Wages, Applications, and Skills*

Ioana Marinescu and Ronald Wolthoff†

September 11, 2012

Abstract

Do higher wages attract more and better applicants? Using data from the popular employment website CareerBuilder.com, we show that higher wages attract better 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 7.5% more applications. Our directed search model shows that these results are consistent with skills being highly job specific. Additionally, the model shows how matching frictions can generate inequality in both 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 and compare vacancies. Whereas in the past job ads were typically spread out over many different news papers, these days most vacancies can be found in a few mouse clicks at zero monetary cost. Background information on employers is also much easier to obtain than before. These developments have likely reduced information frictions and increased the level of competition in this market. In a homogeneous world, this competition would cause a firm that offers more attractive terms of employment than other firms to attract more applicants and to fill its vacancy more easily. Of course, reality is more complex. Workers are heterogeneous in skills and jobs differ in their skill requirements, so the relationship between the terms of employment that a firm offers and the likelihood that it fills its vacancy is not straightforward. For example, hospitals looking for neurosurgeons may very well need longer to fill their vacancy than a local school searching for a janitor, even though the neurosurgeon job pays considerably more.

*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, and the University of Toronto. We are grateful to Alan Manning for comments on a previous version of this draft. 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.

Understanding the relationship between skills, wages, and job queues (or number of applications) provides insights into various questions that are of interest to labor economists. For example, the relation between skills and wages determines the returns to investment in human capital. Further, the link between skills and job queues is informative about how unemployment varies with skill levels. Finally, the relation between wages and job queues is related to the level of competition within a market, with a strong positive impact of the wage on the number of applications implying a more competitive market.

Despite the relevance of these questions, the literature that studies the relationship between skills, wages, and job queues is limited in size. The main cause of this appears to be a lack of data containing all the necessary variables, and in particular job applications. 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 wage that these vacancies pay, the number of applicants that each vacancy attracts, as well as various firm and applicant characteristics.

We use this CareerBuilder data to document a number of new empirical facts. We first consider the economy-wide relation between skills, wages, and job queues and show that higher wage jobs get fewer applicants, but that these applicants are of higher quality. The same patterns emerge when we look within detailed occupation and industry categories, respectively based on the Standard Occupational Classification (SOC) and the North American Industry Classification System (NAICS). It is only when we characterize similar jobs in a narrower way that the results change. In particular, a positive relationship between the wage and job queues is found when similar jobs are defined as those having the same job title chosen by firms when posting their vacancies (e.g. "RN ambulatory surgery", where RN stands for registered nurse). Specifically, a 10% increase in the wage is associated with a 7.5% increase in the number of applicants.

These findings suggest that controlling for job heterogeneity using SOC codes is not sufficient and may lead to spurious results. Compared to even the most detailed SOC codes, job titles contain extra information by including a description of the hierarchy level or work experience required for the job ('junior accountant' versus 'senior accountant') or a description of a specialization ('java programmer' versus 'SQL programmer'). Hence, occupations are narrower than typically assumed and can better be described by a job title than by an SOC code.

In the second part of the paper, we show that the empirical facts we document are consistent with a simple directed search model if we assume that workers are heterogeneous in the type of job that they can do (their 'skill') as well as in their productivity in those jobs. The model reveals that the empirical patterns we uncover are consistent with a limited degree of skill transferability across different types of jobs, i.e. workers trained to do a particular type of job will not be very productive in other types of jobs. Given that a type of job is best defined empirically by a very specific job title, our empirical and theoretical results taken together imply that most labor markets are fairly thin. Our model allows us to understand how the fact that the impact of the wage on job applications switches sign when controlling for job title is related to skill specificity, and why our empirical results are consistent with a high degree of skill specificity.

Besides being able to generate predictions consistent with our empirical findings, our model generates additional empirically testable implications. In particular, the surplus created by a match increases faster across job types 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 or have shorter queues, we show that this relationship is very robust and is only reversed when controlling for the specific job title. 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. 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. Kambourov & Manovskii (2009a). Our findings suggest that skills are not very transferable and labor markets are therefore relatively thin.

