Wages, Applications, and Skills*

Ioana Marinescu and Ronald Wolthoff †

December 5, 2012

Abstract

Do firms that post higher wages attract more and better applicants? Using data from the popular employment website CareerBuilder.com, we document that, indeed, higher wages attract more educated and experienced applicants. Surprisingly, higher wages are associated with fewer applications, and this is robust to controlling for detailed occupation and industry fixed effects. However, within specific job titles, a 10% higher wage is associated with 7.4% more applications. These results are consistent with a directed search model in which skills are two-dimensional and highly job specific. The model has additional testable implications about wages and unemployment across skill levels.

1 Introduction

The rise of employment websites over the past decade has made it much easier for job seekers to find job vacancies and assess their attractiveness. Whereas in the past, job ads were typically spread out across countless newspapers, these days most vacancies can be found with just a few mouse clicks. Background information on employers has also gotten much easier to obtain. These developments have likely reduced information frictions and increased the level of competition in the labor market. Intuitively, this competition among employers implies that there will be more applications to jobs that offer more attractive terms of employment, and in particular higher wages. This paper uses data from an employment website to investigate the relationship between wages and the number of applicants.

However, estimating the relationship between wages and the number of applicants with observational data is fraught with difficulties. After all, workers are heterogeneous in skills, and

*This paper has benefited from feedback during presentations at CREST-Paris, the University of Chicago, the Annual Meeting of the Society of Economic Dynamics 2012, the University of Toronto, the University of Wisconsin-Madison, UC Santa Barbara, the Tinbergen Institute, and the University of British Columbia. We are grateful for comments from Briana Chang, Kerwin Charles, Peter Kuhn, Alan Manning, Craig Riddell, and Chris Taber. We thank Chaoran Chen for excellent research assistance.

†Ioana Marinescu: Harris School of Public Policy, University of Chicago, ioana.marinescu@gmail.com; Ronald Wolthoff: Department of Economics, University of Toronto, ronald.p.wolthoff@gmail.com.

jobs differ in their skill requirements. This heterogeneity can obscure the relationship between wages and the number of applicants. For example, hospitals looking for an ambulatory surgery nurse may very well need longer to fill their vacancy than a local grocery store searching for a cashier, even though the nurse job pays considerably more. This paper will carefully account for worker and job heterogeneity when analyzing the relationship between wages and applications.

The relationship between skills, wages, and applications is central to much of labor economics. Indeed, job applications provide a unique insight into the process through which job matches form in the labor market. Studying job applications can help us answer three broad types of questions. First, what kind of jobs are most attractive to workers, and, in particular, what is the importance of the wage in determining job choice? Second, what kind of jobs do workers think are suitable for them, given their understanding of skill demand by employers? Third, which types of jobs are harder to get given the number and quality of applicants? Providing empirical evidence pertaining to these questions promises to shed new light on some of the fundamental questions in labor economics. For example, the relation between skills and wages determines the returns to investment in human capital. Further, the link between skills and the number of applications is informative about how unemployment varies with skill levels. Finally, the relation between wages and the number of applications is related to the level of competition within a market, with a more competitive market implying a stronger positive impact of the wage on the number of applications. Hence, an analysis of applications and their relationship to wages and skills promises to shed light on many important questions in labor economics.

Despite the relevance of these questions, we know little about the relationship between skills, wages, and applications. The main cause of this appears to be a lack of data containing all the necessary variables. In particular job applications are usually unobserved. This paper overcomes this problem by using data from CareerBuilder.com, the largest employment website in the US. This data set contains detailed information on available vacancies in two large US cities in the beginning of 2011. It includes the posted wage associated with these vacancies, the number of applicants that each vacancy attracts, as well as various firm and applicant characteristics. The very detailed information present in this dataset allows us to perform a careful analysis of the relationship between wages and applications in the presence of job and worker heterogeneity.

In the first part of the paper, we use this CareerBuilder data to document a number of new empirical facts. We start with analyzing the economy-wide relationship between skills, wages, and applications and show that higher wage jobs attract more experienced and more educated applicants. Surprisingly, higher wage jobs also attract fewer applicants. Since the negative relationship between wages and the number of applicants may be due to heterogeneity in job type, we further control for detailed occupation and industry fixed effects, respectively based on the Standard Occupational Classification (SOC) and the North American Industry

Classification System (NAICS). We also experiment with using firm fixed effects. However, the negative relationship between wages and applications is not sensitive to the addition of these controls, suggesting that job heterogeneity is important, even within detailed six-digit SOC codes. Thus, while the relationship between wages and applicant quality is positive, the relationship between wages and the number of applicants is negative and robust to the inclusion of a large number of controls.

We finally attempt to control even more precisely for differences across jobs by using job titles, which are freely chosen by firms when posting their vacancies (e.g. "RN ambulatory surgery", where RN stands for "registered nurse"). In this new specification, we replace SOC fixed effects with job title fixed effects. We again find a positive relationship between wages and the quality of applicants as measured by education and experience, suggesting significant worker heterogeneity even within the very narrowly defined occupations captured by job titles. However, contrary to previous results, we find a positive relationship between wages and the number of applications within job titles: a 10% increase in the wage is associated with a 7.4% increase in the number of applicants. The basic intuition that higher wage jobs attract more applicants therefore seems to hold only when jobs are sufficiently similar. This raises the question: why are jobs similar enough within job titles but not within six-digit SOC codes? To answer this question, we regress applications on six-digit SOC codes, controls, and word fixed effects for each word that appears in a job title. We find that there are two types of words in job titles that significantly impact the number of applications once SOCs are controlled for. The first type of word is a description of the hierarchy level or work experience required for the job (e.g. 'senior'), the second type of word is an indication of a special area or skill requirement (e.g. 'Java' for computer programmers). We conclude that the relationship between wages and applications is positive within a job title but not within an SOC, because job titles contain relevant information about work experience, hierarchy, and specialization that are not included in six-digit SOC codes.

In the second part of the paper, we develop a theoretical model with the objective of explaining why the relationship between wages and applications is positive within job titles but negative when broader job groupings such as SOCs are used. Since job seekers on CareerBuilder seem to direct their applications based on the wage and other job characteristics, we use a directed search model. Further, since we have shown empirically that there exists heterogeneity in worker quality both across and within job titles, we assume that there are two dimensions of worker heterogeneity. The first dimension of quality heterogeneity, 'skill', determines workers' choice of which job title to apply to. The second dimension of quality heterogeneity, 'productivity', determines which wage a worker directs their application to within a job title. Intuitively, there are good and bad cashiers and good and bad nurses, where 'nurse' or 'cashier' are values taken by the 'skill' dimension while good or bad are values taken by the 'productivity' dimension. We thus rely on our empirical analysis to determine what type of model is most adequate, here a directed search model with two dimensions of worker heterogeneity.

The model predicts, in line with the empirical evidence, that higher wage jobs attract more and better applicants (i.e. higher 'productivity') within job titles. On the other hand, across job titles, higher wage jobs attract better (i.e. higher 'skill') but fewer applicants. To obtain the negative relationship between wages and applications across job titles, it is necessary to assume that the surplus from the job match increases faster in skill than the vacancy creation costs. This assumption implies that high skill jobs are more profitable once a match is formed. Therefore, absent hiring constraints, firms would prefer to create high skill jobs. To make firms indifferent between creating a low and a high skill job, it must be that it is harder to hire in the high skill job, i.e. that the number of applicants decreases with skill and therefore also decreases with the posted wage. The model thus explains under what condition we can obtain a negative relationship between wages and applications across job titles.

Finally, we show that, to obtain both a positive relationship between wages and applications within job titles, and a negative relationship within SOCs, it is necessary to assume that skills have very low transferability across different job titles within an SOC. In other terms, the observed pattern of empirical results can only emerge if workers' productivity is drastically reduced when they work in job titles different from their own, and this productivity drop-off occurs even if workers remain within the same six-digit SOC. The model and the empirical results taken together show that, within a job title, workers have a fairly similar productivity, while there are economically significant differences in worker productivity within a six-digit SOC code. The model therefore provides a skill-based explanation for why the relationship between wages and applications is positive within job title but negative within broader groupings such as SOC.

Our model not only yields predictions that are consistent with the facts we wanted to account for, but also generates additional empirically testable implications. In particular, the surplus created by a match increases faster with skill than the vacancy creation cost. Furthermore, high-skilled workers are less likely to remain unemployed and capture a higher share of the surplus than low-skilled workers. Our model thus shows that matching frictions play an important role in understanding inequality in both wages and unemployment across skill levels.

Our findings make three key contributions to the literature. First, our results on the relationship between the wage and the number of applicants that a vacancy attracts are related to work by Holzer et al. (1991) and Faberman & Menzio (2010). Compared to these papers, we employ a data set that is larger and more representative of the entire labor market, instead of mostly focusing on low-skilled jobs. While there are some elements in previous work suggesting that high-wage jobs attract fewer applicants (Holzer et al., 1991; Faberman & Menzio, 2010), we show that this relationship is robust in a large number of specifications, but completely reverses when controlling for specific job titles. Our theoretical model provides a plausible explanation for this pattern in the data. Second, our paper documents the positive relationship between wages and the quality of the applicant pool in a large sample of jobs from the US. Our work is thus complementary to the evidence from a recent working paper (Bó et al., 2012) showing that higher wages attract better applicants to a public-sector job in Mexico[1]. Third, our results shed light on the specificity of human capital, i.e. how transferable skills are from one job to another, adding to a line of research by e.g. Poletaev & Robinson (2008) and Kambourov & Manovskii (2009a). Our empirical findings and theoretical analysis taken together suggest that skills are not very transferable and labor markets are therefore relatively thin.

Further, our work is related to a number of other threads in the literature. The empirical content of our paper contributes to the emerging literature on labor markets on the Internet (see e.g. Kuhn & Shen, 2012, Marinescu, 2012, Brenčič & Norris, 2012, Pallais, 2012), which is nowadays the most important channel for recruitment and job search (Barnichon, 2010). From an empirical point of view, our paper is 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. We therefore learn how wages influence the job matching process rather than final matches, and such an analysis yields interesting new insights on the functioning of the labor market.

The theoretical content of our paper adds to the theoretical job search literature in several ways. Closest to our model is work by Guerrieri et al. (2010) and Chang (2012), which we extend to a dynamic model of the labor market. We show that our equilibrium exhibits similar properties, such as distortion of contracts by incentive compatibility constraints to induce separation of types, even if types are observable. In that sense, our model is also related to work by Inderst & Muller (2002) and Lang et al. (2005). While most of the theoretical literature relies on a single dimension of worker heterogeneity, we show how adding a second one can enrich the model and yield predictions that are closer to our empirical work. Having two dimensions of worker heterogeneity allows us to distinguish between a notion of skill or qualification and a notion of productivity or performance on the job given a set of skills. This distinction is important since it is intuitive that skills are not the only determinant of a worker's value to the firm: there exists good and bad nurses and good and bad cashiers, even though nurses have a higher level of skills than cashiers. When recruiting, firms must pay attention to both dimensions of heterogeneity and set their wages accordingly. We show how this richer heterogeneity structure in a model of adverse selection leads to new insights relative to the existing literature.

This paper proceeds as follows. Section 2 describes the data set and documents the key

[1] Bó et al. (2012) find a positive but insignificant relationship between wages and applications.

empirical facts. In section 3, we discuss a directed search model with two-dimensional heterogeneity that captures these facts. Section 4 discusses the interpretation of our empirical and theoretical results. Section 5 concludes.

2 Empirical Analysis

In this section, we discuss the empirical part of our study. We start by describing the data in section 2.1, before presenting the results in section 2.2 and discussing robustness checks in section 2.3.

2.1 Data

We use proprietary data provided by CareerBuilder.com, the largest US employment website. Some background work (data not shown) was done to compare job vacancies in CareerBuilder.com with data on job vacancies in the representative JOLTS (Job 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 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 the vacancies on CareerBuilder are not perfectly representative of the ones in the US economy as a whole, they form a substantial fraction of the market.

