Who Gets Hired? The Importance of Competition among Applicants

Who Gets Hired? The Importance of Competition among Applicants

Edward P. Lazear, *Stanford University*

Kathryn L. Shaw, *Stanford University*

Christopher T. Stanton, *Harvard University*

Being hired into a job depends not only on one's own skill but also on that of other applicants. When another able applicant applies, a well-suited worker may be forced into unemployment or into accepting an inferior job. A model of this process defines over- and underqualification and provides predictions on its prevalence and on the wages of mismatched workers. It also implies that unemployment is concentrated among the least skilled workers, while vacancies are concentrated among high-skilled jobs. Four data sets are used to confirm the implications and establish that the hiring probability is low when competing applicants are able.

I. Introduction

Two workers, identical in qualifications, apply for a job. There is one job opening, so the employer hires only one applicant while the other remains unemployed. Luck has played a role in determining outcomes, presumably good luck for the one hired and bad luck for the one turned down.

We thank Peter Kuhn for his insightful comments and participants at the University of Chicago, the Copenhagen Business School, and the Stanford Graduate School of Business. Contact the corresponding author, Kathryn L. Shaw, at kathryns@stanford.edu. Information concerning access to the data used in this paper is available as supplementary material online.

[*Journal of Labor Economics*, 2018, vol. 36, no. S1]
 0734-306X/2018/36S1-0012$10.00
Submitted April 22, 2016; Accepted September 18, 2017

Consider another scenario. Two similar although not identical workers apply. The better of the two is offered the job. The slightly inferior worker is told that the position has been filled but that another, lower-paying job is still available. The worker accepts the position, fearing that the alternative might be long-term unemployment. In this case, luck takes the form of job assignment, but one worker enjoys good luck while the other's luck is less favorable.

Neither of these situations is well described by standard theory. Most production technologies are assumed to be smooth, with substitution across worker types and numbers being permitted. But at least some situations in the real world may be closer to a type of technology with some complementarities, where the notion of a job slot makes more sense. Slots have been analyzed before in the economics literature. The tournament structure,[1] by postulating a winner and a loser, invokes the notion of slots. Prior to that, matching models that date back at least to the early 1960s also analyze allocations based on slots.[2] However, these literatures are primarily conceptual and, to our knowledge, lack empirical evidence on the importance of competitor quality in determining job offers.

Job slots give rise to stochastic outcomes, where luck plays an important role. In markets, luck may take many forms, but the luck that is the focus of attention here is that which affects job offers. Central to the analysis is that a given applicant's luck depends on the others who apply for a job at the same time. As in Waldman (2003), a worker has no control over what others do, but the outcome of a job search or tournament process likely depends crucially on the other applicants with whom the worker competes. The Waldman approach, like the one adopted here, emphasizes the technological matching aspect of slots, where some workers are better suited to one job and others to another. Relatedly, the span-of-control papers (e.g., Lucas 1978; Rosen 1982) allocate workers to firms and supervisory and subordinate positions in them on the basis of talent. Comparative advantage determines job assignment in that literature, and that is the mechanism emphasized here.

Although there is precedent in the literature, most labor market analyses neither emphasize nor document empirically that whether a worker obtains a job depends not only on a worker's own skills but on the skills of others. For example, in human capital theory a worker's wage is determined by his own stock of human capital and its market price. In contrast, in the approach of this analysis even workers with high levels of human capital may

[1] See Lazear and Rosen (1981) and the literature that follows, e.g., DeVaro (2006), who makes explicit reference to slots and the competition created by limiting them.

[2] The matching literature that goes back to Gale and Shapley (1962) has the feature that slots are crucial and matching may not be perfect. The algorithms and results of this literature tend to abstract from market considerations and do not provide the specific implications that are the focus of this paper. Closest to imbedding markets into a matching framework is Becker (1993), who considers matching in marriage markets.

be underemployed and may receive low wages because an even better worker took the scarce job slot.

Whether luck in hiring is important in affecting lifetime worker wealth depends on the cost of mobility and on the thickness of markets.[3] If bad luck in job assignment can be undone rapidly by subsequent job search, it may not be of major consequence. But in some markets, like that for academics, where hiring occurs only at scheduled times, luck in finding a job may have long-term effects. The documented importance of cohort effects on wealth at least suggests that hiring luck may be of some consequence (Oyer 2008; Kahn 2010). At a minimum, it is important to understand the way in which the existence of slots and competition for jobs affects outcomes.

In what follows, a model of slots and within-firm job assignment that yields specific, testable implications is presented. Some implications help reconcile puzzles in empirical findings that are not understood in the context of standard economic theory. Specifically, the analysis below produces the following results, which are borne out in the empirical section.

1. The probability of obtaining a job and the job to which a worker is assigned depends on the quality of the competition. Because firms are slot constrained, a worker remains unemployed when other applicants for a job have superior qualifications. But for those others applicants, the worker would have been employed.
2. Less able workers are more likely to be unemployed than more able ones. Although this finding is well established empirically, it is hardly obvious. Usually, markets for more homogeneous products are thicker than those for less homogeneous ones. Low-ability workers are likely to be more similar to one another than are high-ability ones. This puzzle is explained by the fact that high-ability workers are more flexible and can perform satisfactorily in a larger number of jobs.
3. Analogously, vacancy rates are highest in high-paying, high-quality jobs. Most workers can fill lower-quality positions, but higher-quality positions require high-ability applicants who may not be present. This implication, coupled with the first, provides a cross-sectional Beveridge curve. Friction in job search implies mismatch, and here it takes a specific form that has been observed in data. Firms complain that they cannot find workers while at the same time workers cannot find jobs. But the workers who have a difficult time finding jobs are not the ones suited to the jobs that are vacant.
4. As in the earlier search literature, unemployment, at least of the frictional variety, is a consequence of bad luck.

[3] Burdett and Coles (1999) analyze the role of luck in a dynamic context. If workers have the option of looking again, luck becomes less important, and sorting is improved.

5. It is common to speak about a person being overqualified for a position, but what does that mean formally? A clear definition of over- and underqualification is provided, and that definition yields implications about observed wages for workers who find themselves over- or underqualified for a job. Those who are in the wrong job receive wages below that which would be expected had they been lucky enough to find the appropriate job for their skills.
6. "Bumping" creates overqualification. Workers who are better suited to high-level jobs are bumped into lower level ones or unemployment by workers who are even better qualified. Conversely, underqualification results when a firm settles for a low-quality worker because there is no higher-quality worker available to do the high-level job.

II. Model

The goal of the model is to capture the idea that luck is important in getting hired. A worker must encounter a firm that can make use of the worker's skills, which depends on the qualifications of others who are employed by the firm. Key here is the notion of "slots."[4] A firm is not free to simply add workers to increase output. For example, a school might have a given number of classrooms, and if there is already one teacher per classroom it may not be cost-effective to add another teacher to that room.

The use of slots in the model will be shown to be crucial and helpful in understanding the existence of unemployment. Absent the concept of slots, it is difficult to generate unemployment in equilibrium without reverting to some kind of rigidity, the most obvious of which is sticky wages. Search theory uses a weak notion of slots implicitly because whether a worker locates a firm that wants that particular worker's skills is stochastic and based on the idiosyncratic aspects of both the firm and the worker. But the level of abstraction in search theory is generally too high to generate the implications that are required for analyzing the detailed micro data that will be used in this study. As a consequence, the notion of slots and how those slots relate to others already employed or also applying to a firm is explored.

A formal search theoretic literature by Shimer and Smith (2000) and Shimer (2005) comes closest to generating the implications that are provided by this model. In particular, implicit in Shimer and Smith is a matching model. Imagine heterogeneity among firms and workers such that each worker, in equilibrium, is optimally sorted to a particular firm. The equilibrium wage function generates wages conditional on worker types, where in a frictionless world the firm that is the best match bids the most for a worker of every given type. However, if search is costly, then workers will accept jobs that

[4] Some of the earliest related work is by Diamond (1982), who recognized that a worker's ability to find a job depends on the number of others who are searching. Here, a slot is skill specific, and a particular substitution technology is modeled.

are not their perfect match and firms will hire workers who are not their most preferred type. Implicit in this structure is a definition of over- and under-qualification. Workers who end up working at a firm that prefers cheaper, lower-quality workers than themselves are overqualified and less valuable there than they would be were they to find the optimal match. Those who end up working at a firm that prefers better, more expensive workers than themselves are underqualified and are less valuable there than they would be were they to find the optimal match.

Slot models that emphasizes vacancies are also found in the literature, as in Shimer (2005). That model is characterized by one vacancy per firm. Unemployment results mechanically because if more than one applicant arrives at a firm that has only one vacancy, some will not be employed.

Explicitly modeling the structure of slots contrasts with the vast literature on supply-side unemployment, where business cycles induce workers to stay home because their reservation values exceed their productivity. These prior models are consistent with some real-world observations, but not all. For example, one of two otherwise identical workers, both in terms of productivity and alternative uses of time, may be offered a job while the other is forced into unemployment. If workers stay home when the reservation value exceeds the wage, then neither of the identical workers is unemployed or they both are. The notion of slots is particularly helpful here and conforms to standard intuition about job finding. Once a job is already filled, another equally qualified applicant is not offered employment.

It is also desirable that the model does not create unemployment by assumption. Thus, no worker is inherently unemployable in the sense of having ability so low that he can never add positive value. If a worker does not obtain a job, it is because he has encountered bad luck that precludes productive employment because of the workforce composition, not because he is so unproductive that no firm will employ him.

Another goal of the model is that the worker's wage and standard of living be affected by luck that takes the form of being hired or not and on the job assignment if hired. A worker may be overqualified in the sense of being more productive in another job were it available but may accept the one offered because the better job is already filled.

A. Production

The production function has both slot features and complementarity. Smooth production functions give no role to slots. In standard theory, labor utilization is a continuous variable and, despite diminishing marginal productivity, everything occurs smoothly. This is at odds with what is frequently observed in the real world, where positions are discrete and having an open slot is necessary before a worker can be hired.

The assumption of complementarity creates slots but also a reason for having more than one worker in a firm. For example, there might be an advantage to having one firm serve the same client, so that if there is a problem, one cannot blame an outside party for the difficulties encountered.

A useful benchmark against which to compare results of this model is a world where slots do not exist in the sense that a firm can always hire a worker and the resulting effect on output is smooth. In the standard production function, output is

$$Q = Q(q_1, q_2, \dots, q_n), \tag{1}$$

with $Q_i > 0$, $Q_{ii} < 0$, $\forall\, i$. Each i can be thought of as a worker type, and the q reflects the amount of labor of type i that is used. Production is smooth in the sense that increasing the input of labor of type i increases the output of the firm, albeit at declining rates.

Here, a different technology is assumed. Each firm can use at most one worker for each job, which, for simplicity, is assumed to equal two. Again, this is closer to the vacancy literature discussed above, where a vacancy defines a unique slot and the arrival of more than one worker does not result in more output than can be produced by filling that vacancy.

The comparison, then, is not against that which would occur in a frictionless world but rather that which would occur in a world where any number of workers can be accommodated by the firm, just at a decreasing marginal product. Unlike smooth production in equation (1), in a two-slot case, output at the firm takes the form

$$Q(A_1, A_2) = Z(A_1, A_2)[q_1(A_1) + q_2(A_2)] + [1 - Z(A_1, A_2)]q_2(A_1),$$

where A_1 is the ability of the highest-ability worker, A_2 is the ability of the second highest ability worker, and $Z(A_1, A_2)$ equals 1 if it pays to hire workers into both slots and 0 if it pays to hire a worker into only one of the two slots. Below, $Z(A_1, A_2)$ is derived optimally (in lemma 2).

The point here is that even if multiple workers of identical ability apply to the firm, only one of them can be placed in each job. Unlike the production function in equation (1), more labor cannot be applied to job 1. The number of heads, as opposed to hours, is the determining factor. A slot structure may arise when capital is discrete, as in the case of a single position at the control panel in an automated steel plant. If the slot is unfilled, then q_i is defined to be equal to 0.

To make things concrete, suppose that $Q[q_1(A_1), q_2(A_2)]$ is derived from underlying production:

$$q_1(A) = \gamma + \delta A, \tag{2a}$$

$$q_2(A) = \alpha + \beta A. \tag{2b}$$

Think of job 1 as the difficult job and job 2 as the easy job. In the oDesk data examined below, job 1, the difficult job, is programming, and job 2, the easy job, is administrative support. High-ability workers are better in every job, but they have a comparative advantage in job 1. Thus, let $\alpha > \gamma$ and $\delta > \beta$, as shown in figure 1. High-ability workers produce more than low-ability workers in each of the two jobs because both β and δ are positive, but ability has a greater effect in augmenting output in the difficult job than in the easy job.

Let the worker's reservation value be K, thought of as the value of leisure. The value of a worker's output is then R_{q_i}, where R is the price of the product. Define A_0 as that ability such that a worker would have a value of output in the difficult job that just equals the value of the alternative, K. Using equation (2a),

$$A_0 = \frac{K/R - \gamma}{\delta}.$$

Furthermore, define A^* as the ability such that the worker is equally productive in both jobs (as shown in fig. 1), so that $\gamma + \delta A^* = a + \beta A^*$. Thus,

$$A^* = \frac{\alpha - \gamma}{\delta - \beta}.$$

Then,

$$R(\gamma + \delta A) > K \quad \text{for } A > A_0. \tag{3}$$

Any worker with $A < A_0$ would never be hired into the difficult job because his ability is so low that his output would be below the value of not working.

