Who Gets Hired? The Importance of Finding an Open Slot

Edward P. Lazear*

Kathryn L. Shaw**

and

Christopher T. Stanton***

October 30, 2015

*Graduate School of Business and Hoover Institution, Stanford University
** Graduate School of Business, Stanford University
***Harvard Business School, Harvard University

Abstract

Despite seeming to be an important requirement for hiring, the concept of a slot is absent from virtually all of economics. Closest is the macroeconomic studies of vacancies and search, but the implications of slot-based hiring for individual worker outcomes has not been analyzed. A model of hiring into slots is presented. Job assignment is based on comparative advantage. Crucially, and consistent with reality, being hired and assigned to a job depends not only on one's own skill, but on the skill of other applicants. The model has many implications the most important of which are: First, bumping occurs, when one applicant is bumped from a job into a lower paying job or unemployment by another applicant who is more skilled. Second, less able workers are more likely to be unemployed because high ability workers are more flexible in what they can do. Third, vacancies are higher for difficult jobs because easy jobs never go unfilled. Fourth, some workers are over-qualified for their jobs whereas others are underqualified. The mis-assigned workers earn less than they would had they found an open slot in a job that more appropriately matches their skills. Despite that, overqualified workers earn more than the typical worker in that job. These implications are borne out using four different data sets that match the data requirements for of these point and others implied by the model.

Two workers, seemingly identical in qualifications, apply for a job. There is one opening so the employer must choose only one to hire. One worker is employed, the other continues looking. Luck has played a role in determining outcomes, presumably good luck for the one hired and bad for the one turned down.

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 in a poor labor market 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. 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 worker 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. A worker has no control over what others do, but the outcome of a job search process likely depends crucially on the applicants with whom the worker competes.

Whether hiring luck is important in affecting lifetime worker wealth depends on the cost of mobility and on the thickness of markets. 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, hiring luck may have long term effects. Cohort effects on wealth that have be important at least suggest that hiring luck may be of some consequence. At a minimum, it is important to understand the way in which the existence of slots and competition for jobs affect outcomes.

In what follows, a model of slots and within-firm job assignment is presented that yields specific, testable implications. Some help reconcile puzzles that exist in empirical findings, but 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. Less able workers are more likely to be unemployed than the more able. 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 and the low ability workers are likely to be more similar to one another than are the high ability ones. This puzzle is explained by the fact that high ability workers are more flexible and can do a larger number of jobs.

2. Analogously, vacancy rates are highest in the high-paying, high-quality jobs. Most workers can fill the lower quality positions, but higher quality positions require higher ability applicants who may not be present. This implication, coupled with the first, provides cross-sectional Beveridge-curve implications. 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.

3. Unemployment, at least of the frictional variety, is a consequence of bad luck. Because firms are slot constrained, a worker remains unemployed when other applicants for a job have superior qualifications. But for those others, the worker would have been employed.

4. It is common to speak about a person being over-qualified for a position, but what does that mean formally? A clear definition of over- and under-qualification is provided and that definition yields implications about observed wages for workers who find themselves over- or under-qualified 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.

5. "Bumping" creates over-qualification. Workers who are better suited to high level jobs are bumped into lower level ones by even better qualified workers. Conversely, under-qualification results when an unfilled high-level job is filled by a worker who must take that job because there is no higher quality worker available to do it.

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 is the notion of "slots" that is central and absent from most prior standard analyses.[1] 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.[2]

The use of slots in the model is an innovation that 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 worker. But the level of abstraction in search

[1]The earliest related work is Diamond (1982), which 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.

[2]In the extreme, it is described by a perfect-complements technology, where $Q=\min[x, \theta y]$, where x and y are two factors of production and θ is a technology parameter.

theory is 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.

Supply-side unemployment, where the fall in demand during a business cycle induces people to stay home because their reservation values exceed their productivity, is 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. This cannot be explained in standard technologies with smooth production. 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 will not be 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 take the one offered because the better job is already filled.