Our main data set contains all job vacancies posted on CareerBuilder.com in the Chicago and Washington DC Designated Market Areas (DMA) in January and February 2011. A DMA is a geographical region set up by the A.C. Nielsen Company and 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. For each vacancy, we observe the following characteristics: the job title, the salary if specified[2], whether the salary is by hour or by year, the education required, the experience required, the name of the firm, and the number of days the vacancy has been posted for. We normalize all salaries to be expressed in yearly amounts, assuming a full-time work schedule. When a salary range is provided, we take the middle of the interval. The job title is the title of the job posting, as freely chosen by the firm: this is something like "senior accountant". Because job titles are not normalized, there are many unique job titles. We did some basic cleaning to make job titles more comparable, the most important of which was to put every word in lower case and get rid of punctuation signs. We also determined that the first three words are generally the most

[2]We discuss the issue of firms not posting wages in section 2.3.

important ones, so our preferred specifications use the first three words of a job title, but we do extensive robustness checks to assess the consequences of this word restriction[3]. Based on the full content of the job posting, an internal CareerBuilder algorithm assigns an O*NET SOC (Standard Occupational Classification)[4] code to the job posting. Additionally, based on the firm's name, CareerBuilder uses external data sets like Dun & Bradstreet to retrieve the NAICS (North American Industry Classification System) industry code and the number of employees of the firm. This large number of vacancy characteristics makes the CareerBuilder data an attractive source of information compared to existing datasets.

In addition to the vacancy characteristics, we also observe several outcome variables for each vacancy. A worker who searches for a job will typically do so by specifying one or two keywords and a location. CareerBuilder then shows 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, DMA and the name of the firm. Our first outcome variable, the number of views, represents the number of times that a job appeared in a listing after a search. 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. This number of clicks is our second outcome variable. 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 a job ad. From these numbers on job views, clicks and applications, we construct two new variables: the number of applications per 100 views, which is our key outcome of interest, and the number of clicks per 100 views, which we use for robustness checks. We use applications and clicks per 100 views as our main outcome because, in as much as we are interested in workers' choices among known options, we want to correct for heterogeneity in the number of times a job appears in a listing.

In addition to this first dataset on vacancies, we have a second dataset containing a random sample of jobs from the Chicago and Washington DC DMAs in January and March 2011. This data contains the same information as above, but also includes job seeker characteristics, and in particular measures that can serve as proxies for their productivity. Specifically, we observe for those workers their education level (if at least an associate degree) and their amount of general work experience (in bins of 5 years). We will use these job seeker characteristics to analyze the quality of the applicant pool that a firm attracts.

Table 1 shows summary statistics. The average job ad receives almost 6 clicks and a bit more than one application per hundred views. The average yearly salary is $57,323; this number is somewhat higher than the $45,230 US average wage in 2011 (BLS Occupational Employment Statistics). This wage number is obtained after we cleaned the data by removing the bottom and top 0.5% of salaries to eliminate outliers and errors.

[3]See section 2.2 and 2.3 for a more detailed discussion.

[4]See http://www.onetonline.org. We henceforth refer to this classification simply as SOC.

2.2 Empirical Results

We start with examining the association between log wages and the number of applications per 100 views (table 2). Column I presents the simplest possible specification without any controls. In this specification, 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.[5] Clearly, caution is required in interpreting this unexpected result (fewer applicants to higher-paying jobs) since we ignore heterogeneity by not including controls.

To assess to what degree the negative relationship between wages and applications can be explained by a failure to control for relevant variables, subsequent columns in table 2 add a number of controls for job characteristics. In column II, we add the required level of education and experience for the job, and industry and detailed occupation fixed effects (595 six-digit SOC codes). A priori, this should control for most heterogeneity and allow us to compare very similar jobs. Yet, we still get a negative and significant association between the wage and the number of applicants, with a point estimate that is essentially unchanged. In column III, we add firm fixed effects instead of SOC fixed effects to check whether firm heterogeneity explains away the negative relationship between wages and applications. However, the coefficient on the wage remains unchanged when adding firm fixed effects. The key lesson from columns I-III is that there exists a strong negative correlation between the posted wage and the number of applications, even when controlling for many observables. Remarkably, the magnitude of the coefficient on wages is fairly insensitive to the addition of controls, suggesting that the negative association between wages and applications is robust.

Recognizing that SOC codes may not fully capture job heterogeneity, we control for job title fixed effects (column IV). This should allow us to estimate the relationship between wages and applications among even more homogenous groups of jobs. Since firms can freely choose the job titles for their positions, many different ones exist; 4875 in our sample. While adding all these job titles decreases the risk of omitted variable bias, it also reduces the power of statistical tests. Inspection of the job titles revealed that the fourth word is often an indication of geography, such as "business development manager washington" or "customer service representative fairfax". For the estimation in column IV of table 2, we therefore truncate the job title provided by the firm to the first three words[6], reducing the number of unique job titles to 4371. We also omit the SOC codes to avoid multicollinearity.[7] Interestingly,

[5] 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.

[6] In table 3, we will show that using more than the first three words of the job title does not affect the estimated impact of wages on applications.

[7] While there are on average more than 8 job titles per SOC, jobs with the same job title rarely have different SOCs. This is because the SOC code is inferred from the full text of the job post including the job title by a CareerBuilder algorithm. Although two jobs with the same job title could in principle be classified as belonging

controlling for job title fixed effects (finally) reverses the relationship between wages and the number of applications that we have established so far: the coefficient on the wage is now positive and significant and almost as large in absolute value as in column II when we controlled for SOC fixed effects. Finally, in column V, we add both job title and firm fixed effects, which should essentially absorb all of the firm-side heterogeneity in the data. Column V shows that within essentially identical jobs, higher wages are indeed associated with more applicants: the point estimate implies that a 10% increase in the wage is associated with a 7.4% increase in applicants.

We summarize our findings about the relationship between wages and applications as follows:

Empirical Result 1. *Across job titles, vacancies that offer higher wages receive fewer applications.*

Empirical Result 2. *Within a job title, vacancies that offer higher wages receive more applications.*

To analyze how our results depend on the definition of a job title, table 3 shows some alternative specifications. In columns I-V, we investigate the sensitivity of the effect of wages on applications to using fixed effects for the first n words of the job titles, for various values of n. The first column repeats the specification in table 2, col. IV, but instead of using fixed effects based on the first three words of the job title, we include fixed effects for the first word of the job title. Subsequent columns (II - V) use increasingly larger parts of the job title. Column III corresponds to column IV of table 2. The trade-off between omitted variable bias and loss of power is clearly visible. When only the first or the first two words are used (column I and II), the standard error of the wage is relatively small, but so is the amount of variation in the number of applications explained by the model. In addition, the estimated wage coefficient is negative. All this changes when the third word is included, as is demonstrated in column III. Adding the fourth or all remaining words (column IV and V) has virtually no effect on the estimated coefficient or on the R^2. Instead, it only increases the standard error of the estimated wage coefficient, thereby reducing the significance level. Summarizing, these results confirm that the first three words of a job title contain important information that is not captured by six-digit SOC codes.

To better understand what information job titles add relative to SOC codes, the last column of table 3 presents the results of a different exercise, in which we repeat the SOC fixed effects specification from table 2, col II, and we additionally include separate dummy variables for each of the words appearing in a job title truncated to its first three words. Hence, in this specification, we ignore the order and the combinations in which the words appear in the

to different SOCs if the full text of the posts differs significantly, this is uncommon: in practice, the average number of SOCs per job title is 1.4, with a median value of 1.

job title. This is arguably somewhat restrictive, because it assumes that - for example - the word ‘assistant’ has the same effect every time it appears.[8] Not surprisingly, the estimated wage coefficient is smaller than when controlling for job title fixed effects (col. III-V), and is not significant in this specification. Since the inclusion of these word dummies makes the significant and negative relationship between wages and applications disappear, it suggests that individual words in the job title have explanatory power beyond SOC codes. At the same time, the explanatory power of SOC fixed effects and job title word fixed effects taken together is lower than the explanatory power of three-word job title fixed effects.

A benefit of the word fixed effect specification in col. VI is that it allows us to determine which specific words in the job title have explanatory power beyond the detailed SOC codes. The most frequent of these words are listed in table 4. We manually[9] classified these words in three clusters of meaning, using the actual job titles in which these words appear as a guide. Our classification reveals that there are a couple of reasons why job titles are more precise than detailed SOC codes. First, job titles indicate hierarchy or experience level. For example, a job title may say “registered nurse supervisor” or “senior accountant”, even though the first will still fall within the registered nurse SOC (29-1140) while the second will fall within the general SOC for accountants and auditors (13-2010). These kinds of words are listed in the second column of 4. The second reason why job titles are more precise is that SOC codes do not account well for detailed specializations or areas of work. The third and fourth columns list these types of “specialization” words: in the third column, there are a number of different specialties, while the fourth column concentrates specifically on computers or IT. For example, the word “and” is significant because it appears in job titles that require at least two different areas of specialization or skill. In the realm of IT, many words are significant, suggesting that SOC codes do not classify IT jobs at a fine enough level of detail: for example, there is an SOC code for programmers, but Java programmers are a distinct and relevant specialization.

Summarizing, the results in tables 3 and 4 show that job titles are able to capture a large amount of job heterogeneity ignored by six-digit SOC codes, and in particular heterogeneity in seniority/hierarchy and specialization. The analysis of the role played by words in job titles also illustrates the power of using free-form Internet text data to understand 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.

[8] Examples of such jobs include ‘executive assistant’, ‘assistant store manager’, and ’assistant professor’.

[9] We have also used an external dictionary (http://words.bighugelabs.com) to classify words with similar meanings. However, the results of this alternative classification proved to be pretty useless because words that are synonyms in general are often not synonyms in the specific job vacancy context. For example, the external dictionary says that distribution, government and system are synonyms. However, in our job vacancy data, these words are not synonyms because distribution and government are two very different industries, and systems is a computer science term (like in “Windows systems engineer”) that has nothing to do with either the government or distribution industries.

Having established that higher wage jobs attract more applicants within a job title, we now consider whether these jobs also attract higher quality applicants. We use two different measures of the quality of a job's applicant pool: 1) the average level of relevant work experience among applicants (conditional on applicants having at least an associates' degree), and 2) the average education level among applicants expressed in years of education; and. Table 5 displays the results. We find that higher wages are associated with significantly better applicants in terms of experience and education and that this result is robust to the addition of controls, including the job title fixed effects. A 10% increase in the wage is associated with an increase in the experience of the average applicant by 0.15 years, or roughly 1.1% (column II). Similarly, a higher wage significantly increases the average amount of education of the applicants (column III and IV). Overall, we conclude that higher wages attract higher quality applicants, and this relationship is robust to the addition of a large set of controls, including job title fixed effects.

We summarize our findings about the relationship between wages and applicant quality below:

Empirical Result 3. *Across job titles, vacancies that offer higher wages get higher quality applicants.*

Empirical Result 4. *Within a job title, vacancies that offer higher wages get higher quality applicants.*

2.3 Robustness

After presenting the main results, we now turn to some robustness checks to rule out other potential explanations for the patterns we find. We subsequently consider biases that may arise from the definition of some of our key variables, omitted variables, and sample selection.

The first issue that we address is the definition of our dependent variable. 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 2 are qualitatively unaffected by this alternative definition and conclude that our results are robust.

A second issue is that 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 salary, 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 6 explores the relationship between wages and clicks. When no controls are used (col. I), we see a significant and negative association between the wage and clicks per 100 views. When controlling for basic job characteristics (vacancy duration, dummy for salary expressed by hour, DMA and calendar month), firm fixed effects, and job title fixed effects, the coefficient on the wage becomes positive and highly significant, 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 2 can be reproduced for clicks per view, an outcome whose determinants are fully known, makes us more confident about our basic results. A higher wage is generally associated with fewer clicks and 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 that our key results are not driven by omitted variable bias.

The third potential issue is that many jobs do not post wages, meaning that the relationship we estimate is based on a selected sample of jobs that do post a wage[10]. One important reason for the wage not being posted is the use by many companies of Applicant Tracking Systems (ATS) software that keeps track of job postings and applications. This software also sends out the job posting to online job boards such as CareerBuilder. Before sending out the job posting, ATS software typically removes the wage information, even if it was provided by the firm. The use of ATS is likely to be an important explanation for the absence of a posted wage, because about two thirds of jobs are posted through ATS software and this proportion is similar to the proportion of jobs without a posted wage[11].