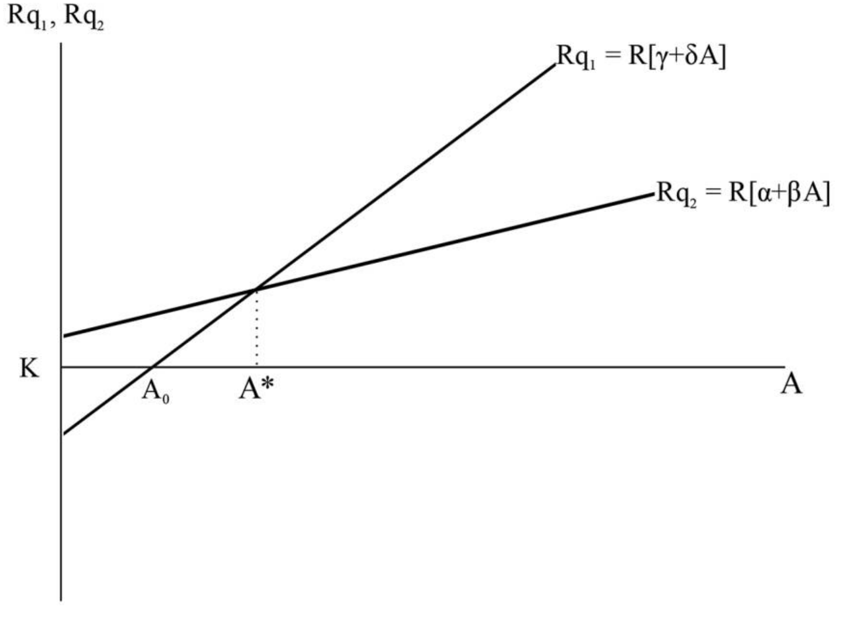

FIG. 1.—Comparative advantage

Also, in keeping with the desire to avoid assuming that no worker is inherently unemployable, at least in normal times, the minimum ability in the population $A_{\min}$ is assumed to be sufficiently high so that

$$R(\alpha + \beta A_{\min}) > K. \tag{4}$$

Finally, to complete the model, assume a single period and that costly search takes the form of allowing each worker to search at one and only one firm. If there are N workers and M firms, each of which is equally likely to be searched by any worker, then the number of workers who arrive at any given firm follows a binomial distribution. The probability of a given worker arriving at any firm is then $1/M$. If h is the number of workers that arrive at a firm, then the density of arrivals is

$$p(h) = \frac{N!}{h!\,(N-h)!}\left(\frac{1}{M}\right)^{h}\left(1-\frac{1}{M}\right)^{N-h}, \tag{5}$$

where $p(h)$ is the proportion of firms that have h applicants, $h = 0, 1, \ldots, N$. Although N is exogenous (perfectly inelastic labor supply is assumed for simplicity), M is derived as part of a competitive equilibrium, described in a later section.

Firms that have zero applicants are uninteresting for the analysis. Those that have one applicant hire that applicant and place her in the job to which her output is highest based on the production function in equations (2a) and (2b). The action all comes from firms that receive two or more applicants because it is only those firms that ever leave workers unemployed.

Because equations (2a) and (2b) imply that output and profit are both (weakly) increasing in worker ability, the firm will always choose to hire at most the two highest-ability workers from among the pool of applicants. Thus, define A_1, A_2 as the ability of the best and second-best applicant, respectively. If more than two workers apply to a firm, then some applicants necessarily are unemployed in the same way that unemployment occurs in the Shimer model. More interesting, however, is the assignment problem among the two best workers and the situations where firms choose optimally to use only one worker even though at least two apply. This is the essence of the slot-based assignment model, and the two lemmas and proposition derived below describe the full array of outcomes for the economy.

Costly search creates bilateral monopoly (the worker has at most one offer and the firm sees two and only two workers), so there is a need to allocate the rents. Although the structure is competitive in the sense that there are many firms and many workers, bargaining opportunities exist once pairing occurs. The bargaining game is not crucial to the model as long as it results in some rent-splitting parameter that is common across firms. Since ex post wage determination is not the focus of this analysis, firms are assumed to commit to a wage schedule when they advertise a job. Firms advertise wage schedules that will later be shown to be consistent with competitive equilibrium.[5] Thus, wages in the two jobs are given

$$w_1(A) = [\lambda(\gamma + \delta A)]R, \tag{6a}$$

$$w_2(A) = [\lambda(\alpha + \beta A)]R, \tag{6b}$$

with $0 \leq \lambda \leq 1$.

It is now possible to derive the implications of the model. First, firms that receive two applicants are no different from those that receive more than two because the firm discards all but the two highest-ability applicants. Thus, the following lemmas relate to firms with two or more applicants.

LEMMA 1: If a firm assigns any worker to the difficult job, it will always be the highest-ability worker, that is, the worker with ability A_1.

Proof: The firm assigns workers to maximize profits, which implies that the better worker should be assigned to the difficult job at a given firm and the poorer worker to the easy job if

$$\gamma + \delta A_1 + \alpha + \beta A_2 - w_1(A_1) - w_2(A_2) > \gamma + \delta A_2 \\ + \alpha + \beta A_1 - w_1(A_2) - w_2(A_1)$$

or if

$$[\gamma + \delta A_1 + \alpha + \beta A_2](1 - \lambda) > [\gamma + \delta A_2 + \alpha + \beta A_1](1 - \lambda),$$

which reduces to

$$\gamma + \delta A_1 + \alpha + \beta A_2 > \gamma + \delta A_2 + \alpha + \beta A_1.$$

Rearranging terms, this requires that

$$(\delta - \beta)(A_1 - A_2) > 0,$$

which must hold because $\delta > \beta$ and $A_1 > A_2$.

LEMMA 2: Both slots are filled if and only if

$$A_2 > \frac{(\beta - \delta)A_1 - \gamma}{\beta}.$$

[5] Reneging on the wage offer is possible but is thought to be analogous to refusing to pay a worker at the end of a pay period after the work has already been done.

Conversely, one and only one slot is filled if

$$A_2 \leq \frac{(\beta - \delta)A_1 - \gamma}{\beta}.$$

Proof: The choice is between hiring two workers or hiring only the best worker and assigning her to the easy job. If the highest-ability worker is best assigned to the difficult job, then there is always gain to hiring the low-ability worker into the easy job by inequality (4). Thus, two slots are filled if and only if

$$\alpha + \beta A_1 - w_2(A_1) < \gamma + \delta A_1 - w_1(A_1) + \alpha + \beta A_2 - w_2(A_2)$$

or if and only if

$$(\alpha + \beta A_1)(1 - \lambda R) < (\gamma + \delta A_1)(1 - \lambda R) + (\alpha + \beta A_2)(1 - \lambda R).$$

This reduces to the condition that

$$A_2 > \frac{(\beta - \delta)A_1 - \gamma}{\beta}.$$

When

$$A_2 \leq \frac{(\beta - \delta)A_1 - \gamma}{\beta},$$

the firm earns more by hiring one worker. One worker is always better than zero because for any $\lambda \leq 1$, the firm never loses money on any worker who is put into the easy job. QED

Recall that $Z(A_1, A_2)$ is an indicator variable that is 1 when a firm fills both slots and 0 when the firm fills only the low-ability slot. Thus, directly from lemma 2,

$$Z(A_1, A_2) = \begin{cases} 1 & \text{if } A_2 > \dfrac{(\beta - \delta)A_1 - \gamma}{\beta}, \\ 0 & \text{otherwise.} \end{cases}$$

It is now possible to state a proposition that provides necessary and sufficient conditions for full employment in the economy.

Proposition 1: Unemployment occurs if for at least one firm $A_2 < [(\beta - \delta)A_1 - \gamma]/\beta$ or if at least one firm sees more than two applicants.

Proof: The second part of the statement is trivial. If a firm can employ only two workers and it receives more than two applications, some of the applicants are denied employment.

The first part of the statement follows directly from lemma 2, which states that one and only one worker is employed when $A_2 < [(\beta - \delta)A_1 - \gamma]/\beta$. The other worker is then unemployed. QED

There are a number of points that come out of this slot-based structure. First, “bumping” occurs. If two high-ability workers, defined as having $A > A^*$, show up at the firm, then the one with the lowest ability is bumped down to the easy job even though he is inherently more productive in the difficult job than the easy job. The worker must do the easy job not because he has a low ability but because the difficult job is best assigned to the even higher ability worker. The worker who is bumped into the easy job earns less than he would have had he been able to secure a difficult job. Similarly, when a firm’s two best applicants are low-ability workers, defined as having $A < A_0$, then the lowest-ability worker is bumped out of a job altogether and ends up being unemployed despite the fact that the firm has two job openings. In this situation, because one job is too difficult for either of the workers to do successfully, it goes unfilled.

Here, the notion of bumping has implications for the probability of employment. When applicants cannot coordinate and firms have a fixed number of positions, some workers will be bumped to unemployment if higher-ability applicants show up for the same slots. Sometimes the inferior applicant will be offered a lower-paying job in the same firm. The data used from oDesk permit testing of the first implication but not of the second.

Second, low-ability workers are more likely to be unemployed than high-ability ones. Again, part of this is trivial. Given that profits are increasing in worker ability in equations (2a) and (2b), a firm that receives more than two applicants discards all but the two best workers. But even firms that receive only two applicants may turn one down; lemma 2 and proposition 1 describe the conditions under which this occurs. Bad luck for low-ability workers takes the form of applying to a firm where at least one applicant is of higher ability but no applicant has sufficiently high ability that the firm wants to employ multiple workers. Because no applicant is assigned to the difficult job, only one worker is hired and other applicants are unemployed. For very low-ability workers, good luck means either that other applicants are of even lower ability or that there is only one higher-ability applicant and that worker has sufficiently high ability to induce the firm to employ both workers (meaning that the high-ability worker has ability greater than $\gamma/(\beta - \delta) + [\beta/(\beta - \delta)]A_2$, which is the condition in lemma 2, rewritten).

The implication that low-ability employees suffer more unemployment is not an obvious one and is, in some respects, counterintuitive. The market for high-ability workers might be thought to be thinner than that for low-ability workers, just as the market for mansions is thinner than that for low-priced development houses.[6] The time on the market for more idiosyncratic goods and services is generally expected to be longer, not shorter, than those for homogenous ones. But high-ability workers are not idiosyncratic. The abil-

[6] See Lazear (1986) for an analysis of pricing, time on the market, and inventories in thick and thin markets.

ity to work in both jobs, as opposed to only one, makes them more employable than the less able. This implication makes intuitive sense. A high-ability worker who cannot find a suitable job has the option of taking a job that is generally filled by a low-ability worker. Especially in tough labor markets, highly educated workers sometimes take jobs that do not make use of their training and skills. But low-ability workers do not have that flexibility. If the only job that is open is one that low-ability workers cannot perform, their only alternative is unemployment. The empirical implication is that workers with low levels of education have higher unemployment rates. This is hypothesis 1 below. While it is generally well known that well-educated workers suffer less unemployment, it is useful to see the magnitudes of these rates and to understand a theoretical logic that is consistent with this observation.

Third, it is possible, although highly unlikely, that the application process is such that no unemployment results. There is nothing in the model that guarantees unemployment. Unemployment is not assumed; it occurs only when there is some bad luck in the world. Under the right distribution of applicants across firms, no unemployment occurs. The unemployment described by this model is of the “frictional” variety, which can result even in very tight labor markets.

Fourth, the jobs dominated by low-ability workers have the lowest vacancy rates. There are never unfilled easy jobs; only difficult jobs sometimes go unfilled. This Beveridge curve–type result (low vacancies with high unemployment) is testable and consistent with occupational difference in mismatch between vacancies and unemployment found in Lazear and Spletzer (2012).[7] Using education as an observable measure of ability, the empirical implication in hypothesis 2 below is that vacancy rates rise with education of the typical worker in the job.

Fifth, workers may be over- or underqualified for jobs but are still profitably employed in those jobs. Recall that A^* is defined as that ability level such that the worker is equally productive in both jobs, given before as $A^* = (\alpha - \gamma)/(\delta - \beta)$. Workers for whom $A > A^*$ prefer to be assigned to the difficult job, which happens if the worker in question is the highest-ability worker of the two who apply. But it is possible, even for a worker whose $A > A^*$, to be the lower-ability worker of the two best to apply to the firm, in which case he will be forced to do the easy job. He is “overqualified” but still successful in the sense that he is more valuable in that task than in taking leisure. Good luck for a high-ability worker (whose $A > A^*$) consists of being paired with a workmate who is of lower ability because the difficult job, which yields higher wages for those whose $A > A^*$, goes to the highest-ability worker. In the oDesk data, this shows up as failing to obtain the job offer. The worker then continues searching either for other jobs at the same

[7] There are chronic “shortages” of workers to fill professional jobs.

firm or for jobs elsewhere. On some occasions, the rejected worker may settle for a lower-level job elsewhere.

An overqualified worker who works in the easy job will be underpaid relative to his expected pay on the higher-skilled job. To see this, use the wage model of equations (6) but simplify to compare wages across ability levels by setting $R = 1$ and $\lambda = 1$. When $A > A^*$ and the worker is in the overqualified job, then the actual wage will be w_2, but the worker would have earned a predicted wage of w_1 if in the job for which he is best qualified. The wage gap—defined as the predicted minus the actual wage, $\hat{W} - W$—is then equal to $(\gamma + \delta A) - (\alpha + \beta A)$. As is evident in figure 1, for $A > A^*$, the value of the wage gap is positive for the overqualified who work in job 2.

Although the overqualified worker earns less than she would were she in the difficult job, she earns more than the typical worker in the easy job because her ability is high relative to those workers and because productivity increases in ability in both jobs. This is a testable implication once an empirical definition of over- and underqualified is established.