Our work is related to a number of threads in the literature. First, our paper contributes to the emerging empirical literature on labor markets on the Internet (see e.g. Kuhn & Shen, 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. By measuring job applications, we use a more sensitive measure of workers' preferences for jobs: a worker may prefer to work for a particular firm and send in an application, but he will not necessarily be hired. In that sense, employment changes are a somewhat indirect measure of the impact of wages on workers' preferences for various jobs. Focusing on the more direct measure of job applications is therefore a promising approach.

Second, our work 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. Unlike these papers, we do not assume that workers' characteristics are unobservable. Nevertheless, we show that the equilibrium exhibits similar properties, such as distortion of contracts by incentive compatibility constraints to induce separation of types. 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 doctors and good and bad nurses, even though doctors have a higher level of skills than nurses. When recruiting, firms must pay attention to both dimensions of heterogeneity and set their wages accordingly.

This paper proceeds as follows. Section 2 describes the data set and documents the key 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 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.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[1], 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 were the crucial ones in most cases, so we define a job title variable based on the first three words.[2] Based on the full content of the job posting, an internal CareerBuilder algorithm assigns an SOC (Standard Occupational Classification) 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.

Besides these characteristics, we 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 variable of interest, 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, and this number of clicks is our second variable of interest. 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, we construct two outcome 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 chose to focus on applications and clicks per 100 views 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 data, we have a second data set 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 additionally we have measures of applicant quality. Specifically,

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

[2] See section 2.3 for a more detailed discussion.

we have the share of applicants with a masters' degree, and the share of applicants with more than 10 years of experience.

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 US average wage in 2010 (see BLS Occupational Employment Statistics), which is consistent with the higher than average education of applicants on the website (see below). 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. The average posting firm has about 19000 employees. Finally, on average 25% of applicants have a masters degree, and 50% of them have 11 years of experience or more.

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 6.6% decline in applications per view. Clearly, caution is required in interpreting this result 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 job characteristics controls. In column II, we add the required education and experience for the job, and industry and detailed occupation fixed effects (595 six-digit SOC codes). In principle, this should control for most heterogeneity and allow us to compare jobs that are very similar. Yet, we still get a negative and significant association between the wage and the number of applicants. 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 large sets of 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 very robust.

Recognizing that SOC codes may not fully capture job heterogeneity, in column IV, we control for job title fixed effects. This should allow us to estimate the relationship between wages and applications among more homogenous groups of jobs. Interestingly, controlling for job title fixed effects completely reverses the pattern 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 indeed are associated with more applicants: the point estimate implies that a 10% increase in the wage is associated with a 7.5% increase in applicants[3].

We summarize our findings 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 better understand why the impact of wages on applications switches sign when using job title fixed effects, it is useful to discuss why six-digit SOC codes are not specific enough compared to job titles. Inspection of the data reveals that there are on average 8.6 job titles per SOC[4]. By exploring job titles, we were able to determine two key reasons why job titles are more precise than detailed SOC codes. The first is that job titles indicate specialties. For example, a registered nurse is coded as SOC 29-1140, but the job title specifies whether she works in oncology or at the ICU. The second reason why job titles are more precise is that SOC codes do not account well for 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). Job titles are therefore able to capture a large amount of job heterogeneity ignored by six-digit SOC codes.

We have just shown that higher wage jobs attract more applicants. But do these jobs also attract higher quality applicants? We measure applicant quality by the share of applicants with 11 years of experience or more, and the share of applicants with a master's degree. We find that higher wages are associated with a significantly higher share of high experience applicants, and the relationship is not sensitive to the addition of controls (table 3). A 10% increase in the wage is associated with a 1.3 percentage point increase in the share of high experience applicants, which represents a 2.6% increase in high experience applicants. Similarly, a higher wage increase the share of applicants with a master's degree (col. III), but this effect falls short of statistical significance when controlling for job title fixed effects (col. IV). However, even with job title fixed effects, the impact of the wage on education is still positive, and

[3]In column V, the results were calculated using the user-generated command felsdvreg in Stata. This does not allow for the calculation of the R^2.