To assess the extent to which our estimates of the impact of the wage on applications and applicant quality is affected by selection bias, we examine whether jobs with a posted wage get more or better applicants than jobs without a posted wage. One way of testing this idea is presented in table 7, in which we examine the relationship between posting a wage and the number of applicants. We find that jobs with a posted wage get a larger number of applicants, but this relationship becomes insignificant when controlling for both job title and firm fixed effects (col. III). Since ATS use is typically determined at the firm level and seems responsible for the non-posting of the wage, it makes sense that the impact of posting a wage is wiped out after controlling for firm fixed effects.[12] Hence, jobs with a posted wage do not get significantly

[10] Further evidence on the incidence of wage posting is available in Hall & Krueger, 2012. In a survey of US workers, 31% of the respondents answered affirmatively to the question of whether they had exact knowledge of their pay before interviewing for their job. However, this does not tell us exactly how frequently job ads contain wage information, for multiple reasons. First, the worker is the unit of observation in their survey, not the firm. Second, the worker may have learned the pay from a different source than the job posting. Third, it is not clear how respondents answer this question if the job ad specified a wage range.

[11] Private communication with CareerBuilder.com.

[12] We performed the same robustness check with clicks per view as the dependent variable and the results are the same: with job title and firm fixed effects, no significant impact of posting a wage exists.

more applicants when controlling for a large number of observables.

Another way of testing this idea is presented in table 8, in which we examine the relationship between posting a wage and the quality of applicants in terms of education and experience. We do not find a relationship between posting a wage and the quality of applicants after controlling for job title fixed effects (cols. II and IV). Overall, we conclude that, as a group, jobs without a posted wage are not different from jobs with a posted wage once we condition on observables. We speculate that jobs without a posted wage are probably able to signal their salary through other means. Another possibility consistent with this pattern is that jobs without a wage offer a roughly average pay, and job candidates correctly infer this when they do not see a posted wage. In both cases, we think that the existence of many jobs without a posted wage is unlikely to bias our results about the relationship between the level of posted wages and the number and quality of applicants.

A fourth and last issue is that sample selection may explain why the relationship between wages and the number of applicants switches sign when using job title fixed effects. After all, with job title fixed effects, the effect of the wage is identified off the job titles with at least two observations. To assess this, we re-estimate the specification with SOC codes instead of job title on a restricted sample with at least two observations per job title. Again, a significantly negative relationship between wages and the number of applicants arises, as shown in table 9. Hence, sample selection due to job titles does not drive our results.

In conclusion, we have found that higher wages are associated with higher quality applicants across the board. Moreover, higher wages are generally associated with fewer applicants. However, within job titles, higher wages are associated with more applicants. We have examined a number of caveats that could affect these results, including the definition of variables, omitted variable bias and sample selection bias and have found that our results are quite robust to these sources of bias.

3 Theory

In this section, we show that the patterns that we find in the data are consistent with a directed search model in which firms with vacancies post wages to attract applications from heterogeneous workers. To explain the empirical patterns both within and across job titles, we extend the existing literature by developing a model in which workers differ in not one but two dimensions. These dimensions are 1) the type of job that they can do, and 2) their productivity in that job. While the model allows for a continuum of types, all its relevant properties hold when there are as few as two types in each dimension, e.g. bad nurses, good nurses, bad cashiers, and good cashiers. We will generally refer to this example to convey the intuition behind the results.

We show that the equilibrium needs to satisfy incentive compatibility constraints to induce separation of worker types. Without incentive compatibility, adverse selection would arise with low-productivity workers applying to high-wage jobs. After deriving the equilibrium, we compare the model's predictions with the empirical findings. This comparison yields several new predictions, including 1) a job type in the model (e.g. a nurse or cashier) can be interpreted as a job title in the data, but not as a six-digit SOC; and 2) the value of the match surplus increases faster with skill than the vacancy creation cost.

3.1 Setting

Consider the steady state in an economy in continuous time with a mass 1 of workers and a positive mass of firms, determined by free entry. Workers and firms live forever, are risk-neutral, and discount the future at rate $r > 0$. Each worker supplies one indivisible unit of labor and each firm has one position, which can be filled by at most one worker. Many different types of jobs exist and each type of job produces a different consumption good. We first discuss workers' characteristics, followed by firms' characteristics, and finally the matching technology.

Workers are heterogeneous in the type of job that they can do, as well as in the amount of output that they produce in that job. We call the type of job that a worker can do his 'skill' and assume that it can be represented by a single index x; nurses and cashiers are different along this dimension. The amount of flow output that a worker produces in job x is called his 'productivity' y; this is related to how good or bad a nurse someone is. Workers are characterized by their type (x, y), drawn from an exogenous distribution $F(x, y)$ when they enter the market for the first time. A worker's type stays constant throughout his career. To simplify exposition, we assume that $F(x, y)$ has full support on $\mathcal{X} \times \mathcal{Y} \equiv [\underline{x}, \overline{x}] \times \left[\underline{y}, \overline{y}\right] \subset (0, \infty)^2$ and that a continuum of workers of each type exists.[13] We will occasionally focus on the limit case in which the heterogeneity in productivity y vanishes (i.e. $\underline{y} \to \overline{y}$), to analyze the scenario in which job types provide a detailed classification of the labor market and the difference in output produced by a good nurse and a bad nurse is small relative to the difference between a a nurse and a cashier. In general, our model with two dimensions of skill heterogeneity will generate distinctly different patterns compared to standard models with one dimension of heterogeneity.

An employed worker of type (x, y) produces good x (corresponding to worker skill x), which is sold at the exogenously given price $p(x)$. This price $p(x)$ is increasing in x. Therefore, the value of the output created by a worker who creates y units (y reflects worker productivity) of good x equals $p(x)\, y$. The worker gets a flow payoff equal to his wage w, while the firm keeps the remainder $p(x)\, y - w$. Steady state unemployment is generated by job destruction shocks

[13] The full support assumption is not essential for any of the results and is rather weak, since the density of workers of a particular type can be arbitrarily close to zero. A continuum of workers of each type is helpful since it allows us to apply standard large-market results.

which destroy existing matches at a rate $\delta > 0$. Unemployed workers obtain a flow payoff $b(x)y$, consisting of unemployment benefits, household production and/or the value derived from leisure.[14] We assume that $b(x)$ is weakly increasing in x and is strictly smaller than the output price $p(x)$ for all x in order to rule out structural unemployment. While the payoff of an unemployed workers is increasing in his skill and productivity by assumption, we will later show that in equilibrium the same holds for the wages of employed workers.

Firms choose the good x that they wish to produce (their 'job type') when they enter the market and create a vacancy. This decision is irreversible. In order to find a worker for their vacancy, firms post job ads when they enter the market. These job ads specify both the job type x and the firm's wage offer w. The firm commits to the wage offer as well as to not hiring workers who cannot produce the required good (i.e. have the wrong x).[15] The firm cannot commit to only hire a particular productivity type y and is willing to hire any worker that provides a higher payoff than continued search. Hence, we assume that workers' characteristics are observable, but the equilibrium analysis will reveal that this is not crucial; the same equilibrium arises if y is unobservable.

Firms incur a flow cost $c(x)$ while having a vacancy, which is increasing in skill x. The cost $c(x)$ may include a fixed component which is independent of x, such as administrative costs or the cost of posting a job ad on the career website. The variable component may reflect the cost of labor involved in recruitment or the cost of acquiring the technology required for production. Importantly, we assume that the surplus $[p(x) - b(x)]\, y$ increases faster than the vacancy creation cost $c(x)$, i.e. $\frac{d}{dx}\frac{p(x)-b(x)}{c(x)}y > 0$. Firms optimally choose what type of vacancy to create and what wage to post given the vacancy creation costs and the expected surplus from a realized match.

All job ads are posted in a central location (the employment website), where they can be observed by workers at zero cost. Hence, search in this economy is directed and unemployed workers decide to which type of job they wish to apply.[16] The matching process is subject to frictions and the number of matches that are formed at a particular job type x and wage w is determined by a matching function. As standard in the literature, we will consider a Cobb-Douglas matching function exhibiting constant returns to scale.[17] As a result, the matching rates solely depend on the ratio $\lambda(x, w)$ of applicants to vacancies (the 'applicant-vacancy ratio') at a job type x and wage w. We generally omit the arguments x and w to simplify

[14]Note that scenarios in which the payoff is independent of skill, i.e. $b(x) = b_0$, or in which the payoff is proportional to market productivity, i.e. $b(x) = b_1 p(x)\, y$, are special cases of this formulation.

[15]We consider the case in which workers can create output in different job types in section 3.4.

[16]We abstract from on-the-job search, but discuss in section 3.3 that this does not affect the qualitative results.

[17]See Petrongolo & Pissarides (2001) for a survey of the literature on the matching function. They conclude: "The stylized fact that emerges from the empirical literature is that there is a stable aggregate matching function of a few variables that satisfies the Cobb-Douglas restrictions with constant returns to scale in vacancies and unemployment." Rogerson et al. (2005) provide a theoretical overview of models featuring a matching function.

notation. Given an applicant-vacancy ratio λ, firms match at a Poisson rate $m(\lambda) = A\lambda^{\alpha}$ for $A > 0$ and $\alpha \in (0, 1)$. For future reference, note that this implies $m'(\lambda) > 0$ and $m''(\lambda) < 0$. Correspondingly, workers match at a Poisson rate $\frac{m(\lambda)}{\lambda} = A\lambda^{\alpha-1}$, which is decreasing in λ. In other words, the probability that a firm matches increases in the applicant-vacancy ratio, while the probability that a worker matches decreases in this ratio. Both firms and workers will take the matching rate into account when determining their equilibrium strategies.

3.2 Equilibrium

In this subsection, we will analyze the workers' and firms' optimal strategies and derive the equilibrium in the economy. We will first rule out the existence of a pooling equilibrium in which multiple types of workers apply to the same firm. Subsequently, we will characterize the separating equilibrium.

As standard in directed search models, workers and firms face a trade-off between matching probability and match payoff: a high wage provides the worker with a high payoff in case of a match, but attracts - ceteris paribus - a lot of applications, which implies a low matching probability for the worker. Symmetrically, a low wage provides firms with a high payoff if they match, but at the cost of a lower matching probability. In addition, firms care about the type of worker that they attract. Given these trade-offs, workers and firms decide at which combination of x and w they want to match.

First, consider the choice of the job type. A worker's choice regarding the type of job for which he wants to search is trivial, since we have assumed that he can only work in one particular job type. A firm can create a vacancy in any job type, but once it has chosen a particular job type x, its profit is independent from the measure of workers and vacancies in other job types. We can therefore first analyze the sub-market formed by workers and firms at a particular job type x in isolation, after which the economy-wide equilibrium follows immediately. Proofs are relegated to the appendix.

Second, within a job type, workers with different productivity levels y compete for the jobs that are posted by firms. As a first result, we show that any two workers who differ in their productivity will not apply to the same job in equilibrium. That is, good and bad nurses will direct their applications to different positions. The intuition for this result is the following. Although all workers care about both wages and matching probabilities, low-productivity and high-productivity workers have different marginal rates of substitution (MRS) between these two factors, because they differ in their outside option $b(x)\, y$. Low-productivity workers have a worse outside option and care therefore at the margin more about matching probabilities (as opposed to wages) than high-productivity workers. Using this fact, we will now show that it is not possible to have an equilibrium where low and high types apply to the same wage.

Consider a situation in which all low-productivity (i.e. low y) and high-productivity workers apply to the same wage and a deviating firm posts a slightly higher wage. Now, suppose that the deviant attracts a slightly higher number of applicants such that low-productivity applicants are exactly indifferent between the deviant and the other firms. In this case, high-productivity workers strictly prefer applying to the deviant high-wage firm. Compared to low-productivity applicants, high-productivity workers value the higher wage more than they are hurt by the larger number of competing applicants. This will increase the number of applicants at the deviant further, and low-productivity workers will ultimately decide to stay away. Hence, all applicants at the deviant will have high productivity, increasing its payoff in a discrete manner compared to the marginal increase in the wage offer, and ultimately making the deviation profitable. The following lemma formalizes this.

Lemma 1. *There exists no equilibrium in which a firm posts a job ad (x, w) and attracts workers of both types (x, y_1) and (x, y_2), for any $y_1, y_2 \in \mathcal{Y}$, and $y_1 \neq y_2$.*

Instead, different worker types must be separated in equilibrium. In terms of our example, some firms post high wages and only attract good nurses, while other firms post low wages and only receive applications from bad nurses.