Consider next workers who are underqualified for jobs but are profitably employed in those jobs. If both applicants to a firm have ability below A^* but the condition of lemma 2—namely, $A_1 > \gamma/(\beta - \delta) + [\beta/(\beta - \delta)]A_2$—is satisfied, the highest-ability worker is assigned to the difficult job. But because his ability is below A^*, his absolute output in the easy job would be higher. In that sense, he is underqualified for the job, producing low output there relative to what he would have produced in the easy occupation. Good luck in that case consists of being paired with a workmate whose ability is even greater because the higher-ability worker is assigned to the difficult job, where his output and wage are lower than they would be were he assigned to the easy job. In this case, being the higher-ability worker is bad luck.

The wage gap for underqualified workers who are in the difficult job but have ability $A < A^*$ remains $\hat{W} - W$ but now is given by $(\alpha + \beta A) - (\gamma + \delta A)$. The appropriate job for underqualified workers is the easy job, whereas the actual job is the difficult job. As before and, in this case, perhaps counterintuitively, the wage gap is predicted to be positive because these workers would be earning more in the easy job for which they are better suited than in the difficult one into which they are thrust. Were wages attached to jobs rather than to workers, a worker who was underqualified would receive more in that job than in his appropriate job. Thus, for both over- and underqualified workers, the wage gap is predicted to be positive, which is stated as hypothesis 3 below.

There is one difference in wage predictions between under- and overqualified workers. Unlike the overqualified worker, the underqualified worker earns less than the typical worker in the difficult job. Again, because productivity increases in ability in both jobs and because his ability is lower than that of the typical worker in the difficult job, his wages are expected to be lower than average for that job. This is also testable.

A general statement is that good luck consists of applying to a firm where the other applicant's ability permits the worker to be assigned to the job in which he has a comparative advantage. The assignment to jobs, existence of unemployment, wages, and profits all depend on the distribution of talents in the population, on the number of slots of each type, and on luck that manifests in the quality of competing applicants.

Sixth, it is quite possible that firms will complain about not being able to find qualified workers while workers simultaneously complain about not being able to find a job. This is a cross-market Beveridge-like result. Vacancies are high among the difficult jobs, and unemployment is more prevalent among the low-ability workers. Vacancies occur in the difficult jobs only when neither of the workers who arrives at the firm has sufficiently high ability to satisfy the employ-two conditions of lemma 2. When that occurs, the highest-ability applicant bumps the other applicants out of their easy job and into unemployment, creating higher unemployment rates among the less skilled workers. This is the issue of mismatch. Programmer jobs go unfilled when all of a firm's applicants have ability that is too low, again as described by lemma 2.[8]

Seventh, over time there has been an increasing return to education and the variance of income has risen. Such time-series implications can be captured by shifts of the skill gradients originally displayed in figure 1. In figure 2, the skill gradients shift upward, but the upward shift is greater for the difficult job than for the easy job. With skill bias in the underlying production technology, the rising return to skills over time is captured by a greater increase in δ (to δ') than in β (to β'). This is consistent with the notion that technological progress is more complementary with skill level in difficult jobs than it is in easy jobs. The idea is that technology has increased the difference between the output of the abler farmer and the less able one, but it has increased the difference between the output of the abler engineer and the less able one by even more.

The implication is that the skilled worker who is in the easy job will get a bigger pay reduction today than he would have in the past. This outcome is not due merely to the rising return to human capital. It results because the gap between productivity in the job for which a worker is appropriately qualified and the one for which he is overqualified has grown over time. It is the interaction between skill-biased technical change and the slot allocation that comes out of this personnel economics model that generates the result. This is also in keeping with a rising variance of wages over time. Among the highly able, there will be some workers in the difficult job, and there will

[8] Again, see Shimer and Smith (2000) for earlier work that derives this implication.

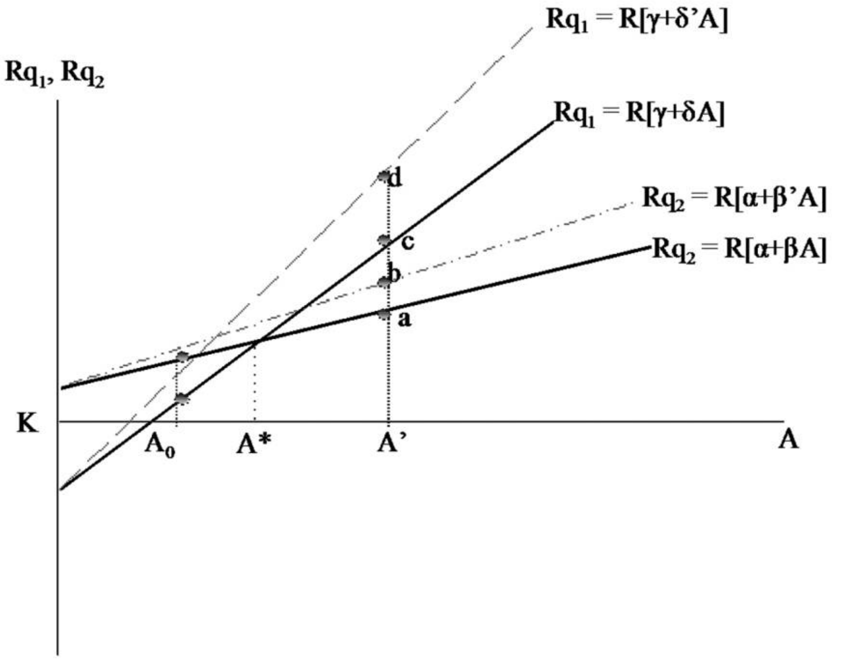

FIG. 2.—Changes in relative returns

be some in the easy job. The pay gap between these two jobs has risen over time, so the variance of pay has risen over time.

B. Competitive Equilibrium and Endogenous Slots

It is important to show that the number of firms and the rent distribution parameter, λ, can be derived from a competitive equilibrium and that the number of firms is consistent with the search technology assumed, namely, that each worker searches at one and only one firm. Allowing free entry among firms, given perfectly inelastic labor supply, produces a zero-profit competitive equilibrium. Intuitively, the margin of adjustment is the stochastic arrival of workers. Were positive profits available more firms would enter, which would lower the probability of an applicant showing up at any given firm. This reduces the highest-order and the second highest order statistics among those firms that obtain at least two applicants and also increases the proportion of firms that obtain zero or one applicant.[9] Entry by more firms means more output, but this also drives up the average cost of output because each firm expects to have a poorer selection of workers. Intuitively, if there were only one firm, it would get every applicant and would

[9] This is a direct property of the binomial distribution that results from this search technology.

hire the best two workers, maximizing efficiency and guaranteeing the lowest possible cost. But one firm means very low output. As the number of firms increases, the expected number of applicants per firm falls, which means that the labor choice is poorer and, consequently, the cost per unit is higher. But more firms mean more output in total.

Equilibrium in the product market obtains with expected price, $E(R)$, determined by the intersection of average cost and product demand. Because labor is the scarce factor and because there is free entry of firms, labor receives all the rents, which means that $\lambda = 1$. The number of firms is set such that expected supply of output equals expected demand for output. Additionally, because arrivals of workers to firms is stochastic, the actual output and therefore realized market price will, in general, differ from the expected amount, but only expectations affect the number of firms in the market because firms choose to operate or not before market output is determined.

Adjustments in product markets come about through variations in the number of firms. Since firms are assumed to be slot constrained, they cannot adjust on the intensive margin by hiring more labor. Thus, the marginal cost of output, which equals the average cost of output, is the cost per unit that results from adding another firm.

Each of N workers can search one and only one firm, so the distribution of arrivals at the various M firms is given by a binomial distribution with $1/M$ being the probability of any given applicant arriving at a specific firm. Define $P(k; 1/M, N)$ as the binomial density that any given firm sees k applicants from the population of N total applicants.[10]

Since the firm keeps at most two applicants and only the best two are relevant, the distributions of the highest-order and the second highest order statistics are inputs into the firm's choice. In general, the density of the rth-order statistic from k arrivals is given by

$$f_k^r(\mathrm{A}) = \frac{k!}{(r-1)!\,(k-r)!}F(A)^{r-1}(1-F(A))^{k-r}f(A) \quad k = 0, \ldots, N. \tag{7}$$

As before, A_1 and A_2 are defined as the best and second-best applicant to come to the firm. Then the density of A_1 is given by

$$f_k^k(A_1) = \frac{k!}{(k-1)!}F(A_1)^{k-1}f(A_1), \tag{8}$$

and the density of A_2 is given by

$$f_k^{k-1}(A_2) = \frac{k!}{(k-2)!}F(A_2)^{k-2}(1-F(A_2))f(A_2). \tag{9}$$

[10] From the binomial, $P(k; 1/M, N) = (N!/(k!\,(N-k)!))(1/M)^k(1-(1/M))^{n-k}$.

The expected output for the firm depends on the number of applicants that the typical firm receives, which depends on the number of firms, M, and the number of workers, N. Given the binomial distribution,

$$
\begin{aligned}
\mathrm{E}(Q) = {} & \\
& P\left(1;\frac{1}{M},N\right)\left[\int_{A^*}^{\infty}(\gamma+\delta A)f(A)\,dA+\int_0^{A^*}(\alpha+\beta A)f(A)\,dA\right] \\
& + \sum_{k=2}^{N}P\left(k;\frac{1}{M},N\right)\left[\begin{aligned}&\int_0^{\infty}\int_{[(\beta-\delta)A_1-\gamma]/\beta}^{\infty}(\gamma+\delta A_1+\alpha+\beta A_2)f_k^k(A_1)f_k^{k-1}(A_2)\,dA_2\,dA_1\\ &+\int_0^{\infty}\int_0^{[(\beta-\delta)A_1-\gamma]/\beta}(\alpha+\beta A_1)f_k^k(A_1)f_k^{k-1}(A_2)\,dA_2\,dA_1\end{aligned}\right].
\end{aligned}
\tag{10}
$$

Some firms end up with only one applicant and therefore only one worker, who is assigned on the basis of comparative advantage. This is shown in the first term. Some firms have two or more applicants. Those firms select the best two and employ either both when the condition of lemma 2 is satisfied or only one when that condition is not met. The expected output from having two or more applicants is given by the second term.

Next, given the technology, average cost is given by the wage bill dividing by output. Just as in equation (10), there are two possibilities. Either the firm gets only one applicant, shown in the first term of equation (11), or the firm gets two or more applicants, shown by the second term. Because all firms are ex ante identical, the expected average cost is the expected cost over the various states, namely, over the different number and quality of applicants that the firm encounters. Thus,

$$
\begin{aligned}
\text{Expected Average Cost} = {} & \\
& P\left(1;\frac{1}{M},N\right)\left[\int_{A^*}^{\infty}\frac{W_1(A)f(A)\,dA}{Q(A,0)}+\int_0^{A^*}\frac{W_2(A)f(A)}{Q(0,A)}\,dA\right] \\
& + \sum_{k=2}^{N}P\left(k;\frac{1}{M},N\right)\left[\begin{aligned}&\int_0^{\infty}\int_{[(\beta-\delta)A_1-\gamma]/\beta}^{\infty}\frac{[W_1(A_1)+W_2(A_2)]}{Q(A_1,A_2)}f_k^k(A_1)f_k^{k-1}(A_2)\,dA_2\,dA_1\\ &+\int_0^{\infty}\int_0^{[(\beta-\delta)A_1-\gamma]/\beta}\frac{W_2(A_1)}{Q(0,A_1)}f_k^k(A_1)f_k^{k-1}(A_2)\,dA_2\,dA_1\end{aligned}\right].
\end{aligned}
\tag{11}
$$

There is free entry of firms, which guarantees zero profits:

$$
\text{Expected Average Cost} = \mathrm{E}(R). \tag{12}
$$

Expected average cost must equal the expected price.

Next, ex ante supply must equal ex ante demand in the product market, so

$$M\mathrm{E}(Q) = D(\mathrm{E}(R)), \tag{13}$$

where $D(R)$ is product demand.

Finally, the actual price, R, is determined by spot market supply and demand, so

$$\sum_m Q_m = D(R), \tag{14}$$

where Q_m is defined as the actual output of firm m. Note that M is determined ex ante, but both the actual Q and R are determined ex post.

Because there are no fixed costs, the zero profit condition requires that $\lambda = 1$. If λ were less than 1, then firms would earn ex post profits that would not be offset by any costs. This nails down the wage functions, given in equations (6a) and (6b).

The system of equations (10)–(14) is five equations that uniquely determine the five unknowns, $\mathrm{E}(Q)$, Average Cost, $\mathrm{E}(R)$, actual R, and M. Actual output, Q, is merely equal to the sum of realized output by each of the M firms, which is on the left side of equation (14).

III. Related Literature

Research on hiring has recently increased due to the availability of new data, but little has been done that resembles the market-level approach taken here. In the past, those studying hiring needed to work with firms to obtain personnel records. The advent of online job boards and online contracting firms has changed this. With these data, much more is known about the types of workers firms seek and those who are hired.[11] But even those data would not answer the questions posed here because most firms provide data only on those hired, not on all applicants.[12] The key implication of the slot-based model is that the probability of employment is based not only on the applicant's qualifications but also on the qualifications of others who apply for the job.

Early results from other data sources support some of the implications of the model. Kudlyak, Lkhagvasuren, and Sysuyev (2012) study a job appli-

[11] See Agrawal et al. (2013), Agrawal, Lacetera, and Lyons (2012), Autor (2001), Brenčič (2010), Gee (2015), Ghani, Kerr, and Stanton (2014), Kuhn and Mansour (2014), Kuhn and Shen (2013a, 2013b), Kuhn and Skuterud (2004), Marinescu and Wolthoff (2015), Nakamura et al. (2009), Stanton and Thomas (2016, 2017), and Pallais (2014). One relevant result is that employers have increased their minimal skill demand in response to the increase in job seekers during the Great Recession (Modestino, Shoag, and Balance 2015).