[4]Rarely, jobs with the same job title have different SOCs. This is because the SOC code is inferred from the full text of the job post by a CareerBuilder algorithm, which means that two jobs with the same job title may be classified as belonging to different SOCs on the basis of the full text of the job post. In practice, the average number of SOC per job title is 1.4, with a median value of 1.

not significantly different from the estimate in column III. The lack of significance for the education result when controlling for job titles can be explained by two factors. First, the sample size is quite small given the number of job titles. Second, the level of education is likely to be more narrowly defined by the job title than the level of experience: it seems more likely that more experienced workers will apply to slightly more junior jobs in their area if the pay is good, rather than applying to a job that requires a different educational attainment. Overall, we conclude that higher wages attract higher quality applicants, and this relationship is robust to the addition of a large set of controls.

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 that we find. We subsequently consider biases that may arise from the definition of some of our key variables, omitted variables, and sample selection, but we conclude that they are unlikely to play a role in our analysis.

The first potential issue with our analysis is that the definition of some key variables may be driving our results. We discuss here the definition of a job title and the choice of applications per job view as an outcome. With respect to the definition of job title, we have truncated the job title provided by the firm to the first three words. The fourth word was often an indication of geography, such as "business development manager washington" or "customer service representative fairfax". In the regressions with job title fixed effects (Table 2, col. IV), restricting to three words leads us to using 4371 unique job titles instead of 4875 in the unrestricted version. While we lose power and significance levels are therefore lower, we find that the results from Table 2, col. IV are quantitatively unaffected if we do not truncate the job title to the first three words. With respect to the choice of applications per job view as an outcome, we have made this choice to correct for heterogeneity in job views across jobs. An alternative choice for an outcome variable is simply the log number of applicants for each job. We find that our key results from Table 2 are qualitatively unaffected by this alternative definition of job queues. We conclude that our results are robust to alternative definitions of job titles or job queues.

The second potential issue with of our analysis is that omitted variable bias could potentially 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 4 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), industry dummies, 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 2.8% 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 caveat to our analysis 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. 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.

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. First, in table 5, 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.[5]

[5] 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.

Hence, jobs with a posted wage do not get significantly more applicants when controlling for a large number of observables.

Second, in table 6, we examine the relationship between posting a wage and the quality of applicants in terms of education and experience. We do not find any relationship between posting a wage and the quality of applicants. 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.

Finally, one may wonder whether the fact that the relationship between wages and the number of applicants switches sign when using job title fixed effects is caused by sample selection. 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 7. Hence, sample selection does not drive our results.

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 doctors, and good doctors. 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-type workers applying to high-type jobs. This result may seem surprising since workers' types are not private information, but is easily explained by the frictional nature of the matching process. Because of the search frictions, a firm trying to match with a particular worker type will find it optimal ex post to hire the first worker who is "good enough" instead of waiting for their preferred worker type to show up.

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 doctor) 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. The amount of flow output that a worker produces in job x is called his 'productivity' y. 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.[6] 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 doctor and a bad doctor is small relative to the difference between a doctor and a nurse.

An employed worker of type (x, y) produces good 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 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

[6] 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.

unemployment is generated by job destruction shocks 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. 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.

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).[7] 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. 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. $\frac{d}{dx}\frac{p(x)-b(x)}{c(x)}y > 0$.

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.[8] 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.[9] As a result, the matching rates solely depend on the ratio $\lambda(x,w)$ of applicants to vacancies (the 'queue length') at a job type x and wage w. We generally omit the arguments x and w to simplify notation. Given a queue length 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 $\frac{m(\lambda)}{\lambda} = A\lambda^{\alpha-1}$[10].