In order to sustain such an equilibrium, two incentive compatibility constraints must be satisfied: the good nurses must not want to apply to the jobs aimed at the bad nurses, and vice versa. One can show that while the incentive compatibility constraint for the good nurses is automatically satisfied, the incentive compatibility constraint for the bad nurses binds. In order to keep bad nurses away from the jobs for good nurses (i.e. prevent adverse selection), the matching rate at these high wage jobs must be sufficiently low, such that the bad nurses - who care relatively more about matching probability - prefer the low-wage jobs with higher matching probability. This incentive compatibility constraint for bad nurses increases the wage of good nurses and decreases their matching probability relative to a world without bad nurses.

So far, we have only distinguished between low-productivity and high-productivity workers. Of course, with a continuum of productivity types, not two but a continuum of wages will be offered. Each wage w attracts a particular productivity type y and a particular applicant-vacancy ratio λ, determined by the free-entry condition. Each combination of w, λ, and y must satisfy the incentive compatibility constraint for all other worker types. As in the two productivities case, the sub-market for the lowest productivity type is undistorted (i.e. the same as in a world without other types). Incentive compatibility constraints then determine how quickly λ increases as a function of y for the remaining sub-markets. The following lemma formalizes this.

Lemma 2. *In any equilibrium, a unique set of wages is posted within each type of job x. Each wage attracts workers of a particular productivity type y. The applicant-vacancy ratio for the* *least productive workers is determined by the unique solution to*

$$\underline{y} = \frac{r + \delta + m'(\lambda)}{m(\lambda) - \lambda m'(\lambda)} \frac{c(x)}{p(x) - b(x)}, \tag{1}$$

while the applicant-vacancy ratios for the remaining types are determined by the differential equation

$$\frac{d\lambda}{dy} = \frac{1}{r + \delta} \frac{[\lambda(r + \delta) + m(\lambda)] m(\lambda) p(x)}{[m(\lambda) - \lambda m'(\lambda)](p(x) - b(x)) y - [r + \delta + m'(\lambda)] c(x)}. \tag{2}$$

The corresponding wages follow from

$$w = p(x) y - \frac{(r + \delta) c(x)}{m(\lambda)} \tag{3}$$

and a worker's value of unemployment is given by

$$rV_U(y) = \frac{\lambda(r + \delta) b(x) y + m(\lambda) w}{\lambda(r + \delta) + m(\lambda)}. \tag{4}$$

In this lemma, equation (3) specifies the relation between the wage and the applicant-vacancy ratio in an arbitrary sub-market as implied by the free-entry condition. Equation (1) defines the solution in the (undistorted) sub-market for the lowest productivity type. Finally, equation (2) specifies how fast the applicant-vacancy ratio must increase across the sub-markets to satisfy the incentive compatibility constraints.

>From this lemma, it is only a small step to the economy-wide equilibrium. We therefore omit the proof of the following proposition which establishes the existence of a unique market equilibrium.

Proposition 1. *A unique market equilibrium exists. The equilibrium satisfies lemma 2 for each job type x.*

To summarize, we can treat each skill x as a separate job market. Within each skill, firms post a continuum of wages corresponding to the heterogeneity in worker productivity y. A separating equilibrium arises in which high-wage jobs attract high-productivity applicants and low-wage jobs attract low-productivity applicants.

3.3 Empirical Content

The simple model presented above provides several testable predictions regarding the relationship between productivity, wages, the number of applications and payoffs. In this subsection, we discuss a few of these predictions and show that they match the empirical facts obtained in the data. Proofs are again relegated to the appendix.

The first prediction relates wages and applicant quality within a job type. Within a particular job type x, various wages are being posted with low wages attracting applications from low-productivity workers and high wages attracting applications from high-productivity workers. This immediately yields a positive relationship between productivity and wage in each job type.

Model Prediction 1. *Within a job type, a positive correlation exists between the wage that a firm posts and the productivity of the applicants to its vacancy.*

The second prediction concerns the relationship between the wage and the number of applicants within a job type. The number of applicants that a firm attracts is ultimately determined by the free entry condition:

$$m(\lambda)\frac{p(x)\,y-w}{r+\delta}-c(x)=0$$

Since the entry cost is the same for all firms in a job type x, firms must obtain the same expected payoff for a vacancy in equilibrium, so that $m(\lambda_1)(p(x)\,y_1-w_1)=m(\lambda_2)(p(x)\,y_2-w_2)$ for any $y_1, y_2 \in \mathcal{Y}$ and the corresponding wages and applicant-vacancy ratios. Hence, there must exist a negative relationship between a firm's matching probability (or its number of applicants) and its payoff from a match. Firms attracting high-productivity workers create more output ($y_2 > y_1$) but also pay higher wages ($w_2 > w_1$) than firms with low-productivity workers. In the appendix, we show that the latter effect dominates, i.e. wages increase faster than output, such that $\lambda_1 < \lambda_2$. In other words, since the difference between wages and output is smaller in jobs targeting high-productivity workers, firms receive a lower flow payoff after hiring these workers. So, to make firms indifferent between low-productivity and high-productivity workers, it must be easier to match with high-productivity workers. Hence, a vacancy targeted at high-productivity workers attracts in expectation more applications.

Model Prediction 2. *Within a job type, a positive correlation exists between the wage that a firm posts and the number of applications it receives.*

This positive equilibrium relationship between wages and applications reflects the way in which the labor market prevents adverse selection of low-productivity workers into jobs for high-productivity workers within a job type (see the previous subsection for a discussion of the incentive compatibility constraints). If adverse selection were not a concern, a negative relationship between wages and applications would arise in each job title, as shown in the proof of lemma 2.[18] Empirically, we find a positive relationship between wages and applications

[18] The intuition is that in such a world the applicant-vacancy ratio λ determines which fraction of the surplus created by a match goes to the firm. If λ were constant across y, this fraction would be constant. Since the created surplus is larger for larger values of y, firms attracting high-productivity workers would obtain a higher payoff. This violates the free-entry condition. Additional entry would take place in those sub-markets, reducing applicant-vacancy ratios. For additional details, see the proof of lemma 2.

within job title. This result suggests that firms face adverse selection of workers within a job title.

Summarizing, the model yields two predictions concerning the relationship between wages and the number and quality of applicants within a job type: a higher wage is associated with more and better applicants. These predictions line up when a job type in the model is interpreted as a job title in the data: a positive relationship was found in the empirical analysis between wages and the number of applications, as well as between wages and indicators of the worker's productivity such as education and experience. By contrast, the predictions of the model do not hold in the data if we assume that a job type is a six-digit SOC code: empirically, there is a negative relationship between wages and the number of applicants within six-digit SOC codes. Hence, the data and the model jointly suggest that a job type corresponds to a job title rather than a six-digit SOC code.

Having explored the model's predictions within job type, we now consider the model's predictions across job types. First, we analyze the relationship between wages and skill. As mentioned before, we focus on the case in which the degree of heterogeneity in productivity y is small relative to the heterogeneity in skill x. Here, we start with the limit case in which there is no heterogeneity in productivity, i.e. $\underline{y} = \overline{y} = y$. One can then show that wages increase with skill x (i.e. nurses earn more than cashiers) if the value of output $p(x)$ increases faster than the vacancy creation cost $c(x)$, as we assumed. Hence, without productivity heterogeneity, jobs that pay more attract applications with a higher level of skill.

With heterogeneity in productivity, the relationship between wages and skill no longer holds one-to-one, since good cashiers may earn more than bad nurses. However, the positive relationship between wages and skill levels survives as long as the degree of heterogeneity in productivity y is sufficiently small, i.e. $\underline{y}$ is sufficiently close to $\overline{y}$. Hence, we can derive the following empirical prediction.

Model Prediction 3. *Across job types, a positive correlation exists between the wage that a firm posts and the skill of its applicants.*

Second, we consider the relationship between wages and the number of applicants, initially omitting heterogeneity in productivity, i.e. $\underline{y} = \overline{y} = y$. In that case, one can show that w and λ are negatively related, i.e. vacancies for nurses receive fewer applications than vacancies for cashiers. The main intuition is as follows: if there were as many applications per vacancy for nurses as for cashiers, firms posting vacancies for nurses would make a higher profit. This cannot be an equilibrium, given free entry of firms. Instead, it has to be that nursing jobs get fewer applications (nurses are harder to recruit), so that profits are equalized between both types of jobs.

Mathematically, lemma 2 reveals that, without heterogeneity in productivity y, a firm's payoff is equal to a fraction of the surplus $[p(x) - b(x)]\, y$, minus the vacancy creation cost.

To be precise, given free entry, we have that:

$$V_V = \frac{m(\lambda) - \lambda m'(\lambda)}{r + \delta + m'(\lambda)} \left[p(x) - b(x)\right] y - c(x) = 0.$$

The fraction of surplus that goes to the firm depends on the number of applicants λ that it attracts. If the number of applicants were constant across job types x (i.e. nurse jobs receive as many applications per vacancy as cashier jobs), the fraction of the surplus going to the firm would be constant as well. Since by assumption the surplus $[p(x) - b(x)]\,y$ increases faster in skill x than the vacancy creation costs, with an equal number of applicants across job types, firms posting a high-skill job (e.g. nurse) would make a larger profit than firms posting a low-skill job (e.g. cashier). In equilibrium, firms must get equal payoffs across job types, so firms must pay a higher fraction of the surplus to workers in high-skill jobs, which is the case if the competition for workers is high and there are few applications, i.e. λ low. Hence, without heterogeneity in productivity, a negative relationship arises in equilibrium between the wage that a firm posts and the number of applicants that it attracts.

With heterogeneity in productivity y, the relationship between wages and applicants is no longer straightforward, in particular because of the positive relationship between wages and applicants within a job type. However, as long as the degree of productivity heterogeneity is sufficiently small, this effect is dominated and the negative correlation between wages and applications survives. The model therefore yields the following prediction.

Model Prediction 4. *Across job types, a negative correlation exists between the wage that a firm posts and the number of applications that it receives.*

The two predictions of the model across job types again line up perfectly with the empirical results: if we do not control for job heterogeneity by including job title fixed effects, we find a significant negative relationship between wages and the number of applications (Prediction 4), and a significant positive relationship between wages and indicators of skill (Prediction 3), such as education and experience.

3.4 Transferable Skills

In this section, we relax the assumption that a worker with a particular skill set x can only produce output in one type of job. Instead, we allow him to also be productive in other job types. We show that the equilibrium described above survives as long as the transferability of skills across job types is not too high. To simplify the exposition, we will again abstract from heterogeneity in productivity, by assuming that $\underline{y} = \overline{y} = y$.

When workers can produce output in more than one type of job, they may be inclined to apply to jobs that do not perfectly match their skill but offer higher wages or lower applicant-vacancy ratios. Whether this will occur in equilibrium depends on the extent to which firms are willing to hire them for those jobs, which in turn depends on their productivity in those jobs. Hence, we need to specify how productive someone trained to be a cashier would be as a nurse relative to the productivity of someone with a nursing degree, and vice versa. In other words, we need to specify how transferable skills are across different types of jobs.

Note first that given the nature of the baseline equilibrium described in proposition 1, we do not need to consider downward deviations in application behavior, i.e. applications to jobs that require less skill than the worker possesses. Since wages are lower and applicant-vacancy ratios are higher in low-skill jobs, the worker will never find such a deviation profitable. This is true even if the worker's skills are perfectly transferable and he would create the same amount of output y as someone trained to do those lower-skill jobs. On the other hand, since wages are higher and applicant-vacancy ratios shorter in higher-skill jobs, upward deviations (applications to jobs that require more skill than the worker possesses) are clearly profitable if skills are perfectly transferable. In order to maintain the baseline equilibrium, we therefore need an upper bound on the degree of skill transferability across jobs, and this upper bound has to imply less than full transferability. There are various ways in which one can formalize how a worker's productivity may decrease if he works in jobs for which he does not have the required skill. We consider two of the more natural ways to formalize skill transferability and show that qualitatively they give the same result.

