For example, listing "character" skills might be a signal of lower-paying jobs where obedience is required, even if character skills themselves are valuable in all jobs. We leave questions such as these for future work.

2.3   Variation in Skill Requirements

Appendix Table A1 provides summary statistics for these measures. 37 percent of professional jobs specify a cognitive skill requirement and 36 percent specify a social requirement. A quarter of ads specify both requirements. Frequency of appearance for the other job skills ranges from one-ninth of ads (project management) to nearly a third (character and software). Skill requirements are a bit more prevalent in the firm sample, and especially for ads in the Compustat-matched sample.

There is also substantial variation in skill requirements across ads. Figure 2 illustrates variation in skill requirements that can be accounted for by the other characteristics found in the ads. We regress the probability that an ad posts a given skill on our most detailed set of controls: 6-digit occupation, MSA, and firm fixed effects as well as education and experience requirements.[17] We estimate this regression on the firm sample, which restricts to firms with at least 10 ads and postings in at least 2 MSAs and 2 occupations, to limit noise in the firm fixed effects.

The bottom bar (solid dark blue) plots the variance in the fitted values for the occupation fixed effects. Occupation typically accounts for a fairly small fraction of the total variance in skill requirements, between 2 and 8 percent, except in finance and software, where it accounts for nearly 20 percent. This is intuitive given the nature of the skill requirements, and it gives us confidence that our measures capture real variation in employer demands for the indicated skill.

The next bar (pattern) plots the variance in fitted values for the remaining controls except firm effects. Controls for location, and the more typical skill measures (education and experience) tend to account for roughly 15 percent of the total variation. This is because,

---

[17]Education controls include an indicator for whether the ad has any education requirement and the number of years required if there is a requirement (otherwise zero), and experience controls are analogous.

Figure 2: Variances of Skill Requirements
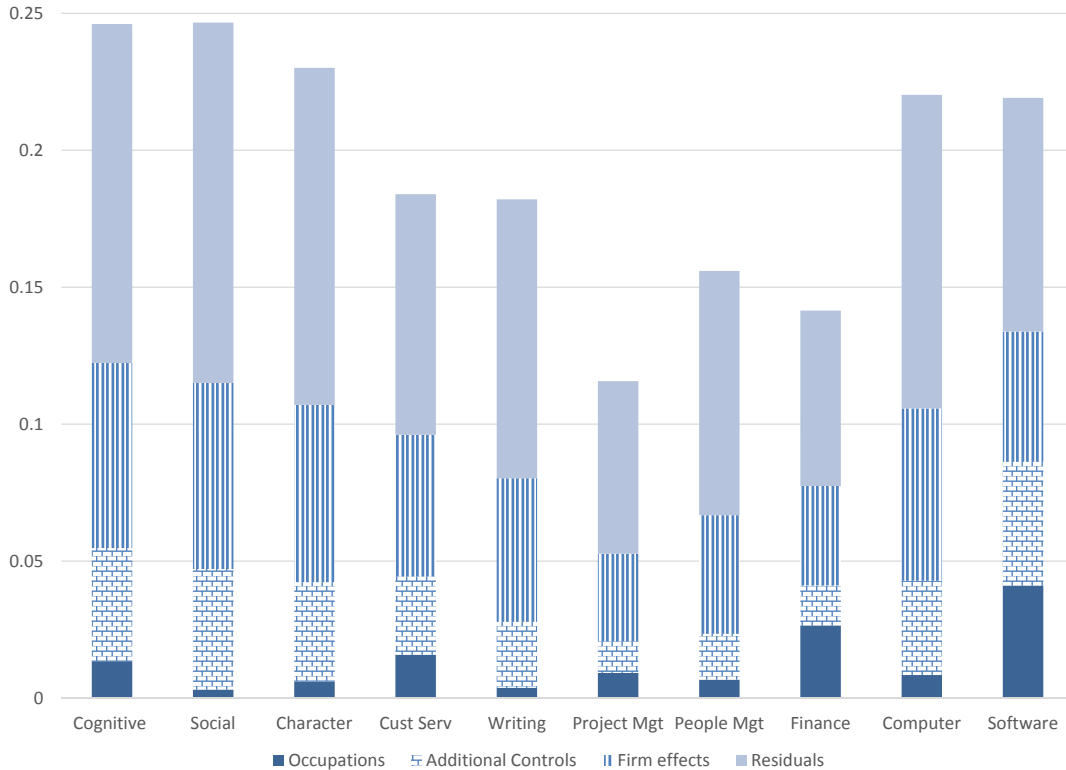
NOTES: Based on the firm sample. We regress an indicator for whether an ad has the skill requirement on occupation (6 digit) fixed effects, additional controls (MSA fixed effects and education and experience requirements) and firm fixed effects. Bars plot variances of fitted values based on specified controls or the residuals.

as we explain below, almost all of the skill measures are positively correlated with education and experience requirements and because the skill requirements in a location tend to reflect the skill of employed workers in the location (see Hershbein and Kahn 2016).

