NBER WORKING PAPER SERIES

OPENING THE BLACK BOX OF THE MATCHING FUNCTION: THE POWER OF WORDS

Ioana Marinescu
Ronald Wolthoff

Working Paper 22508
http://www.nber.org/papers/w22508

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016

This paper has benefited from feedback during presentations at CREST-Paris, the University of Chicago, the Annual Meeting of the Society of Economic Dynamics, the University of Toronto, the University of Wisconsin-Madison, UC Santa Barbara, the Tinbergen Institute, the University of British Columbia, the Bureau of Labor Statistics, the NBER Macro Perspectives Summer Institute. We are grateful for comments from Briana Chang, Kerwin Charles, Francis Kramarz, Peter Kuhn, Alan Manning, Craig Riddell, and Chris Taber. We thank Chaoran Chen and Sanggi Kim for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w22508.ack

Opening the Black Box of the Matching Function: the Power of Words
Ioana Marinescu and Ronald Wolthoff
NBER Working Paper No. 22508
August 2016
JEL No. J31,J63,J64

ABSTRACT

How do employers attract the right workers? How important are posted wages vs. other job characteristics?Using data from the leading job board CareerBuilder.com, we show that most vacancies do not post wages, and, for those that do, job titles explain more than 90% of the wage variance. Job titles also explain more than 80% of the across-vacancies variance in the education and experience of applicants. Finally, failing to control for job titles leads to a spurious negative elasticity of labor supply. Thus, our results uncover the previously undocumented power of words in the job matching process.

Ioana Marinescu
Harris School of Public Policy
University of Chicago
1155 East 60th Street
Chicago, IL 60637
and NBER
ioanamarinescu@uchicago.edu

Ronald Wolthoff
University of Toronto
ronald.p.wolthoff@gmail.com

1 Introduction

A blossoming research agenda in macroeconomics analyzes business cycle fluctuations and public policy through the lens of the matching function (Petrongolo & Pissarides, 2001), which converts a given number of vacancies and unemployed workers into a resulting number of hires. This approach has been fruitful for improving our understanding of macro issues (Yashiv, 2007). Yet, there is limited empirical evidence on the micro foundations of the matching function, even though theory shows that these foundations are important for aggregate outcomes including unemployment and efficiency (Rogerson et al., 2005). Outstanding questions include the source of search frictions, the role of heterogeneity, and the strategies used by workers and firms during the matching process. Answering these questions requires opening the black box of the matching function.

Early empirical work on the matching process, e.g. by Barron et al. (1985, 1987), focused on its later stages, i.e. the screening and interviewing of applicants. This early work recognizes that workers are heterogeneous and that different firms may employ different strategies to find the best possible match. Of course, a firm's pool of applicants is not exogenous and can be influenced by the way in which the firm advertises its position, which raises the question of how employers advertise the nature of their jobs, and how workers react to these signals. The theoretical search and matching literature explores these questions extensively. In particular, directed search models assign a key role to posted wages as an instrument for firms to attract the right pool of applicants (see e.g. Moen, 1997; Shimer, 2005a; Eeckhout & Kircher, 2010).

Yet, empirical evidence about employers' search strategies is limited. Holzer et al. (1991) show that the wages that firms pay affect the number of applicants that they attract, but the data in Holzer et al. (1991) does not capture whether firms communicated wages to potential applicants. Recent evidence from employment websites indicates that, typically, only a small fraction of job ads specifies a wage (see e.g. Kuhn & Shen, 2012; Brencic, 2012; Brown & Matsa, 2014). This raises the question of what other strategies firms use to attract applicants and how successful these strategies are.

In this paper, we answer this question by using a new data set from CareerBuilder.com, a leading online job board. CareerBuilder.com contains about a third of all US vacancies. Furthermore, the jobs and job seekers found on CareerBuilder.com are fairly representative of the US labor market. We use a data set of all the vacancies posted in Chicago and Washington, DC at the beginning of 2011. For each vacancy, we observe the information that firms provide in their job ads. We also have information on the pool of applicants that each job ad attracts, in particular the number of applicants and their education and work experience.

Using this data, we study how firms advertise their positions and how this affects the number and quality of applicants that they attract. We find that a job characteristic typically not considered in the economics literature plays a critical role. This piece of information is the *job title* of the vacant position as chosen by the employer, e.g. "senior accountant" or "network administrator". We demonstrate that the words in the job title communicate important information about a position, and that workers use them to direct their search. Even among the minority of vacancies that post a wage, job titles play a far more important role than wages in explaining the number and quality of applicants that a vacancy attracts. Thus, our results uncover the previously undocumented power of words in the search and matching process. We now describe our results in more detail.

First, we analyze employers' advertising behavior. We show that all job ads specify a job title, which is prominently featured on CareerBuilder's website, while only a subset of the ads clearly specifies other characteristics a worker may care about, such as a wage (20%). A natural interpretation of this fact is that listing these other characteristics would provide little information in addition to the job title. To support this interpretation, we show that even when firms post wages, job titles explain more than 90% of the wage variance. By contrast, six-digit SOC codes, the most detailed occupational classification commonly used by economists, can only explain a third of the wage variance. Thus, employers advertise their jobs using the power of words embodied in the job title.

Second, we show that job titles play an important role in aiding workers' search decisions. Indeed, job titles explain more than 80% of the across-vacancies variation in the average education and experience of applicants. This shows that workers use job titles to direct their search. If a job ad specifies a wage, workers use this information as well. Everything else equal, a higher wage increases the number of applicants as well as their education and experience levels. Strikingly, when job titles are *not* controlled for, higher-wage vacancies attract significantly *fewer* applicants. This negative relationship between wages and the number of applicants is counterintuitive because it implies a negative elasticity of labor supply. Yet, this negative relationship holds even when controlling for detailed occupations (SOC). It is only when we control for job titles that we find that a 10% increase in wages is associated with a 7% *increase* in the number applicants. Job titles are thus crucial to explain the education and experience levels of applicants, and it is necessary to control for job titles to recover a positive elasticity of applications with respect to the posted wage.

Third, we analyze the reasons why job titles explain posted wages so much better than existing occupational classifications (SOC codes). One might think that the larger explanatory power of job titles is mechanical, because there are many more job titles than SOC codes (several thousands vs. 840). However, job titles do contain important information that is not captured by SOC. Indeed, the adjusted R-squared in regressions with job titles is almost as high as the R-squared, showing that job titles have explanatory power above and beyond just adding further degrees of freedom. We show that, relative to the detailed SOC occupations, job titles better reflect the hierarchy, level of experience, and specialization of different jobs. For example, job titles separate accountants into "staff accountants," "senior accountants," and "directors of accounting", with the last two earning more than the first. Likewise, job titles separate sales representatives into lower-wage "inside sales" and higher-wage "outside sales," and information technology administrators are divided between lower-wage "network administrators" and higher-wage "systems administrators." Thus, words in the job title are powerful predictors of the posted wage.

Our overarching conclusion is that words in job titles play a fundamental role in the initial stages of the search and matching process and are key to understanding labor market outcomes. We make three main contributions to the literature. Our first contribution is to document the role of job titles in employer search. Our second contribution is to verify many of the predictions of directed search models, and to show how specifically workers direct their search. Finally, our third contribution is to demonstrate the power of words in explaining the variance in posted wages.

First, although job titles have been used to analyze career paths and promotions within firms (see e.g. Lazear, 1992), we are – to the best of our knowledge – the first to analyze their role in the search and matching process. By showing that firms advertise jobs primarily through job titles rather than wages, we add to two areas of research: the employer search literature analyzing firms' strategies for finding the right match (Barron et al., 1985, 1987) as well as the emerging empirical literature on online job search (see e.g. Kuhn & Shen, 2012; Brenčič & Norris, 2012; Pallais, 2012; Faberman & Kudlyak, 2014; Marinescu, 2014; Gee, 2014).

Second, our finding that job titles and (if present) posted wages explain the number and type of applicants that a firm attracts validates directed search models as realistic models of the labor market. Indeed, in directed search models, employers compete for workers by posting job characteristics, and workers direct their search to their preferred jobs. Thanks to the competition among employers, directed search models typically lead to efficient outcomes (see Rogerson et al. (2005) for a review of the literature). In the basic directed search model, a firm that posts a higher wage attracts more applicants (see e.g. Moen, 1997), while in a setting with heterogeneous workers, higher wages may also attract better applicants (see e.g. Shi, 2001). Prior literature did not find strong evidence of this expected positive relationship between wages and the number of applicants (Holzer et al., 1991; Faberman & Menzio, 2010; Bó et al., 2012).

In contrast, we find a positive relationship, but *only within a job title*. Therefore, controlling for job titles is crucial in validating the prediction of directed search models that higher wages attract more applicants. We also document a positive relationship between wages and the quality of the applicant pool, consistent with evidence from the public sector in Mexico (Bó et al., 2012). Even when we include the majority of jobs that do not post wages, we find that job titles explain more than 80% of the variation in the quality of applicants that a vacancy attracts. This shows that wages are not necessary for workers to direct their search, as illustrated by e.g. the theoretical model of Menzio (2007). Thus, job titles play a key role in explaining how workers direct their search as well as the role of wages in directed search.[1]

Third, we add to the literature on the causes and consequences of the wage variance. On the empirical side, the literature that decomposes the wage variance (e.g. Abowd et al., 1999; Woodcock, 2007; Abowd et al., 2002; Andrews et al., 2008; Iranzo et al., 2008; Woodcock, 2008) focuses on realized wages. It finds that unobserved characteristics captured by worker and firm fixed effects together explain most of the variance in realized wages (see e.g. Table 4 in Woodcock, 2007). We show that job titles explain as much of the variance in posted wages as worker and firm fixed effects explain in realized wages. This suggests that *observable* job characteristics play an important role in explaining the wage variance. On the theoretical side, a recent paper (Hornstein et al., 2011) argues that a large residual wage variance is inconsistent with basic search models (i.e. without on-the-job search). We show that the wage variance is strongly overestimated if one does not control for job titles. Posted wages, while having the same cross-sectional variance as realized wages in the Current Population Survey, are almost fully determined by job titles. This result suggests that job seekers face limited wage dispersion for given job characteristics and can help explain why unemployed job seekers accept the majority of job offers (Krueger & Mueller, 2014; Wolthoff, 2015).

Our results have broader theoretical and empirical implications. On the theoretical side, our findings speak to the importance of heterogeneity as captured by job titles for understanding the labor market: heterogeneity may be an important source of frictions that the standard matching function does not explicitly model. If workers and firms are very different, it becomes important to find just the right match, and this can explain that it takes time and effort to locate the right trading partner. Some theoretical models have explored the role of heterogeneity in workers and firms: examples of search models with two-sided heterogeneity include Shimer & Smith (2000), Shimer (2005a) and Eeckhout & Kircher (2010). In a related working paper (Marinescu

[1] Our results are also related to the literature on the elasticity of labor supply to the individual firm (see Manning, 2011, for a review). While this literature examines how changes in firm-level employment relate to wage levels, we analyze how wages influence the number of job applications a firm receives, an important first factor in the process that leads to final matches.