The first approach is to assume that workers incur a deterministic penalty in their productivity when working in jobs that do not match their skill, where the magnitude of the penalty depends on the distance in skill level. Consider a worker of type x_i who, instead of applying to a wage w_i with an applicant-vacancy ratio λ_i at his own skill level, applies to a job of type $x_d > x_i$ with wage w_d and applicant-vacancy ratio λ_d. If this worker gets the job x_d, he will produce $\theta(x_d, x_i)\, y$ units of output, where $\theta(\cdot)$ captures the degree of transferability. It equals 1 for $x_d = x_i$ and is is strictly decreasing in the distance between the job types, i.e. $\frac{\partial \theta}{\partial x_d} < 0$ and $\frac{\partial \theta}{\partial x_d} > 0$ for all $x_d > x_i$.[19] The firm posting the vacancy x_d is willing to hire this worker upon meeting as long as $p(x_d)\, \theta(x_d, x_i)\, y > w_d$, i.e. the value of this worker's output is higher than the wage. The baseline equilibrium therefore survives if

$$\theta(x_d, x_i) < \frac{w_d}{p(x_d)\, y}$$

for all $x_d > x_i$ and associated w_d. The right-hand side of this condition represents the share of the surplus going to workers in jobs of type x_d, which we have shown to be increasing in x_d in equilibrium. Since $\theta(x_d, x_i)$ is assumed to be decreasing in x_d, this implies that the condition is satisfied for all x_d if it is satisfied for $x_d \to x_i^+$, i.e.$\lim_{x_d \to x_i^+} \theta(x_d, x_i) < w_i / p(x_i)\, y$. In other terms, if a worker is not productive enough in a job that requires marginally more skill than his own, then he will for sure not be productive enough in jobs that require an even higher skill

[19] Note that the baseline model corresponds to $\theta(x_d, x_i) = 0$ for all $x_d > x_i$.

level. Quantitatively, since the share of the surplus going to workers is strictly less than one, i.e. $w_i/p(x_i)\, y < 1$, the condition above implies that workers must incur a large productivity loss even in jobs that only require marginally more skill than their own, and will therefore never get hired by these firms. However, assuming that a worker will never get hired in a job that requires just slightly more skill than his own may seem unrealistic. We therefore also consider a second approach.

Our second approach to modeling skill transferability assumes that workers may be productive in jobs with skill requirements higher than their own, but this can only happen with some probability less than one. Specifically, suppose that the worker will produce y units of output with probability $\tau(x_d, x_i)$ and zero units of output with probability $1-\tau(x_d, x_i)$, where $\tau(\cdot)$ captures the extent to which the worker can transfer his skill from job type x_i to x_d and satisfies the same properties as $\theta(\cdot)$. Given this structure, the firm will hire the worker upon meeting with probability $\tau(x_d, x_i)$, implying a matching rate $\tau(x_d, x_i)\,\lambda(x_d)$ for the worker in this submarket. Hence, whether jobs of type x are attractive to workers of type x_i depends on the shape of $\tau(x_d, x_i)$. The following lemma derives a condition on $\tau(x_d, x_i)$ such that workers only apply to jobs that exactly match their skill and the baseline equilibrium survives.

Lemma 3. *When skills are partially transferable, the baseline equilibrium described in proposition 1 survives if*

$$\tau(x_d, x_i) < T(x_d, x_i) \equiv \frac{\lambda_d (r+\delta)[w_i - b(x_i)\, y]}{\lambda_i (r+\delta)[w_d - b(x_i)\, y] + m(\lambda_i)[w_d - w_i]} \frac{m(\lambda_i)}{m(\lambda_d)} \text{ for all } x > x_i. \tag{5}$$

While the expression for the bound $T(x_d, x_i)$ is not very intuitive, it is straightforward to confirm that $T(x_d, x_i)$ is decreasing in the skill requirement of the alternative job x_d and equal to 1 for $x_d = x_i$. Hence, the condition $\tau(x_d, x_i) < T(x_d, x_i)$ implies that the transferability of skill must decrease sufficiently quickly in x_d. In other words, if a worker considers job titles with higher and higher skill level relative to his own, he finds that his skills are less and less transferable. Such a condition makes sense, as we expect workers to be less and less likely to perform well as they are asked to do tasks well above their qualification. However, contrary to what happens in the deterministic specification of skill transferability we examined above, the worker has a positive chance of getting hired if he (out of equilibrium) were to apply to a vacancy that does not match his skill. For the equilibrium described in proposition 1 to survive when using this second modeling approach, the degree of transferability of skills across any two job titles must be small enough, and it must decline with the difference in skill between these two job titles.

Our baseline model was able to account for the empirical facts by assuming that each worker is only productive in a specific job type x and therefore will only apply to jobs of type x. In this section, we have shown that, if we allow workers to be productive in jobs with skills different from their own, we need a limited degree of skill transferability across job types to account for the empirical results. This suggests that, empirically, skills are indeed not very transferable, and the degree of transferability decreases with the difference in skill level between jobs.

Overall, our model yields predictions that are consistent with our key empirical results. Namely, higher wages are associated with better applicants both within and across job types. Higher wages are associated with fewer applicants across job types, but more applicants within job types. In the next section, we discuss to what extent this model contributes to our understanding of empirical facts, and whether alternative models could also explain our empirical findings.

4 Discussion

Because of the broad range of issues our results speak to, we devote this section to the discussion of our results in a number of contexts. We first discuss the empirical implications of our model. We then analyze how our model can explain the low estimated elasticity of applications with respect to the wage, and how we can understand imperfect competition in the labor market. Finally, we re-evaluate some of the key assumptions of our model and discuss alternative assumptions.

4.1 Empirical Implications of the Model

In order to generate the empirical patterns we uncovered in the data, we made two important assumptions in our theoretical analysis. First, skills are not very transferable across different types of jobs. Second, the surplus created by a match increases faster with skill than the cost of creating a vacancy. We now discuss the empirical relevance of these assumptions.

The central assumption that skills are not very transferable across job requiring different skill levels insures that workers only apply to their own job title, such that job titles constitute closed labor markets in the model. Given our empirical results, this assumption means that workers' skills have limited transferability, even within a six-digit SOC code. This suggests that six-digit SOC codes do not capture skill specificity well enough.

This result has a number of empirical implications. First, the empirical literature in labor economics has traditionally used broader classifications of occupations than six-digit SOC codes (e.g. Kambourov & Manovskii, 2009b, Poletaev & Robinson, 2008).[20] Our results imply that this ignores a significant amount of heterogeneity and, therefore, that occupations are narrower than typically assumed[21]. Second, the fact that there is limited skill transferability

[20] For example, the Dictionary of Occupational Titles - used in the referenced studies - distinguishes between 564 detailed occupations, compared to 840 detailed occupations in the six-digit SOC.

[21] If six-digit SOC codes include too broad of a range of worker skills compared to job titles, this can explain

across jobs with different skill levels has important implications for e.g. the measured degree of occupational mobility, occupational mismatch, frictional wage dispersion, and the cost of job loss. Our results suggest that switching to an even slightly different job likely has high costs, and that search frictions are likely to be very important in these relatively thin markets.

Even though our model and empirical results taken together imply that skills are not readily transferable, workers do occasionally switch occupation in real life, e.g. because they gain experience through learning-by-doing (from junior accountant to senior accountant) or because of occupation-specific productivity shocks (from construction worker to truck driver). Our model does not account for such occupational switches because they are not the focus of our analysis. Instead, the model is designed to understand the labor market prospects of workers with a given skill set at a given moment in time.[22]

Our other central assumption, i.e. that surplus increases faster with skill than the the cost of creating a vacancy, is crucial in yielding a negative relationship between wages and applicant-vacancy ratios across job titles. This assumption implies that if it were equally easy to fill low-skilled and high-skilled jobs, firms would find it more profitable to create high-skilled jobs. Therefore, to make firms indifferent between both types of jobs, high-skilled jobs must be harder to fill, i.e. applicant-vacancy ratios are smaller in high-skilled jobs. Whether the surplus of a match indeed increases faster with skill than the the cost of creating a vacancy is an empirical question on which little conclusive evidence is available. Nevertheless, it is relatively unlikely that the vacancy creation cost increases as fast as the surplus because of fixed components in this vacancy creation cost. To mention a simple example, the cost of posting a vacancy on CareerBuilder.com is independent of the wage or the required skill level. Although there is little systematic empirical evidence on how the cost of posting a vacancy varies with skill, the assumption that the surplus increases faster with skill than cost of posting a vacancy seems *prima facie* plausible.

To further assess the credibility of the assumption that the surplus of a match indeed increases faster with skill than the the cost of creating a vacancy, one can consider its implications. In our model, the assumption implies that high-skilled workers have lower unemployment rates and capture a higher share of the surplus than low-skilled workers. Both these predictions are supported by empirical evidence. In fact, lower unemployment rates for high-skilled or more educated workers are a robust and virtually undisputed finding (see e.g. Elsby et al. 2010).[23] There is also independent empirical evidence supporting the pre-

why the estimated impact of the wage on applications is negative within SOC but positive within job title. Consistent with this interpretation of the data, we find that the negative relationship between wages and applications is more pronounced in SOCs where there is a larger degree of heterogeneity in the experience level of applicants (results available upon request).

[22]Note however that directed search models can easily be extended to include productivity shocks or learning (see e.g. Gonzalez & Shi, 2010; Menzio & Shi, 2011).

[23]More ambiguity exists about the cause of the lower unemployment rates. For example, Elsby et al. (2010) find that more educated workers are less likely to enter unemployment but not more likely to exit unemployment.

diction of our model that high-skilled workers capture a higher share of the surplus. Martins (2009) shows that employer rents are lower in firms with a higher share of high-skilled workers because wages increase faster than productivity. Galindo-Rueda & Haskel (2005) show that high-skilled workers participate more in rent-sharing than low-skilled workers. Our model is thus able to generate accurate empirical predictions beyond the basic facts for which we wanted to account, and this strengthens the plausibility of the assumption that the surplus of a match indeed increases faster with skill than the the cost of creating a vacancy.

4.2 Competition and the Wage Elasticity of Applications

In a perfectly competitive model with homogenous workers and firms, there is a unique equilibrium wage level. If a firm deviates and offers a slightly higher wage, all workers will switch to this deviant firm. This simple setup leads to the intuition that higher wages should have an infinitely large effect on the number of applications. Clearly, this is not what we find in the data since our strongest positive effect implies an elasticity of applications with respect to wages below one: a 10% increase in the wage is associated with a 7.4% increase in applications. This number can shed light on the degree to which the labor market is perfectly competitive, as we will explain in the remainder of this section.

In the simplest perfectly competitive model with homogenous workers and firms, there is a unique wage in equilibrium. In order to generate more than one wage in the data and to be able to empirically estimate the impact of an increase in wages on the number of applicants, one must assume some heterogeneity or out-of-equilibrium behavior. Therefore, if in fact we observe more than one wage, this already rules out the simplest model, and we can therefore no longer directly use it to interpret the data.

In our model, we deviate from the simplest competitive model by introducing both matching frictions and worker heterogeneity. Matching frictions alone do not generate wage dispersion (see Moen, 1997), but already explain why an (out-of-equilibrium) increase in the wage only has a limited effect on the number of applicants. The reason is that matching frictions introduce a second dimension (besides the wage) to the desirability of a job: the matching probability. If all workers were to apply to a deviant firm offering a marginally higher wage than the equilibrium level, each one of them would get the job with probability zero. Therefore, workers have to trade off the wage against the matching probability, and this limits the impact that a wage increase can have on the number of applicants. By introducing worker heterogeneity on top of matching frictions, our model generates equilibrium wage dispersion within a job title. Workers have to trade-off the wage and the matching probability, knowing that they compete against workers with various levels of productivity. The key intuition is that lower productivity workers do not apply to high wage jobs because they know that high

On the other hand, using an instrumental variable strategy, Riddell & Song (2011) find that more educated workers have shorter unemployment durations, consistent with the prediction of our model.

productivity workers, who have better outside options, are more willing to endure the high level of competition (i.e. many applications) that is associated with these high wage jobs. In our model, an increase in the wage yields more and better applicants, but at the implicit cost of driving away the lower productivity applicants, so the impact of an increase in the wage cannot be infinite. Overall, our model accounts for the low elasticity of applications with respect to the wage within a job title by introducing both matching frictions and heterogeneity in worker productivity.