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

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

[9] 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.

[10] 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.

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.

Consider the choice of the job type first. 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.

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 cannot apply to the same job in equilibrium. That is, good and bad doctors 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-type (i.e. low y) and high-type workers apply to the same wage and a deviating firm posts a slightly higher wage. Now, suppose that the deviant attracts a queue length such that low-type applicants are indifferent between the deviant and the other firms. For the low-type applicants to be indifferent, the wage of the deviant must be slightly higher and the queue slightly longer than at other firms; however, the queue must not be so long as to make the deviant job unattractive to low-type applicants. In this case, high-type workers strictly prefer applying to the deviant high-wage firm because, compared to low-type applicants, they value the higher wage more than they are hurt by the longer queue. This will increase the queue length at the deviant further, and low-type 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 doctors, while other firms post low wages and only receive applications from bad doctors. In order to sustain such an equilibrium, two incentive compatibility constraints must be satisfied: the good doctors must not want to apply to the jobs aimed at the bad doctors, and vice versa. One can show that while the incentive compatibility constraint for the good doctors is automatically satisfied, the incentive compatibility constraint for the bad doctors binds. In order to keep bad doctors away from the jobs for good doctors (i.e. prevent adverse selection), the matching rate at these high wage jobs must be sufficiently low, such that the bad doctors - who care relatively more about matching probability - prefer the low-wage jobs with higher matching probability. This incentive compatibility constraint for bad doctors increases the wage of good doctors and decreases their matching probability relative to a world without bad doctors.

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 queue length 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-type 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 queue length 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 queue lengths 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 queue length 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 queue length 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.*

3.3 Empirical Content

The simple model presented above provides several testable predictions regarding the relationship between productivity, wages, queue lengths 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.

Consider first 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 queue lengths. 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.*

Comparing the predictions of the model with the empirical results, we see that they 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 if we assume that a job type is a six-digit SOC code: indeed, there is a negative relationship between wages and the number of applicants within six-digit SOC codes. Hence, the labor market is characterized by a high degree of specialization and consists of a very large number of narrow sub-markets.

The equilibrium patterns reflect the way in which the labor market (i.e. the market within a job type x) prevents adverse selection of low-productivity workers into jobs for high-productivity workers. Except in the sub-market for the least productive workers, wages and queue lengths are generally larger than in a world without adverse selection. If adverse selection were not a concern, a counterfactual negative relationship between wages and applications would arise in each job title, as shown in the proof of lemma 2.[11] The empirical relationship between wages and applications therefore provides information on the structure of the labor market.

Next, we consider the model's predictions across job types. First, we analyze the relationship between wages, skill and human capital. As mentioned before, we focus on the case in

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

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 higher-ranked job types pay higher wages (i.e. doctors earn more than nurses) if the value of output $p(x)$ increases faster than the vacancy creation cost $c(x)$, as we assumed.

With heterogeneity in productivity, the relationship between wages and skill no longer holds one-to-one, since good nurses may earn more than bad doctors. 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.*

Next, 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 doctors receive fewer applications than vacancies for nurses. The main intuition is as follows. Lemma 2 reveals that, without heterogeneity, a firm's payoff is equal to a certain fraction of the surplus $[p(x) - b(x)]\, y$, minus the vacancy creation cost. To be precise, given free entry, we have that:

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

The fraction of net surplus that goes to the firm depends on the queue length λ that it attracts. If queues were constant across job types, the fraction would be constant as well. In that case, firms posting a high-skill job would make a larger profit than firms posting a low-skill job since the surplus $[p(x) - b(x)]\, y$ increases faster than the vacancy creation costs. In equilibrium, firms must get equal payoffs across job types, so firms must pay a higher fraction of net 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.

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.