& Wolthoff, 2012), we present an equilibrium search model with two-dimensional heterogeneity that rationalizes some of the stylized facts of this paper. Overall, our results suggest that incorporating heterogeneity in both workers and jobs is a fruitful avenue for search-theoretical research on macroeconomic issues.

On the empirical side, we introduce a new occupational classification which is based on employers' own description of their jobs rather than researchers' interpretation. We show that this new classification improves in important ways on existing occupation classifications (SOC) and has important implications for how we understand labor markets. We expect that this classification will prove to be a useful research tool in a wide variety of contexts. For example, it may help to shed more light on the gender and race wage gap (Blau, 1977; Groshen, 1991; Blau & Kahn, 2000), inter-industry wage differentials (Dickens & Katz, 1986; Krueger & Summers, 1986, 1988; Murphy & Topel, 1987; Gibbons & Katz, 1992), the specificity of human capital (Poletaev & Robinson, 2008; Kambourov & Manovskii, 2009), and occupational mobility and worker sorting (Groes et al., 2015).

This paper proceeds as follows. Section 2 describes the job board and the data set. We analyze the job ads that firms post in section 3 and the applicant pool that a job ad attracts in section 4. Section 5 provides robustness checks and discussion, after which section 6 concludes.

2 Data

Background. We use proprietary data provided by CareerBuilder.com, which is the largest online job board in the United States, visited by approximately 11 million unique job seekers during January 2011.[2] While job seekers can use the site for free, CareerBuilder charges firms several hundred dollars to post a job ad on the website for one position for one month. A firm that wishes to keep the ad online for another month is subject to the same fee, while a firm that wishes to advertise multiple positions needs to pay for each position separately, although small quantity discounts are available (see CareerBuilder, 2013). At each moment in time, the CareerBuilder website contains over 1 million jobs.

Search Process. A firm posting a job is asked to provide various pieces of information. First, it needs to specify a job title, e.g. "senior accountant," which will appear at the top of the job posting as well as in the search results. CareerBuilder encourages the firm to use simple, recognizable job titles and avoid abbreviations, but firms are free to choose any job title they

[2] See comScore Media Metrix (2011). Monster.com is similar in size, and whether Monster or CareerBuilder is larger depends on the exact measure used.

desire. Further, the firm provides the full text of the job ad, a job category and industry, and the geographical location of the position. Finally, the firm can specify education and experience requirements as well as the salary that it is willing to pay.

A job seeker who visits CareerBuilder.com sees a web form which allows him to specify some keywords (typically the job title), a location, and a category (broad type of job selected from a drop-down menu). After providing this information[3], the job seeker is presented a list of vacancies matching his query, organized into 25 results per page. For the jobs that appear in the list, the job seeker can see the job title, salary, location, and the name of the firm. To get more details about a job, the worker must click on the job snippet in the list, which brings him to a page with the full text of the job ad as well as a "job snapshot" summarizing the job's key characteristics. At the top and bottom of each job ad, a large "Apply Now" button is present, which brings the worker to a page where he can send his resume and his cover letter to the employer.

Sample. Our data set consists of vacancies posted on CareerBuilder in the Chicago and Washington, DC Designated Market Areas (DMA) between January and March 2011. A DMA is a geographical region set up by the A.C. Nielsen Company that consists of all the counties that make up a city's television viewing area. DMAs are slightly larger in size than Metropolitan Statistical Areas and they include rural zones. We observe all vacancies posted in these two locations during January and February 2011 and we observe a random subsample of the vacancies posted in March 2011.

Variables. The CareerBuilder data is an attractive source of information compared to existing data sets, in particular due to the large number of variables that it includes. For each vacancy, we observe the following *job characteristics*: the job title, the salary (if specified), whether the salary is hourly or annual, the education level required for the position, the experience level required for the position, an occupation code, and the number of days the vacancy has been posted. The occupation code is the detailed, six-digit O*NET-SOC (Standard Occupational Classification) code.[4] CareerBuilder assigns this code based on the full content of the job ad using O*NET-SOC AutoCoder, the publicly available tool endorsed by the Bureau of Labor Statistics.[5] This procedure is consistent with the approach of commonly used labor market

[3] It is not necessary to provide information in all three fields.

[4] See `http://www.onetcenter.org/taxonomy.html`. We henceforth refer to this classification simply as SOC.

[5] See `http://www.onetsocautocoder.com/plus/onetmatch`.

surveys, such as the Current Population Survey (CPS).[6] We further observe the following *firm characteristics*: the name of the firm, an industry code, and the total number of employees in the firm. CareerBuilder uses external data sets, such as Dun & Bradstreet, to match the two-digit NAICS (North American Industry Classification System) industry code and the number of employees of the firm to the data.

In addition to these characteristics, we also observe several *outcome variables* for each vacancy. Our first outcome variable, the number of views, represents the number of times that a job appeared in a listing after a search. The second outcome variable, the number of clicks, is the number of times that a job seeker clicked on the snippet to see the entire job ad. Finally, we observe the number of applications to each job, where an application is defined as a person clicking on the "Apply Now" button in the job ad.

From these numbers, we construct two new variables that reflect applicant behavior: the number of applications per 100 views, and the number of clicks per 100 views. These measures correct for heterogeneity in the number of times a job appears in a listing, allowing us to analyze applicants' choices among known options.

For a random subset of the vacancies of January and March 2011, we also observe some *applicant characteristics*. Specifically, we observe the number of applicants broken down by education level (if at least an associate degree) and by general work experience (in bins of 5 years). We will use these job seeker characteristics as proxies for worker productivity to analyze the quality of the applicant pool that a firm attracts.

Cleaning. We express all salaries in yearly amounts, assuming a full-time work schedule. When a salary range is provided, we use the middle of the interval for most of the analysis, but we perform robustness checks in section 5. To reduce the impact of outliers and errors, we clean the wage data by removing the bottom and top 0.5% of the salaries.

Because job titles are free-form, many unique ones exist and the frequency distribution is highly skewed to the right. To improve consistency, we cleaned the data. Most importantly, we formatted every title in lower case and removed any punctuation signs, employer names, or job locations. In most of our analysis, we restrict attention to the first four words of a job title. As we will discuss, because this assumption has minimal impact on the number of unique job titles in our sample, our results are not sensitive to it.

Representativeness. Some background work (data not shown) was done to compare the job vacancies on CareerBuilder.com with data on job vacancies in the representative JOLTS (Job

[6] This means that misclassification is unlikely to be a larger problem in the CareerBuilder data than in the CPS. See Mellow & Sider (1983) for an analysis of inconsistencies in occupational codes in the CPS.

Openings and Labor Turnover Survey). The number of vacancies on CareerBuilder.com represents 35% of the total number of vacancies in the US in January 2011 as counted in JOLTS. Compared to the distribution of vacancies across industries in JOLTS, some industries are overrepresented in the CareerBuilder data, in particular information technology; finance and insurance; and real estate, rental, and leasing. The most underrepresented industries are state and local government, accommodation and food services, other services, and construction.

While CareerBuilder data is not representative by industry, in most other respects it is representative of the US labor market, as documented by Marinescu & Rathelot (2014). Using a representative sample of vacancies and job seekers from CareerBuilder.com in 2012, they show that the distribution of vacancies across occupations is essentially identical (correlation of 0.95) to the distribution of vacancies across all jobs on the Internet as captured by the Help Wanted Online data. Furthermore, the distribution of unemployed job seekers on CareerBuilder.com across occupations is similar to that of the nationally representative Current Population Survey (correlation of more than 0.7). Hence, the vacancies and job seekers on CareerBuilder.com are broadly representative of the US economy as a whole, and they form a substantial fraction of the market.

Descriptive Statistics. Table 1 shows the summary statistics for our sample. The full sample consists of more than 60,000 job openings by 4,797 different firms. On average, each job was online for 16 days, during which it was viewed as a part of a search result 6,084 times, received 281 clicks, and garnered 59 applications. Per 100 views, the average job receives almost six clicks and approximately one application.[7]

Only a minority of job ads include an explicit experience requirement (0.3%) or an explicit education requirement (42%). When specified, these requirements appear in the "job snapshot" box at the end of the full job ad, but they do not appear in the job snippet that job seekers first see in the search results. Therefore, employers may choose not to fill in education and experience requirements if they feel that the overall job description is sufficiently informative.

We observe a posted wage for approximately 20% of the jobs. When present, the wage does appear in the job snippet as part of the search results. The average posted yearly salary is just over $57,000, and we will show in more detail below that posted wages on the website have the same distribution as the wages of full-time workers in the Current Population Survey. In the smaller sample with posted wages, there are 1,369 unique firms. Finally, we observe the average applicant quality for approximately 2,300 positions. The average applicant has between 16 and 17 years of education (conditional on holding at least an associate degree) and just over

[7] Keep in mind that the average of ratios does not necessarily need to equal the ratio of averages.

13 years of work experience.

3 Job Ads

We start our empirical analysis by studying firms' decisions regarding the job ads that they post. We provide evidence that job titles are the key way in which firms communicate the characteristics of their vacancies (including the wage).

3.1 Job Titles

All job ads in our sample specify a job title. This job title is prominently featured on the employment website and is the main piece of information that workers use to search the CareerBuilder database. This suggests that job titles play a key role in the search and matching process. Since an analysis of job titles is—to the best of our knowledge—novel in the literature, we start by providing some descriptives.

The full sample of more than 60,000 job openings contains 22,009 unique job titles. Truncation to the first four words marginally reduces this number to 20,447. In the subsample of jobs with posted wages, the corresponding numbers are 4,669 and 4,553, respectively. In Table 2, we list the ten most common job titles (after truncation), both for the full sample and for the subsample of jobs that post wages. Note that the most common job titles are typically at most three words long. We also show the most common job titles if we truncate the job title to the first two words or the first word. Figure 1 provides a more comprehensive overview of frequent job titles in the form of a word cloud, in which the size of a job title depends on its frequency.

Inspection of the table and the figures reveals that job titles often describe occupations, e.g. "administrative assistant," "customer service representative," or "senior accountant." This raises the question of how job titles compare to other definitions of occupations, in particular the occupational classification (SOC) of the Bureau of Labor Statistics. Since our data includes SOC codes, we can explicitly make this comparison. Perhaps the most obvious difference between job titles and SOC codes concerns their variety: in our full sample, the number of unique job titles is more than 25 times the number of unique SOC codes.[8] In other words, job titles provide a finer classification. For example, they distinguish between "inside sales representative" and "outside sales representative," between "executive assistant" and "administrative assistant," and between "senior accountant" and "staff accoutant"—while each of the two jobs in these pairs

[8]20,447 versus 762, to be exact. In the subsample with posted wages, the difference is smaller, but still a factor of eight (4,553 versus 594). Note that the SOC classification distinguishes 840 occupations in total, some of which do not appear in our data.

would have the same SOC code. While some of the larger variety in job titles might be due to noise in employers' word choice, we will show in the following sections that distinctions between job titles are economically significant.

3.2 Wages

While all job ads contain a job title, only 20% of vacancies in our sample post a wage. We interpret this as a sign that a job title provides sufficiently precise information on compensation such that, in equilibrium, firms consider it unnecessary to also post a wage.[9] To provide evidence for this idea, we first analyze firms' decision to post a wage or not, and then firms' decision of how much to offer, conditional on posting a wage.