[12] Burks et al. (2015) show that firms source job applicants using current employee referrals. Hoffman, Kahn, and Li (2015) work with one large job applicant testing firm to identify optimal hiring processes.

cations website and show that when applicants begin their search, their education level explains sorting to job postings: some job vacancies attract applications from highly educated workers and some do not. However, as search proceeds over time, education becomes a weaker predictor of which jobs applicants seek. Over time, the average job seeker begins to apply to jobs that attracted only the less educated in the first week of search. Thus, the highly educated have a broader range of job choices than the less educated and over time will bump the less educated from jobs as they search. This is in keeping with the result in the search literature that wage demands decline with the duration of unemployment.[13]

There is a very large literature on labor demand that is also somewhat related.[14] Prominent in this literature is the empirical research on skill-biased technical change. Overall, the rising introduction of information technologies in the workplace has resulted in rising returns to education and greater demand for workers who perform nonroutine cognitive tasks.[15] This literature would be consistent with the model here, in which more educated workers are demanded across a variety of jobs because the more educated can perform a range of jobs that require cognitive skills. But that could always have been true. In the literature on technical change, the absolute advantage of more skilled workers rises over time.

Other related literature involves vacancy dynamics. Andrews et al. (2008) show, using data from the United Kingdom, that vacancies for nonmanual work are less likely to be filled. Van Ours and Ridder (1991) examine data from the Dutch Bureau of Statistics. They find that jobs that require more education fill more slowly and that vacancy flows are more sensitive to the business cycle for low-education openings. Van Ours and Ridder (1992) find that higher education requirements are associated with longer vacancy durations.

An important additional related literature touches on the nature of how firms post jobs. This literature has implications for the within-firm assignment results, driven by the assumption that multiple jobs are available at the firm to which the worker applies. In particular, the notion is that a worker implicitly applies for multiple jobs at a firm and that a firm receives more than one applicant per job. The latter is surely correct, but Peter Kuhn has shown, using Chinese internet job-posting data, that the typical (modal) number of jobs per

[13] See Rogerson, Shimer, and Wright (2005) for a review of the search theory literature.

[14] Educational effects on hiring arise in the job market signaling literature in which workers invest in education to signal worker quality for the hiring decision (Spence 1973). Altonji and Pierret (2001) show that education as a sorting device for workers diminishes with experience.

[15] For early papers, see Abowd et al. (2007), Autor, Katz, and Krueger (1998), Bresnahan, Brynjolfsson, and Hitt (2002), Katz and Murphy (1992), and Autor, Levy, and Murnane (2003).

ad is one.[16] His data do reveal, however, that the average number of jobs available per advertised job is well above one and appears to be closer to two.

In this model, the worker applies to the firm and accepts the best job that is offered. It is useful to ask whether a worker's application is general or whether it is directed at one specific job. It is difficult to answer this question directly, but some introspection, along with audit studies of the job application process, may be helpful in assessing the assumption of a general application. There are some jobs that are quite specific, and it seems reasonable to assume that the worker applies to only that job. An example is an assistant professorship at a given department. But even here, an applicant might consider a postdoctoral fellowship if the assistant professorship has already been filled. In more typical jobs, applicants often do not even know what the job entails until going to the job interview. Consider, for example, an individual who applies for a managerial position at a major retailer. Without knowing the details of the firm's hierarchy and the duties of each position, it is unlikely that the worker is applying for the advertised job and that one only. More likely is that the individual is applying for a managerial position at the firm and is willing to consider any potentially suitable job offered.

The sociology literature offers some support for the notion that individuals apply to firms rather than to a specific job at the firm, despite the vacancy being advertised as specific to a particular position. In a field experiment studying discrimination, Pager, Bokowski, and Western (2009) document that minority applicants are more likely to be steered or channeled to jobs that are not customer facing even though they apply for the same positions as nonminority candidates. For example, black applicants who applied for server jobs at restaurants were more likely than whites to have busser or runner jobs suggested instead. Also consistent with the model, Pager et al. (2009) document cases of applicants being channeled into better jobs, presumably because no other suitable applicants for those jobs arrived.

Again, it is not necessarily the case that an applicant who suffers bad luck and is not offered a job ends up accepting an inferior job at the same firm. In the model, an applicant may be forced into unemployment by a superior applicant.

IV. Data Sets

The predictions detailed below are tested using data from four sources. The Conference Board provides data on vacancies, while the Current Population Survey (CPS) and the Panel Study of Income Dynamics (PSID) provide data on wages, education, and occupation over time. oDesk provides data on hiring and job applicants to assess how the pool of applicants influences who is hired.

[16] Peter Kuhn, discussion slides, National Bureau of Economic Research conference, Stanford University, November 2015.

The Conference Board conducts a monthly survey of online job postings. The Conference Board Help Wanted OnLine data come from jobs posted on 16,000 internet job boards, corporate job boards, and smaller job sites that serve niche markets. The Conference Board has two measures of job postings. One is "new ads," which are ads posted for the first time in the previous month. The second is "total ads," which are new ads plus ads reposted from the previous month. The data are available by occupation from 2006 to 2014 and are aggregated to the annual level by averaging monthly data. There are 846 observations for 9 years over 94 occupations. The goal in using the Conference Board data is to create a variable that is similar to the vacancy rate measured by occupation, so that vacancy rates can be related to occupational skill levels. Using these Conference Board series, there are more vacancies when jobs go unfilled more than 1 month. Therefore, the measure of vacancies used below is the "unfilled jobs ratio," which is the ratio of unfilled ads to total ads, where unfilled ads is the difference between new ads and total ads. Table 1 shows that the average unfilled jobs ratio across occupations and years is 0.47.

The second data set used below is the March CPS, which provides information on wages and personal characteristics of CPS respondents. Data are obtained from 1975 to 2013. The sample is restricted to men who work full time (defined as more than 35 hours per week) and are between the ages of 25 and 54, with a total sample size of 866,432 observations across the years. Mean values are given in table 1. Wages are defined as real annual earnings, expressed in 2013 dollars. Education is defined as years of education, and this measure varies with the survey year because different surveys catego-

Table 1
Summary Statistics

Variable	Observations	Mean	SD	Minimum	Maximum
Conference Board data:					
Mean Education by Occupation	846	14.03	1.62	9.17	18.13
Unfilled Jobs Ratio	846	.470	.097	.076	.707
CPS data:					
Age (years)	870,665	39.05	8.27	25	54
Education (years)	870,665	13.61	2.76	0	19
Yearly Earnings	866,432	66,849.35	58,876.55	1.09	1,845,631
PSID data:					
Age (years)	28,255	40.77	10.73	25	65
Education (years)	28,106	13.22	2.70	0	17
Tenure (weeks)	27,010	49.25	82.59	0	780
Hourly Wage	28,255	25.013	13.74	.0005	154.74

NOTE.—The sample for the Conference Board data is 9 years (2006–2014) for 94 occupations. The sample for the Current Population Survey (CPS) data is the number of men aged 25–54 for March during the years 1975–2013. The sample for the Panel Study of Income Dynamics (PSID) data is men aged 25–65 from 1968 to 2010. Wages are expressed in 2013 dollars.

rize education groups differently over time. Consequently, all educational groupings are converted into years of education to make them comparable. Another variable used extensively is occupation. The CPS changes its occupational definitions over time. The one used here is the 1990 occupational code, which is largely carried forward to 2013 and backward to 1975. This occupation variable is at a relatively fine level, with an average of 343 occupations delineated by this variable.[17]

The third data set is the PSID, which is an unbalanced panel from 1968 through 2010. The data set follows 5,382 men between the ages of 25 and 65, with an average of 5.25 years of data per person. Mean values for the 28,255 observations are given in table 1. The PSID originates with a sample in 1968 and then introduces new respondents into the sample as parents have children who become respondents. Note that the definition of occupation is coarse in the PSID compared with the other data sets, with only 25 occupations defined.

The fourth data set is from the online labor market oDesk. It is used to quantify the extent of luck in hiring on the extensive margin at the job-opening level. oDesk.com (rebranded as Upwork.com after a merger with their largest competitor, Elance.com) is an online labor market for outsourced services. As of the beginning of 2014, oDesk had processed more than $1.3 billion in contracts (Zhu et al. 2015). The oDesk platform allows employers to post jobs, hire the online applicants, make payments to these globally distributed remote workers, and monitor workers with proprietary project management software.

The oDesk data provide a unique opportunity to study the role of luck in finding jobs because the transactions data contain records of employers' hiring along with the entire set of applications that employers receive. In these data, bad luck for an applicant takes the form of being bumped out of a job into unemployment by a different applicant.[18] Using information on different job categories, it is also possible to study how luck varies based on underlying skill requirements for a given job. When an employer posts a job opening, the task category is selected, along with a job title and a description of the work to be done remotely. Applicants then submit a short cover letter and their electronic resume as displayed on a profile maintained on oDesk;

[17] There are a total of 384 unique 1990 codes throughout the data set. However, not all occupations are present for all years. In earlier years, occupational codes were less precise. For example, no one falls into the Legislator occupational code until 1982 because occupational definitions before 1982 were too imprecise to place individuals into this category. There are 293 occupational codes in the data set for 1975 and a high of 373 occupational codes in 1994.

[18] The oDesk data are best used to study questions at the job-opening level. Inference about workers' careers is more difficult because oDesk captures only online activity, while workers' outside options are not observed. For this reason, little can be said about workers' movement across job categories or about assignment to different jobs.

importantly, they also bid an hourly wage.[19] For workers who have worked on oDesk before, there is a public evaluation (score of 1 to 5) of past performance done by employers. Stanton and Thomas (2016, 2017), Ghani et al. (2014), Horton (2010), Pallais (2014), and Agrawal et al. (2012, 2013) describe oDesk marketplace institutions in more detail.[20]

Table 2 provides summary information about jobs posted on oDesk by different task category. There are nine main categories of jobs posted on oDesk. The data used cover the period from January 2008 to June 2010.[21] The two largest job categories are administrative support and web development. Taking wages as a proxy for skills, variation in average wages paid across categories is evident in the data. Variation is also evident in within-job category quantiles of the wage distribution; the 10th and 90th percentiles within each job category are also displayed. The administrative support job category has lower wages than web development at the mean and across quantiles of the wage distribution.

V. Empirical Results

The main points from the theoretical model are used in sequence to develop a number of empirical implications.

> Hypothesis 1: Bumping occurs. If multiple high-ability workers show up at the firm, then the highest-ability worker gets the job offer and the other applicants remain unemployed or accept a different job.

The most obvious and important implication of the analysis is that the probability of receiving a job offer depends on the quality of competing applicants. The oDesk data are used to test for this. In this online job site, the employer posts a job opening, and applicants respond to that opening with their resumes and wage bids. The employer hires one of the applicants or none at all.

The goal of working with the oDesk data is to estimate how the arrival of other workers changes the hiring probability for an individual applicant. Because characteristics of applicants and jobs are known, it is possible to assess how hiring varies with worker quality, rival applicant quality, and the skill requirements of a job. In other contexts or data, it may be feasible to examine

[19] Employers may also search for worker profiles directly and invite individual workers to apply.

[20] For additional details on the global distribution of work on Upwork/oDesk, see Horton, Kerr, and Stanton (2017). For additional details about job matching on Upwork/oDesk, see Horton (2017), who studies matching frictions and algorithmic recommendations to help alleviate them.

[21] Use of data starting after 2008 ensures that changes in the market in the post-recession period can be captured by a low-dimensional trend. Further details about the sample composition are given in the note to table 2.

Table 2
Summary Statistics by Job Category, oDesk Data

Job Category	Vacancies	Mean Applications per Vacancy	Mean Probability That Vacancy Is Filled	Mean Log Hourly Wage for Those Hired	10th Percentile of Log Wage	90th Percentile of Log Wage	SD of Log Wage
Administrative support	14,968	55.56	.28	1.06	.10	2.05	.73
Business services	2,335	26.57	.20	1.74	.51	2.88	.89
Customer service	1,371	60.52	.14	1.48	.69	2.41	.71
Design and multimedia	10,465	26.53	.25	2.24	1.37	3.00	.64
Networking and information systems	2,530	19.46	.20	2.68	1.72	3.51	.71
Sales and marketing	10,402	21.49	.20	1.50	.55	2.59	.77
Software development	10,287	18.35	.16	2.66	2.08	3.22	.52
Web development	48,021	26.50	.22	2.41	1.72	3.00	.54
Writing and translation	10,502	20.05	.21	1.82	.80	2.81	.82

Note.—Shown are summary statistics for oDesk vacancies posted between January 2008 and June 2010. Only jobs paying hourly contracts are included, and summary statistics on wages are for workers who are hired. Jobs must have at least one worker-initiated application to be included in the sample. For the purposes of estimating a worker quality index in table 3, the choice set must be observed. For this reason, vacancies that overlap with other vacancies posted by the same employer are dropped.

how bumping varies with the number of applicants to a job. However, the focus here is on worker and other applicant quality rather than the number of applicants because on oDesk there tends to be many more applicants than available slots.

Applicant characteristics are multidimensional, but hiring requires employers to rank applicants against one another, collapsing characteristics into a single index. Because hiring decisions are observed, data on employers' revealed preference are used to construct an index of worker attractiveness that allows workers to be quality ranked. Stanton and Thomas (2016) provide the probability model for estimating employers' weights on worker characteristics for the purposes of forming this index. Some details about the model are provided here, while additional details are included in the appendix.