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 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 shorter queues. 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 nurse would be as a doctor relative to the productivity of someone with an MD 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 queues are longer in those, 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 queues 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 transferability, 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 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 a queue λ_i at his own skill level, applies to a job of type $x_d > x_i$ with wage w_d and queue 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 $x_d > x_i$.[12] Clearly, 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 labor share 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 level. Quantitatively, since the labor share 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 at firms that only require marginally more skill than their own, and will therefore never get hired by these firms.

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 consider a second specification in which the output created by the match is a random variable, which is realized when the worker and the firm meet and stays constant for the duration of the match (if one is formed). 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 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

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

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. We conclude that, for the equilibrium described in proposition 1 to survive, 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

In this section, we discuss a number of issues related to the interpretation of our results. 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.

First, the 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. Yet, the empirical literature in labor economics has traditionally used broader classifications of occupations (e.g. Kambourov & Manovskii, 2009b, Poletaev & Robinson, 2008).[13] Our results imply that this ignores a significant amount of heterogeneity and, therefore, that occupations are narrower than typically assumed. This has important implications for e.g. the measured degree of occupational mobility, occupational mismatch, or frictional wage dispersion.

Although skills are not very 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.[14] The fact that we need large skill heterogeneity and limited skill transferability to yield correct empirical predictions has important implications for labor mobility 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.

Second, the assumption that surplus increases faster with skill than the the cost of creating a vacancy is crucial in yielding a negative relationship between wages and job queues 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. queues are shorter 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 unfortunately little conclusive evidence is available. Nevertheless, it seems reasonable to think that, even though recruiting a high-skilled worker is more costly, it is unlikely that this 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 does not contradict the evidence.

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 im-

[13]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.

[14]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).

plications. 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 very robust and virtually undisputed finding (see e.g. Elsby et al. 2010).[15] There is also independent empirical evidence supporting the prediction 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 assumptions in the model.

4.2 Competition and the elasticity of applications with respect to the wage

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 a 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% 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, the equilibrium predicts a unique wage. 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 does not have an infinitely large 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.

[15] 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. 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.

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 also introducing heterogeneity, our model generates equilibrium wage dispersion. 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 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. Then a marginal increase in the wage by a particular firm may not affect the worker's application decision, if he knows that his chances of being hired by that firm are low due to a bad realization of the match-specific productivity component. 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 what 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 queue lengths 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.[16]

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 queue length 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. In-

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

stead, it is enough that high-skilled workers are significantly more productive than 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 job queues. We first document a number of new empirical facts. Using data from CareerBuilder.com, the largest US employment website, we show that firms that pay higher wages attract better 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. 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% increase in the number of applicants.

We explain these 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. We argue that 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 implies that it is very important in empirical applications 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.

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 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.

Tables

	obs.	mean	s.d.	min	max
Master degree	2,282	0.273	0.255	0	1
High experience	2,379	0.506	0.265	0	1
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 1: Summary statistics

	I	II	III	IV	V
Log yearly wage	-0.770***	-0.642***	-0.710***	0.593**	0.876***
	(0.052)	(0.075)	(0.087)	(0.302)	(0.283)
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	No	No	No	Yes	Yes
Observations	11,708	11,708	11,708	11,708	11,708
R^2	0.017	0.133	0.363	0.480	

Robust standard errors in parentheses (except in col. V). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 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 2: Effect of log wage on the number of applications per 100 views.

	High Experience		Master Degree	
	I	II	I	II
Log yearly wage	0.136***	0.130*	0.173***	0.096
	(0.010)	(0.068)	(0.010)	(0.077)
Controls	No	Yes	No	Yes
NAICS (2 digits)	No	Yes	No	Yes
Job title	No	Yes	No	Yes
Observations	1,755	1,300	1,696	1,257
R^2	0.091	0.817	0.154	0.821

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. High experience = fraction of applicants with at least 11 years of experience; master degree = fraction of applicants with a master degree. 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 3: Effect of log wage on the quality of applicants

	I	II
Log yearly wage	-1.045***	1.582***
	(0.089)	(0.375)
Controls	No	Yes
NAICS (2 digits)	No	Yes
Job title	No	Yes
Observations	11,694	11,694
R^2	0.01	0.568

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.