3.2.1 Wage Posting

Table 3 displays our results regarding whether firms post a wage or not. Using a linear probability model, we find that both job titles (column I) and firm fixed effects (column II) have high explanatory power for the decision to post a wage: they each explain around 70% of the variance in wage-posting behavior. Including both simultaneously essentially explains all of the variation in job posting behavior (the R^2 is 0.93 in column III), while including additional job characteristics does not substantially improve the model fit (column IV). Finally, in column V, the left-hand side variable is the firm fixed effect estimated in column I. We find that job titles explain most of the firm fixed effects. That is, the types of jobs that firms advertise, as embodied in the job titles in their ads, account for systematic differences in wage-posting behavior across firms.

3.2.2 Wage Variance

To determine which observable characteristics are most important in explaining posted wages, we regress log wages on various sets of controls. The results of this exercise are presented in Table 4. Following most of the literature on realized wages, we start by regressing the posted wages on firm fixed effects. This explains about 50% of the variance (column I). As we demonstrate in column II, job titles do a better job: fixed effects for (the first four words of) the job

[9]An additional explanation might be the fact that some companies use Applicant Tracking Systems (ATS) software that keeps track of job postings and applications. This software typically removes the wage information by default before sending out the job posting to online job boards such as CareerBuilder (private communication with CareerBuilder.com). However, this is unlikely the full story. First, the fact that most job ads do not advertise wages is consistent with the worker survey data of Hall & Krueger (2012) and evidence from job boards in other countries where ATS may be less common (Brencic, 2012; Kuhn & Shen, 2012). Second, if firms really cared about advertising wages, it would be suboptimal for them to rely on software that removes this information.

titles explain more than 90% of the wage variance. This result is not due to the fact that there are many unique job titles. To demonstrate this, we consider two measures of fit, the adjusted R^2 and the Akaike Information Criterion (AIC), which, unlike R^2, account for the number of controls. As table 4 shows, both measures favor the specification with job title fixed effects over the one with firm fixed effects. We also perform a permutation test in which we re-estimate the specification with job title fixed effects (column II) 1000 times with randomly re-assigned wages. The average adjusted R^2 is 0 in this case, confirming that our results are not simply the result of the large number of job titles.[10]

Specifications that include both firm and job title fixed effects as well as other job characteristics can explain nearly all of the variation in posted wages (column III and IV). Hence, firm fixed effects explain most of the wage variance that remains after controlling for job titles, implying that within job title variation in posted wages can largely be explained by firm fixed effects. However, since job titles already explain 90%, the additional explanatory power compared to job titles alone is relatively small. This suggests that job titles are the most important determinant of differences in posted wages across vacancies, and that job titles might explain the differences in posted wages across firms as well. To test this latter hypothesis directly, we regress the firm fixed effects estimated in column I on job title fixed effects. We find that job titles can explain about 90% of the variance in firm fixed effects (column V). Hence, we conclude that most of the systematic differences in posted wages across firms can be explained by firms posting jobs with different job titles.

One might wonder whether the fact that some firms post a wage range, rather than a single wage, affects this conclusion. This is not the case. First, as we show in section 5, our results are robust to alternative ways of dealing with wage ranges. Second, a natural interpretation of wage ranges is that firms plan to condition the wage that they eventually pay on the type of the worker that they hire. This means that the range does not actually contribute to the wage dispersion that workers of a particular type face.[11]

Another way of assessing the degree of wage variance within a job title is to consider the mean-min ratio, introduced by Hornstein et al. (2011). They show that calibration of basic search models to the US labor market implies a gap between the mean and the lowest wage in the labor market of at most 5%, corresponding to a mean-min ratio of less than 1.05. If the degree of wage dispersion were larger, the model would imply that workers should not accept a new job as quickly as is observed in the data, but should search longer to find a higher-paying

[10]These results are available upon request.

[11]See Shimer (2005a) for an example of a model in which each firm posts a menu of wages, one for each possible type of worker. Hence, a worker of a given type who gets hired by a given firm does not face any uncertainty regarding the wage that he will receive.

job.

In the working paper version, Hornstein et al. (2007) demonstrate that the mean-min ratio in the US Census typically exceeds 2. However, since they calculate the mean-min ratio within SOC codes, our results suggest that they are overestimating the degree of frictional wage dispersion. We show in appendix B that the the mean-min ratio within a job title is much smaller than within an SOC. Indeed, it can be as low as 1.22, which means that controlling for job titles resolves more than 90% of the original discrepancy between the predictions of the calibrated model and the data. Hence, our results indicate that job seekers accept jobs relatively quickly because the degree of wage dispersion within a job title is limited. That is, as long as job seekers are aiming for a specific job title, they cannot realistically hope to find a job with well above-average pay.

Overall, our results on the determinants of posted wages as well as wage-posting behavior imply that knowing the job title is enough to determine the level of the offered wage and whether the wage will be explicitly posted. Job titles are so informative that firms may not need to post the wage for a position. Job seekers familiar with an industry should be able to infer the wage from the job title.

3.2.3 Comparison with CPS

Our conclusion that there is little residual wage variation is at odds with much of the literature, which attributes an important role to frictional wage dispersion and/or worker and firm fixed effects. However, we analyze posted (i.e., offered) wages rather than realized (i.e., accepted) wages. We now study the extent to which this difference in posted vs. realized wages explains our results. To do so, we exploit a commonly used data set with realized wages, the CPS.

Wage Distribution. To compare posted wages on CareerBuilder.com to earned wages in the labor market, we use data from the basic monthly CPS from January and February 2011. We restrict the CPS data to employed individuals in the Chicago and Washington, DC MSAs, such that the sample covers approximately the same time frame and geographic area as the CareerBuilder data.

The results of the comparison between the two data sets are displayed in Figure 2. The upper panel shows that the posted wage distribution is more compressed than the realized wage distribution. However, the CareerBuilder data does not properly distinguish full-time from part-time jobs: in particular, hourly wages are converted to full-time equivalent. Furthermore, CareerBuilder data does not account for sporadic patterns of employment that could occur for some workers in the CPS. Therefore, posted wages mostly capture full-time work. Another difference between the CPS and CareerBuilder.com is that CPS earnings are top-coded. To make the two data sets more comparable, we restrict the CPS data to workers who work full-time and whose earnings are not top-coded, and the CareerBuilder data to earnings levels that are not top-coded in the CPS. With these restrictions, the distribution of realized wages in the CPS and the distribution of posted wages in CareerBuilder are almost identical (Figure 2, lower panel). Therefore, we conclude that the distribution of posted wages on CareerBuilder.com is very close to the distribution of realized wages in the CPS for full-time workers.[12]

Effect of Occupations. To further compare the CareerBuilder data to the CPS, we investigate the effect of occupational controls in both data sets. Although we do not observe job titles in the CPS, we can control for occupations via the SOC codes. Table 5 presents three wage regressions for the CPS with increasingly finer occupation controls. We use the CPS weights for the outgoing rotation group. In column I, we regress log weekly earnings on the most aggregated classification (*major* occupations), distinguishing 11 different occupations. This explains approximately 15% of the variation in the wages. Column II and III show the specifications with 23 *minor* and 523 *detailed* occupations[13], respectively. This increases the (adjusted) R^2. The most detailed occupational classification available in the CPS explains slightly over a third of the wage variance (column III), leaving about two thirds of the wage variance unexplained.

Having shown that SOC codes can explain about a third of the variance in wages in the CPS, we now analyze to what degree SOC codes can explain the variance in *posted* wages. In columns IV, V, and VI, we use the CareerBuilder sample and run the same specifications as in columns I, II, and III. The results in terms of the explained wage variation are strikingly similar to the CPS sample: major occupations explain about 15% of the variance in posted wages and detailed occupations explain slightly over a third of the variance. This similarity between the explanatory power of occupations when using the CPS and the CareerBuilder data sets supports the idea that occupations are equally important in explaining realized wages and posted wages.

While the most detailed SOC codes available in the CPS distinguish between 523 occupations, the CareerBuilder data includes a much more detailed measure of occupations: job titles. As we have shown, job titles can explain more than 90% of the variance in posted wages (column VII). Since the explanatory power of occupations is essentially the same in the CPS

[12]Of course, the two distributions do not need to coincide. It is straightforward to specify a search model where realized wages first-order stochastically dominate posted wages. The goal here is merely to show that posted wages at CareerBuilder are similar to wages in the labor market as a whole.

[13]The CareerBuilder data uses the SOC 2000 classification while CPS uses Census occupational codes based on SOC 2010. To address this difference in classification, we converted SOC 2000 to SOC 2010 and then to Census codes. Because SOC 2010 is more detailed than the SOC 2000, a small number of Census codes had to be slightly aggregated. In Table 5, the same occupational classifications are used for both CareerBuilder and CPS data.

and CareerBuilder samples, it is possible that job titles, were they available in the CPS, would explain most of the variance in realized wages. This is an exciting area for future research.

3.3 Word Analysis

To better understand the explanatory power of job titles, in this section we estimate an additional specification for the wage variance as well as for the wage-posting decision. In this exercise, we take the residuals of each dependent variable, after controlling for detailed SOC codes, and we regress them on a set of dummy variables for each word appearing in the job title. Compared to job title fixed effects, this specification is restrictive because it ignores the order and combinations in which words appear in a job title. It assumes, for example, that the word "assistant" has the same effect every time it appears.[14] Yet, this specification allows us to determine which words are most important. Because the analysis of wage determinants is more familiar, we start with examining which words are associated with higher wages. We then move on to analyze which words are associated with a higher probability of posting a wage.

3.3.1 Wage Variance

The results of the regression of the wage residuals on the word dummies are presented in column II of Table 6. Not surprisingly, this specification explains a smaller share of the wage variation than the specification with job title fixed effects (column I). That said, it still explains about half of the within-SOC variance in posted wages.

Next we explore which words are most important in explaining wages. In Table 7, we list words that appear at least 100 times and are significant at least at the 5% level in the specification from column II in Table 6. We checked the job titles in which these words appear and classified the words into three categories. The first column includes words that signal the expected experience level of the holder of the job title. Within an SOC code, job titles that include the words "manager" or "senior" have higher than average posted wages, whereas wages are lower than average for titles that include the words "specialist" or "junior". For instance, within the SOC code 13-2011 ("Accountants and Auditors"), accounting managers and senior accountants earn more than accounting specialists and junior accountants. In the second column are words that denote specialties or skills. For example, within the SOC code 41-3099 ("Sales Representatives, Services, All Other") or 41-4012 ("Sales Representatives, Wholesale and Manufacturing, Except Technical and Scientific Products"), inside sales jobs, which require employees to contact customers by phone, pay less than outside sales jobs, where employees

[14] Examples of such jobs include "executive assistant," "assistant store manager," and "assistant professor.".

must travel and meet face-to-face with customers.[15] Finally, the third column is similar to the second column, but focuses on computer skills and specialties. For example, within SOC code 15-1071 ("Network and Computer Systems Administrators"), network administrators earn less than systems administrators.