The third bar (vertical lines) plots the variance in fitted values for the firm fixed effects. Interestingly, we find that there is substantial systematic variation across firms in their tendency to specify these skill requirements. They account for nearly 30 percent of the total variation for most of the skill measures. Firms clearly have large systematic differences in their recruiting strategies, and possibly differences in skill utilization, even controlling for the distribution of occupations and MSAs they post in.

Finally, there is substantial remaining variation across ads in the propensity to specify these skill requirements, even within this very detailed set of controls. The residual variances tend to be about half of the total variation in requirements.

Table 2 presents bivariate correlations for each pair of skill requirements, weighted by ads, in the firm sample. There are two main takeaways from Table 2. First, cognitive skill and social skill requirements are positively correlated with each other, with years of education and experience required, and with most of the other eight job skills. This suggests that they are general skills deemed important by employers across a wide range of occupations and

labor markets. Second, the correlations between education and experience requirements and job skills are mostly positive but relatively small. This suggests that there is wide variation in employer skill demands even after accounting for education and experience requirements.

Table 2: Correlations of Skill Requirements

	Education	Experience	Cognitive	Social	Character	Cust Serv	Writing	Proj Mgmt	Ppl Mgmt	Financial	Computer	Software
Years of Education Required	1.00											
Years of Experience Required	0.30	1.00										
Cognitive	0.20	0.37	1.00									
Social	0.05	0.25	0.64	1.00								
Character	-0.06	0.14	0.59	0.69	1.00							
Customer Service	-0.27	-0.38	-0.03	0.17	0.14	1.00						
Writing	0.12	0.24	0.57	0.52	0.52	-0.07	1.00					
Project Management	0.20	0.57	0.55	0.45	0.39	-0.20	0.39	1.00				
People Management	-0.05	0.01	0.35	0.34	0.38	0.13	0.30	0.27	1.00			
Financial	0.02	0.21	0.43	0.35	0.37	-0.04	0.36	0.38	0.39	1.00		
Computers (General)	-0.06	0.27	0.52	0.52	0.54	-0.02	0.50	0.40	0.24	0.41	1.00	
Software (Specific)	0.26	0.61	0.36	0.25	0.11	-0.33	0.24	0.50	-0.06	0.02	0.27	1.00

Notes: Table shows ad-weighted bivariate correlations across all skill measures at the firm level using the firm sample. See table 1 for skills definitions.

3 Skill Demands and Worker and Firm Outcomes

The goal of this paper is to understand whether pay differentials across markets and firms are associated with differences in demand for skills. In particular, are firms and markets that pay more also more likely to require employees to have cognitive skills, social skills or both? Differences in demand for these skills and associated returns in terms of wages and firm performance may reflect variation in production technology, especially within narrowly defined occupations. Since we do not measure individual productivity or wages directly in our dataset, we use three proxies. The first is the average wage 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 the types of skills demanded. 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.

In thinking about the relationship between job skills and wages, a few examples may be instructive. First, consider computer programmers in Washington, DC versus Manchester-Nashua, NH. Hourly wages for computer programmers are about 25 percent higher in Washington DC compared to New Hampshire. 35 percent of computer programmer vacancies in Washington, DC require social skills, compared to only 21 percent in Manchester-Nashua, and firms with computer programmer vacancies in Washington, DC have about 10 percent higher revenue per worker than their New Hampshire counterparts.

Hourly wages for management analysts are about 25 percent higher in San Jose, CA compared to San Diego. 84 percent of management analyst vacancies in San Jose, CA require cognitive skills, compared to only 75 percent in San Diego, CA, and firms with management analyst vacancies in San Jose are also about 2 percentage points more likely to be publicly traded.