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. Customers may prefer that the same software outsourcer provide both programming and data entry services so that if there is a problem, one outsourcer cannot blame the other for the difficulties encountered. One possible structure would allow a firm to have n slots, with output equal to

(1) $Q = Q(q_1, q_2, \ldots, q_n)$.

with $Q_{ii}<0$, and $Q_{ij} > 0 \; \forall \; i, j$.

To make things simple and specific to the data that is used later, let n=2 and imagine that the two slots in the firm correspond to data entry and programming (or low-level and high-level

programming). Output is positive even if only one of the two tasks is done, but is better to have both done in the same firm. Additionally, suppose that

(2) a. $q_1 = (\gamma + \delta A_1)$
b. $q_2 = (\alpha + \beta A_2)$

where A_i is the ability of the worker who occupies slot i. 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 zero.

An example would be to allow output in (1) to be given by

$$Q = q_1 + q_2 + q_1 q_2$$

where q_i is given by (2), but this form is not required for any of the results to hold.

Think of job 1 as the difficult job and job 2 as the easy job. In the current example, job 1, the difficult job, is programming and job 2, the easy job, is data entry. 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 assumed to be 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 Rq_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 (2a),

$A_0 = (K/R - \gamma) / \delta$.

Further, define A* as the ability such that the worker is equally productive in both jobs (as shown in figure 1), so that

$\gamma + \delta A^* = \alpha + \beta A^*$

Thus, $A^* = (\alpha-\gamma) / (\delta-\beta)$.

Then,

(3) $R(\gamma + \delta A) > K$ for $A > A_0$

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

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

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

Finally, to complete the model, assume a single period and that costly search takes the form of allowing each firm that receives any applicants to interview exactly two workers from the pool of potential employees. Initially, assume that if there are M firms, then let there be 2M

workers so that there are just enough slots in the economy to employ the entire workforce. Later, the number of slots will be shown to be derivable from a structure that yields the number of slots as part of competitive equilibrium.[3]

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, once pairing occurs, bargaining opportunities exist. Consistent with the notion that slots are important, the worker is paid for what he does on the job to which he is assigned, not for his ability. Therefore, define the wage that goes to workers in job i at firm m as $\lambda\, q_{im}$ with $0 \leq \lambda \leq 1$ so

(5) a. $w_1(A) = [\, \lambda_1 (\gamma + \delta A) \,]\, R$
b. $w_2(A) = [\, \lambda_2 (\alpha + \beta A)]\, R$

The determination of λ can be thought of as the outcome of some (any) bargaining game between the worker and firm. It is can be taken as exogenous, but below, λ is derived as the outcome of a competitive equilibrium. Because workers are the scarce factor and firms are assumed to enter freely, the equilibrium value of λ is 1.

Implications

Full employment may result under the (rare) perfect circumstance that firm draws have a particular realization. Label A_{1m} the ability of the higher ability worker of the two workers that firm *m* encounters and label A_{2m} the ability of the lower ability worker of the two workers that firm *m* encounters.

Lemma 1: If any worker is assigned to the difficult job, it will always be the highest ability worker, that is, the worker with ability A_{1m}.

Proof:

The better worker should be assigned to the difficult job at firm *m* and the poorer worker to the easy job at firm m if

$\gamma + \delta A_{1m} + \alpha + \beta A_{2m} > \gamma + \delta A_{2m} + \alpha + \beta A_{1m}$

[3]This search structure is somewhat contrived. A more general structure would be to allow there to be 2M workers and M* firms, where each worker is allowed to search at one and only one firm. Then the number of applicants that a firm receives would follow a binomial distribution, where k is the number of applicants, 2M is the number of trials, and 1/M* is the probability that a worker applies to any one specific firm. The probability of receiving k applicants is then

$$\frac{n!}{k!(n-k)!}\left(\frac{1}{M^*}\right)^k\left(1-\frac{1}{M^*}\right)^{n-k}.$$

The more general structure adds complexity without much insight because instead of analyzing one case, where there are two applicants, it is necessary to analyze n cases where then number of applicants varies between 0 and 2M.

Additionally, the number of slots that comes out of the rigid structure, where each firm gets two applicants, is not generally in the competitive equilibrium equal to one-half the number of workers who search for jobs.

or if

$(\delta - \beta)(A_{1m} - A_{2m}) > 0$

which must hold because $\delta > \beta$ and $A_{1m} > A_{2m}$.