In the hiring probability model, the employer who posts job opening i observes a vector of measured characteristics for applicant j, X_j. The employer forms a quality index from these characteristics for the purposes of finding the best employee to be hired. This quality index, labeled q_j, takes the form

$$q_j = \exp(X_j\beta).$$

The parameter vector β allows for a mapping of multiple worker characteristics into a linear index. The variables X, in the quality index, are those listed in table A1, except the wage variables.

The employer also observes hourly wage bids for each of the J applicants. The employer therefore trades off higher productivity with the hourly wage. This leads to a hiring rule in which the employer chooses the worker who produces the most output per hourly wage subject to the best applicant's wage-adjusted productivity being greater than the employer's outside (no-hire) option. The objective when hiring on job opening i is $\max_{\{j\in J_i\cup 0\}}((q_j \exp(\varepsilon_j))/w_j)$, where q_j is the expected quality of worker j, w_j is the wage bid for worker j on opening i, and ε_j is assumed to be a purely idiosyncratic type I extreme value error.[22]

The assumption of extreme value errors means that a logarithmic transformation of the employer's objective gives a conditional logit form for the probability that applicant j is hired. After taking logs, the probability that worker j is hired is

$$\frac{\exp(X_j\beta - \alpha\log(w_j))}{1 + \sum_j \exp(X_j\beta - \alpha\log(w_j))}. \tag{15}$$

Consistent estimation of β and α is required for the purposes of estimating how workers of different quality affect bumping of other workers. How-

[22] The employer's choice set includes J_i applicants indexed by $\{q_j, \varepsilon_j, w_j\}_{j=1}^{J_i}$ and a no hire option, $\{0, \varepsilon_0, 0\}$.

ever, the ability for bids to adjust to job characteristics, omitted individual characteristics, or, most importantly, the extent of competition may bias estimates of the quality index and alter conclusions about the extent of bumping. Therefore, a technique to account for the endogenous portion of bids, instrumenting log bids with workers' local exchange rates, is used. Applicants' local labor market opportunities are denominated in their own currency, while contracts on oDesk are denominated in dollars. With any friction that limits immediate adjustment of local prices to exchange rate parity, appreciation of the dollar relative to the local currency shifts the dollar-denominated relative wage between online and offline work. This shift in exchange rates is expected to change equilibrium bidding behavior. Details about the implementation of the control function approach of Petrin and Train (2010), along with further details about model estimation, are contained in the appendix.

After an attempt to account for endogenous bids, the choice model is estimated separately for the two largest job categories, administrative support and web programming. The parameter estimates and the choice model are used to estimate the sensitivity of the hiring probability for a given worker to the characteristics of other workers who apply for a job.[23]

The estimates are used to form a linear index of worker quality that comes from employers' revealed preference. The revealed preference index is the fitted value $X_{jt}\hat{\beta}_{\text{JobCategory}}$. For the purposes of analyzing bumping, applicants are ranked on the basis of the index.

Given this setup, it is possible to test hypothesis 1, namely, that bumping occurs. The probability of bumping is hypothesized to depend on the required skill level for the job category, the worker's own quality, and the quality of subsequent applicants. This is taken to data in the following way. First, workers themselves are ranked relative to the distribution of quality. A worker is said to be of good quality if he or she is above the median quality index for a job category; otherwise, the worker is classified as bad quality. Second, when considering the effect of other applicants, a worker is said to be lucky if the next applicant to arrive has a lower index value. A worker is unlucky if the next applicant to arrive has a higher index value.

The probability of hiring is then estimated as a function of the characteristics of the next worker to apply to an opening; the probability is allowed to vary on the basis of the job category and the worker's own quality.[24] For

[23] The parameters of the choice model are not of direct interest, so they are reported in table A2. However, the basic results are sensible: the estimated parameter values show that employers value applicants with better feedback scores and those with past experience. For some parameter values, the estimated weights on worker characteristics are larger when log wage bids and the control function enter the model, suggesting that wage bids may be positively correlated with characteristics that employers value.

[24] Use of the next applicant eliminates the concern that workers may be strategic in applying if some part of the set of prior applicants is observable.

example, the experiment for applicant 1 is whether applicant 2 has a better or worse quality index; this is repeated for applicant 3 matched to applicant 4 and for applicant 5 matched to applicant 6, taking into account the job category and whether each worker is of good or bad quality. For the primary analysis, only the first several applicants are used to balance observations from employers that hire quickly with those that wait to receive many applications.[25] Results are then reported (in subsequent columns of the table) using the first 20 applicants (who are paired as the first six were) to check robustness.

Panel A of table 3 displays the results. The primary finding is that the likelihood of being hired depends on the quality of other applicants. Good applicants are uniformly more likely to get a job when the next applicant to arrive is of lower quality. Although seemingly obvious, we are unaware of prior empirical evidence on this point.

For both administrative support and web programming, the change in hiring probability due to luck is substantial. In administrative support, the hiring probability falls by about 38% (col. 1) for a good applicant if the next applicant to arrive is even better. The estimated sensitivity to the quality of other applicants changes most when controls for price are omitted. After allowing for wage adjustment (col. 2), the effect remains substantial, at about 18%. In web programming, the decline is also about 18% after controlling for wage bids (col. 5). Another way to state this is that a good worker's chances of being hired decline if the next applicant to arrive has a more favorable quality index. Adding the additional applicants to reach 20 applicants does not change the conclusion (cols. 3, 6).

For bad applicants (those below the median of the quality index), luck also plays a role, but the level change in hiring probability due to luck is smaller than it is for good applicants; these workers are very unlikely to be hired, and with multiple applicants to a job conditioning on only the identity of the next applicant when computing luck has a smaller effect on hiring probabilities.

The results in table 3 corroborate the basic assumptions on which the model is constructed. Hiring depends not only on a given applicant's characteristics but also on the characteristics of the others with whom he or she competes for the job. The job to which an applicant is assigned, if any, depends on the competition. Although this seems obvious at the most intuitive level, it is, as far as we are aware, the first evidence of its kind that establishes the relative nature of the hiring process. If a better applicant is present, the worker is given a lower-quality job or none at all. It is also consistent with the view that slots are a fundamental part of the hiring process.

[25] A spline with applicant order is included in the characteristics, X. Only worker-initiated applicants are considered for the purposes of these calculations. Ninety-two percent of the first six applications to web programming jobs are initiated by workers.

Table 3
Hiring Probabilities Based on Who Else Applies for a Job

	Sample, Number of Applicants					
	Administrative Support, n = 6 (1)	Administrative Support, Including Bids and Control Function, n = 6 (2)	Administrative Support, Including Bids and Control Function, n = 20 (3)	Web Programming, n = 6 (4)	Web Programming, Including Bids and Control Function, n = 6 (5)	Web Programming, Including Bids and Control Function, n = 20 (6)
	A. Hiring Probabilities by Own and Future Applicant Type					
Good Applicant and Better Next Applicant	.0111	.0140	.0113	.0119	.0122	.0114
Good Applicant and Worse Next Applicant	.0180	.0171	.0137	.0167	.0150	.0140
Proportion difference	−.383	−.181	−.175	−.287	−.187	−.147
	(.014)	(.029)	(.024)	(.011)	(.014)	(.01)
Bad Applicant and Better Next Applicant	.00109	.0117	.00982	.00504	.00938	.00847
Bad Applicant and Worse Next Applicant	.00251	.00989	.00849	.00674	.0110	.00977
Proportion difference	−.566	.183[a]	.157[a]	−.252	−.147	−.133
	(.202)	(.042)	(.028)	(.014)	(.02)	(.015)

	B. Hiring Probability Regressions with Employer and Applicant Position Fixed Effects (Baseline Is a Bad Applicant with a Worse Next Applicant)					
Good Applicant and Better Next Applicant	−.00803	−.00320	−.00242	−.00500	−.00280	−.00272
	(.0002)	(.0004)	(.0003)	(.0002)	(.0002)	(.0001)
Good Applicant	.0161	.00694	.00472	.00981	.00503	.00455
	(.0004)	(.0012)	(.0009)	(.0004)	(.0005)	(.0004)
Bad Applicant and Better Next Applicant	.00150	.00131[a]	.000819[a]	−.00166	−.00115	−.00125
	(.0001)	(.0003)	(.0002)	(.0001)	(.0002)	(.0001)
Constant	.00259	.0102	.00881	.00693	.0102	.00957
	(.0002)	(.0012)	(.0009)	(.0002)	(.0005)	(.0004)
Number of observations	27,008	27,008	87,994	73,459	73,459	227,709
R^2	.72	.628	.641	.717	.599	.605

NOTE.—The sample is described in the text. In both panels, the dependent variable is the hiring probability from the conditional logit model. A good or bad applicant is defined as one above or below the median for the set of worker characteristics, X, times coefficients, β. The index does not include the control function and the log bid. Whether the next applicant is better or worse is coded using the index of applicant quality. All standard errors (in parentheses) are calculated using block-bootstrap replications to account for parameter uncertainty in forming $X\beta$. In panel B, regression results are presented from models regressing the hiring probability on the measures of paired applicant quality along with employer-by-applicant position fixed effects. Coefficients are additive relative to the constant. The excluded category is a bad applicant. Employer-invited applications are excluded.

[a] The difference in signs between col. 1 (without accounting for bids) and cols. 2 and 3 (accounting for bids) is driven by bids not being included in the quality index used to rank applicants. The sign in cols. 2 and 3 is negative when the bid and control function are included in the quality ranking.

However, there is a potential concern with many of these estimates due to the applicant seeing the job description, which is unobserved by the econometrician. Any sorting based on the job description is not a problem for the conditional logit estimates behind the quality index because pairwise comparisons between workers identify the parameters. However, sorting may be problematic for assessing an employers' comparison of a worker and the no-hire option. Sorting may also muddle the interpretation of luck: are some workers unlucky due to other applicants' arrival, or are workers inappropriately applying to jobs for which they are likely to be unqualified, taking a chance on themselves? For example, when it is a tough job, mostly good applicants apply so the first applicant is less likely to get the job because it is only for highly skilled people. This is like low-quality students who take a shot at applying to Stanford. They have a small chance of getting in, not because they were unlucky to be followed by a high-quality applicant but rather because the quality of the applicant pool reflects the overall difficulty of getting into the program in the first place.

The skill categories are broad, and it is possible that the jobs and their requirements differ significantly even within the administrative or programming category. Obviously, lacking further information little can be done to address this concern. However, to the extent that a given employer posts similar jobs within each category, including employer fixed effects cleans out some of the unobserved variation. Models with employer fixed effects restrict comparison to within-employer differences in the composition of applicants across job openings. To make the presentation comparable to panel A, the comparison must be across job openings for a given employer. To do this, the estimated hiring probability is regressed on employer-by-applicant position fixed effects and indicators for the quality of workers who arrive subsequently. The equation is

$$\hat{p}_{ij} = \alpha + \beta_1 \mathrm{Good}_{jo} \times \mathrm{Better}_{o+1} + \beta_2 \mathrm{Good}_{jo} + \beta_3 \mathrm{Bad}_{jo} \times \mathrm{Better}_{o+1} + e \times o + \epsilon,$$

where o indexes the order of application to job opening i, Good and Bad are indicators for a worker above or below the median of the quality index, and Better indicates that the next worker to arrive is ranked higher than worker j who is applicant o. The equation also includes fixed effects for the employer e by applicant order o. The idea behind these results is to measure whether a high-quality first applicant to job A is more likely to get the job if the second applicant is of low quality compared with a similar high-quality first applicant to job B when an even higher quality second applicant appears. By including employer by applicant order fixed effects in this comparison, this specification removes any systematic unobserved differences in employer quality.

The results, which now include the applicant order fixed effects, are presented in panel B of table 3. The results are presented as regression output with block-bootstrapped (by employer) standard errors. For applicants above the median of the quality index, the results are similar to those in panel A with the inclusion of employer-by-applicant fixed effects. For example, in column 2 a good applicant to an administrative support position has a hiring probability of about 0.017 (calculated by adding the coefficient on Good Applicant and the Constant term) if the next applicant is ranked lower, but this probability falls to 0.014 if the next applicant is better. This decline is on the order of about 19%. The decline is also about 19% in column 5 for web programming. Overall, these results suggest that tests for the presence of luck are not being confounded by permanent unobserved attributes about jobs that vary with employers.

It seems clear from the oDesk results that hypothesis 1 is mostly confirmed.[26] These results provide strong evidence that at least in the context of this one platform where all applicants' characteristics can be observed, the probability of obtaining a job offer depends on the quality of the competition.

> Hypothesis 2: Low-ability workers are more likely to be unemployed than high-ability ones. High-ability workers can work both the easy job and the difficult job, which makes them more employable. The model predicts that workers with sufficiently high ability can never be unemployed.[27]

Figure 3 shows the unemployment rate by education using Bureau of Labor Statistics (BLS) information. As expected, there is a considerable increase in unemployment as education falls. To reiterate, although the fact is not a new one, the pattern requires explanation. Low skill does not mean unemployable in the same sense that low-priced, lower-quality goods are not more likely to stay on the shelves of a store longer than are high-priced, higher-quality goods. Indeed, in most cases the reverse is true. The reason that the low skilled are more likely to suffer unemployment than the highly skilled is that the highly skilled are capable of doing a larger variety of jobs, whereas the less skilled can do many fewer. This means that if an applicant encounters a situation where other applicants are better suited to the job than he, the high-ability applicant may be offered another job when the low-ability applicant is not.

[26] The exception is the difference in subsequent applicant quality for bad applicants in administrative support. These results are driven by wage adjustment; including wages as part of the quality index makes the sign consistent with the other estimates.