Taken together, the frequent words listed Table 7 explain about 20% of the wage residuals after accounting for SOC (Table 6, column III). Therefore, at least 20% of the within-SOC wage variance is due to differences in seniority as well as in skills and specialties. When we regress the wage residuals on the frequent words that denote seniority from the first column of Table 7, we find that these words alone explain about 14% of the within-SOC variance in posted wages (Table 6, column IV). Overall, we find that an important reason why job titles explain wages better than SOC codes is their ability to reflect differences across jobs in management/seniority and skills/specialties.

Figure 3 provides a more complete overview of the words that are associated with either higher or lower wages. The size of the words depends on their frequency, while the intensity of the color represents the magnitude of their effect. We can classify the words in this figure roughly the same way was the frequent words from table 7. For example, "president" and "intern" indicate very different levels of seniority and have opposing effects on the wage, just as one would expect. "Scientist" and "retail" are examples of skills/specialties leading to a higher and a lower wage, respectively. Finally, in terms of computer skills, the word "linux" is associated with higher wages, while the generic term "computer" leads to lower wages.

3.3.2 Wage Posting

The words that significantly increase or decrease the probability that a job ad contains a wage are displayed in Figure 4. Unlike the figure for the wage level, this figure does not show a clear pattern. In particular, both "high-wage" words and "low-wage" words (from Figure 3) can predict a higher probability of posting a wage. For example, if we consider words indicating seniority, then both "manager" (higher wage) and "junior" (lower wage) increase the probability that a wage is present in the ad, while "chief" (higher wage) and "representative" (lower wage) decrease this probability. If we consider words indicating specialization, then both "web" (higher wage) and "retail" (lower wage) increase the probability of posting a wage, while both

[15] The most frequent word among those that are associated with higher wages in the second column is "-". This is not a typo; this character typically separates the main job title from additional details about the job. These additional details were deemed important enough for the firm to specify them in the job title. Presumably, all other things equal, a more specialized job comes with a higher pay. Some examples of the use of "-" are: "web developer - c# developer - net developer - vb net developer" or "web developer - ruby developer - php developer - ror pearl java".

"-" (higher wage) and "associate" (lower wage) decrease the probability of posting a wage.

Hence, the result of Brencic (2012) that jobs with higher skill requirements are less likely to post wages does not hold within SOC codes in our data set. This suggests that such an effect, if present, operates across SOC codes. Indeed, if we regress the wage-posting dummy on educational requirements alone, we find that jobs that require a high school degree are significantly more likely to post a wage than jobs requiring a college degree.

In summary, our results indicate that job titles capture a large amount of job heterogeneity ignored by detailed SOC codes. In particular, job titles reflect heterogeneity in seniority as well as in skills and specialization that SOC codes do not. Our word analysis also illustrates that free-form Internet text data has the power to deepen our understanding of job heterogeneity in the labor market. While text analysis has been used in other areas of economics, in particular news analysis (e.g. Gentzkow & Shapiro, 2010), we show a promising new use of such text analysis for the study of labor markets.

4 Applicants

After analyzing job ads, we now turn to workers' application behavior. In particular, we analyze which job characteristics affect the number and the average quality of the applicants that a vacancy attracts. We again find that job titles play a fundamental role.

4.1 Number of Applicants

We start with examining the relationship between log wages and the number of applications per 100 views[16]. The results of this exercise are presented in table 8.

Without controls, we find that there is a significant negative association between the wage and the number of applicants a vacancy gets: a 10% increase in the wage is associated with a 6.3% decline in applications per view (column I).[17] Adding controls such as job characteristics as well as detailed occupation fixed effects (column II) and firm fixed effects (column III) increases the R^2, but it does not change the negative relationship between wages and the number of applicants. The coefficient on the wage and its significance remain essentially unchanged as more controls are added, indicating a 5.8% decline in the number of applications in response to a 10% increase in the wage offer.

[16]We use the number of applications per 100 job views to correct for heterogeneity in the number of views across jobs. An alternative choice for the outcome variable is simply the logarithm of the number of applicants for each job. We find that our key results from Table 8 are qualitatively unaffected by this alternative outcome definition.

[17]A 10% increase in the wage decreases the number of applications per 100 views by $0.770 \log(1.1) = 0.073$, which is a 6.3% decline compared to the sample average of 1.168.

Since job titles turned out to be important in explaining the wages that firms post, it is reasonable to expect that they also play a role in workers' application decisions. Therefore, in columns IV and V, we repeat the specifications of columns II and III, respectively, while replacing the SOC code fixed effects with job title fixed effects. This allows us to estimate the relationship between wages and the number of applicants among even more homogeneous groups of jobs.

Controlling for job title fixed effects instead of SOC codes fundamentally changes the results. In particular, it reverses the negative relationship between wages and the number of applications seen in columns I, II, and III. The coefficient on the wage is now positive and significant. That is, within job title, higher wages are associated with more applicants, as one would expect on the basis of economic theory. When we include firm fixed effects to absorb even more firm-side heterogeneity (column V), the same pattern emerges, although both the coefficient and the standard error increase somewhat. The point estimate implies that a 10% increase in the wage is associated with a 7.4% increase in the number of applicants per 100 views.

Comparison and Interpretation. The adjusted R^2 and the AIC strongly favor the specifications with job titles. This suggests that the negative relationship between wages and applications found when using SOC codes is spurious. In other words, jobs with the same SOC can still differ considerably in their characteristics. These differences directly affect both the wage and the number of applicants, rendering estimates obtained without controlling for job title heterogeneity inconsistent. Admitteldy, job titles do not capture *all* heterogeneity. Yet, the fact that they reverse the sign on the wage coefficient indicates that they capture a crucial component of heterogeneity, and that they are important for understanding the relationship between wages and applicant behavior.

The fact that a positive relationship between wages and applications arises after controlling for relevant heterogeneity has another important implication. It suggests that workers are able to target their applications to jobs that are better matches. Perhaps this is not surprising given the design of CareerBuilder.com, which makes it easy for job seekers to quickly compare a large number of vacancies. Nevertheless, this result is important because it validates models of directed search (e.g., Moen, 1997; Shimer, 1996) as realistic models of labor market search, relative to canonical random search models (e.g., Pissarides, 1985; Mortensen & Pissarides, 1994). While these models have strikingly different implications for labor market outcomes, empirical evidence on either side of the debate is scant. The limited evidence that exists presents a mixed or a negative association between wages and applications (see Holzer et al., 1991; Faberman & Menzio, 2010). As these studies use specifications with controls for occupations and/or sectors but not for job titles, our analysis suggests that omitted variable bias might be driving their results.

4.2 Quality of Applicants

To examine the quality of the applicants that a vacancy attracts, we use two different measures of the quality of an applicant pool: i) the average level of work experience among applicants, and ii) the average education level among applicants, expressed in years of education.

Average Experience. Table 9 displays the results for the average experience level. We find that, across all jobs, job titles explain most of the variance in the average experience level of applicants (column I): the R^2 is 0.95. Among jobs with a posted wage, job titles also explain most of the variance in the average experience level of applicants across jobs (column II). While wages also explain variation in the average experience level of applicants, they have much lower explanatory power than job titles (compare the R^2 in column II vs. column III). Adding the wage on top of job titles barely improves the model fit (compare the R^2 in column II vs. column IV). The specification in column IV, which has the highest explanatory power of all our specifications, indicates that a 10% increase in the wage is associated with an increase in the experience of the average applicant by 0.12 years, or roughly 1%. While higher wages do attract more experienced applicants within a job title, job titles alone account for most of the variation in applicant experience across jobs.

Average Education. In table 10, we focus on the education level of the average applicant. The results are very similar to what we found when explaining the experience of the average applicant. That is, job titles explain most of the variation in the average level of education of applicants across jobs (column I and II). Wages add little explanatory power once we control for job titles (compare columns III and IV), even though higher wages do attract more educated applicants within a job title (column IV).

Interpretation. Job titles explain most of the variation in the experience and education levels of applicants across jobs. Our results suggest that more educated and more experienced workers do not, for the most part, earn higher wages for doing the same job better. Instead, higher levels of education and experience allow workers to access job titles with higher pay. The sorting of workers into different job titles is likely to play an important role in explaining the impact of education and experience on wages.

While the data limitations and the small magnitude of the effects necessitate caution, the fact that wages are associated with significantly better applicants in each specification suggests a real effect. This result gives further credence to the idea that search in this market is directed rather than random. That is, low-skill and high-skill workers sort themselves into low-wage and high-wage jobs, respectively.

4.3 Word Analysis

4.3.1 Number of Applicants

We now discuss words that are associated with a larger or smaller number of applicants per view within SOC code (Figure 5). Many of the words that predict a greater number of applicants per view are also words that predict lower wages (compare with Figure 3). These include words denoting lower seniority and experience, such as "assistant", "support", "specialist", "coordinator", "entry", and "junior". As for words denoting specialities, we can see, for example, that lower-wage inside sales jobs receive more applications than the average job in the same SOC.

Conversely, many of the words that predict a lower number of applicants per view within SOC are words that also predict higher wages. The two word clouds are remarkably similar. Words that denote higher seniority and management positions such as "manager", "senior", and "director" are associated with a smaller number of applicants. Words that are associated with higher paying specialities or areas, such as "developer", "engineer", and "linux", have a lower number of applicants.

Overall, examining the words that predict wages and words that predict the number of applicants enlightens us on the negative relationship between wages and applicants within SOC. Within SOC, words in the job title associated with higher wages predict a lower number of applicants per view, and vice versa for words in the job title associated with lower wages. Substantively, jobs with higher seniority or managerial responsibilities tend to pay higher wages and attract fewer applicants. Similarly, specialties with higher pay tend to attract fewer applicants.

4.3.2 Quality of Applicants

We now turn to the quality of applicants and examine words that predict higher or lower education and experience within an SOC code. Because the sample size for this exercise is much smaller, we report the results in Table 11 rather than a word cloud.

Words that predict higher experience or higher education tend to be words that also predict higher wages, and vice versa for words that predict lower experience or education. For example, words that indicate higher seniority or management such as "manager" and "senior" are associated with higher experience, while words like “director” and “chief” are associated with both higher experience and higher education. Lower education and experience are associated with certain specialties. The example of “rn" (registered nurse) is interesting, as it is associated with both lower experience and lower education. This is explained by the fact that, within SOC 29-1111 (“Registered Nurses”), “rn” indicates a lower level job compared to job titles where “rn” does not appear, such as “nurse manager,” “nurse clinician,” and “director of nursing.”

Overall, examining the behavior of applicants reveals that words in the job title are powerful predictors of the number and quality of applicants that a vacancy receives. Job titles not only explain firms’ wage-posting behavior, but they also play a key role in directing workers’ search.

5 Robustness and Discussion

5.1 Robustness and Further Results

We have shown that job titles explain most of the variance in posted wages. One reason for this large explanatory power may be that there are many unique job titles in the sample. However, restricting the sample to job titles that appear at least twice does not change the results (compare Table 4 to Table A.1 in the appendix). Thus, even when focusing on job titles that appear in at least two job postings, we find that job titles explain more than 90% of the variance in posted wages.