Table 4: Effect of log wage on the number of clicks per 100 views

	I	II	III
Posted a wage	0.478***	0.172**	0.091
	(0.031)	(0.081)	(0.071)
Controls	No	Yes	Yes
NAICS (2 digits)	No	Yes	Yes
Firm effects	No	No	Yes
Job title	No	Yes	Yes
Observations	61,051	61,050	61,050
R^2	0.006	0.498	

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

	High Experience		Master Degree	
	I	II	I	II
Posted a wage	0.006	-0.037	0.005	-0.016
	(0.013)	(0.072)	(0.013)	(0.072)
Controls	No	Yes	No	Yes
NAICS (2 digits)	No	Yes	No	Yes
Job title	No	Yes	No	Yes
Observations	2,379	1,774	2,282	1,704
R^2	0.000	0.790	0.000	0.830

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. High experience = fraction of applicants with at least 11 years of experience; master degree = fraction of applicants with a master degree. 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 wage posting on the quality of applicants

	I	II	III
Log yearly wage	-0.702***	-0.579***	-0.530***
	(0.067)	(0.115)	(0.104)
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	5,609	5,609	5,609
R^2	0.012	0.169	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.

Table 7: Effect of log wage on the number of applications per 100 views, using job titles with at least two observations

A Proofs

A.1 Proof of Lemma 1

Proof. Consider first the Bellman equations. Suppose that in equilibrium a firm posts a job ad (x, w) and attracts a queue length $\lambda(x, w)$, 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[17]

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

and

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

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

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

and

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

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 a queue length $\lambda^* = \lambda(x, w^*)$ 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 in the queue, 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, marginally higher than w^*, obtains a strictly positive payoff V_V^d. Workers are willing to

[17] To simplify notation, the dependence of the value functions and the queue length on x and w is suppressed.

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 queue length λ_d that the deviant attracts is such that a high-type 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-type applicants are indifferent between w^* and $w_d > w^*$, it must be that the deviant job has a slightly longer queue length, i.e. $\lambda_d > \lambda^*$. In this situation, low-type applicants face a tradeoff between a low wage w^* and a short queue λ^* or a slightly higher wage w_d and longer queue λ_d. Since low-type 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-type 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 queue λ_d. Note that since the deviant offers a marginally higher wage, he must attract a marginally longer queue for high-type 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.

A.2 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 queue 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 queue length that it attracts will be determined by this equation. Alternatively, we can think of the firm as choosing a queue length, 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 queue length 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 queue length 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 queue length λ determines which fraction of the surplus created by a match goes to the firm. If the queue length λ were constant across y, this fraction would be constant. Since the created surplus is larger for larger values of y, firms attracting high-type workers would obtain a higher payoff. This violates the free-entry condition. Additional entry would take place in those sub-markets, reducing the queue lengths.

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 14, the payoff of such a worker equals

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

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

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

When queue lengths stay the same as in the scenario with commitment, i.e. (15) holds, the left hand side simplifies to

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

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 queue length $\lambda_2 < \lambda_1$.

Equilibrium therefore requires the queue length at high-type 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

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

A.3 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 queue lengths 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. □

A.4 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 the queue length 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. □

A.5 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 a queue 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 queue length 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

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

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

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

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. □

A.6 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 queue length is decreasing in skill, $\frac{d\lambda}{dx} < 0$, the queue length is decreasing in the wage. Because of continuity, a negative correlation between wages and applications then exists for a sufficiently small degree of heterogeneity. □

A.7 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 queue 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 a queue 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 attracting a queue λ.[18] 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 doctor can do as well as a nurse in a nursing 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 $x < x$ 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,

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

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. $\square$

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

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.

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.

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.

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.