[27] Lemma 2 yields a sufficient condition for never suffering unemployment. The most stringent form of the condition occurs when the lowest-ability worker has ability equal to zero. Then, as long as A_1 exceeds $-\gamma/(\delta - \beta)$, it is certain that both workers are employed. Thus, any worker with ability greater than $-\gamma/(\delta - \beta)$ can be certain that he will be at a firm that employs both applicants and can never suffer unemployment.

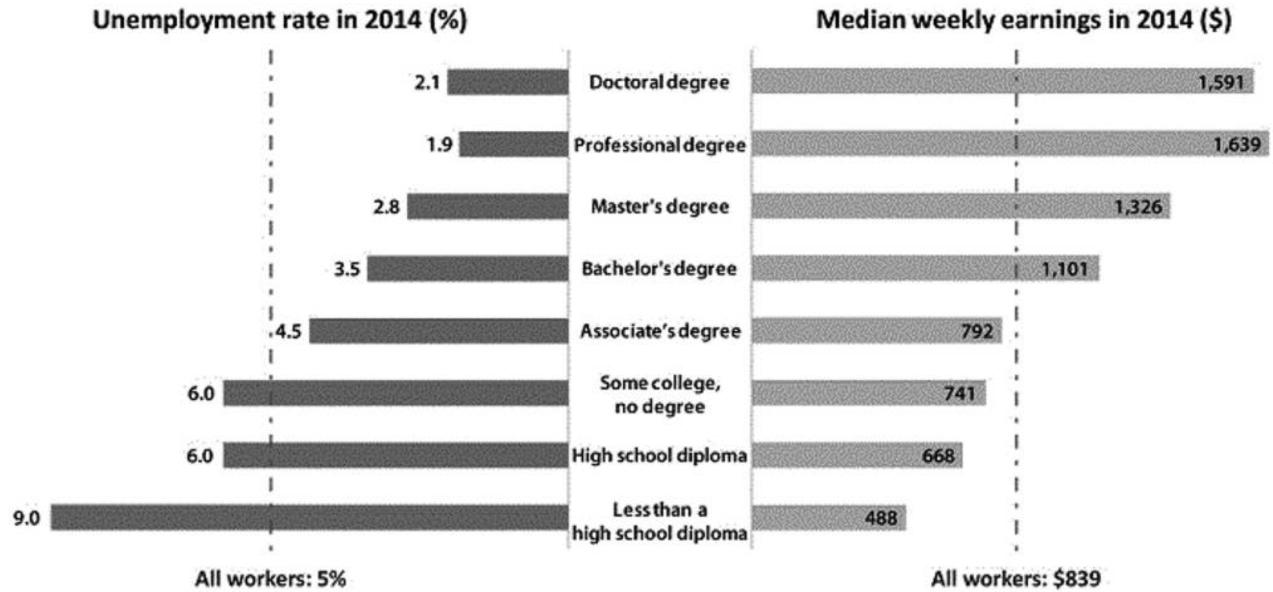

FIG. 3.—Earnings and unemployment rates by educational attainment. Data are for persons aged 25 and over. Earnings are for full-time wage and salary workers. Source: Current Population Survey, US Bureau of Labor Statistics, employment projections, April 2, 2015 (http://www.bls.gov/emp/ep_chart_001.htm). A color version of this figure is available online.

Hypothesis 2 also states that high-ability workers should have less unemployment because they can work in the easy and difficult jobs: it is true in the CPS data that the highly educated work a wider range of jobs. The CPS data for 1975–2013 are used to calculate the mean education by occupation (for 343 occupations) and the variance of education for each occupation. The result is that the higher the mean education by occupation, the lower the variance of education—implying that difficult occupations employ only the highly educated. Easy occupations employ a wider range of less educated and highly educated people.

> Hypothesis 3: Vacancy rates are highest for the high-skilled jobs. The reason is that easy jobs are never unfilled: a high- or low-ability worker can take them. The difficult jobs may go unfilled if no worker with sufficiently high ability arrives at the firm.

As described above, the Conference Board vacancy rate in these data is the unfilled jobs ratio, equal to the unfilled job postings divided by the total job postings. The hypothesis is that this rate rises with skill level. The Conference Board data include the Standard Occupational Classification three-digit occupation, of which there are 94. An occupational skill level is attached to each occupation by going to the March CPS data and calculating the average education for each of these occupations by year, a variable labeled Occupational Education: highly skilled occupations are those occupied by highly educated people. The unit of analysis is an occupation-year: there are 94 occupations times 9 years, or 846 observations.

The first test of hypothesis 3 is in table 4. The regression that is estimated is the Conference Board measure of vacancies (i.e., the unfilled job postings ratio) as a function of the level of education in each occupation (the Occupational Education), with this Occupational Education variable permitted to have separate coefficients for each year of data. Regression results show that an increase in the Occupational Education results in a higher unfilled jobs ratio. Job postings stay unfilled longer when the jobs are more skilled. Each Occupational Education × Year variable is an independent test of the hypothesis, so the fact that all 9 years produce significant results is strong confirmation that unfilled vacancies are higher in the occupations with the highest levels of education. The coefficients are sizable: moving from a high school–educated occupation to a college-educated occupation increases unfilled jobs by about 15% over a mean of 47%.

The second test of hypothesis 3 is in table 5 using oDesk online vacancy data. The dependent variable is whether a posted job has been filled. It is regressed on pay as a measure of job skill as well as employer fixed effects and time fixed effects. Employer effects help to remove unobserved employer differences in familiarity with the platform or differences in unobserved employer attractiveness. The two measures of pay used to proxy the skill level

Table 4
Conference Board Online Job Postings Regression Results

	Unfilled Jobs Ratio
Occupational Education × Year 2006	.040***
	(.0070)
Occupational Education × Year 2007	.038***
	(.0067)
Occupational Education × Year 2008	.037***
	(.0065)
Occupational Education × Year 2009	.036***
	(.0065)
Occupational Education × Year 2010	.035***
	(.0063)
Occupational Education × Year 2011	.034***
	(.0062)
Occupational Education × Year 2012	.036***
	(.0063)
Occupational Education × Year 2013	.037***
	(.0062)
Occupational Education × Year 2014	.036***
	(.0062)
Constant	−.044
	(.094)
N	846
R^2	.4144

NOTE.—The dependent variable is the percentage of online job postings that are unfilled in an average month by year. The Occupational Education level is the mean education level from Current Population Survey data for the 94 occupations. The education average for 2014 is imputed from 2013 data. Regression is weighted by the number of observations in each occupation. Standard errors (in parentheses) are clustered by occupation.

*** $p < .01$.

of the job are the mean wage in the job category and the 90th percentile of wages in the job category. The probability of filling a job is negatively related to both measures. Table 5 also displays this same regression with the coefficients on the mean wage permitted to vary with the main job categories in the data (col. 3): these results also show that the probability of filling a job is negatively related to skill (i.e., wage) for each job category.

By way of background, note that in table 2 applications per vacancy are highest in the largest low-skill tasks (as measured by the mean or the 90th percentile of wages). Hiring rates are also higher in jobs with low skill requirements, as shown by a comparison of column 3 for administrative support with web programming or software development.

In sum, using measures of job vacancies for jobs posted online, vacancies rise with skill level, which is consistent with the model presented here.

The combination of the results for hypotheses 2 and 3 produces a cross-sectional analogue of the Beveridge curve. The Beveridge curve is usually applied to the economy as a whole over time and reveals that periods of high

Table 5
Probability of Filling a Job in oDesk Data

	All Job Categories		
	(1)	(2)	(3)[a]
Mean wage in job category	−.0449***		
	(.00642)		
90th percentile of wages in category		−.0849***	
		(.0189)	
Administrative support omitted (baseline)			
Design and multimedia			−.0529***
			(.0127)
Networking and information systems			−.0684***
			(.0204)
Sales and marketing			−.0940***
			(.0119)
Software development			−.141***
			(.0138)
Web development			−.0761***
			(.01000)
Writing and translation			−.0893***
			(.0121)
Firm effects	Y	Y	Y
Time effects	Y	Y	Y
Number of job openings	110,881	110,881	110,881
Number of employers	58,753	58,753	58,753

[a] The independent variables are the mean wage for each major job category.
*** $p < .01$.

unemployment are also periods with low vacancy rates. The cross-sectional version of that point is that occupations that have high vacancies tend to have low unemployment rates and vice versa. This does not follow directly, however, because the vacancy data are for jobs, whereas the unemployment data are for workers. Still, the education levels relate to the workers who are occupants of the jobs, even in the Conference Board data, so it is reasonable to conclude not only that highly educated workers have low unemployment rates but that they are also found in jobs with high vacancy rates.

HYPOTHESIS 4: The over- and underqualified have a positive observed wage gap, $\hat{W} - W$, where $\hat{W}$ is the wage that the worker would receive were he placed in the job in which he has an absolute advantage.

The PSID is used to test this hypothesis. To do so, it is necessary to define over- and underqualified in the PSID data along with an individual's usual occupation. Consider, for example, an individual who is in one occupation, the usual occupation (say, physician), for most of her life and then switches to a less skilled occupation (say, retailing). This could reflect life-cycle choice, where the highly skilled person decides to take an easier job as she moves gradually into retirement. It could also reflect an involuntary move that results from a primary job loss that forces the worker to accept another job. Either case is consistent with the formal specification in the model where the worker's productivity, $\gamma + \delta A$, is higher in the usual occupation than is productivity, $\alpha + \beta A$, in the unusual one. The prediction is that her predicted wage $\hat{W}$ should be higher in her usual job than is the actual wage W in the unusual one for which she is overqualified.

An opposite example is also possible. Consider a journeyman machinist who has spent almost his entire career in that job. Now suppose that his plant closes and he is forced to find another job. Unable to find another machinist job, he locates a clerical position in a start-up. He is not well suited to that position, but because the start-up can find no one better to fill the job, it hires the former machinist. His productivity as machinist, $\alpha + \beta A$, exceeds his productivity in the clerical position, $\gamma + \delta A$, so he is underqualified for the clerical job. He has an absolute advantage as a machinist but works as a clerk because he can find nothing that suits his skill set and the firm that hires him can find no one better. This is bad luck. The worker is underqualified for the clerical position and should earn less there than he did as a machinist. Once again, the predicted wage, based on his skills and assignment to his appropriate job—in this case, machinist—should exceed what he earns as a clerk in the start-up.

Related to these examples, over- and underqualification and the usual occupation were defined in the following way. For each individual, the modal occupation was determined, defined as the occupation in which the worker spent the most years. A worker was deemed to be in an "unusual" occupation if the occupation held during that year differs from the modal occupation.[28] About one-fifth of the observations fit this definition of being unusual. The worker in an unusual occupation was defined as overqualified if the mean wage of her usual occupation exceeded the mean wage of the unusual occupation. Conversely, the worker in the unusual occupation was defined as underqualified if the mean wage of her unusual occupation exceeded the mean wage of her usual occupation. This resulted in 9.7% of the observations being classified as underqualified and 10.3% being classified as overqualified.

There are two implications for the panel data tests. First, as before, the wage in the usual occupation should exceed that received in the unusual occupation for which the worker is either over- or underqualified: the wage gap, $\hat{W} - W$, should be positive for both subsamples. In figure 4, the worker with ability A' who is assigned to the easy job is overqualified. She would have earned the amount that corresponds to point 1 were she in the appropriate job (the difficult one), but she earns only the amount that corresponds

[28] Observations for which there were two or more modal occupations were dropped.

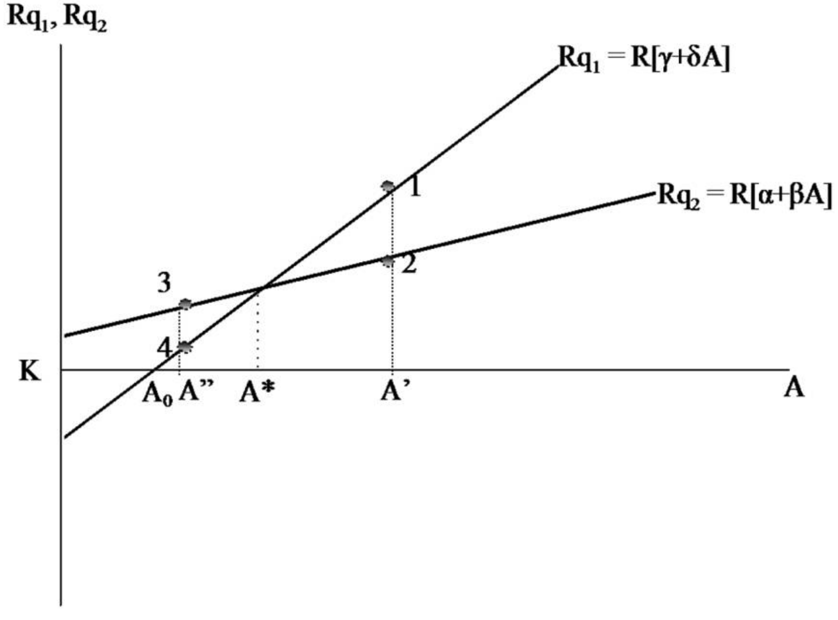

FIG. 4.—Over- and underqualification

to point 2. Second—and important—even though overqualified workers receive less in the job for which they are overqualified than in their usual job, they should still earn more than the typical worker in the easy job. In figure 4, although point 2 lies below point 1, it lies above point 3, which yields the wage of the typical worker in the easy job. Even in the easy job, output increases in ability, so her wage should be higher than the median for that occupation.