Our analysis of the wage variance has shown that job titles essentially determine posted wages. However, firms often post a wage range rather than a single wage, and we have focused so far on explaining the midpoint of this range. In Table A.2, we show that job titles are just as powerful in explaining the minimum of the range (column I) and the maximum of the range (column III). We also find, again, that SOC codes explain less than 40% of the variance in the minimum or the maximum offered wage (columns II and IV). Finally, we investigate the power of job titles in explaining how large the wage range is. We define the wage range as the maximum minus the minimum divided by the midpoint. We divide the range by the midpoint to adjust for the fact that higher wage jobs may also have larger absolute ranges. This range variable takes the value of zero when only one wage value is posted. Remarkably, we find that job titles have high explanatory power for wage ranges as well: they explain about 80% of the variance in the wage range. By contrast, SOC codes only explain about 20% of the variance in the wage range. We conclude that job titles explain most of the variance in the minimum, the midpoint and the maximum of the posted wage range, as well as in the size of the posted wage range.

We have shown that the *first four* words in the job title explain most of the variance in posted wages. How sensitive is the explanatory power of job titles to the number of words that are used? Table A.3 examines the explanatory power of the first n words of the job title for various values of n. The first word of the job title already has a great deal of explanatory power: first word fixed effects explain about 60% of the wage variance. Astonishingly, the first word of the job title has greater explanatory power than the most detailed occupational classification that can be used in the CPS (see Table 5, column VI). Using the first three words of the job title significantly improves the explanatory power of the model, with an R^2 of 0.93. Using the first four words only slightly improves the explanatory power compared to using the first three words. Finally, using all words in the job title essentially does not add any explanatory power compared to using the first four words. These results show that the first four words of the job title convey almost all of the information that is relevant for posted wages, and justify our choice of using the first four words to define the job title.

We have analyzed the impact of wages on the number of applicants and have shown that this impact is only estimated to be positive when job titles are controlled for. However, omitted variable bias could contaminate the relationship between wages and the number of applications. Since we cannot control for the full text of the job ad, we may be missing information that is relevant for the worker's application decision. To assess whether this is the case, we turn to an examination of the impact of the wage on clicks per 100 views. Recall that when a job is listed as a snippet on the result page, only the posted wage, job title, firm, and DMA are listed. The applicant must click to see more details. Hence, we have all the variables that can drive the applicant's click decision, eliminating the scope for omitted variable bias.

Table A.4 explores the relationship between wages and clicks per 100 views: the results are similar to those obtained when applications per 100 views is the dependent variable (Table 8). When no controls are used (column I), we see a significant and negative association between the wage and clicks per 100 views. When controlling for basic job characteristics, firm fixed effects, and job title fixed effects, the coefficient on the wage becomes positive and highly significant (column V), implying that a 10% increase in the wage is associated with a 2.9% increase in clicks per 100 views. The fact that the qualitative results in Table 8 can be reproduced for clicks per view, an outcome whose determinants are fully known, improves our confidence in our basic results. A higher wage is generally associated with fewer clicks and fewer applications per view. It is only within job title that a higher wage results in more clicks and more applications per view. Finding a reversal in the relationship between wages and clicks when controlling for job titles provides further credibility to our results for the number of applications, and confirms the important role of job titles in understanding applicant decisions.

5.2 Discussion

Directed Search. Labor market frictions are important to understand economic phenomena like unemployment or the distribution of wages (Rogerson et al., 2005). In the search and matching literature, such frictions are explicitly modelled. There are two main ways of modeling meetings between workers and firms: random search and directed search. Typically, random search leads to inefficient outcomes with unemployment being inefficiently low or high, while directed search leads to efficient outcomes. The key reason why directed search is more efficient than random search is that directed search allows for greater *competition* between employers. In directed search models (also called competitive search), workers can observe job characteristics and direct their applications to a specific subset of jobs. Therefore, employers have an incentive to compete on characteristics in order to attract workers. By contrast, in random search models, workers and firms meet randomly, and sequentially. Thus, parties to a bilateral worker-firm meeting have some monopoly power because going to another meeting is costly. If the random search model is coupled with bargaining over wages, then the outcome of the model is inefficient unless the workers' bargaining power satisfies the Hosios (1990) condition. By contrast, in directed search models, competition typically leads to the efficient outcome.

Nowadays, the overwhelming majority of vacancies are posted on the Internet (Barnichon, 2010). Furthermore, Internet job search is associated with higher job finding rates than offline job search (Kuhn & Mansour, 2014). The informational structure of job boards (including CareerBuilder) is such that directed search is the more realistic assumption. Indeed, workers are able to (i) observe the characteristics of a number of jobs[18] and (ii) decide which jobs to apply to based on these characteristics. Therefore, there is clearly competition between employers. However, this does not tell us exactly how directed search works in this environment, and the theoretical literature has developed several types of directed search models.

In the most stylized version of directed search, workers and firms are homogeneous. Firms post wages and workers direct their applications based on observed wages (see e.g. Moen, 1997). Such a model predicts that higher wage jobs should attract more applications: firms are willing to pay higher wages because they can fill their vacancies faster thanks to the higher number of applicants.[19] This prediction that higher wage jobs attract more applicants is verified in our data, but only within job title. This suggests that it is only within job title that workers and jobs are homogeneous enough to satisfy the prediction of the model. Conversely, this implies

[18] In fact, Acemoglu & Shimer (1999) show that it can be enough for workers to observe the characteristics of just *two* random jobs in order to achieve efficiency in a directed search model.

[19] With one application per worker, this relationship is off the equilibrium path. In equilibrium, a positive relationship between posted wages and the number of applicants can for example be obtained by allowing for multiple applications per worker as in Galenianos & Kircher (2009).

that there is much job and worker heterogeneity across job titles, even within a given SOC.

Directed search models can be extended to capture such job and worker heterogeneity (see e.g. Shi, 2001; Shimer, 2005a; Eeckhout & Kircher, 2010). In these models, firms use wages to induce the efficient sorting of different types of workers across jobs.[20] If firms cannot communicate job characteristics and workers must search randomly (see e.g. Shimer & Smith, 2000), then sorting will typically not be perfect and inefficiencies will arise (see related literature discussion in Shimer, 2005a). In our data, workers observe job characteristics before applying, and we find evidence of sorting on wages. Specifically, we find that higher wage jobs attract better applicants (higher education and more experience) both across and within a job title.

Importantly, directed search models do not necessitate posting the wage. Indeed, the crucial ingredient of directed search is the presence of information about jobs that allows for competition between firms to attract workers. In the theoretical literature, there are at least two examples of directed search models that do not have posted wages yet still yield efficient outcomes. Acemoglu & Shimer (1999) show that posting a bargaining parameter can be enough to achieve efficiency. Menzio (2007) develops a model of cheap talk where firms post non-binding signals about their wages.

In our empirical work, we find that only a minority of firms post wages. Yet, this is consistent with directed search if firms can convey information about their jobs using other signals. In particular, we show that the job title is an important mechanism for firms to communicate information about their jobs. Indeed, job titles explain most of the wage variance among jobs with a posted wage. Furthermore, across *all* jobs, including those that do not post wages, job titles explain more than 80% of the variation in the education and experience of applicants, which shows that job titles are used by workers to direct their search.

Overall, our evidence is consistent with many predictions of directed search models. Even though most jobs do not post wages, we have shown that employers are able to communicate important job characteristics through job titles, which is consistent with the directed search framework. Indeed, the key feature of directed search is competition between employers based on the advertisement of job characteristics. Such competition is the reason why directed search models typically lead to an efficient level of unemployment. In the next subsection, we discuss in more detail the informational content of job titles, a key instrument used by employers to advertise their job characteristics.

[20] The exact relations between wages, the number of applicants and the quality of applicants often depend on the fine details of the model.

Informational Content of Job Titles. We have found that job titles can explain most of the variance in posted wages and in the quality of applicants that a vacancy receives, while SOC codes explain a much lower share of the variance. We interpret our empirical results as a implying that the SOC occupational classification masks important heterogeneity across jobs in required skills and/or job duties. This heterogeneity makes jobs differentially attractive to different types of workers. However, one may wonder whether our results are consistent with alternative interpretations. For example, could it be the case that firms use job titles simply to signal or justify higher wages without there being any differences in required skills or job duties? We provide two arguments against job titles being mere signals for wages.

First, words in the job title that are predictive of higher wages are typically also indicative of differences in job duties and/or skills. Clearly, a "java programmer" needs different skills than a "C++ programmer", and does not just get paid more for the same work. Among the most common job titles (Table 2, panel B), we find higher paid "outside sales" jobs and lower paid "inside sales" jobs (Table 7). These two jobs have different job duties: outside sales jobs involve travelling to meet customers face-to-face, while inside sales jobs involve talking to customers by phone from a fixed location.[21] "Management" or "senior" job titles pay more but also typically involve managing people on top of other job duties. For example, to take some of the most common "senior" job titles (Table 2), "senior accountants" supervise "junior accountants" or "staff accountants." Likewise, "senior financial analysts" supervise "junior financial analysts."[22] Furthermore, senior jobs also typically have higher relevant experience requirements, implying that the worker needs higher skill to be hired for the job.[23] Hence, the job title is not simply a signal for higher pay. The words in the job title can contain real information about job duties and/or required skills.

Second, we can rely on our results regarding applicant behavior to show that job titles with different wages attract applicants with different skills. As we have shown, different job titles attract applicants with systematically different levels of education and experience (Table 9 and Table 10). Moreover, we have shown that, within SOC code, higher-paid jobs attract fewer applicants (Table 8, col. II), and it is only within job title that higher-paid jobs attract more applicants. This provides additional support for the idea that jobs with different job titles have

[21]For a comparison of these two jobs, see e.g. Wikipedia.org (`http://en.wikipedia.org/wiki/Sales`) or Monster.com (`http://career-advice.monster.com/job-search/career-assessment/inside-sales-or-outside-sales/article.aspx`).

[22]Investopedia.com, a website devoted to information about investing, makes clear the distinctions between the job duties of financial analysts and accountants at various levels of seniority. See `http://www.investopedia.com/financial-edge/0213/a-day-in-the-life-of-a-financial-analyst.aspx` and `http://www.investopedia.com/terms/j/junior-accountant.asp`

[23]A familiar example from academia is the distinction between assistant and associate professor: to be hired as an associate professor, a candidate must have a much stronger publication record.

different job duties or require different skills. After all, if all jobs within an SOC code had the same job duties and skill requirements irrespective of their job title, then we would expect a positive relationship between wages and the number of applicants. Yet, this is not what we observe in the data: the relationship between wages and the number of applicants is negative, even controlling for SOC codes. Hence, we can conclude that there are substantial differences in job duties or skill requirements across different jobs within an SOC code, and workers self-select as a function of job characteristics. For example, higher-skill workers apply relatively more often to higher-skilled, higher-paying jobs, and vice versa for lower skilled workers. By contrast, the relation between posted wages and the number of applicants is positive within a job title, indicating that job heterogeneity in duties and skill requirements within a job title is much smaller than across job titles within a given SOC. Hence, applicant behavior is consistent with higher-paying job titles requiring higher skills (or different job duties) rather than just paying more for the same job duties.

Overall, the evidence from our analysis of both job posting and applicant behavior supports the interpretation that the SOC classification misses important differences in required skills and/or job duties among jobs with different job titles. Jobs are substantially more heterogeneous than what is captured by the most detailed occupational classification.