However, besides matching frictions and worker heterogeneity in productivity, there could be two other reasons why the elasticity of applications with respect to the wage is low. First, there could be heterogeneity in workers' tastes for specific employers. Monopsony models of the labor market are built on this idea (see e.g. Bhaskar et al. 2002). One key example is distance: workers prefer to work for employers who are closer to where they live. Therefore, an employer has an easier time attracting workers who live relatively close. A marginal increase in the wage will not attract all workers, but only those workers who were formerly indifferent between the firm and its competitors based on the distance from the workers' residence to the firm. Second, there could be match-specific productivity. Suppose that this match-specific component is known to the worker when he decides to which firm to apply. If a worker knows that his chances of being hired by a particular firm are low due to a low match-specific productivity component, then a marginal increase in the wage by this particular firm may not affect the worker's application decision. So, either worker heterogeneity in taste for specific jobs or match-specific productivity can decrease the elasticity of applications with respect to the wage.

Our model does not include worker heterogeneity in taste or match specific productivity because our data does not allow us to measure these dimensions. On the other hand, we do have measures of worker productivity and skills, such as education and experience, and this is why our model concentrates on this type of worker heterogeneity. The empirical elasticity of applications with respect to the wage is likely to be low both for reasons that we include in our model, i.e. matching frictions and worker heterogeneity in productivity, and for reasons we do not explicitly model, i.e. worker heterogeneity in taste for specific jobs and match-specific productivity.

In conclusion, our model is a parsimonious way of explaining a number of stylized facts we have documented, but other mechanisms may contribute to explaining the quantitative estimate of the elasticity of applications with respect to the wage. What is clear is that the simplest model of a perfectly competitive labor market fails to account for the empirical relationship between wages and applications, and it is necessary to introduce matching frictions or worker heterogeneity to account for the data.

4.3 Alternative Models

In order to yield empirically correct predictions about the relationship between wages and applications within and across job titles, we make specific assumptions. Here, we discuss some alternative assumptions one may make to explain the empirical results.

Within a job title, we rely on worker heterogeneity and adverse selection in order to generate a positive association between wages and the number of applicants. One can also generate such a positive relationship using on-the-job search. For example, Delacroix & Shi (2006) analyze a directed search model with on-the-job search and show that wage dispersion arises in equilibrium. Unemployed workers apply to low wages, while employed workers apply to higher wages. Since firms that pay higher wages obtain a smaller match payoff, those firms must match with larger probability in order to obtain equal profit, implying that wages and applicant-vacancy ratios are positively related. We do not use their model for interpreting our empirical findings, because it assumes that workers are homogeneous and therefore cannot explain the positive correlation between wages and productivity within a job title. On-the-job search can be introduced in our model, but complicates notation considerably and would only strengthen the positive relationship between wages and applications.[24]

Across job titles, we use the assumption that the surplus increases faster with skill than the vacancy cost to explain the negative relationship between wages and applications. Alternatively, one could attempt to explain this relationship by arguing that there are fewer applications in high-skilled jobs because there is a disequilibrium between the supply and demand of high-skilled versus low-skilled workers. In our model, the distribution of worker types is irrelevant because of free entry: firms will just create more jobs if there are more workers of a particular type, until the equilibrium applicant-vacancy ratio is reached. The demand for various types of workers is relevant but can be captured by the price of the output $p(x)$. For example, if the demand for high-skilled workers relative to low-skilled workers goes up, $p(x)$ simply increases faster with x. In this case, high-skilled jobs generate an even higher surplus and therefore, to make firms indifferent between creating high-skilled and low-skilled jobs, it must become more difficult to recruit in these high-skilled jobs. Hence, applications in high-skilled jobs decline even further following an increase in the derivative of $p(x)$ with respect to x. Our model therefore does not disagree with the view that high demand for high-skilled workers is the reason why we see fewer applicants in high skilled jobs. On the other hand, our model elegantly demonstrates that it is not necessary to have a disequilibrium between the demand and supply of skills to generate a negative relationship between wages and applications[25]. Instead, it is enough that high-skilled workers are significantly more productive than

[24] See also Menzio & Shi (2010, 2011) for models of on-the-job search and worker heterogeneity.

[25] The idea that within most six-digit SOCs there are 'not enough' high skilled workers compared to demand seems empirically quite implausible. Yet one would have to make an assumption of this sort to explain the negative relationship between wages and applications within SOC by a disequilibrium between supply and

low-skilled workers; and the more productive they are, the smaller the number of applications received by high-skilled jobs compared to low-skilled jobs.

Overall, our results and model speak to a broad range of empirical questions, including skill specificity, inequality between workers of different skill levels, and imperfect competition in the labor market. Our model is a parsimonious way of capturing our key empirical results. At the same time, the model generates additional empirically valid predictions, which contributes to its plausibility.

5 Conclusion

In this paper, we analyze the relationship between skills, wages and the number of applications. First, we document a number of new empirical facts. Using data from CareerBuilder.com, the largest US employment website, we show that jobs that post higher wages attract more experienced and more educated applicants. The relationship between the wage that a firm posts and the number of applicants that it attracts crucially depends on how we define a class of similar jobs. Economy-wide, higher wage jobs attract fewer applicants, and this continues to be the case if one controls for industry and/or occupation, where we capture occupations by six-digit SOC codes. However, when controlling for the job title as specified by the firm in the job ad, the sign of the relationship reverses: within a job title, a 10% increase in the wage is associated with a 7.4% increase in the number of applicants. This implies that it is very important to control for detailed job titles, or otherwise be aware of the substantial degree of heterogeneity in job characteristics that exists even within a six-digit SOC code.

Second, we explain these empirical patterns with a directed search model in which workers are heterogeneous in both the type of job that they can do and their productivity within that job type. In equilibrium, different sub-markets for each worker type arise. The combination of data and theory suggests that skills are very specific in the sense that they cannot easily be transferred from one job title to another. This finding has important implications for a number of questions in labor economics, including occupation specific human capital, job mobility and the cost of job loss.

Our model and empirical analysis also speak to the importance of matching frictions to explain the functioning of the labor market. First, we note that the elasticity of applications with respect to the wage is less than one, so it is very small indeed compared to what it would be in the simplest competitive model of the labor market, which would predict this elasticity to be essentially infinite. In our model, matching frictions and heterogeneity in worker productivity can explain why the elasticity of applications with respect to the wage is not infinite. As we pointed out in our discussion, other mechanisms could also contribute to explaining a low elasticity of applications with respect to the wage. Disentangling the relative

demand of skills.

importance of these different theoretical mechanisms is beyond the scope of this paper, but future empirical research on this topic would be enlightening. Second, our model shows that matching frictions can play an important role in explaining the inequality in both wages and unemployment across skills. In our model, the difference in wages and unemployment across skills are two sides of the same coin. In a frictional labor market, if high-skilled workers are sufficiently productive relative to their hiring cost, equilibrium dictates that they are less likely to be unemployed and at the same time get a higher share of the surplus than low-skilled workers. The inequality in surplus sharing between low-skilled and high-skilled workers exacerbates the wage inequality arising directly from productivity differences. Exploring how productivity differences between low-skilled and high-skilled workers jointly determine wage and unemployment inequality seems to be a promising avenue for future empirical research.

References

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

Bhaskar, V., Manning, A., & To, T. (2002). Oligopsony and monopsonistic competition in labor markets. *The Journal of Economic Perspectives*, 16(2), 155–174. ArticleType: research-article / Full publication date: Spring, 2002 / Copyright © 2002 American Economic Association.

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.

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

Chang, B. (2012). Adverse selection and liquidity distortion in decentralized markets. mimeo.

Delacroix, A. & Shi, S. (2006). Directed search on the job and the wage ladder. *International Economic Review*, 47(2), 651–699.

Elsby, M. W., Hobijn, B., & Sahin, A. (2010). *The Labor Market in the Great Recession*. Working Paper 15979, National Bureau of Economic Research.

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

Galindo-Rueda, F. & Haskel, J. (2005). Skills, workforce characteristics and firm-level productivity: Evidence from the matched abi/employer skills survey. mimeo.

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

Gonzalez, F. M. & Shi, S. (2010). An equilibrium theory of learning, search, and wages. *Econometrica*, 78(2), 509–537.

Guerrieri, V., Shimer, R., & Wright, R. (2010). Adverse selection in competitive search equilibrium. *Econometrica*, 78(6), 1823–1862. mimeo.

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.

Inderst, R. & Muller, H. M. (2002). Competitive search markets for durable goods. *Economic Theory*, 19, 599–622.

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

Kambourov, G. & Manovskii, I. (2009b). Occupational specificity of human capital. *International Economic Review*, 50(1), 63–115.

Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Princeton University Press.

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

Lang, K., Manove, M., & Dickens, W. (2005). Racial discrimination in labor markets with posted wage offers. *American Economic Review*, 95(4), 1327–1340.

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. (2012). Online job search and unemployment insurance during the great recession. *working paper*.

Martins, P. S. (2009). Rent sharing before and after the wage bill. *Applied Economics*, 41(17), 2133–2151.

Menzio, G. & Shi, S. (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, 145(4), 1453–1494.

Menzio, G. & Shi, S. (2011). Efficient search on the job and the business cycle. *Journal of Political Economy*, 119, 468–510.

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

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.

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.

Riddell, W. & Song, X. (2011). The impact of education on unemployment incidence and re-employment success: Evidence from the us labour market. *Labour Economics*, 18, 453–463.

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

Tables

Table 1: Summary statistics

	obs.	mean	s.d.	min	max
Years of education	2,282	16.63	1.35	14	24
Years of experience	2,379	13.28	5.13	2.5	26
Yearly wage	11,900	57,323	31,690	13,500	185,000
Applications per 100 views	61,051	1.168	2.570	0	100
Clicks per 100 views	60,979	5.640	5.578	0	100
Employees	61,135	18,824	59,280	1	2,100,000

Table 2: Effect of log wage on the number of applications per 100 views.

	I	II	III	IV	V
Log yearly wage	-0.770***	-0.642***	-0.710***	0.582**	0.912**
	(0.052)	(0.075)	(0.087)	(0.278)	(0.440)
Controls	No	Yes	Yes	Yes	Yes
NAICS (2 digits)	No	Yes	No	Yes	No
SOC (6 digits)	No	Yes	Yes	No	No
Firm effects	No	No	Yes	No	Yes
Job title (3 words)	No	No	No	Yes	Yes
Observations	11,708	11,708	11,708	11,708	11,708
R^2	0.017	0.133	0.363	0.476	-

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controls include a constant, vacancy duration, dummy for salary expressed per hour, required education and experience, log number of employees of posting firm, designated market area, and calendar month. Column V was calculated using the user-generated command felsdvreg in Stata, which does not allow for the calculation of the R^2.

Table 3: Robustness of the effect of log wage on the number of applications per 100 views.

	I	II	III	IV	V	VI
Log yearly wage	-0.411***	-0.068	0.582**	0.620*	0.586	0.102
	(0.100)	(0.198)	(0.278)	(0.350)	(0.360)	(0.146)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
NAICS (2 digits)	Yes	Yes	Yes	Yes	Yes	Yes
SOC (6 digits)	No	No	No	No	No	Yes
Job title	1 word	2 words	3 words	4 words	All words	No
Word effects	No	No	No	No	No	Yes
Observations	11,708	11,708	11,708	11,708	11,708	11,708
R^2	0.185	0.376	0.476	0.498	0.500	0.315

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controls include a constant, vacancy duration, dummy for salary expressed per hour, required education and experience, log number of employees of posting firm, designated market area, and calendar month.

Table 4: Words that significantly affect the number of applications per 100 views.

Sign of word coefficient	Management / seniority	Special area / skills	Computer
Negative (fewer applications)	<u>Senior</u> Supervisor Management	And Maintenance Restaurant Truck Driver Therapist	<u>Engineer</u> Developer Software Java Security
Positive (more applications)	Specialist Administrative Technician	Service Project Inside Warehouse	

Based on specification VI of table 3. Words are included when they appear at least 100 times and are significant at the 10% level. Words are ordered by frequency and underlined when they appear at least 500 times.