Conversely, underqualified workers not only receive less in the job for which they are underqualified than in their usual job but also receive less than the typical worker in the difficult job. In figure 4, the underqualified worker is one who has ability level A'' but works in the difficult job. Instead of receiving the wage that corresponds to point 3, he receives the lower wage that corresponds to point 4. An underqualified worker also receives less than the typical worker in the job for which he is underqualified because output increases in ability and his ability is low for that occupation. Therefore, his wage should be lower than the median for that occupation, where the wage that he receives at point 4 is lower than the wage that the typical worker in the difficult job receives at point 1.

Log wages are used because of skewness in levels. The estimating equation to create a person-specific estimated counterfactual log wage is

$$\log \hat{W}_{it} = b_{0i} + t + e_{it}. \qquad (16)$$

The predicted log wage is initially a function of just time and person fixed-effects, b_{0i}. However, in the empirical specifications tested the model was also reestimated, adding variables relating to age, tenure, and education as well as the individual fixed effects and time dummies of equation (16). The results were not altered, as detailed below, because the additional variables (age, tenure, and education) vary little after controlling for the person fixed effect.

The coefficients in equation (16) are obtained by estimating the regression on the subsample of qualified workers (who are neither over- nor underqualified for their jobs, as defined above). No under- or overqualified job spells influence the results. The predicted log wage for over- and underqualified workers uses these coefficients from equation (16) to construct the counterfactual wages for these workers. The predicted log wage include the coefficients and person fixed effects estimated from the usual occupation regression, providing an estimate of what the worker would receive were he in his normal job.[29]

The point should be reiterated that if the predicted counterfactual wage is not estimated carefully, then the means of the estimated wage gap, $\log \hat{W}_{it} - \log W$, for over- and underqualified workers will not be accurate tests of the implications of the model. For example, were cross-sectional data used, individuals who are deemed to be overqualified for their jobs may be in those jobs because of some unobserved ability component that is not captured by measured variables. A person with a PhD who is working as an administrative assistant may be in that job because she is not qualified to be a professor. Thus, it is important that person fixed effects be included in the regression. Since panel data are needed to estimate equation (16), the smaller data set of the PSID is used instead of the larger data set of the CPS.[30]

The resulting wage gap, $\log \hat{W}_{it} - \log W$, is predicted to be positive for both over- and underqualified workers. For example, a worker who is overqualified has $A > A^*$ in figure 1 but is working in the easy job and earns only $\alpha + \beta A$ instead of the appropriate and higher $\gamma + \delta A$, as predicted by equation (16). Conversely, a worker who is underqualified has $A < A^*$ but

[29] Two regressions were run to predict counterfactual wages in the usual job. The first specification included age, age^2, and tenure as well as person and year fixed effects. The second included only year and person fixed effects. The R^2 was almost identical because once fixed effects are included, only time-varying education and aging contributes to the regression, the latter being captured mostly by year effects.

[30] Another possible concern is that person fixed effects do not account for time-varying perceptions of worker ability due to either learning or workers' skill acquisition. However, the results reported below are similar when restricting the analysis to workers who are over 30 or over 35 years of age, suggesting that any early-career moves between occupations due to learning or skill acquisition are not driving the results.

works in the difficult job, earning only $\gamma + \delta A$ instead of the appropriate and higher $\alpha + \beta A$.

Summarizing, there are four patterns that should appear in the results. First, those who are overqualified for their jobs should earn less in that job than their estimated earnings were they properly placed in the job in which they are more productive. Second, those who are underqualified for their jobs should also earn less in that job than they would be estimated to earn were they properly placed in the job in which they are more productive. Third, although the overqualified earn less than they would in their proper jobs, they should earn more than the typical worker in the job for which they are overqualified because output increases in ability in all jobs. Fourth, not only do the underqualified earn less than they would in their proper jobs but they also should earn less than the typical worker in the job for which they are underqualified because their abilities are lower than those typical for the difficult job.

The results are reported in table 6. In panel A, column 1 reports the results for those job spells that correspond to underqualification, and column 2 reports the results for those that correspond to overqualification.[31] The first row reports that the average log wage gap between the predicted wage and the actual wage is positive, in accordance with the predictions, for both groups. The number reported is the average log wage gap across all person-years that fit the definition of under- or overqualification. Because individuals who are incorrectly assigned have lower productivity than they would have were they in their appropriate jobs, their wages are below that predicted. The average log wage gap is positive and statistically significant for the overqualified but, although positive, is not statistically significant for the underqualified. The row "Fraction above zero" shows that 56% of those in jobs for which they are underqualified receive wages in those high-level jobs that are below that which they are predicted to have earned had they been assigned to the appropriate, albeit lower-skilled, job.

Hypothesis 4 concerns the gap between predicted and actual wages for an individual. An additional implication of this hypothesis concerns the difference in actual log wages and occupation-cell average log wages for the under- and overqualified. The prediction is that actual log wages should be lower than the mean log wage in the occupation for those who are underqualified and should be higher than the mean log wage for the overqualified. Underqualified workers are poorer than the typical worker in that job and, consequently, earn less than the average worker in that job. Overqualified workers are better than the typical worker in the job and, consequently, earn more than the average worker in that job. Panel B of table 6 reports the average

[31] Individuals are dropped from the sample if their wage is an outlier, measured as their wage being greater than 3 standard deviations from the occupational average wage.

Table 6
Panel Study of Income Dynamics (PSID) Average Hourly Log Wage Gap for Workers Who are Over- or Underqualified for Their Job

	Work Spells Corresponding to a Worker Being in a Job for Which He Is Underqualified	Work Spells Corresponding to a Worker Being in a Job for Which He Is Overqualified
	A. Predicted Log Wage − Actual Log Wage, $\hat{W} - W$	
Mean	.008	.078
	(.008)	(.008)
Fraction above zero	56.4	64.5
Number of observations	2,530	2,657
	B. Actual Log Wage − Mean Log Wage in Inappropriate Occupation	
Mean	−.175	.058
	(.010)	(.010)
Number of observations	2,732	2,907

Note.—Data are from the PSID. Sample sizes differ due to missing data for some regressions in panel A. Predicted log wages come from a model with individual fixed effects, tenure, age, and age^2 estimated on 19,397 person-year observations for individuals in their usual occupation. There are 4,385 unique individuals in the model. Standard errors are in parentheses.

difference between the received log wage and the mean log wage of workers who are in the occupation, most of whom are there appropriately. Both predictions are borne out, and the differences are statistically significant.

Finally, the PSID data allow an assessment of which workers are in jobs for which they are over- and underqualified. Figure 5 presents the results of local polynomial regressions that flexibly characterize the probability of over- and underqualification as a function of the log wage in the usual occupation. According to the theory, it is those in the middle of the distribution who are likely to be either over- or underqualified, while over- and underqualification varies over the distribution. The probability of overqualification is increasing with skill up to a point and then declines, which is intuitive. The more able the worker is, the more likely that a random job assignment results in overqualification. Underqualification is declining with the usual log wage beyond some point. This, too, is intuitive. The more able the worker is, the less likely that a random job assignment results in underqualification.

The U shapes also make sense. Job assignment is not random. Extremely high-wage workers are unlikely to be overqualified because they are unlikely to be bumped down to low-skilled jobs. Conversely, those with very low wages are unlikely to be underqualified because they are so low ability that they are rarely assigned to higher-skilled jobs. These inverted U-shaped patterns are exactly what comes out of the theory exposited above. The patterns are neither obvious nor predicted by other models, as far as we are aware.

Hypothesis 5: When the return to skills rises over time, there are increasingly adverse wage consequences of mismatch. This result depends on

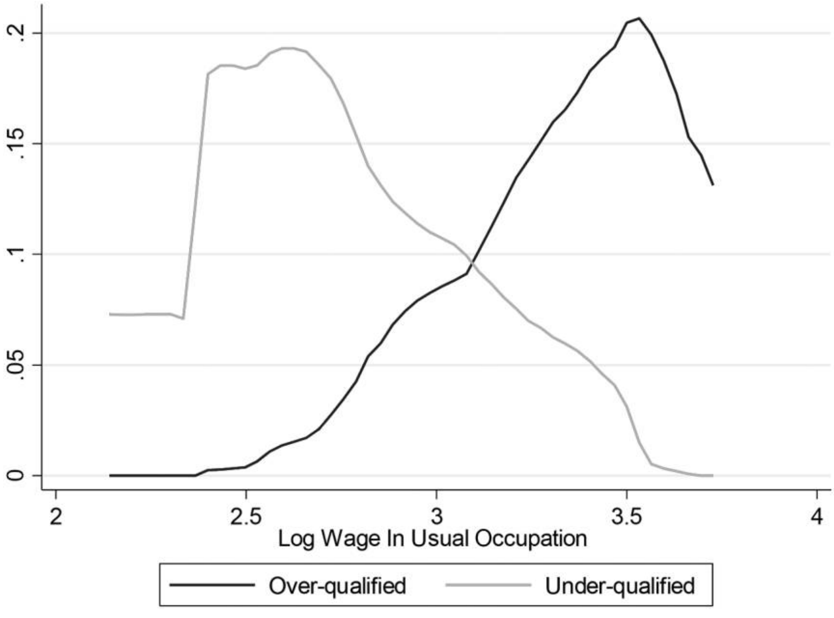

FIG. 5.—Probability of over- and underqualification by log wage in usual occupation. A color version of this figure is available online.

the assumption, made earlier, that the nature of skill-biased technical change steepens the relation between wages and ability more in difficult jobs than in easy jobs.

It is well known that over the past 30 years the return to education has risen. It is natural to expect, as argued earlier, that the return to ability has increased more in high-skilled jobs than in low-skilled ones. This implies that the difference between δ' and δ is greater than the difference between β' and β, as shown in figure 2. As a result, the variance of pay is greater today for the highly able than it was in the past because the wage loss for taking the easy job in the past was $b - a$; today, the wage loss is $d - c$. Thus, as the return to skills has risen, there is a rising variance of pay for the highly able.

Using the CPS data, the equation to test this is a simple regression:

$$\sigma_{it} = b_0 + b_1 \text{Year}_{it} + b_2 \text{OccupationalSkill}_{it} + b_3 \text{OccupationalSkill}_{it} \times \text{Year}_{it} + e_{it}, \tag{17}$$

where the variance of pay is calculated for each occupation i and for each year t, resulting in a data set of 12,733 observations (for 39 years times an average of 326 occupations). The variable OccupationalSkill is the median education for that occupation each year. The first implication is that $b_1 > 0$ because the rising return to skills increases the variance of earnings over

Table 7
Rising Mismatch Over Time as the Variance of Income within Occupations Rises Over Time

	Dependent Variable: Within-Occupation Variance of Income	
	(1)	(2)
Year	770.4***	−273.5
	(73.6)	(200.0)
OccupationalSkill	.86***	−34.5***
	(.071)	(7.54)
OccupationalSkill × Year		.018***
		(.0038)
N	12,733	12,733
R^2	.5209	.5415

NOTE.—The sample size for the Current Population Survey data by occupation is 39 years (1975–2013) times an average of 326 occupational groups per year. The number of occupations varies between years, as in earlier years the occupational categories were broader. For example, there are 293 occupations in 1975, compared with 373 in 1994. The dependent variable is the variance of income within occupational group by year. OccupationalSkill is the average education in that occupation. Regression observations are weighted by the number of observations per occupation. In some years, there is only one observation for some occupations; these observations are dropped. Standard errors are in parentheses.
*** $p < .01$.

time. The second implication is that $b_3 > 0$ because the variance of earnings rises more for the highly able, as suggested in figure 2.

Regression results in table 7 are consistent with both implications. The variance of pay has risen over time, but it has risen most for the highly skilled.

VI. Conclusion

A worker's skills alone do not determine the job into which he or she is hired or, indeed, whether the worker is hired at all. The existence of job slots that firms post means that even qualified workers may not be hired or may not be assigned to the job for which they are best suited because there is a superior applicant for the posted position.

Although this idea is intuitive, it has not been modeled in a way that allows its implications to be explored in individual-based data. The model and analysis presented herein not only provides many specific predictions on what should be observed in hiring and job assignment but also tests and validates those predictions using four different data sets.

First, the probability of being hired depends not only on an applicant's skills but also on the skills of the competition. This is verified using oDesk data.

Second, bumping occurs, which happen when a worker takes a job for which he is not well suited but receives the offer because his skills are superior to those of other applicants except the one who gets the job that he prefers. The model provides clear definitions of "overqualification" and "underqualification" that have specific empirical meaning. Using these definitions, the PSID data provide evidence that over- and underqualification occurs and that the wages that are received in those jobs are consistent with the predictions of the model. Specifically, overqualified workers receive lower pay in the job for which they are overqualified than they would in their appropriate positions. Underqualified workers receive lower pay in the job for which they are underqualified than they would in their appropriate positions, but the results for the underqualified are not statistically significant. Additionally, as predicted, overqualified workers receive more than the average worker in that job, and underqualified workers receive less than the average worker in the job for which they are underqualified.

Third, less able workers are more likely to be unemployed because more able workers are capable of performing a wider variety of jobs. The model provides this as an implication, and, not surprisingly, the implication is found in data provided by the BLS.

Fourth, vacancy rates are higher in jobs that require high levels of skill. The lower-skilled jobs can be filled by almost all workers, but only the smaller group of high-ability workers are more able to perform the high-skilled jobs. The Conference Board data on vacancy rates confirm this prediction.

Fifth, the variance of pay rises over time (1975–2013) for the higher skilled. It is well known that the return to education has risen. The variance of pay rises with education because the highly educated who are lucky enough to obtain the appropriate job enjoy a larger wage premium over the less educated than they did in the past.

One caution: The structure assumed throughout allows luck to play an important role because the workers have only one shot at a job search. To the extent that search is relatively cheap in some markets, the results suggested by the model and some of the findings may be less applicable.