Wage Dispersion. Because jobs within an SOC are heterogeneous, wage dispersion within an SOC is not pure frictional wage dispersion. Indeed, due to differences in skill requirements, a given worker would not be hired in each job within an SOC code with equal probability. Similarly, differences in job duties across jobs imply that jobs within an SOC are not equally desirable for a given worker. Hence, from the point of view of each worker, the distribution of wages that he might earn is more compressed and the degree of frictional wage dispersion is smaller than if all jobs within an SOC were identical. Therefore, the presence of job heterogeneity within an SOC supports our observation in section 3.2.2 that measuring wage dispersion within an SOC overestimates pure wage dispersion.

These results are important because wage dispersion can be interpreted as a measure of search frictions in a labor market. If the degree of wage dispersion is small, a worker who receives a wage offer has little incentive to turn this offer down to continue to search for a better one. In that sense, our results can help explain an important stylized fact of many labor markets: the majority of job offers are accepted by unemployed workers (Krueger & Mueller, 2014; Wolthoff, 2015).

6 Conclusion

In this paper, we have used new data from CareerBuilder.com to show that the words contained in job titles play a key role in the search and matching process. Indeed, all employers post a job title while only 20% of jobs have a posted wage. Even among jobs with a posted wage, job titles explain 90% of the variance in posted wages. Furthermore, job titles explain more than 80% of the variance in the education and experience levels of applicants that a vacancy attracts. Job titles are much more specific than the most detailed SOC occupational classification used in the Current Population Survey. Without such a specific measure of occupations, one can obtain spurious results. For example, we find that, within an SOC code, higher paying jobs attract *fewer* applicants; the relationship between posted wages and the number of applicants only becomes positive within job titles.

A key reason why job titles explain wages and applicant behavior better than SOC codes is that words in job titles capture important distinctions between jobs along the seniority/management and the skills/specialization dimensions. Overall, our results show that words in job titles play a much more important role than posted wages in the first stages of the search and matching process: employers use job titles to advertise their jobs, and workers use job titles to direct their search.

Our results show that job titles are a powerful tool to describe job characteristics, and perform much better than SOC codes across a variety of dimensions. Our findings thus open fruitful avenues for future research to better understand a variety of labor market issues, and in particular wage differentials and human capital investment.

References

Abowd, J. M., Creecy, R. H., & Kramarz, F. (2002). *Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data*. Longitudinal Employer-Household Dynamics Technical Paper 2002-06, Center for Economic Studies, U.S. Census Bureau.

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

Acemoglu, D. & Shimer, R. (1999). Holdups and Efficiency with Search Frictions. *International Economic Review*, 40(4), 827–849.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(3), 673–697.

Barnichon, R. (2010). Building a composite Help-Wanted Index. *Economics Letters*, 109(3), 175–178.

Barron, J. M., Bishop, J., & Dunkelberg, W. C. (1985). Employer search: The interviewing and hiring of new employees. *Review of Economics and Statistics*, 67, 43–52.

Barron, J. M., Black, D. A., & Loewenstein, M. A. (1987). Employer size: The implications for search, training, capital investment, starting wages, and wage growth. *Journal of Labor Economics*, 5, 76–89.

Blau, F. D. (1977). *Equal pay in the office*. Lexington, Mass.: Lexington Books.

Blau, F. D. & Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic Perspectives*, 14(4), 75–100.

Bó, E. D., Finan, F., & Rossi, M. (2012). *Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service*. Working Paper 18156, National Bureau of Economic Research.

Brencic, V. (2012). Wage posting: evidence from job ads. *Canadian Journal of Economics*, 45(4), 1529–1559.

Brenčič, V. & Norris, J. B. (2012). Employers' on-line recruitment and screening practices. *Economic Inquiry*, 50(1), 94–111.

Brown, J. & Matsa, D. (2014). Boarding a sinking ship? An investigation of job applications to distressed firms. mimeo.

CareerBuilder (2013). Job Postings.

comScore Media Metrix (2011). comScore Media Metrix Ranks Top 50 U.S. Web Properties for January 2011.

Dickens, W. T. & Katz, L. F. (1986). *Interindustry Wage Differences and Industry Characteristics*. Working Paper 2014, National Bureau of Economic Research.

Eeckhout, J. & Kircher, P. (2010). Sorting and decentralized price competition. *Econometrica*, 78, 539–574.

Faberman, J. R. & Menzio, G. (2010). Evidence on the relationship between recruiting and the starting wage. mimeo.

Faberman, R. J. & Kudlyak, M. (2014). *The Intensity of Job Search and Search Duration*. Working Paper 14-12, Federal Reserve Bank of Richmond.

Galenianos, M. & Kircher, P. (2009). Directed search with multiple job applications. *Journal of Economic Theory*, 114, 445–471.

Gee, L. (2014). The more you know: How applicant pool information affects application rates. mimeo.

Gentzkow, M. & Shapiro, J. M. (2010). What drives media slant? evidence from U.S. daily newspapers. *Econometrica*, 78(1), 35–71.

Gibbons, R. & Katz, L. F. (1992). Does unmeasured ability explain inter-industry wage differentials? *Review of Economic Studies*, 59(3), 515–35.

Groes, F., Manovskii, I., & Kircher, P. (2015). The U-shapes of occupational mobility. *Review of Economic Studies*, forthcoming.

Groshen, E. L. (1991). The structure of the female/male wage differential: Is it who you are, what you do, or where you work? *The Journal of Human Resources*, 26(3), 457–472.

Hall, R. E. & Krueger, A. B. (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics*, 4(4), 56–67.

Holzer, H., Katz, L., & Krueger, A. (1991). Job queues and wages. *Quarterly Journal of Economics*, 106, 739–768.

Hornstein, A., Krusell, P., & Violante, G. L. (2007). *Frictional Wage Dispersion in Search Models: A Quantitative Assessment*. Working Paper 13674, National Bureau of Economic Research.

Hornstein, A., Krusell, P., & Violante, G. L. (2011). Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review*, 101(7), 2873–2898.

Hosios, A. J. (1990). On the Efficiency of Matching and Related Models of Search and Unemployment. *The Review of Economic Studies*, 57(2), 279–298.

Iranzo, S., Schivardi, F., & Tosetti, E. (2008). Skill dispersion and firm productivity: An analysis with employer-employee matched data. *Journal of Labor Economics*, 26(2), 247–285.

Kambourov, G. & Manovskii, I. (2009). Occupational mobility and wage inequality. *Review of Economic Studies*, 76, 731–759.

Krueger, A. B. & Mueller, A. I. (2014). *A Contribution to the Empirics of Reservation Wages*. Working Paper 19870, National Bureau of Economic Research.

Krueger, A. B. & Summers, L. H. (1986). *Reflections on the Inter-Industry Wage Structure*. Working Paper 1968, National Bureau of Economic Research.

Krueger, A. B. & Summers, L. H. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica*, 56(2), 259–93.

Kuhn, P. & Mansour, H. (2014). Is Internet Job Search Still Ineffective? *The Economic Journal*, 124(581), 1213–1233.

Kuhn, P. & Shen, K. (2012). Gender discrimination in job ads: Evidence from china. mimeo.

Manning, A. (2011). Chapter 11 - imperfect competition in the labor market. In *Handbook of Labor Economics*, volume Volume 4, Part B (pp. 973–1041). Elsevier.

Marinescu, I. (2014). The general equilibrium impacts of unemployment insurance: evidence from a large online job board. *Working paper*.

Marinescu, I. & Rathelot, R. (2014). Mismatch unemployment and the geography of job search. mimeo.

Marinescu, I. & Wolthoff, R. P. (2012). Wages, applications and skills. mimeo available online at `http://federation.ens.fr/ydepot/semin/texte1213/MAR2013WAG.pdf`.

Mellow, W. & Sider, H. (1983). Accuracy of response in labor market surveys: Evidence and implications. *Journal of Labor Economics*, 1(4), pp. 331–344.

Menzio, G. (2007). A theory of partially directed search. *Journal of Political Economy*, 115(5), 748–769.

Moen, E. R. (1997). Competitive search equilibrium. *Journal of Political Economy*, 105, 385–411.

Mortensen, D. T. & Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, 61, 397–415.

Murphy, K. & Topel, R. (1987). Unemployment, risk, and earnings: testing for equalizing wage differences in the labor market. In K. Lang & J. Leonard (Eds.), *Unemployment and the Structure of Labor Markets*. Basil Blackwell Inc.

Pallais, A. (2012). Inefficient hiring in entry-level labor markets.

Petrongolo, B. & Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, 39, 390–431.

Pissarides, C. A. (1985). Short-run equilibrium dynamics of unemployment, vacancies and real wages. *American Economic Review*, 75, 676–690.

Poletaev, M. & Robinson, C. (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.

Rogerson, R., Shimer, R., & Wright, R. (2005). Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4), 959–988.

Shi, S. (2001). Frictional assignment I: Efficiency. *Journal of Economic Theory*, 98, 232–260.

Shimer, R. (1996). *Essays in Search Theory*. PhD thesis, Massachusetts Institute of Technology.

Shimer, R. (2005a). The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy*, 113(5), 996–1025.

Shimer, R. (2005b). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review*, 95(1), 25–49.

Shimer, R. & Smith, L. (2000). Assortative matching and search. *Econometrica*, 68(2), 343–369.

Wolthoff, R. P. (2015). Applications and interviews: Firms' recruiting decisions in a frictional labor market. mimeo.

Woodcock, S. D. (2007). *Match Effects*. Discussion Paper dp07-13, Department of Economics, Simon Fraser University.

Woodcock, S. D. (2008). Wage differentials in the presence of unobserved worker, firm, and match heterogeneity. *Labour Economics*, 15(4), 771–793.

Yashiv, E. (2007). Labor search and matching in macroeconomics. *European Economic Review*, 51(8), 1859 – 1895.