Table 5: Effect of log wage on the average quality of applicants

	Average years of experience		Average years of education	
	I	II	III	IV
Log yearly wage	2.714***	1.602**	0.757***	0.245**
	(0.203)	(0.629)	(0.045)	(0.122)
Controls	No	Yes	No	Yes
NAICS (2 digits)	No	Yes	No	Yes
Job title (3 words)	No	Yes	No	Yes
Observations	1,755	1,300	1,696	1,257
R^2	0.238	0.961	0.282	0.971

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Average education / experience is calculated over the sample of applicants to each job. The regressions are weighted by the number of applicants using Stata's analytic weights. Controls include a constant, vacancy duration, dummy for salary expressed per hour, required education and experience, log number of employees of posting firm, designated market area, and calendar month.

Table 6: Effect of log wage on the number of clicks per 100 views

	I	II	III	IV	V
Log yearly wage	-1.045***	-0.546***	-0.777***	1.636***	1.738***
	(0.089)	(0.126)	(0.163)	(0.350)	(0.421)
Controls	No	Yes	Yes	Yes	Yes
NAICS (2 digits)	No	Yes	No	Yes	No
SOC (6 digits)	No	Yes	Yes	No	No
Firm effects	No	No	Yes	No	Yes
Job title (3 words)	No	No	No	Yes	Yes
Observations	11,694	11,694	11,694	11,694	11,694
R^2	0.011	0.161	0.382	0.564	-

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controls include a constant, vacancy duration, dummy for salary expressed per hour, designated market area, and calendar month. Column V was calculated using the user-generated command felsdvreg in Stata, which does not allow for the calculation of the R^2.

Table 7: Effect of wage posting on the number of applications per 100 views.

	I	II	III
Posted a wage	0.478***	0.186***	0.109
	(0.031)	(0.072)	(0.095)
Controls	No	Yes	Yes
NAICS (2 digits)	No	Yes	Yes
Firm effects	No	No	Yes
Job title (3 words)	No	Yes	Yes
Observations	61,051	61,050	61,050
R^2	0.006	0.495	

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controls include a constant, vacancy duration, dummy for salary expressed per hour, required education and experience, log number of employees of posting firm, designated market area, and calendar month. Column V was calculated using the user-generated command felsdvreg in Stata, which does not allow for the calculation of the R^2.

Table 8: Effect of wage posting on the average quality of applicants

	Average years of experience		Average years of education	
	I	II	III	IV
Posted a wage	-0.704**	-0.317	-0.231***	-0.124
	(0.311)	(1.297)	(0.069)	(0.120)
Controls	No	Yes	No	Yes
NAICS (2 digits)	No	Yes	No	Yes
Job title (3 words)	No	Yes	No	Yes
Observations	2,379	1,774	2,282	1,704
R^2	0.007	0.947	0.013	0.967

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Average education / experience is calculated over the sample of applicants to each job. The regressions are weighted by the number of applicants using Stata's analytic weights. Controls include a constant, vacancy duration, dummy for salary expressed per hour, required education and experience, log number of employees of posting firm, designated market area, and calendar month.

Table 9: Effect of log wage on the number of applications per 100 views, using job titles with at least two observations

	I	II	III
Log yearly wage	-0.676***	-0.509***	-0.472***
	(0.066)	(0.115)	(0.110)
Controls	No	Yes	Yes
NAICS (2 digits)	No	Yes	No
SOC (6 digits)	No	Yes	Yes
Firm effects	No	No	Yes
Job title	No	No	No
Observations	5,810	5,810	5,810
R^2	0.011	0.168	0.400

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Controls include a constant, vacancy duration, dummy for salary expressed per hour, required education and experience, log number of employees of posting firm, designated market area, and calendar month.

Proofs

Proof of Lemma 1

Proof. Consider first the Bellman equations. Suppose that in equilibrium a firm posts a job ad (x, w) and attracts an applicant-vacancy ratio $\lambda(x, w)$, potentially consisting of various types of workers. The value of unemployment $V_U(\cdot)$ and the value of employment $V_E(\cdot)$ for these workers must then satisfy the following Bellman equations[26]

$$rV_U(y) = b(x)\,y + \frac{m(\lambda)}{\lambda}\left(V_E(y) - V_U(y)\right) \tag{6}$$

and

$$rV_E(y) = w - \delta\left(V_E(y) - V_U(y)\right). \tag{7}$$

Likewise, the value of a vacancy V_V and the value of a filled job V_J for the firm must satisfy

$$rV_V = -c(x) + m(\lambda)\left(V_J - V_V\right) \tag{8}$$

and

$$rV_J = p(x)\,E[y] - w - \delta\left(V_J - V_V\right), \tag{9}$$

where $E[y]$ represents the expected productivity of the worker that the firm will hire, which is determined by the worker's endogenous application decisions.

To show that pooling of worker types cannot occur in equilibrium, suppose that some firms post a particular wage w^* and attract an applicant-vacancy ratio $\lambda^* = \lambda(x, w^*)$ consisting of various types of workers. Let y_2 denote the productivity of the highest type of worker who applies to the firm with positive probability. Further, let y_1 denote an arbitrary lower type of applicant, i.e. $y_1 < y_2$. The payoffs for these workers can be derived by solving equation (6) and (7) and evaluating the result in w^* and λ^*. This yields

$$rV_U^*(y) = \frac{\lambda^*(r+\delta)\,b(x)\,y + m(\lambda^*)\,w^*}{\lambda^*(r+\delta) + m(\lambda^*)}.$$

The firms posting the wage w^* obtain a payoff

$$V_V^* = -c(x) + m(\lambda^*)\,\frac{p(x)\,E[y] - w^*}{r+\delta},$$

which equals zero because of the free-entry condition. □

We will now derive a contradiction by showing that a deviant firm that posts a wage w_d,

[26] To simplify notation, the dependence of the value functions and the applicant-vacancy ratio on x and w is suppressed.

marginally higher than w^*, obtains a strictly positive payoff V_V^d. Workers are willing to apply to this deviant if the expected payoff that it provides is at least equal to the expected payoff provided by the other firms. We will show that this implies that the deviant will only attract applicants of high-productivity type y_2. In other words, if the number of applications λ_d that the deviant attracts is such that a high-productivity applicant y_2 is indifferent between the deviant and firms posting w^*, workers of lower type y_1 strictly prefer firms posting w^* over the deviant. The reason for this result is as follows. If high-productivity applicants are indifferent between w^* and $w_d > w^*$, it must be that the deviant job has a slightly higher number of applicants, i.e. $\lambda_d > \lambda^*$. In this situation, low-productivity applicants face a tradeoff between a low wage w^* and a low applicant-vacancy ratio λ^* or a slightly higher wage w_d and a higher applicant-vacancy ratio λ_d. Since low-productivity workers have a worse outside option, they value matching probability relative to the wage more than higher type workers, and therefore (w^*, λ^*) is more attractive. This means that low-productivity workers stay away from the deviant, and the deviant only receives applications from high type workers.

The payoff of a worker of type y_2 who applies to the deviant equals

$$rV_U^d(y_2) = \frac{\lambda_d (r+\delta) b(x) y_2 + m(\lambda_d) w_d}{\lambda_d (r+\delta) + m(\lambda_d)}.$$

These workers are indifferent between the deviant and a firm posting w^* if and only if $rV_U^d(y_2) = rV_U^*(y_2)$, which defines a relationship between the deviant's wage offer w_d and his expected number of applicants λ_d. Note that since the deviant offers a marginally higher wage, he must attract marginally more applicants for high-productivity applicants to be indifferent, i.e. $\lambda_d > \lambda^*$. Solving $rV_U^d(y_2) = rV_U^*(y_2)$ for w_d yields

$$w_d = rV_U^*(y_2) + \frac{\lambda_d (r+\delta)}{m(\lambda_d)} \left[rV_U^*(y_2) - b(x) y_2 \right].$$

This expression can be used to eliminate the wage w_d from the payoff of a worker of type y_1 who applies to the deviant, i.e.

$$\begin{aligned} rV_U^d(y_1) &= \frac{\lambda_d (r+\delta) b(x) y_1 + m(\lambda_d) w_d}{\lambda_d (r+\delta) + m(\lambda_d)} \\ &= rV_U^*(y_2) - \frac{\lambda_d (r+\delta)}{\lambda_d (r+\delta) + m(\lambda_d)} b(x)(y_2 - y_1). \end{aligned}$$

Whether workers of type y_1 are willing to apply to the deviant depends on whether this expression is larger or smaller than $rV_U^*(y_1)$. The difference between these two payoffs equals

$$rV_U^d(y_1) - rV_U^*(y_1) = \left[\frac{\lambda^* (r+\delta)}{\lambda^* (r+\delta) + m(\lambda^*)} - \frac{\lambda_d (r+\delta)}{\lambda_d (r+\delta) + m(\lambda_d)} \right] b(x)(y_2 - y_1).$$

This difference is negative, since

$$\frac{d}{d\lambda}\frac{\lambda(r+\delta)}{\lambda(r+\delta)+m(\lambda)}=(r+\delta)\frac{m(\lambda)-\lambda m'(\lambda)}{[\lambda(r+\delta)+m(\lambda)]^2}>0$$

and $\lambda_d > \lambda^*$.

Hence, the deviant firm will indeed only attracts workers of type y_2. As a result, the output that it will produce in a match increases discretely, which makes the marginal increase in the wage offer profitable. The value of a vacancy to the deviant is:

$$V_V^d=-c(x)+m(\lambda_d)\frac{p(x)y_2-w_d}{r+\delta}>0.$$

Therefore, starting from a (partially) pooling equilibrium, a deviant firm offering a marginally higher wage makes a positive profit by attracting only high productivity workers. This implies that such an equilibrium is not sustainable and a fully separating equilibrium must arise.

Proof of Lemma 2

Given separation of productivity types, we can write the Bellman equations for a firm attracting type-y workers as

$$rV_V(y)=-c(x)+m(\lambda)(V_J(y)-V_V(y)) \tag{10}$$

when the firm has a vacancy, and

$$rV_J(y)=p(x)y-w-\delta(V_J(y)-V_V(y)). \tag{11}$$

when the firm is matched with a worker. The workers' Bellman equations remain unchanged from (6) and (7).

In order to derive the equilibrium of our model, we first analyze a situation in which firms commit to hiring only a particular productivity type. In that case, workers will only apply to jobs targeted at their type and incentive compatibility constraints are redundant. Hence, we can solve for the equilibrium wage of each productivity type independently.

Denote the equilibrium payoff of an unemployed worker of type y again by $rV_U^*(y)$. If this worker applies to a particular wage w, then this wage w and the corresponding applicant-vacancy ratio λ must satisfy

$$w=rV_U^*(y)+\lambda(r+\delta)\frac{rV_U^*(y)-b(x)y}{m(\lambda)}, \tag{12}$$

as follows from (6) and (7). A firm deciding what wage to post realizes that the number of applicants that it attracts will be determined by this equation. Alternatively, we can think of the firm as choosing an applicant-vacancy ratio, after which (12) determines the wage. Since this is analytically slightly easier, we follow that approach.

First, solve (10) and (11) to obtain

$$V_V = -c(x) + m(\lambda)\frac{p(x)y - w}{r+\delta}. \tag{13}$$

Substitute (12) then to eliminate w and take the first order condition with respect to λ. This yields after some rewriting

$$rV_U^*(y) = \frac{(r+\delta)b(x)y + m'(\lambda)p(x)y}{r+\delta+m'(\lambda)}.$$

Substituting this back into V_V gives

$$V_V = -c(x) + \frac{m(\lambda) - \lambda m'(\lambda)}{r+\delta+m'(\lambda)}[p(x) - b(x)]y, \tag{14}$$

which must equal zero.

Equations (13) and (14) provide us with two expressions which pin down the equilibrium relationship between w, λ and y within a job type. Given a productivity level y, the equilibrium applicant-vacancy ratio λ is determined by

$$y = \frac{r+\delta+m'(\lambda)}{m(\lambda) - \lambda m'(\lambda)}\frac{c(x)}{p(x) - b(x)} \tag{15}$$

and the corresponding equilibrium wage w follows from

$$w = p(x)y - \frac{(r+\delta)c(x)}{m(\lambda)}. \tag{16}$$

Note that (15) implies a negative relationship between productivity and the applicant-vacancy ratio in this scenario, because

$$\frac{dy}{d\lambda} = m''(\lambda)\frac{r+\delta+m(\lambda)/\lambda}{[m(\lambda) - \lambda m'(\lambda)]^2}\frac{c(x)}{p(x) - b(x)} < 0.$$

The intuition for this result is as follows. As can be seen in (14), the applicant-vacancy ratio λ determines which fraction of the surplus created by a match goes to the firm. If λ were constant across y, this fraction would be constant. Since the created surplus is larger for larger values of y, firms attracting high-productivity workers would obtain a higher payoff. This violates the free-entry condition. Additional entry would take place in those sub-markets, reducing the applicant-vacancy ratios.