Lemma 1 merely says that workers should be assigned according to comparative advantage. The worker of ability A_1 is better in every job, but has a comparative advantage in the difficult job.

Another preliminary result is useful.

Lemma 2: Both slots are filled if and only if

$A_1 > \gamma/(\beta-\delta) + [\beta/(\beta-\delta)]A_2$

or alternatively, if and only if

$A_2 > [\,(\beta-\delta)A_1 - \gamma\,]/\beta$

Proof: The choice is between hiring two workers or hiring only the best worker and assigning him 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 (4). Thus, two slots are filled if and only if

$\alpha + \beta A_1 < \gamma + \delta A_1 + \alpha + \beta A_2$

or iff

$A_1 < \gamma/(\beta-\delta) + [\beta/(\beta-\delta)]A_2$. |||

It is now possible to state conditions for full employment in the economy. A number of intermediate steps are useful.

Proposition 1 A firm fills both of its slots if $A_{1m} > A^*$.

Proof:

First note that because both profits and wages are non-decreasing in output, the output-maximizing assignment also maximizes both wages and profits.

Lemma 1 states that if any worker is assigned to the difficult job, it is the highest ability worker defined as $A_{1m.}$ Since $A_{1m.} > A_0$, output of A_{1m} in the difficult job exceeds the reservation value by (3). Furthermore, since $\gamma+\delta A > \alpha + \beta A$ for $A>A^*$, the high ability worker is assigned to the difficult job. Therefore A_{1m} is employed and assigned to the difficult job.

Then the only issue then is whether A_{2m} should be assigned to the easy job or not hired at all. But (4) guarantees that output of every worker in the easy job is greater than the reservation value so it is profit increasing to employ A_{2m} in the easy job. |||

Proposition 2 Full employment occurs if $A_{1m} > A^* \;\forall$ m.

Proof: By Prop. 1, since $A_{1m} > A^*$ at all M firms, every firm employs every worker that it interviews, which guarantees that all 2M workers are employed. |||

It is also possible to state the converse of Prop. 2, under which unemployment results.

Proposition 3 Unemployment occurs if and only if for at least one firm, $A_2 < [\ (\beta-\delta)\ A_1 - \gamma\ ] / \beta$

Proof:

This 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. |||

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 lowest ability of the two is bumped down to the easy job, even though he is inherently more productive in the difficult job than in the easy job. The worker must do the easy job, not because he is low ability, but because the difficult job is best assigned to the even-higher ability worker. The worker who is bumped earns less than he would have had he been able to secure a difficult job. Similarly, when two low ability workers, defined as having $A<A_0$, show up at the firm, then the lowest ability worker is bumped out of a job altogether and ends up being unemployed. Second, low ability workers are more likely to be unemployed than high ability ones. Unemployment requires that the best of the firm's applicants has ability lower than A^*. Bad luck for low ability workers takes the form of applying to a firm where the other applicant is of higher ability, but not sufficiently high (as specified in lemma 2) that the firm wants to employ both. Because neither applicant can do the difficult job, only one is hired and the other is unemployed. For very low ability workers, good luck means either that the other applicant is of even lower ability or that the other applicant has ability sufficiently high to induce the firm to employ both workers, meaning that the high ability worker has ability greater than $\gamma/(\beta-\delta)+[\beta/(\beta-\delta)]A_2$.

The second implication of this model, 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 for low priced development houses.[4] 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 ability to work in both jobs, as opposed to only one, makes them more employable. 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 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.

The unemployment implications go a step further. In this stylized model, high ability workers, i.e., those with $A>A_0$, are never unemployed. The wages they receive and the job to which they are assigned depends on the ability of the other worker who is employed at the firm, but it always pays to employ workers with $A>A_0$, irrespective of the other worker's ability. The worker with the highest ability is assigned to the difficult job, but the fact that A exceeds A_0

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

means that worker can perform the difficult or easy job, producing positive output in whichever job he is assigned.[5]

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 describes by this model is of the "frictional" variety, which can result even in very tight labor markets.