Our hypothesis is that while most computer programmer vacancies require cognitive skills, computer programming jobs 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 may engage in more complex problem-solving and analysis. These are only illustrative examples, and of course 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:

$$(1) \qquad log(wage)_{om} = \alpha + \overline{Skill}_{om}\beta' + controls + \epsilon_{om}$$

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 experiment with a few different sets of controls, always including the more standard skill measures, education and experience.[18] Additional controls are added successively and described below. 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) on the full sample. Column 1 summarizes a sparse model that includes the 10 skill measures and the education and experience variables. The coefficient of 0.11 on cognitive skills is statistically significant at the less than 1 percent level. The magnitude of the coefficient implies that a 10 percentage point increase in the share of job vacancies requiring cognitive skills is associated with 1.1 percent higher wages. Alternatively, a one standard deviation increase in the propensity to specify a cognitive skill requirement across occupation-MSAs (0.17) increases wages by nearly 2 percent. We estimate that adding any cognitive skill requirement is associated with a roughly equivalent wage gain to increasing the education or experience requirement by one year. The coefficient on social skills is much larger in magnitude and also statistically significant at the less than 1 percent level. A standard deviation increase (0.12) in demand for social skill is associated with a 5 percent wage increase.

Column 2 adds the share of vacancies in each MSA-occupation cell with *both* a cognitive skill and social skill requirement. Like Weinberger (2014) and Deming (2016), 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 (found in column 1) is explained by vacancies that ask for *both* types of skills. The coefficient implies that a one standard deviation increase (0.10) in the share of vacancies requiring both cognitive and social skills increases wages by 14 percent. Surprisingly, we find a negative wage return for ads that require only cognitive or only social skills, though this effect reverses with controls. We also report 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

---

[18]We control for the average years of education (experience) required in the cell, as well as the share of ads with any education (experience) requirement. For MSA-occupation cells with no education (experience) requirement, the years variable is set to 0 and the share of ads with any requirement will equal 0, thus indicating this imputation.

Table 3: Average Wages and Skill Requirements

	Dependent variable: Log(mean wages) in MSA-Occupation					
	(1)	(2)	(3)	(4)	(5)	(6)
Cognitive	0.113***	-0.413***	0.245***	0.181***	0.0792***	0.0465***
	(0.00908)	(0.0166)	(0.00784)	(0.0139)	(0.00873)	(0.0122)
Social	0.429***	-0.0919***	0.301***	0.236***	0.0517***	0.0202
	(0.0155)	(0.0206)	(0.0121)	(0.0167)	(0.00966)	(0.0127)
Both required		1.319***		0.157***		0.0760***
		(0.0349)		(0.0278)		(0.0198)
Years of Education	0.131***	0.129***	0.0764***	0.0765***	0.00865***	0.00873***
	(0.000770)	(0.000763)	(0.000844)	(0.000844)	(0.000995)	(0.000995)
Years Experience	0.160***	0.161***	0.0848***	0.0849***	0.0318***	0.0318***
	(0.00120)	(0.00118)	(0.00120)	(0.00120)	(0.00102)	(0.00102)
Base controls			x	x		
Detailed Controls					X	X
F-stat (cog and social)	553.1	855.0	1004	680.4	69.66	51.35
F-stat (all 10 skills)	1874	2054	612.6	560.1	59.93	55.83
MSA-Occ Cells	56,611	56,611	56,611	56,611	56,611	56,611
R-squared	0.702	0.710	0.846	0.846	0.940	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. Base controls include MSA characteristics from the ACS, four-digit SOC occupation fixed effects, and the share of ads in the MSA-occupation that are in each of the 2-digit NAICS industries. Detailed controls include MSA and six-digit SOC occupation fixed effects and the industry shares. Observations are from the full sample, weighted by number of ads in the MSA-occupation cell. See table 1 for skills defintions.

reject that these groups of coefficients are equal to zero ($p = 0.000$).

Columns 3 and 4 of Table 3 add a base set of controls: four-digit SOC occupation fixed effects, MSA characteristics from the ACS and the distribution of ads across two-digit NAICS industries. These controls help adjust for a range of alternative explanations for the positive correlation between skill requirements and wages. First, as noted, firms may post skill requirements in cities that have more skilled workers, which may also be more expensive cities where wages are higher. Second, occupation and industry controls adjust for a range of factors including differences in representation in online job ads, the propensity to not post requirements because skills are assumed, and the propensity to post requirements as a way of signalling something about the job other than skill utilization in production. For example, jobs for a salesperson may in general not specify a social skill requirement because it is always assumed, but within ads for salespeople, the prevalence of key words may signal a slightly different focus.

Even within these controls, the pattern of results is similar to the raw correlations. While the coefficient on cognitive skill requirements increases and the coefficient on social skills decreases, both remain positive and statistically significant at the less than 1 percent level, and the interaction between the two skills in Column 4 is also positive and statistically significant. We see positive returns for each skill individually that doubles when the two are required together. We can again 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). Cognitive and social skill requirements are positively correlated with occupation-MSA wages, with the majority of the return concentrated among ads that specify both requirements.