As argued in the main text, commitment to a particular productivity type is not credible. We will now show that elimination of commitment changes the equilibrium outcomes by creating incentives for workers of a particular type y_1 to act like a different type $y_2 \neq y_1$ and apply to jobs (w_2, λ_2) instead of jobs (w_1, λ_1). Analogous to the proof of lemma 17, the payoff of such a worker equals

$$rV_U(y_2|y_1) = \frac{(r+\delta)\left[\lambda_2 b(x) y_1 - c(x)\right] + m(\lambda_2) p(x) y_2}{\lambda_2 (r+\delta) + m(\lambda_2)}.$$

The worker will choose which type y_2 to mimic in order to maximize this payoff. The first order condition of $rV_U(y_2|y_1)$ with respect to y_2 is, after some manipulation, equivalent to

$$\begin{aligned} & (r+\delta)\frac{\left[m(\lambda_2) - \lambda_2 m'(\lambda_2)\right]\left[p(x) y_2 - b(x) y_1\right] - \left[r+\delta+m'(\lambda_2)\right] c(x)}{\left[\lambda_2(r+\delta) + m(\lambda_2)\right]^2}\frac{d\lambda_2}{dy_2} \\ = \; & \frac{m(\lambda_2) p(x)}{\lambda_2(r+\delta) + m(\lambda_2)} \end{aligned} \tag{17}$$

When applicant-vacancy ratios stay the same as in the scenario with commitment, i.e. (15) holds, the left hand side simplifies to

$$(r+\delta)\frac{\left[m(\lambda_2) - \lambda_2 m'(\lambda_2)\right] b(x)(y_2 - y_1)}{\left[\lambda_2(r+\delta) + m(\lambda_2)\right]^2}\frac{d\lambda_2}{dy_2}.$$

This expression clearly becomes zero when the worker truthfully reveals his type ($y_2 = y_1$, $\lambda_2 = \lambda_1$),. However the right hand side of (17) remains positive, rendering the considered equilibrium infeasible. Instead, the solution to the first order condition implies that the worker wants to mimic a worker of type $y_2 > y_1$ and apply to firms with applicant-vacancy ratios $\lambda_2 < \lambda_1$.

Equilibrium therefore requires the the number of applicants at high-wage jobs to be larger than the value implied by (15), in order to discourage less productive workers from applying there. The incentive compatibility constraint is satisfied if(17) holds with equality for $y_2 = y_1 = y$ and $\lambda_2 = \lambda_1 = \lambda$, which implies

$$\frac{d\lambda}{dy} = \frac{1}{r+\delta}\frac{\left[\lambda(r+\delta) + m(\lambda)\right] m(\lambda) p(x)}{\left[m(\lambda) - \lambda m'(\lambda)\right]\left[p(x) - b(x)\right] y - \left[r+\delta+m'(\lambda)\right] c(x)}.$$

Proof of Prediction 1

Proof. Within a job type x, wages are determined by (3). Consider their first derivative with respect to productivity y, i.e.

$$\frac{dw}{dy} = p(x) + \frac{(r+\delta)\,c(x)}{m^2(\lambda)} m'(\lambda) \frac{d\lambda}{dy}.$$

As discussed in the proof of Lemma 2, the incentive compatibility constraint of lower type workers binds in equilibrium, which increases the applicant-vacancy ratios above the value implied by (15). Hence

$$y > \frac{r+\delta+m'(\lambda)}{m(\lambda)-\lambda m'(\lambda)} \frac{c(x)}{p(x)-b(x)},$$

such that $\frac{d\lambda}{dy}$ as specified in equation (2) is strictly positive. Hence, $\frac{dw}{dy} > 0$ and higher wages attract more productive workers. □

Proof of Prediction 2

Proof. The result can be obtained in a similar way as the proof of prediction 1. The derivative of the wage with respect to λ equals

$$\frac{dw}{d\lambda} = p(x) \frac{dy}{d\lambda} + \frac{(r+\delta)\,c(x)}{m^2(\lambda)} m'(\lambda).$$

Since $\frac{d\lambda}{dy} > 0$, this derivative is strictly positive and high wages attract more applicants. □

Proof of Prediction 3

Proof. To study the relationship between wages and skill if heterogeneity in productivity is sufficiently small, consider the limit case in which $\underline{y} = \overline{y} = y$. In that case, only one wage w is offered in each job type. This wage satisfies (3) and attracts an applicant-vacancy ratio λ determined by the solution to

$$y = \frac{r+\delta+m'(\lambda)}{m(\lambda)-\lambda m'(\lambda)} \frac{c(x)}{p(x)-b(x)}. \tag{18}$$

To analyze the relationship between the wage and skill, we first consider how the applicant-vacancy ratio λ depends on skill x. Implicit differentiation of the above expression with respect to x yields

$$\frac{d\lambda}{dx} = -\frac{1}{m''(\lambda)} \frac{[m(\lambda)-\lambda m'(\lambda)]^2}{\lambda(r+\delta)+m(\lambda)} \frac{c'(x)[p(x)-b(x)] - [p'(x)-b'(x)]\,c(x)}{[c(x)]^2} y. \tag{19}$$

This derivative is strictly negative because $m''(\lambda) < 0$ and because an increasing markup implies

$$\frac{d}{dx}\frac{p(x)-b(x)}{c(x)}=\frac{c'(x)[p(x)-b(x)]-[p'(x)-b'(x)]c(x)}{[c(x)]^2}>0.$$

Using equation (3), the derivative of the wage with respect to x equals

$$\frac{dw}{dx}=p'(x)y-(r+\delta)\frac{c'(x)}{m(\lambda)}+(r+\delta)\frac{c(x)m'(\lambda)}{m^2(\lambda)}\frac{d\lambda}{dx}.$$

Substituting equation (19) in this result gives after some manipulation

$$\begin{aligned}\frac{dw}{dx} &= \frac{r+\delta+m'(\lambda)}{m(\lambda)-\lambda m'(\lambda)}\frac{p'(x)c(x)}{p(x)-b(x)}-(r+\delta)\frac{c'(x)}{m(\lambda)}\\ &+(r+\delta)\frac{r+\delta+m'(\lambda)}{\lambda(r+\delta)+m(\lambda)}\frac{m'(\lambda)}{m''(\lambda)}\frac{m(\lambda)-\lambda m'(\lambda)}{m^2(\lambda)}\frac{[p'(x)-b'(x)]c(x)-c'(x)[p(x)-b(x)]}{p(x)-b(x)}\end{aligned}$$

Note that $\frac{r+\delta+m'(\lambda)}{m(\lambda)-\lambda m'(\lambda)}>\frac{r+\delta}{m(\lambda)}$ and $p'(x)>p'(x)-b'(x)$, such that

$$\begin{aligned}\frac{dw}{dx} &> [r+\delta+m'(\lambda)]\left[\frac{1}{m(\lambda)-\lambda m'(\lambda)}+\frac{r+\delta}{\lambda(r+\delta)+m(\lambda)}\frac{m'(\lambda)}{m''(\lambda)}\frac{m(\lambda)-\lambda m'(\lambda)}{m^2(\lambda)}\right]\\ &\times\frac{[p'(x)-b'(x)]c(x)-c'(x)[p(x)-b(x)]}{p(x)-b(x)}.\end{aligned}$$

The first and last factor are positive, such that a sufficient condition for wages to be strictly increasing in skill is

$$\frac{r+\delta}{r+\delta+m(\lambda)/\lambda}\frac{m'(\lambda)}{\lambda m''(\lambda)}\left[\frac{m(\lambda)-\lambda m'(\lambda)}{m(\lambda)}\right]^2\geq -1.$$

Evaluating this in the firm's matching rate $m(\lambda)=A\lambda^\alpha$ yields

$$\frac{(1-\alpha)(r+\delta)}{r+\delta+A\lambda^{\alpha-1}}\leq 1,$$

which is satisfied for all feasible parameter values. This establishes the result for the limit case $\underline{y}=\overline{y}=y$. Because of continuity, it follows that a positive correlation between wages and skill exists for a sufficiently small degree of heterogeneity. □

Proof of Prediction 4

Proof. Without heterogeneity in productivity, the proof follows immediately from earlier results. Since wages are increasing in skill, $\frac{dw}{dx}>0$, and the applicant-vacancy ratio is decreasing in skill, $\frac{d\lambda}{dx}<0$, the applicant-vacancy ratio is decreasing in the wage. Because of continuity, a negative correlation between wages and applications then exists for a sufficiently small degree of heterogeneity. □

Proof of Lemma 3

Proof. We focus on the case in which $\underline{y} = \overline{y} = y$, such that one wage w is offered in each job type in the baseline equilibrium. This wage satisfies (3) and attracts a number of applications λ as determined by (18). To analyze whether this equilibrium survives when workers can partially transfer their skill to other types of jobs, consider a worker of type x_i who – instead of applying to a wage w_i with corresponding λ_i – evaluates the payoff from a 'one-time' deviation in his application behavior by applying to jobs of type $x \neq x_i$ offering a wage w and an applicant-vacancy ratio λ.[27] Denote the worker's value of unemployment by $V_U(x, x_i)$ and his value of employment by $V_E(w, x_i)$. These values satisfy the following Bellman equations

$$rV_U(x, x_i) = b(x_i)\, y + \frac{m(\lambda)}{\lambda}\tau(x, x_i)\left(V_E(w, x_i) - V_U(x_i, x_i)\right) \tag{20}$$

and

$$rV_E(w, x_i) = w - \delta\left(V_E(w, x_i) - V_U(x_i, x_i)\right). \tag{21}$$

Solving these equations gives

$$rV_U(x, x_i) = b(x_i)\, y + \frac{m(\lambda)}{\lambda}\tau(x, x_i)\,\frac{w - rV_U(x_i, x_i)}{r+\delta}.$$

First, consider lower-ranked jobs, i.e. $x < x_i$. For these jobs, $\tau(x, x_i) = 1$, i.e. the worker can perfectly transfer their skill: this implies that, for example, a nurse can do as well as a cashier in a cashier job. The derivative of $rV_U(x, x_i)$ with respect to x then equals

$$\frac{drV_U(x, x_i)}{dx} = -\frac{m(\lambda) - \lambda m'(\lambda)}{\lambda^2}\,\frac{w - rV_U(x_i, x_i)}{r+\delta}\,\frac{d\lambda}{dx} + \frac{m(\lambda)}{\lambda}\,\frac{1}{r+\delta}\,\frac{dw}{dx}.$$

This derivative is strictly positive for any $x < x_i$ since $\frac{d\lambda}{dx} < 0$ and $\frac{dw}{dx} > 0$, as shown in the proof of prediction 3. Hence, $rV_U(x|x_i) < rV_U(x_i|x_i)$ for all $x < x_i$ and workers never want to apply to jobs that require less skill than they possess, even if skills are perfectly transferable.

Next, consider higher-ranked jobs, i.e. $x > x_i$. A necessary and sufficient condition to guarantee that workers do not want to apply to these jobs is that $rV_U(x|x_i) < rV_U(x_i|x_i)$ or, equivalently,

$$\begin{aligned}
\tau(x, x_i) \quad &\leq \quad \frac{r+\delta}{w - rV_U(x_i, x_i)}\,\frac{\lambda}{m(\lambda)}\left[rV_U(x_i, x_i) - b(x_i)\, y\right], \\
&= \quad \frac{\lambda(r+\delta)\left[w_i - b(x_i)\, y\right]}{\lambda_i(r+\delta)\left[w - b(x_i)\, y\right] + m(\lambda_i)\left[w - w_i\right]}\,\frac{m(\lambda_i)}{m(\lambda)}
\end{aligned}$$

[27] Note that it is sufficient to consider deviations at a single point in time by the Unimprovability Principle, see e.g. Kreps (1990).

for all $x > x_i$ and corresponding w and λ.The right-hand side of this condition equals 1 for $x \to x_i$ and is decreasing in x. $\square$