Four, the jobs dominated by low ability workers have the lowest vacancy rates. There are never unfilled easy jobs; only the 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).[6] Using education as an observable measure of ability, the empirical implication in Hypothesis 2 below is that vacancy rates rise with education.

Fifth, workers may be over-qualified or under-qualified 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 in the firm, in which case he will be forced to do the easy job. He is "over-qualified," 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 whose 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 over-qualified job, the worker will be underpaid relative to his expected pay on the higher skilled job. To see this, use the wage model of equation (5), but simplify to compare wages across ability levels by setting R=1 and λ=1. When $A > A^*$ and the worker is in the over-qualified 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 exactly 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 reside in job 2.

Although the over-qualified 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

[5]The more general structure, described in an earlier footnote, would permit some unemployment among high ability workers. This would occur if more than two high-ability workers apply to a firm that has only two slots. But the likelihood of being unemployed declines in ability because hiring the highest ability worker is always a profit-improving strategy.

[6]There are chronic "shortages" of workers to fill professional jobs.

those workers and because productivity increases in ability in both jobs. This is a testable implication once an empirical definition of over- and under-qualified is established.

Consider next workers who are under-qualified for jobs, but 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 under-qualified 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 is lower than were he assigned to the easy job. In this case, being the higher ability worker is bad luck.

The wage gap for under-qualified workers who are in the difficult job but have ability A<A* remains W^ -W, but now is given by $(\alpha + \beta A) - (\gamma+\delta A)$. The appropriate job for under-qualified workers is the easy job, whereas the actual job is the difficult job. As before, 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.
For both over-qualified and under-qualified workers, the wage gap is predicted to be positive, which is stated as Hypothesis 3 below.

Unlike the over-qualified worker, the under-qualified 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 should 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 an absolute (not just comparative) advantage. The assignment to jobs, existence of unemployment, wages and profits all depend on the distribution of talents in the population, the number of slots of each type and on luck that takes the form of the distribution of applicants across firms.

Six, 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 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 (a sufficient condition being that all have ability less than A_0) Furthermore, workers with $A<A_0$, are perfectly able to fill the data entry job, but some cannot find jobs because there are too many applicants of low ability at the firm to which they apply. Even if the distribution of talent in the economy could, given proper matching luck, result in ample supply of qualified applicants and full employment, a random selection of applicants generally results in some mismatch.

Seventh, over time, the variance of income rises with education. One well known fact is that there has been a rising return to education over time {add references}. This has implications for the skill gradients 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. Output q_1 in the difficult job is

today higher at higher ability levels that it was in the past – δ has shifted up to δ'. Output q_2 in the easy job has also shifted up as a function of ability, with the slope of the line rising from β to β'. By assumption, the rising return to skills over time has resulted in a greater increase in δ than in β. The assumption is consistent with the notion that technological progress is more complementary with skill level (skill biased) in difficult jobs than it is in easy jobs. The idea is that technology has increased the difference between the output of the more able farmer and the less able one, but it has increased the difference between the output of the more able 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 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 appropriate qualified and the one for which he is over-qualified 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 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.

Competitive Equilibrium and Endogenous Slots

The number of firms (M) and therefore slots (2M) was simply assumed to be given in the prior discussion. In this section, that assumption is altered allowing the number of firms to be endogenous and not necessarily equal to one-half the number of workers. It is shown that the allowing the number of firms to vary in this fashion produces a zero-profit competitive equilibrium.

Given the technology, adjustments in product markets come about through variations in the number of firms, rather than in output per firm. Since firms cannot adjust on the intensive margin by hiring more labor, the marginal cost of output is the cost per unit that results from adding another firm, i.e., the average cost. Denote the joint density of a particular firm's applicants by $f(A_1, A_2)$. Then, the probability p that the firm fills both slots is, by Lemma 2, the probability that $A_1 > A_2$ and that $A_2 > [\,(\beta-\delta)\,A_1 - \gamma\,]\,/\,\beta$ plus the probability that $A_2 > A_1$ and that $A_1 > [\,(\beta-\delta)\,A_2 - \gamma\,]\,/\,\beta$. Define $H(A_1, A_2)$ as an indicator that is 1 if $A_1 \geq A_2$ and 0 otherwise.