The results in Table 3 show clearly that data on the average skill requirements of jobs in local labor markets have explanatory power beyond what is available in conventional data sources. Labor markets paying relatively high wages to a particular occupation are likely to advertise for relatively higher cognitive and social skill requirements. We still find a positive return for cognitive skills, and it nearly triples when social skills are also required. 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. Appendix Table A3 shows that these results are very similar in the firm sample.

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:

$$(2) \qquad firm\_perf_f = \alpha_0 + \overline{Skill}_f\beta' + \bar{I}_f^o + \bar{X}_f\gamma' + \bar{I}_f^m + \theta_n + \epsilon_f$$

$\overline{Skill}_f$ 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. We alternately include our base set of controls and more detailed controls, as in Table 3. Columns 1-4 present results where the outcome is an indicator variable for whether the firm is publicly traded, and is estimated on the firm sample. We find evidence of a positive association between being publicly traded postings ads with cognitive skill and social skill requirements. This result holds with both the base set of controls (Column 1), and when controlling for occupation, industry and MSA fixed effects (Column 3). The coefficients on cognitive skill requirements are positive in Column 1 and Column 3, but only statistically significant when more detailed controls are included. The coefficients on social skill requirements are larger than the results for cognitive skills, and moreover they are positive and statistically significant at the less than 1 percent level in both specifications. In column 1, the coefficients imply that a one standard deviation increase in the share of vacancies with a social skill requirement (the standard deviation across firms is about 0.2 for both skills) is associated with a 3.2 percentage point increase in the probability that a company is publicly traded. This is a 10 percent increase relative to the baseline 30 percent probability that an ad is posted to a publicly traded firm.

Column 2 and Column 4 add the share of vacancies that require both types of skills. We see strong evidence of complementarity in both specifications. A one standard deviation increase in the probability that both skills are required (0.16) is associated with a 6 percentage point increase in the probability of being publicly. 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.

Table 4: Firm Outcomes and Average Skill Requirements

	Publicly Traded				Log(Revenue per Worker)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive	0.0131	-0.170***	0.0318**	-0.136***	0.469***	0.624***	0.379***	0.0761
	(0.0122)	(0.0180)	(0.0129)	(0.0185)	(0.117)	(0.190)	(0.136)	(0.218)
Social	0.162***	0.0165	0.0934***	-0.0364**	0.218**	0.348**	0.239*	-0.00813
	(0.0114)	(0.0155)	(0.0115)	(0.0154)	(0.105)	(0.164)	(0.123)	(0.185)
Both required		0.365***		0.328***		-0.268		0.531*
		(0.0262)		(0.0260)		(0.259)		(0.298)
Years of Education	-0.00212	-0.00141	-0.00242*	-0.00203	0.00423	0.00312	0.00979	0.00974
	(0.00134)	(0.00134)	(0.00135)	(0.00135)	(0.0222)	(0.0222)	(0.0266)	(0.0266)
Years Experience	0.0236***	0.0239***	0.0125***	0.0128***	0.0851***	0.0839***	0.119***	0.120***
	(0.00150)	(0.00150)	(0.00157)	(0.00157)	(0.0144)	(0.0145)	(0.0182)	(0.0182)
Base Controls	X		X		X		X	
Detailed Controls		X		X		X		X
F-stat (cog and social)	110.2	138.1	41.93	81.19	12.43	8.644	6.560	5.432
F-stat (all 10 skills)	181.6	183.1	130.3	133.2	10.96	10.06	4.072	3.993
# Firms	85,695	85,695	85,695	85,695	3,622	3,622	3,622	3,622
R-squared	0.246	0.248	0.330	0.332	0.511	0.511	0.736	0.737

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

Notes: Observations are at the firm level, weighted by number of ads posted by the firm. 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 firm 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. Years of education and experience equal 0 if the firm 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. Base controls include two-digit NAICS industry fixed effects and the ad-weighted distributions of four-digit occupation fixed effects and MSA characteristics from the ACS. Detailed controls included industry fixed effects and the ad-weighted distributions of MSA and six-digit SOC occupation fixed effects. See table 1 for skills definitions.

Columns 5 through 8 report results for log revenue per worker, among the 3,622 firms in our data that are publicly traded. Since this restriction limits our sample to the 30 percent of adds that are posted by a publicly-traded firm, we view these results with caution. Still, it is interesting to note that the correlation between skill requirements and revenue per worker, a proxy for firm productivity, holds even within a small sample of large, publicly traded firms. Column 5 shows that firms with higher revenue per worker also have significantly higher cognitive skill and social skill requirements. 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 in Column 8, 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.[19]

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 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 differences in skill utilization, perhaps indicating that some firms take better advantage of modern technologies (which may require cognitive skills, social skills, or both) and with commensurate productivity gains. Naturally, we cannot rule out that these correlations are driven by unobserved aspects of labor markets and firms. However, we add to the previous literature by showing that these positive correlations hold within a rich set of controls, and most notably, 6-digit occupation fixed effects.