In sum, a model that explicitly incorporates slots and allows for other applicants to compete with an individual for a job is not only intuitive but provides many testable implications that are confirmed empirically.

Appendix

Estimation of Employer Choice Model

The employer choice model accounts for wage bid endogeneity by using an exchange rate instrument. With this instrument, the control function approach of Petrin and Train (2010) is used to estimate the parameters governing the probability that an individual worker is hired. The idea is to control for the endogenous portion of wage bids directly in the choice model. The endogenous portion of the wage bid is that part of wages that is correlated with unobserved productivity of the worker. By construction, the endogenous portion is orthogonal to the fitted values of the first-stage regression of bids on the instruments. The first-stage regression is

$$\log(w_{jt}) = a_1 + Z_{jt}\gamma_1 + X_{jt}\gamma_2 + u_{jt},$$

where Z_{jt} is the log of the dollar to local currency exchange rate in month t for worker j after netting out currency time trends.[32] The first-stage results are shown in table A1.

Let $\mathrm{CF}_{jt} = \hat{u}_{jt}$ denote the fitted residuals from this first-stage regression for worker j on job opening i during month t. Petrin and Train's (2010) approach includes CF_{jt} in the conditional logit model that maps applicant characteristics to hiring probabilities. The control function captures that part of wage bids that is orthogonal to the "good" variation in wage bids caused by the instruments. This orthogonal component includes unobserved applicant quality.

Separating these sources of variation in bids allows for consistent estimates of α, the parameter that governs price-related substitution patterns. By putting in the residuals, the model allows different loading on the control function and on the log wage. The resulting choice probabilities used to form the likelihood are

$$p_{j|i} = \frac{\exp\left(X_{jt}\beta - \alpha \log\left(w_{ij}\right) + \mathrm{CF}_{jt}\psi\right)}{1 + \sum_{J}^{J_i} \exp\left(X_{jt}\beta - \alpha \log\left(w_{ij}\right) + \mathrm{CF}_{jt}\psi\right)},$$

and the log likelihood to be maximized is $\sum_i \sum_j d_{ij} \log\left(p_{j|i}\right)$, where d_{ij} is an indicator for the option chosen. The sum over j in the log likelihood includes the no-hire (outside) option.

[32] One may worry about worker sorting on the instrument itself through the participation margin of applying for jobs. See Stanton and Thomas (2017) for more detail. For consistency with the specification in Stanton and Thomas (2017), we also include a "tightness" instrument that is the log of the weekly applicant to vacancy arrival rate in a job category; this instrument is better suited to specifications where job categories are pooled.

Table A1
First-Stage Regression of Log Hourly Bids on Exchange Rate Instruments

	Sample: Log Hourly Wage Bid	
Dependent Variable	Administrative Support (1)	Web Programming (2)
Log Dollar to Local Exchange Rate		
(monthly, net of currency time trends)	.208***	.0571***
	(.0335)	(.0104)
Log Wage on Last Hire	.302***	.800***
	(.00153)	(.00754)
No Prior Observed Wage	−.396***	.499***
	(.00389)	(.00717)
Inexperienced Worker	.230***	.0349*
	(.0398)	(.0197)
Good English Skills	−.142***	.0379***
	(.00411)	(.00260)
BA or Higher Degree	−.00140	.00740***
	(.00217)	(.00118)
Prior Experience and No Feedback	−.158***	−.0418**
	(.0397)	(.0195)
Feedback	−.166***	−.0824***
	(.0421)	(.0201)
Feedback^2	.0383***	.0240***
	(.0135)	(.00638)
Feedback^3	−.00174	−.00136**
	(.00135)	(.000629)
Number of observations	365,600	455,123
R^2	.317	.382
F-statistic on excluded instruments	76.37	44.74

NOTE.—The sample restricts the data in table 2 to include only job openings posted by employers that have hired two or more previous workers. Robust standard errors are in parentheses. Models include fixed effects for eight country groups. The last country group includes many countries with very small application shares. All models also contain a calendar time trend, a trend interacted with the country groups, an indicator for an employer-initiated application, an indicator for agency affiliation and its interaction with worker experience, and a piecewise-linear spline with four knots for the application number. The measure for the log dollar to local currency exchange rate is net of currency time trends. An additional excluded instrument captures the log arrival rate of applicants to jobs in the job category over that week.

* $p < .10$.
** $p < .05$.
*** $p < .01$.

Table A2
Conditional Logit Parameter Estimates

	Sample			
	Administrative Support (1)	Administrative Support, with Control Function (2)	Web Programming (3)	Web Programming, with Control Function (4)
Log Wage on Last Hire	−.232	2.893	−.217	6.977
	(.0287)	(1.257)	(.0405)	(2.328)
No Observed Last Wage	−.348	−4.446	−3.810	.599
	(.0963)	(1.678)	(.356)	(1.461)
Inexperienced Worker	−1.162	2.057	3.415	3.703
	(1.013)	(1.648)	(.776)	(.872)
Good English Skills	.172	−1.052	.154	.497
	(.114)	(.625)	(.0680)	(.147)
BA or Higher Degree	.0832	.0179	.0261	.0952
	(.0467)	(.068)	(.0325)	(.04)
Prior Experience and No Feedback	−.938	−1.686	.405	−.0303
	(1.013)	(1.465)	(.693)	(.642)
Feedback	−1.284	−2.013	−.117	−.914
	(1.079)	(1.491)	(.684)	(.693)
Feedback^2	.361	.421	.0355	.266
	(.344)	(.443)	(.210)	(.213)
Feedback^3	−.0250	−.00900	.00191	−.0115
	(.0338)	(.041)	(.0203)	(.02)
Log Hourly Wage Bid		−10.24		−8.978
		(4.198)		(2.877)
Control Function		9.546		8.268
		(4.201)		(2.884)
Number of job openings	5,727	5,727	15,741	15,741
Pseudo R^2	.5105	.5174	.5465	.5499

NOTE.—The sample is described in the note for the first-stage regression. Estimates come from a conditional logit model that includes an option not to hire an applicant. When the control function is included, standard errors (in parentheses) come from using block-bootstrap replications of the entire procedure. Other controls are described in the table note for the first-stage regression.

References

Abowd, John M., John Haltiwanger, Julia Lane, Kevin L. McKinsey, and Kristin Sandusky. 2007. Technology and the demand for skill: An analysis of within and between firm differences. NBER Working Paper no. 13043, National Bureau of Economic Research, Cambridge, MA.

Agrawal, Ajay, John Horton, Nicola Lacetera, and Elizabeth Lyons. 2013. Digitization and the contract labor market: A research agenda. NBER Working Paper no. 19525, National Bureau of Economic Research, Cambridge, MA.

Agrawal, Ajay, Nicola Lacetera, and Elizabeth Lyons. 2012. How do online platforms flatten markets for contract labor? Working paper, University of Toronto.

Altonji, Joseph G., and Charles R. Pierret. 2001. Employer learning and statistical determination. *Quarterly Journal of Economics* 116, no. 1:313–50.

Andrews, M. J., S. Bradley, D. Stott, and R. Upward. 2008. Successful employer search? An empirical analysis of vacancy duration using micro data. *Economica* 75, no. 299:455–80.

Autor, D. H. 2001. Wiring the labor market. *Journal of Economic Perspectives* 15, no. 1:25–40.

Autor, D. H., L. F. Katz, and A. B. Krueger. 1998. Computing inequality: Have computers changed the labor market? *Quarterly Journal of Economics* 113, no. 4:1169–213.

Autor, D. H., Frank Levy, and Richard J. Murnane. 2003. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics* 118, no. 4:1279–334.

Becker, Gary S. 1993. *A treatise on the family*. Cambridge, MA: Harvard University Press.

Brenčič, V. 2010. Do employers respond to the costs of continued search? *Oxford Bulletin of Economics and Statistics* 72, no. 2:221–45.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics* 117, no. 1:339–76.

Burdett, Kenneth, and Melvyn G. Coles. 1999. Long-term partnership formation: Marriage and employment. *Economic Journal* 109 (June): F307–F334.

Burks, Stephen V., Bo Cowgill, Mitchell Hoffman, and Michael Housman. 2015. The value of hiring through employee referrals. *Quarterly Journal of Economics* 130, no. 2:805–39.

DeVaro, Jed. 2006. Internal promotion competitions in firms. *Rand Journal of Economics* 37, no. 3:521–42.

Diamond, Peter. 1982. Wage determination and efficiency in search equilibrium. *Review of Economic Studies* 49, no. 2:217–27.

Gale, D., and L. S. Shapley. 1962. College admissions and the stability of marriage. *American Mathematical Monthly* 69, no. 1:9–15.

Gee, Laura. 2015. The more you know: Information effects on job application rates by gender in a large field experiment. Working paper, Tufts University.

Ghani, Ejaz, William Kerr, and Christopher Stanton. 2014. Diasporas and outsourcing: Evidence from oDesk and India. *Management Science* 60, no. 7:1677–97.

Hoffman, Mitchell, Lisa Kahn, and Danielle Li. 2015. Discretion in hiring. NBER Working Paper no. 21709, National Bureau of Economic Research, Cambridge, MA.

Horton, John. 2010. Online labor markets. In *Workshop on Internet and Network Economics*, 515–22.

———. 2017. The effects of algorithmic recommendations: Evidence from a field experiment. *Journal of Labor Economics* 35, no. 2:345–85.

Horton, John, William Kerr, and Christopher Stanton. 2017. Digital labor markets and global talent flows. NBER Working Paper no. 23398, National Bureau of Economic Research, Cambridge, MA.

Kahn, Lisa. 2010. The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics* 17, no. 2:303–16.

Katz, Lawrence F., and Kevin M. Murphy. 1992. Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107, no. 1:35–78.

Kudlyak, Marianna, Damba Lkhagvasuren, and Roman Sysuyev. 2012. Systematic job search: New evidence from individual job application data. FRB Richmond Working Paper no. 12-03R, Federal Reserve Bank of Richmond, Richmond, VA.

Kuhn, Peter, and Hani Mansour. 2014. Is Internet job search still ineffective? *Economic Journal* 124, no. 581:1213–33.

Kuhn, Peter, and Kailing Shen. 2013a. Do Chinese employers avoid hiring overqualified workers? Evidence from an Internet job board. *Research in Labor Economics* 37:1–30.

———. 2013b. Gender discrimination in job ads: Evidence from China. *Quarterly Journal of Economics* 128, no. 1:287–336.

Kuhn, Peter, and Mikal Skuterud. 2004. Internet job search and unemployment durations. *American Economic Review* 94, no. 1:218–32.

Lazear, Edward P. 1986. Retail pricing and clearance sales. *American Economic Review* 76, no. 6:14–32.

Lazear, Edward P., and Sherwin Rosen. 1981. Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89, no. 5:841–64.

Lazear, Edward, and James Spletzer. 2012. The US labor market: status quo or a new normal? *Annual proceedings of the Jackson Hole Federal Reserve Conference*.

Lucas, Robert E., Jr. 1978. On the size distribution of business firms. *Bell Journal of Economics* 9, no. 2:508–23.

Marinescu, Ioana, and Ronald Wolthoff. 2015. Opening the black box of the matching function: The power of words. NBER Working Paper no. 22508, National Bureau of Economic Research, Cambridge, MA.

Modestino, Alicia, Daniel Shoag, and Joshua Balance. 2015. Upskilling: Do employers demand greater skills when workers are plentiful? Unpublished manuscript, Northeastern University.

Nakamura, Alice, Emi Nakamura, Kathryn Shaw, Richard B. Freeman, and Amanda Pyman. 2009. Jobs online. In *Studies of labor market intermediation*, ed. David H. Autor, 27–65. Chicago: University of Chicago Press.

Oyer, Paul. 2008. The making of an investment banker: Stock market shocks, career choice, and lifetime income. *Journal of Finance* 63, no. 6: 2601–28.

Pager, Devah, Bart Bokowski, and Bruce Western. 2009. Discrimination in a low-wage labor market: A field experiment. *American Sociological Review* 74, no. 5:777–99.

Pallais, Amanda. 2014. Inefficient hiring in entry-level labor markets. *American Economic Review* 104, no. 11:3565–99.

Petrin, Amil, and Kenneth Train. 2010. A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research* 47, no. 1:3–13.

Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. Search-theoretic models of the labor market: A survey. *Journal of Economic Literature* 43:959–88.

Rosen, Sherwin. 1982. Authority, control, and the distribution of earnings. *Bell Journal of Economics* 13, no. 2:311–23.

Shimer, Robert. 2005. The cyclical behavior of unemployment and vacancies. *American Economic Review* 95, no. 1:25–49.

Shimer, Robert, and Lones Smith. 2000. Assortative matching and search. *Econometrica* 68, no. 2:343–69.

Spence, Michael. 1973. Job market signaling. *Quarterly Journal of Economics* 87, no. 3:355–74.

Stanton, Christopher, and Catherine Thomas. 2016. Landing the first job: The value of intermediaries in online hiring. *Review of Economic Studies* 83, no. 2:810–54.

———. 2017. Information frictions and observable experience. Working paper, Harvard Business School.

van Ours, J., and G. Ridder. 1991. Cyclical variation in vacancy durations and vacancy flows. *European Economic Review* 35, no. 5:1143–55.

———. 1992. Vacancies and the recruitment of new employees. *Journal of Labor Economics* 10, no. 2:138–55.

Waldman, Michael. 2003. Ex ante versus ex post optimal promotion rules: The case of internal promotion. *Economic Inquiry* 41, no. 1:27–41.

Zhu, Feng, Rory McDonald, Marco Iansiti, and Aaron Smith. 2015. Upwork: Reimaging the future of work. Harvard Business School case no. 616027.