Then

$$p = \int_0^{\infty} \int_{\frac{(\beta-\delta)A_1-\gamma}{\beta}}^{\infty} H(A_1, A_2)\, f(A_1, A_2)\, dA_2\, dA_1 + \int_0^{\infty} \int_{\frac{(\beta-\delta)A_2-\gamma}{\beta}}^{\infty} [1 - H(A_1, A_2)]\, f(A_1, A_2)\, dA_1\, dA_2 \tag{6}$$

Additionally, expected output at the firm is then

(7)

$$E(Q) = \int_0^{\infty} \int_{\frac{(\beta-\delta)A_1-\gamma}{\beta}}^{\infty} H(A_1, A_2) Q(A_1, A_2) f(A_1, A_2) dA_2 dA_1 +$$
$$\int_0^{\infty} \int_{\frac{(\beta-\delta)A_2-\gamma}{\beta}}^{\infty} [1 - H(A_1, A_2)] Q(A_2, A_1) f(A_1, A_2) dA_1 A_2 +$$
$$\int_0^{\infty} \int_0^{\frac{(\beta-\delta)A_1-\gamma}{\beta}} H(A_1, A_2) Q(0, A_1) f(A_1, A_2) dA_2 dA_1 +$$
$$\int_0^{\infty} \int_0^{\frac{(\beta-\delta)A_2-\gamma}{\beta}} [1 - H(A_1, A_2)] Q(0, A_2) f(A_1, A_2) dA_1 A_2$$

Then the expected average cost per unit of output is then

$$AC = \int_0^{\infty} \int_{\frac{(\beta-\delta)A_1-\gamma}{\beta}}^{\infty} H(A_1, A_2) \frac{w_1(A_1) + w_2(A_2)}{Q(A_1, A_2)} f(A_1, A_2) dA_2 dA_1 +$$
$$\int_0^{\infty} \int_{\frac{(\beta-\delta)A_2-\gamma}{\beta}}^{\infty} [1 - H(A_1, A_2)] \frac{w_1(A_2) + w_2(A_1)}{Q(A_2, A_1)} f(A_1, A_2) dA_1 dA_2 +$$
(8)
$$\int_0^{\infty} \int_0^{\frac{(\beta-\delta)A_1-\gamma}{\beta}} H(A_1, A_2) \frac{w_2(A_1)}{Q(0, A_1)} f(A_1, A_2) dA_2 dA_1 +$$
$$\int_0^{\infty} \int_0^{\frac{(\beta-\delta)A_2-\gamma}{\beta}} [1 - H(A_1, A_2)] \frac{w_2(A_2)}{Q(0, A_2)} f(A_1, A_2) dA_1 dA_2$$

Because there is free entry, the profits must equal zero, so

(9) E(R)=AC.

Note that R is a random variable because output is stochastic because it depends on the actual matching that occurs in the economy.

The L-shaped (ex ante) marginal cost and the fact that all 2M laborers enter the labor force, the number of firms in the economy is determined by

(10) $M\,E(Q) = D(E(R))$

where D(R) is the market demand curve for output. The actual price, R, is determined by

(11) $\sum_m Q_m = D(R)$,

where Q_m is defined as the actual ex post output of firm m. Note that M is determined ex ante, but both the actual Q shown in (13) and R are determined ex post. The supply of output is perfectly inelastic in the relevant range, but that is only for convenience.[7]

With free entry of firms, labor is the scarce factor so in competitive equilibrium, λ equals 1. All rents go to the workers because if there were ex ante rents left to firms, the entry condition in (10) would induce more firms to enter and those firms would demand additional labor until the wages were bid up to exhaust all profit.

The system of five equations, eqq. (7) - (11), uniquely determine the five unknowns, E(Q), AC, E(R), actual R, and M.[8]

Related Literature

Research on hiring has mushroomed due to the availability of new data, but little has been done that resembles the approach taken here. In the past, those studying hiring needed to work with firms to obtain hiring data. But even that would not guarantee success, because firms kept data on who they hired, not on who applied.[9] The advent of online job boards and online contracting firms has changed this. Much more is known about the types of workers firms seek and who is hired.[10]

There is a very large literature on labor demand that could be broadly thought of as papers on hiring (but not on the hiring process) and thus on the skill demand that is the focus of this paper.[11] 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

[7] An upward-sloping curve, either resulting from differentiated labor (in terms of alternative value of time) or from different management skill would not eliminate competitive equilibrium, nor would it prevent the determination of all prices and wages.