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 variation in worker pay and firm performance, even when we include detailed controls for occupation, industry, and location. Variation in firm pay, controlling for worker quality, accounts for an important fraction of the rise in income inequality over the last 30 years (Barth, Bryson, Davis and Freeman (2016), Card Heining and Kline (2016), and Song et al. (2015)). If firms also vary systematically in their propensity to ask for

---

[19]It may seem surprising that the average years of education required by the firm is not positively correlated with firm performance. However, as noted, we also control for the share of ads that have any education requirement. This variable is highly correlated with the years required, since firms demanding more skilled workers are both more likely to post any requirement and also post for more years conditional on having a requirement. We find that the coefficient on the share of ads with any requirement is positive and usually highly significant, even while the number of years required, conditional on any, is not.

particular skills in job vacancies, then perhaps skill requirements can also account for some share of the variation in pay across firms.

Based on the relationship between skill requirements and pay and performance established above, we can calculate how much of the variation in pay across firms can be accounted for by any variation in skill requirements. Consider equation 3, which expresses $log(wages)$ for an occupation, o, MSA, m, and firm, f. β^f are firm fixed effects and *controls* are typical controls found in other datasets (our base controls). This regression is estimated at the firm-occupation-MSA level and is weighted by the number of ads in a cell.

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

The first column of Table 5 gives a general sense of variability in $log(wages)$ that can be attributed to firms. Here we present the ad-weighted standard deviations of β^f in the raw data (column 1), including our base set of controls (column 2), and including our detailed controls (column 3). 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 (0.19 log points). From column 2, we find that half of this variation remains when we include base controls. However, naturally, full controls for MSA, and 6-digit occupation absorb most of the variation in MSA-occupation level wages (column 3). This is a limitation of the fact that we cannot link wages to actual job postings.

The remaining rows of Table 5 summarize firm-level variation for the remaining variables. We provide standard deviations for the firm-level outcomes though, since we cannot separately identify control variables, we omit estimates in columns 2 and 3. A firm that is one standard deviation more likely to be publicly traded is 46 percentage points more likely, while a firm with one standard deviation higher revenue per worker is producing about double (0.83).

The table also shows variation in the average propensity to specify each job skill, and again, as noted in Section 2.3, we find substantial variation across firms. The standard deviation of firm effects for skill requirements is about 0.2 (20 percentage points) for both cognitive skill and social skill, and the variation shrinks only a little even after including detailed controls. The pattern is very similar for the other eight job skills, although the variation is generally smaller for most of them (especially after including detailed controls in Column 3). Thus, Table 5 shows substantial heterogeneity in firm skill demands even within occupation, industry and MSA.

To illustrate the importance of employer skill demands, we combine the positive correlation between skill requirements and job outcomes shown in Section 3 with the variation

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

	(1) No Controls	(2) Base Controls	(3) Detailed Controls
Log Hourly Wages	0.190	0.101	0.027
Publicly Traded	0.459	--	--
Log Revenue per Worker	0.827	--	--
Cognitive	0.203	0.176	0.168
Social	0.201	0.190	0.186
Cognitive and Social	0.162	0.149	0.145
Character	0.188	0.172	0.167
Customer Service	0.180	0.160	0.149
Writing	0.154	0.143	0.140
Project Management	0.106	0.098	0.081
People Management	0.125	0.122	0.116
Financial	0.141	0.101	0.091
Computers (General)	0.185	0.168	0.163
Specific Software	0.244	0.172	0.136

Notes: We regress the variable in each row on firm fixed effects and specified controls. The table reports standard deviations of the firm fixed effects, weighted by the number of postings to each firm. Base controls include MSA characteristics, 4-digit occupation fixed effects and industry fixed effects. Detailed controls include MSA, 6-digit occupation and industry fixed effects. Specifications including controls are omitted for Publicly Traded and Log Revenus per Worker, since they vary only at the firm level.

documented in Table 5 in a simple variance decomposition.[20] The firm fixed effects (β^f in equation 3) incorporate unobserved systematic differences in earnings across firms. We express these as a function of a vector of skill demands ($\overline{skill}_f$) and a residual (ν_f) in equation 4. δ is the coefficient on the residual and α is a vector of coefficients on the skills variables.

$$(4) \qquad \beta^f = \overline{skill}_f \alpha' + \delta \nu_f$$

Plugging equation 4 into 3 yields

$$(5) \qquad log(wages)_{omf} = \overline{skill}_f \alpha' + \delta \nu_f + controls + \epsilon_{omf}$$

This model provides a way to decompose the variance of log wages across firms into skill requirements and other components. In particular, the variance in the firm effects, $Var(\beta^f)$, is equal to $Var(\overline{skill}_f \alpha' + \delta \nu_f)$. 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.[21] We also use the variance-covariance matrix of the skills variables as an input. Importantly, while variation in firm

---

[20]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.