Table 1: Summary Statistics

	obs.	mean	s.d.	min	max
Job characteristics					
Yearly wage	11,715	57,323	31,690	13,500	185,000
Required experience	168	3.20	2.80	0.50	10.00
Required education	25,931	14.88	1.88	12.00	24.00
Vacancy duration	61,135	15.67	8.86	0.00	31.00
Firm characteristics					
Number of employees	61,135	18,824	59,280	1	2,100,000
Outcome variables					
Number of views	61,135	6,084.02	6,133.50	0	262,160
Number of clicks	61,135	280.97	312.11	0	7,519
Number of applications	61,135	59.35	121.68	0	4,984
Clicks per 100 views	60,979	5.64	5.58	0	100
Applications per 100 views	61,051	1.17	2.57	0	100
Applicant characteristics					
Years of education	2,282	16.63	1.35	14	24
Years of experience	2,379	13.28	5.13	2.5	26

Source: CareerBuilder.com

Table 2: Ten most common job titles

Panel A: all job titles					
First 4 words	Freq.	First 2 words	Freq.	First 1 word	Freq.
customer service representative	273	customer service	895	senior	4,176
administrative assistant	242	sales representative	865	sales	2,450
project manager	221	director of	580	director	1,080
sales representative	218	project manager	553	customer	1,034
customer service openings in	188	entry level	476	medical	897
sales representative customer service	188	administrative assistant	395	project	883
staff accountant	184	outside sales	374	business	870
outside sales representative	176	inside sales	307	manager	801
senior accountant	166	business development	279	rn	742
full time retail sales	150	business analyst	267	account	654
Panel B: job titles with a posted wage					
First 4 words	Freq.	First 2 words	Freq.	First 1 word	Freq.
customer service representative	120	customer service	260	senior	810
staff accountant	98	administrative assistant	154	sales	369
administrative assistant	93	outside sales	150	customer	277
senior accountant	92	senior accountant	116	administrative	191
executive assistant	63	staff accountant	110	accounting	172
outside sales representative	62	inside sales	106	outside	166
senior financial analyst	56	director of	96	director	146
controller	54	entry level	86	medical	146
financial analyst	49	executive assistant	85	executive	143
chiropractic technician	48	accounts payable	77	account	132

Source: CareerBuilder.com

Figure 1: Job titles on CareerBuilder.com

All Jobs:

Jobs with Posted Wage:

Note: Job titles are truncated to the first four words. Only job titles that appear in at least 10 job postings are represented. Word cloud created using www.tagul.com. Tagul uses word frequency to determine the size of the words.
Source: CareerBuilder.com

Table 3: Explaining wage posting behavior

VARIABLES	I Posts wage	II Posts wage	III Posts wage	IV Posts wage	V Firm f.e.
Job title f.e.	Yes***		Yes***	Yes***	Yes***
Firm f.e.		Yes***	Yes***	Yes***	
Job characteristics				Yes***	
Observations	61,132	61,135	61,132	61,132	61,132
R^2	0.747	0.697	0.928	0.933	0.765
Adj. R^2	0.619	0.672	0.884	0.892	0.647
AIC	-21,907	-11,056	-93,107	-97,856	-48,551

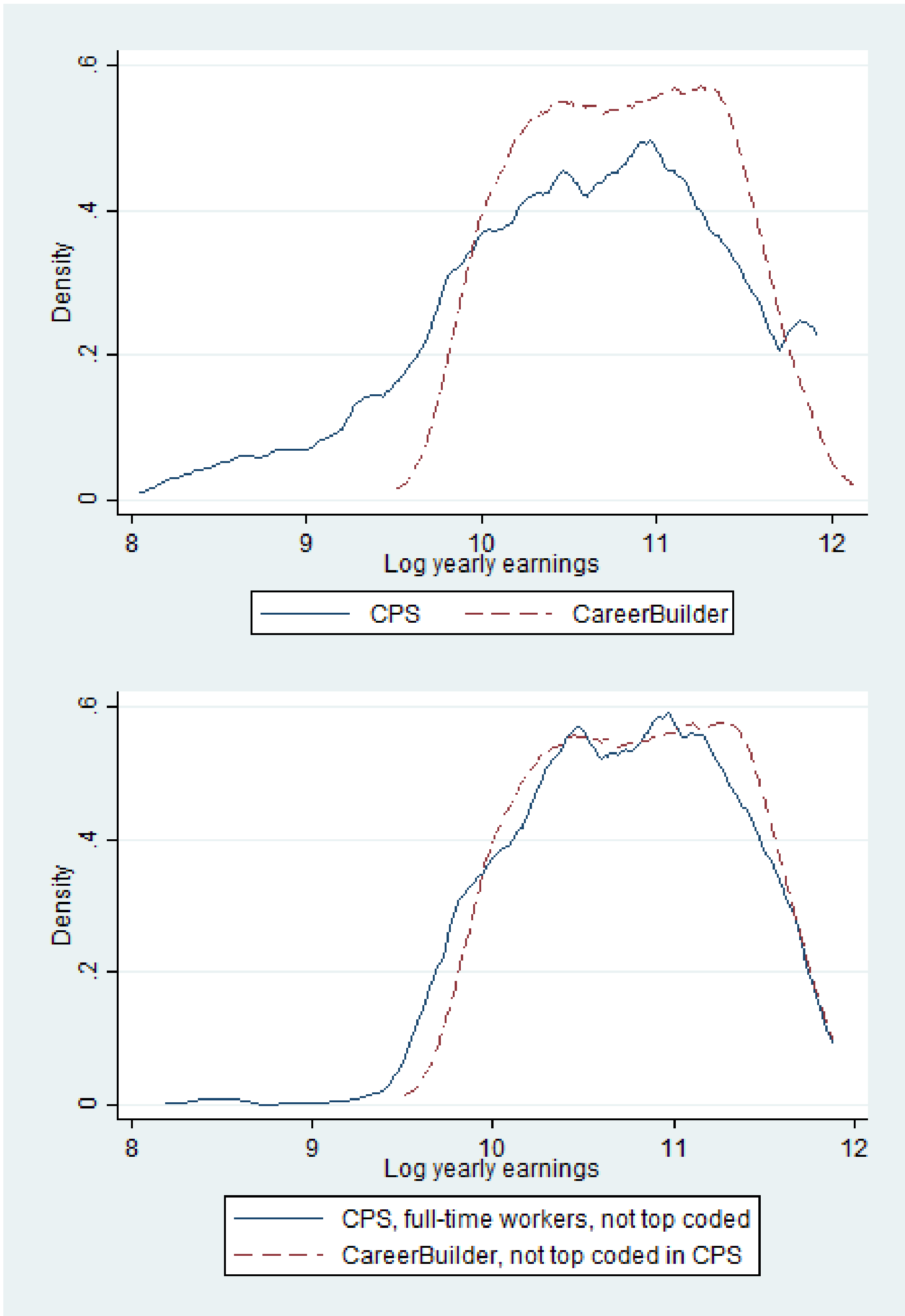
Note: Linear probability model. In columns I-IV, the dependent variable is log yearly posted wage. In column V, the dependent variable is the firm effect estimated in column I. Stars next to "Yes" show the level of significance of the F-test for the joint significance of that group of controls: $***p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$. Job characteristics include vacancy duration, a dummy for salary expressed per hour, required education and experience, designated market area, and calendar month.
Source: CareerBuilder.com

Table 4: Explaining the variance in posted wages

VARIABLES	I Posted wage	II Posted wage	III Posted wage	IV Posted wage	V Firm f.e.
Job title fixed effects		Yes***	Yes***	Yes***	Yes***
Firm fixed effects	Yes***		Yes***	Yes***	
Job characteristics				Yes***	
Observations	11,715	11,715	11,715	11,715	11,715
R^2	0.558	0.902	0.973	0.975	0.902
Adj. R^2	0.499	0.840	0.952	0.956	0.839
AIC	9,900	-14,639	-21,758	-22,778	-17,895

Note: In columns I-IV, the dependent variable is log yearly posted wage. In column V, the dependent variable is the firm effect estimated in column II. Stars next to "Yes" show the level of significance of the F-test for the joint significance of that group of controls: $***p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$. Job characteristics include vacancy duration, a dummy for salary expressed per hour, required education and experience, designated market area, and calendar month.
Source: CareerBuilder.com

Figure 2: The distribution of earnings in the CPS vs. posted wages on CareerBuilder.com



Note: A small number of outliers (log yearly earnings < 8) has been omitted for the CPS data.
Source: Current Population Survey and CareerBuilder.com

Table 5: Using SOC codes fixed effects to explain wages: CPS vs CareerBuilder data

	CPS			CareerBuilder			
	Major	Minor	Detailed	Major	Minor	Detailed	Job Titles
	I	II	III	IV	V	VI	VII
Observations	1,587	1,587	1,587	10,465	10,465	10,465	10,465
R^2	0.149	0.195	0.480	0.144	0.167	0.412	0.943
Adj. R^2	0.144	0.184	0.362	0.143	0.166	0.387	0.907
AIC	4,369	4,280	3,587	15,414	15,125	11,487	-12,925

Note: In columns I-III, the dependent variable is log weekly earnings. In columns IV-VII, the dependent variable is log yearly posted wage. Columns II and IV control for major occupation groups fixed effects. Columns II and V control for minor occupation groups fixed effects. Columns III and VI control for detailed occupation groups fixed effects. Column VII controls for job title fixed effects. The specifications in columns IV-VII only use jobs for which an SOC code was present.
Source: Current Population Survey and CareerBuilder.com

Table 6: Using words to explain within SOC wage variation

VARIABLES	I Wage resid.	II Wage resid.	III Wage resid.	IV Wage resid.
Job title f.e.	Yes			
Words in job title f.e.		Yes		
Frequent words f.e.			Yes	
Frequent words denoting experience and seniority f.e.				Yes
Observations	11,715	11,715	11,715	11,715
R^2	0.871	0.571	0.226	0.136
Adj. R^2	0.790	0.490	0.222	0.135
AIC	-10,752	7,088	10,390	11,597

Note: The dependent variable is wage residuals after a regression of log yearly posted wage on detailed SOC codes fixed effects. f.e. stands for fixed effects. Frequent words are those listed in Table 7. Frequent words denoting experience and seniority are those in the first column of Table 7.
Source: CareerBuilder.com

Table 7: Words that explain wage residuals after detailed SOC fixed effects

Sign of word coefficient	Job level: seniority / management	Specialization / Skills	Computer
Negative (lower wages)	representative	accountant	network
	assistant	account	
	specialist	project	
	associate	medical	
	entry	marketing	
	coordinator	quality	
	support	inside	
	clerk	bilingual	
	operator	office	
	part-time	advisor	
	staff	receptionist	
	junior	recruiter	
Positive (higher wages)	manager	-	engineer
	senior	sales	developer
	executive	consultant	systems
	director	administrative	software
	management	business	architect
	supervisor	outside	web
	of	with	net
	ii	nurse	java
	lead	maintenance	it
	to	health	
		hr	
		or	
		controller	
		auditor	

Note: The words included are significant at the 5% level in explaining the residuals after a regression of the posted wage on SOC codes fixed effects (Table 6, col. II). Words are included when they appear at least 100 times. Words are ordered by frequency and underlined when they appear at least 500 times.
Source: CareerBuilder.com

Figure 3: Words that predict wages within a given SOC code

Higher Wages:

[0.044;0.060]; [0.077;0.200]; [0.202;1.576] log points

Lower Wages:

[-1.101;-0.169]; [-0.167;-0.104]; [-0.101;-0.063] log points

Note: Words that are significant at the 5% level in explaining the residuals after a regression of the posted wage on SOC codes fixed effects (Table 6, column II) and appear at least 10 times. The big rectangle is "-", which typically separates the main job title from additional details. Word cloud created using www.tagul.com. The size of a word represents its frequency, while the color represents the tercile of its coefficient, weighted by frequency.
Source: CareerBuilder.com

Figure 4: Words that predict probability of posting a wage within a given SOC code

Higher Probability:

[0.015;0.042]; [0.046;0.107]; [0.109;0.859] percentage points

Lower Probability:

[-0.720;-0.048]; [-0.046;-0.032]; [-0.027;-0.027] percentage points

Note: The words included are significant at the 5% level in explaining the residuals after a regression of the "Posts wage" dummy on SOC codes fixed effects and appear at least 10 times. The big rectangle is "-", which typically separates the main job title from additional details. Word cloud created using www.tagul.com. The size of a word represents its frequency, while the color represents the tercile of its coefficient, weighted by frequency.
Source: CareerBuilder.com