[21]For example, to understand the fraction of the variance in β^f attributed to cognitive and social skills, we calculate $Var(\overline{skill}_f \alpha' + \delta \nu_f)$ using the estimated coefficients for cognitive social and the interaction but set use a value of 0 for all other components of α and for δ.

Table 6: Decomposing Firm Effects in Wages on Skill Demands

	Log (Wages)			
	(1)	(2)	(3)	(4)
Total Standard Deviation of Firm Effect	0.190	0.190	0.084	0.025
Share attributed to skills:				
Total	11.6%	33.9%	20.9%	6.3%
Social and Cognitive Skills	11.6%	5.5%	4.7%	1.3%
Other Skills		11.3%	7.3%	0.6%
Education and Experience		17.1%	8.9%	4.5%
Residual	88.4%	66.1%	79.1%	93.7%
Additional Skills		X	X	X
Base Controls			X	
Detailed Controls				X
# Firms			85,695	

Notes: Base controls are MSA characteristics and 4-digit occupation fixed effects. Detailed controls are MSA and 6-digit occupation fixed effects. 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. Education and experience include both years required and the share of ads that have any requirement. We regress the firm fixed effect in wages on the firm fixed effect for each of the skills measures (and controls if included). We use coefficients and the variance-covariance matrix of the skills to fit the share of the variance in wages that can be attributed to various components (by fitting variances with the other coefficients set to 0).

fixed effects in both wages and skill requirements is in part driven by measurement error, the correlations in the vector α are unlikely to include such measurement error.

We estimate equations 3-5 on a disaggregated MSA-occupation-firm-level dataset. Results are presented in 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 are collinear with firm fixed effects β^f) and column 4 ads detailed controls.

From column 1, the total standard deviation of the firm wage effect without controls is 0.19 (the same number as in column 1, row 1 of table 5). 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. In Column 2, we also allow the remaining skill measures to account for firm-level variation in wages. When we do, we find that the share attributed to cognitive and social skills falls in half to 5.5 percent because these two skills are correlated with other skill measures. Education and experience requirements account for the largest fraction, 17 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 base controls (column 3), the firm wage gap narrows by about half (as we saw in table 5). 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 accounting for

Table 7: Decomposing Firm Performance Outcomes on Skill Demands

	Publicly Traded		Log Revenue per Worker	
	(1)	(2)	(3)	(4)
Total Standard Deviation of Firm Effect	0.459	0.459	0.685	0.685
Share attributed to skills:				
Total	7.2%	13.2%	14.8%	21.4%
Social and Cognitive Skills	7.2%	1.7%	14.8%	9.4%
Other Skills		3.8%		3.1%
Education and Experience		7.7%		8.9%
Residual	92.8%	86.8%	85.2%	78.6%
Additional Skills		X		X
# Firms	85,695		3,622	

Notes: See table 6.

this variation, with nearly half accounted for by education and experience requirements, and a third by the other eight skill measures. In both columns 2 and 3, we find that the set of 10 skills account for roughly 15 percent of the variation in firm pay, while cognitive and social skills account for 5 percent.

The fourth column includes the detailed controls. As noted above, controls for MSA and 6-digit occupation fixed effects absorb almost all of the variation in firm pay. This is because firm variation in our pay measure comes only from variation in the MSA-occupations firms post in. As such, this exercise naturally has limitations. However, the results in table 6 show that skills measures are important in accounting for firm-level variation in these aggregated wages. This suggests they might also be important for job-level wages.

We can also do a form of this decomposition for the firm performance measures. These measures vary only at the firm level so we cannot include controls. We instead estimate an analogous version of equations 3-5 with no controls or residuals. We can still report the fraction of the variation in performance across firms that can be attributed to variation in skill requirements.

We find that the skill measures account for a modest fraction of the variation across firms in performance. The 10 skills account for 5 percent of the variation in the probability of being publicly traded (Column 2), and roughly 12 percent of the variation in log revenue per worker. For the latter, the bulk of the explanatory power comes from cognitive and social skill requirements. Education and experience requirements contribute substantial explanatory power as well. Overall, we can account for 13 and 20 percent of the variation in whether a firm is publicly traded and log revenue per worker, respectively, with the full set of 10 skills and the more standard education and experience requirements.

This exercise is only illustrative since we do not have ad-level variation in wages or productivity. 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, we find

that firm-level variation in demand for the job skills explains substantial fractions of firm-level variation in these outcomes. Cognitive and social skill requirements alone account for about 5 percent of the variation in log(wages) and 10 percent of the variation in log(revenue per worker). These results suggests the job skill measures we introduce in this paper provide useful information about the variation in skill utilization across firms and why they have diverse outcomes.

5   Conclusion

In this paper we use data from online job vacancies posted 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 10 distinct "job skills" such as cognitive skills, social skills, management and finance. We show that the prevalence of advertisements for high-level 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 these 10 skill requirements can explain about 12 percent of the variance in wages across firms (as measured by average wages in the MSA-occupations they tend to post in), after residualizing on occupation fixed effects and MSA characteristics. These skills account for between 5 and 12 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. Moreover, 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 9 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 (Barth, Bryson, Davis and Freeman (2016), Card Heining and Kline (2016), and Song et al. (2015)). Our results show that some of the variation in pay across firms in the cross-section can be accounted for by variation in demand for cognitive and social skills. We cannot measure production

technologies directly in the BG data, nor do we have variation in skill demands over a long time horizon. However, by linking variation in pay in the cross-section to tasks that are most associated with rising inequality over time, we hope to provide suggestive evidence on this front. 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. (2014), *Polanyi's paradox and the shape of employment growth*, National Bureau of Economic Research.

[4] 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.

[5] 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.

[6] Autor, David H., Frank Levy, and Richard J. Murnane (2003), "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 118(4): 1279-1333.

[7] 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.

[8] 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.

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

[10] Borghans, Lex, Bas ter Weel, and Bruce A. Weinberg (2014), "People Skills and the Labor-Market Outcomes of Underrepresented Groups," *Industrial and Labor Relations Review*, 67(2): 287-334.

[11] Brynjolfsson, E. and McAfee, A. (2014) The second machine age: Work, progress and prosperity in a time of brilliant technologies, W. W. Norton & Company, New York, NY.

[12] 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.

[13] 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.

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

[15] 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.

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

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

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

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

[20] 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.

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

[22] Hershbein, Brad and Lisa B. Kahn (2016), "Do Recessions Accelerate Routine-Biased Technological Change," NBER Working Paper #22762.

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

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

[25] Lu, Qian (2015), "The End of Polarization? Technological Change and Employment in the U.S. Labor Market", mimeo.

[26] Marinescu, Ioana (2017), "The General Equilibrium Effects of Unemployment Insurance: Evidence from a Large Online Job Board," *Journal of Labor Economics*, forthcoming.

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

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

[29] 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.

[30] Sasser Modestino, Alicia, Daniel Shoag and Joshua Ballance (2016), "Upskilling: Do Employers Demand Greater Skill When Workers Are Plentiful?" mimeo.

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

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

[33] 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: Full Sample				Panel B: Firm Sample				Panel C: Compustat-Matched			
	Mean	St Dev	Min	Max	Mean	St Dev	Min	Max	Mean	St Dev	Min	Max
Share of Ads with Skill Requirement:												
Cognitive	0.370	0.170	0	1	0.418	0.203	0	1	0.487	0.173	0	1
Social	0.363	0.120	0	1	0.422	0.201	0	1	0.501	0.172	0	1
Cognitve and Social	0.206	0.099	0	1	0.249	0.162	0	1	0.311	0.152	0	1
Character	0.299	0.123	0	1	0.344	0.188	0	1	0.411	0.154	0	1
Writing	0.200	0.095	0	1	0.229	0.154	0	1	0.258	0.132	0	0.99
Customer Service	0.197	0.142	0	1	0.236	0.180	0	1	0.260	0.154	0	1
Project Management	0.117	0.122	0	1	0.130	0.106	0	1	0.162	0.094	0	0.81
People Management	0.151	0.095	0	1	0.184	0.125	0	1	0.212	0.110	0	0.92
Financial	0.157	0.192	0	1	0.165	0.141	0	1	0.196	0.120	0	0.99
Computer (general)	0.288	0.135	0	1	0.316	0.185	0	1	0.370	0.155	0	1
Software (specific)	0.319	0.279	0	1	0.316	0.244	0	1	0.354	0.205	0	1
Education	0.588	0.161	0	1	0.659	0.218	0	1	0.721	0.168	0	1
Experience	0.578	0.131	0	1	0.623	0.201	0	1	0.671	0.170	0	1
Years School, conditional	15.74	1.35	12	21	15.66	0.92	12	21	15.59	0.75	12	19.22
Years Experience, conditional	4.10	1.31	0.08	15	4.14	1.49	0.08	15	4.57	1.39	0.31	10.80
MSA-Occ median wage	41.97	15.60	8.15	135.00	42.21	7.26	11.96	107.95	43.74	6.17	17.71	77.06
Publicly Traded	NA				0.303	0.459	0	1	1.000	0.000	1.000	1.000
Revenue per Worker ($ millions)					0.473	3.329	0.000	562	0.473	3.329	0.000	562
# Occupations Posted in					98.22	59.45	1	246	133.95	50.08	2	228
# MSAs Posted in					106.19	101.75	1	370	172.82	104.13	1	354
# MSA-Occ cells	56,611				55,241				45,154			
# Ads per MSA-Occ cell (unweighted)	793	4911	1	351149	467	2862	1	177690	173	1063	1	59501
# Firm cells					85,695				3,622			
# Ads per Firm cell (unweighted)					301	246	1	254518	2138	8519	5	162640
# Ads, total	44,891,978				25,803,549				7,743,967			