Table 8: The impact of wages on the number of applicants per 100 views

	I	II	III	IV	V
Log(Posted Wage)	-0.770***	-0.642***	-0.710***	1.268***	0.947*
	(0.052)	(0.075)	(0.087)	(0.373)	(0.517)
Job characteristics		Yes	Yes		Yes
SOC f.e.		Yes	Yes		
Firm f.e.			Yes		Yes
Job title f.e.				Yes	Yes
Observations	11,708	11,708	11,708	11,708	11,708
R^2	0.017	0.133	0.363	0.464	0.584
Adj. R^2	0.0165	0.0835	0.235	0.123	0.268
AIC	61,152	59,754	57,200	54,049	52,042

Note: The dependent variable is the number of applications divided by the number of job views divided by 100. "f.e." stands for "fixed effects". SOC fixed effects are for detailed codes. Job characteristics include vacancy duration, a dummy for salary expressed per hour, required education and experience, designated market area, and calendar month. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: CareerBuilder.com

Table 9: Explaining applicants' average experience

	I	II	III	IV
	All jobs	Jobs with a posted wage	Jobs with a posted wage	Jobs with a posted wage
VARIABLES	Years exp.	Years exp.	Years exp.	Years exp.
Log(Posted Wage)			2.174***	1.242*
			(0.203)	(0.708)
Job title f.e.	Yes	Yes		Yes
Observations	2,379	1,755	1,755	1,755
R^2	0.948	0.956	0.238	0.958
Adj. R^2	0.803	0.824	0.238	0.833
AIC	4,934	3,303	8,293	3,212

Note: The dependent variable is the average number of years of experience among the applicants to each job. The regressions are weighted by the number of applicants to each vacancy using Stata's analytic weights. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: CareerBuilder.com

Table 10: Explaining applicants' average education

	I	II	III	IV
	All jobs	Jobs with a posted wage	Jobs with a posted wage	Jobs with a posted wage
VARIABLES	Years educ.	Years educ.	Years educ.	Years educ.
Log(Posted Wage)			0.757***	0.313*
			(0.045)	(0.173)
Job title f.e.	Yes	Yes		Yes
Observations	2,282	1,696	1,696	1,696
R^2	0.961	0.961	0.282	0.963
Adj. R^2	0.845	0.840	0.281	0.849
AIC	-2,115	-1,699	3,239	-1,803

Note: The dependent variable is the average number of years of education among the applicants to each job. The regressions are weighted by the number of applicants to each vacancy using Stata's analytic weights. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: CareerBuilder.com

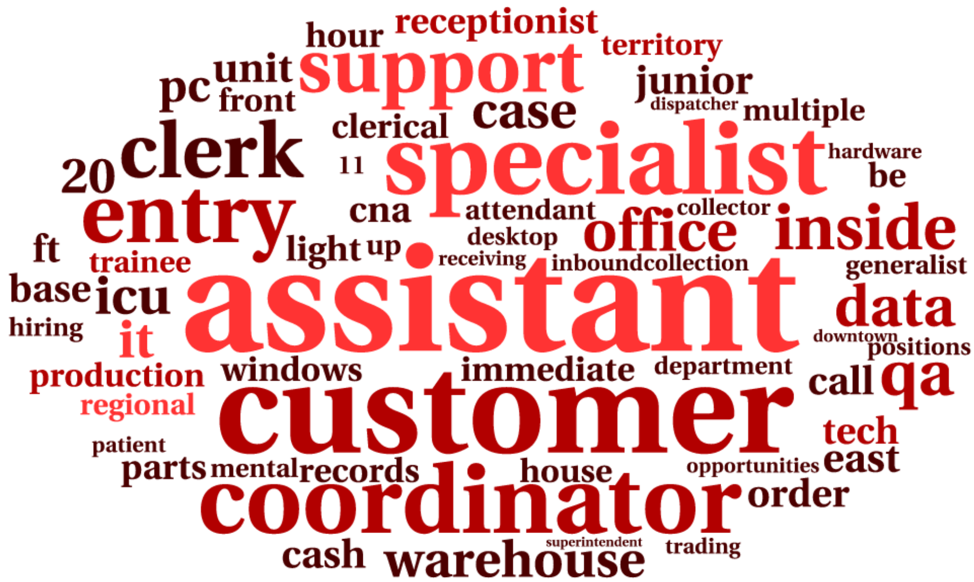
Table 11: Words that predict higher or lower experience and education of applicants within an SOC code

Experience +	Experience -	Education +	Education -
manager	rn	director	rn
senior	web	developer	customer
director	center	nurse	services
executive	insurance	it	needed
of	loan	net	warehouse
retail	3	controller	healthcare
management		research	license
supervisor		performance	
controller		desk	
design		agent	
consulting		summer	
dba		vice	
chief		forklift	
asp		distribution	
		hvac	
		chief	

Note: Words that appear at least 10 times and that are significant at the 5% level in explaining the residuals after a regression of the average education or average experience of applicants on SOC codes fixed effects. Words are ordered by frequency and underlined when they appear at least 100 times. Source: CareerBuilder.com

Figure 5: Words that predict the number of applicants per view within a given SOC code

Higher Number:



[0.154;0.251]; [0.269;0.482]; [0.517;4.974] appl. per 100 views

Lower Number:

[-3.920;-0.377]; [-0.373;-0.301]; [-0.299;-0.164] appl. per 100 views

Note: Words that are significant at the 5% level in explaining the residuals after a regression of the number of applicants per view on SOC codes fixed effects and appear at least 10 times. The big rectangle is "-", which typically separates the main job title from additional details. Word cloud created using www.tagul.com. The size of a word represents its frequency, while the color represents the tercile of its coefficient, weighted by frequency.
Source: CareerBuilder.com

Online Appendix

Appendix A Robustness

Table A.1: Explaining the variation in posted wages: sample restricted to job titles that appear at least twice

	I	II	III	IV	V
VARIABLES	Posted wage	Posted wage	Posted wage	Posted wage	Firm f.e.
Job title f.e.	Yes***		Yes***	Yes***	Yes***
Firm f.e.		Yes***	Yes***	Yes***	
Job characteristics				Yes***	
Observations	10,467	10,467	10,467	10,467	10,467
R^2	0.937	0.568	0.969	0.972	0.891
Adj. R^2	0.908	0.509	0.952	0.956	0.841
AIC	-11,650	8,507	-18,156	-19,063	-11,856

Note: In columns I-IV, the dependent variable is log yearly posted wage. In column V, the dependent variable is the firm effect estimated in column I. Stars next to "Yes" show the level of significance of the F-test for the joint significance of that group of controls: $***p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$. Job characteristics include vacancy duration, a dummy for salary expressed per hour, required education and experience, designated market area, and calendar month.
Source: CareerBuilder.com

Table A.2: Explaining the variation in posted wages: minimum wage offered, maximum wage offered, and wage range

	Min. offered wage		Max. offered wage		Wage range	
	I	II	III	IV	V	VI
Job title f.e.	Yes		Yes		Yes	
SOC f.e.		Yes		Yes		Yes
Observations	11,717	11,717	12,383	12,383	11,898	11,898
R-squared	0.941	0.399	0.943	0.386	0.861	0.236
Adj. R-squared	0.904	0.367	0.908	0.354	0.772	0.196
AIC	-14,347	12,891	-12,721	16,778	-29,113	-8,849

Note: "Wage range" is the maximum offered wage minus the minimum offered wage divided by the midpoint of the range.
Source: CareerBuilder.com

Table A.3: Posted wages: the explanatory power of job titles and how it varies with truncating the job title after the first *n* words

VARIABLES	I Posted wage	II Posted wage	III Posted wage	IV Posted wage	V Posted wage
Job title f.e.	1 word	2 word	3 words	4 words	All words
Observations	11,715	11,715	11,715	11,715	11,715
R^2	0.610	0.865	0.925	0.944	0.946
Adj. R^2	0.568	0.817	0.885	0.909	0.910
AIC	8,416	-4,010	-10,968	-14,359	-14,726

Note: All columns include job title fixed effects, but the definition of job title is different in each column. In column V, all words in the job title are used to define the job title. In columns I-IV, the first *n* words are used to define the job title.
Source: CareerBuilder.com

Table A.4: The impact of posted wages on clicks per 100 views

	I	II	III	IV	V
Log(Posted Wage)	-1.045***	-0.597***	-0.711***	2.035***	1.930***
	(0.089)	(0.130)	(0.167)	(0.399)	(0.454)
Job characteristics		Yes	Yes		Yes
SOC f.e.		Yes			
Firm f.e.			Yes		Yes
Job title f.e.				Yes	Yes
Observations	11,694	11,694	11,694	11,694	11,694
R^2	0.011	0.168	0.389	0.564	0.643
Adj. R^2	0.011	0.121	0.267	0.287	0.371
AIC	72,956	71,012	68,453	63,366	62,012

Note: The dependent variable is the number of clicks divided by the number of job views divided by 100. "f.e." stands for "fixed effects". SOC fixed effects are for detailed SOC codes. Job characteristics include vacancy duration, a dummy for salary expressed per hour, required education and experience, designated market area, and calendar month. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Source: CareerBuilder.com

Appendix B Frictional Wage Dispersion

Hornstein et al. (2011) calculate the mean-min ratio from a Mincer regression using the US census. When defining occupations as a combination of an SOC code and a geographic area, they obtain a ratio of 2.74 (see Table 2 in their paper, for full-time, full-year workers).

We first reproduce this number in the CareerBuilder data. Maintaining the same definition of an occupation, controlling for vacancy duration, a dummy for salary expressed per hour, required education and experience, and calendar month gives a mean-min ratio of 2.44, which is only slightly lower than the estimate of Hornstein et al. (2007). Restricting the set of controls to required education and experience only makes little difference and yields a mean-min ratio of 2.47 (see row 1 in Table B.5).

Subsequently, we change the definition of an occupation to a combination of a job title and a geographic area. In line with our results in section 3.2, this has a large impact on the estimate of the degree of wage dispersion. Depending on the set of controls, we find a mean-min ratio of 1.22 or 1.40 (see row 2 in Table B.5).

Table B.5: Frictional wage dispersion: the mean / min ratio

	Census Mincer regression Mean / min	CareerBuilder Full controls Mean / min	CareerBuilder Restricted controls Mean / min
Occupation and geographic area	2.74	2.44	2.47
Job title and geographic area		1.22	1.40

To assess how these estimates compare to the values by basic search models, we repeat the calibration of Hornstein et al. (2011) with aggregate data from our sample period. Using the methodology of Shimer (2005b) on 2011 BLS data, we set the job-finding rate λ_u^* equal to 0.191 and the job-destruction rate σ equal to 0.022. Maintaining the assumptions that the interest rate r is 0.0041 and the replacement rate ρ equals 0.4, this yields a ratio

$$Mm = \frac{\frac{\lambda_u^*}{r+\sigma} + 1}{\frac{\lambda_u^*}{r+\sigma} + \rho} = 1.078.$$

Hence, while the degree of wage dispersion in the data is still larger than what is implied by the calibration, controlling for job titles rather than SOCs reduces the gap by no less than 90%. We therefore conclude that job titles play a crucial role in reconciling the calibration results with empirical measures of frictional wage dispersion.