Notes: Panel A includes professional BG ads from 2010-2015. Panel B 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. Panel C restricts to firms that can be matched to Compustat. Summary statistics are weighted by number of ads, unless otherwise noted. Skill requirements come from BG data; see table 1 for description. Median wage (from the OES) and Revenue per Worker (from Compustat) are the average from 2010-2015 in the MSA-occupation or firm cell, respectively.

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	0.46	0.33	0.35	0.22	0.08	0.07	0.15	0.84	0.44	0.31
Computer Programmers	0.42	0.33	0.21	0.21	0.13	0.12	0.11	0.05	0.27	0.82
Computer User Support Specialists	0.37	0.38	0.31	0.20	0.33	0.07	0.09	0.04	0.57	0.45
Computer and Info. Systems Analysts	0.47	0.48	0.42	0.23	0.20	0.34	0.35	0.27	0.30	0.56
Financial Analysts	0.88	0.45	0.44	0.22	0.09	0.09	0.08	0.84	0.58	0.37
Managers, Financial	0.51	0.42	0.39	0.21	0.23	0.10	0.32	0.66	0.35	0.21
Managers, General Operations	0.40	0.43	0.42	0.21	0.21	0.12	0.38	0.32	0.30	0.13
Managers, Sales	0.33	0.49	0.39	0.19	0.70	0.07	0.25	0.18	0.30	0.11
Lawyers	0.31	0.29	0.21	0.25	0.07	0.04	0.09	0.12	0.12	0.05
Loan Officers	0.32	0.42	0.36	0.30	0.41	0.01	0.10	0.28	0.38	0.06
Management Analysts	0.83	0.47	0.41	0.30	0.14	0.27	0.16	0.24	0.38	0.40
Registered Nurses	0.19	0.25	0.15	0.07	0.38	0.00	0.11	0.02	0.11	0.03
Overall Professional Occupation Average	0.37	0.36	0.30	0.20	0.20	0.12	0.15	0.16	0.29	0.32

Notes: Postings-weighted averages based on our full sample of Burning Glass professional ads. See table 1 for job skills definitions.

Table A3: Average Wages and Skill Requirements – Firm Sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Log(mean wages) in MSA-Occupations					
Cognitive	0.123***	-0.299***	0.164***	0.150***	0.0471***	0.0286***
	(0.00886)	(0.0155)	(0.00749)	(0.0125)	(0.00704)	(0.00994)
Social	0.300***	-0.0788***	0.144***	0.132***	0.0333***	0.0167*
	(0.0127)	(0.0170)	(0.00995)	(0.0135)	(0.00749)	(0.00980)
Both required		0.945***		0.0288		0.0402***
		(0.0286)		(0.0220)		(0.0153)
Years of Education	0.125***	0.123***	0.0580***	0.0580***	0.00486***	0.00490***
	(0.000795)	(0.000790)	(0.000840)	(0.000841)	(0.000836)	(0.000836)
Years Experience	0.165***	0.167***	0.0620***	0.0620***	0.0208***	0.0208***
	(0.00114)	(0.00113)	(0.00120)	(0.00120)	(0.000957)	(0.000957)
Base controls			x	x		
Detailed Controls					X	X
F-stat (cog and social)	450.8	669.5	418.2	279.4	41.01	29.63
F-stat (all 10 skills)	2073	2218	406.8	369.9	30.45	28.31
MSA-Occ Cells	55,241	55,241	55,241	55,241	55,241	55,241
R-squared	0.696	0.701	0.856	0.856	0.943	0.943

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

Notes: See tables 1 and 3. Observations are from the firm sample.