[8] The search technology, however, remains exogenous, by assuming, even in this competitive equilibrium, that search effort adjusts such that each firm sees two and exactly two workers.

[9] Burks, Cowgill, Hoffman, and Housman (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.

[10] Agarwal, Horton, Lacetera, and Lyons (2013), Agrawal, Lacetera and Lyons (2012), Autor (2001), Brencic, (2009), Gee (2015), Ghani, Kerr, and Stanton (2014), Kuhn and Mansour (2004), Kuhn and Shen, (2013a, b), Kuhn, and Skuterud (2004), Marinescu and Wolthoff (2015), Stanton and Thomas (2015a,b), 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)).

[11] Education 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.

rising returns to education and greater demand for workers who do non-routine cognitive tasks. [12] 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 do a range of jobs that require cognitive skills.

Some results from search theory indirectly support the bumping implication. Kudlyak, Lkhagvasuren, and Sysuyev (2012) study a job applications website and show that education initially predicts job applications, but as search on the website continues, applicants seek jobs at all education levels. Thus, well-educated workers begin to bump the less-educated workers. This is in keeping with the search result that wage demands decline with the duration of unemployment.[13]

Other related literature involves vacancy dynamics. Andrews, Bradley, Stott, and Upward (2008) show that vacancies for nonmanual work are less likely to be filled in U.K. data. van Ours and Ridder (1991)[14] examine data from the Dutch Bureau of Statistics. They find that jobs that require more education fill more slowly and that vacancy flow is more sensitive to the business cycle for low-education openings. van Ours and Ridder (1992) also find that higher education requirements are associated with longer vacancy durations.

Data Sets

The predictions detailed below are tested using data from four sources. The Conference Board provides data on vacancies, 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 with hiring and job applicants to assess how the pool of applicants and influences who is hired.

The Conference Board does a monthly survey of online job postings. The Conference Board Help Wanted OnLine (HWOL) is 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" that are ads posted for the first time in the previous month. The second is "total ads" that are new ads plus ads reposted from the previous month. The data is available by occupation by year, from 2006-2014, where yearly data is the average of the monthly data. There are 846 observations for 9 years times 94 occupations. The goal in using these Conference Board data is to create a variable that is similar to the vacancy rate, but is measured by occupation so that vacancy rates can be related to occupational skill levels. Using these Conference Board series, there are more job vacancies when jobs go unfilled more than one 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

[12] For early papers, see Abowd, Haltiwanger, Lane, McKinsey, and Sandusky (2007), Autor, Katz, and Krueger (1998), Bresnahan, Brynjolfsson, and Hitt (2002), Katz, and Murphy (1992).

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

[14] van Ours, and Ridder (1991), "Cyclical variation in vacancy durations and vacancy flows," *European Economic Review,* 35: 1143–1155.

new ads and total ads. Table 1 shows the average unfilled jobs ratio across occupations and years is .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. 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 categorize education groups differently over time. Consequently, all educational groupings are converted into a 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. The occupation variable used here is at a relatively fine level. There are an average of 343 occupations delineated by this variable.[15] Table 1 shows that average earnings are $66,849 in 2013 dollars.

The third data set is the PSID from 1968 through 2010. The dataset follows 5,382 men between the ages of 25 and 65, with an average of 5.25 years of panel data per person. The PSID originates with a sample in 1968 and then introduces new respondents into the sample as parents have children who become respondents. Table 1 shows that the average wage is $25/hour in 2010 dollars. Note that the definition of occupation is coarse in the PSID as compared with the other data sets. There are only 25 occupations defined in the PSID.

To detect direct evidence of luck, data are available from online labor market oDesk on job postings, applicants for the job and new hires. oDesk.com (recently 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 over $1.3 billion in contracts (Zhu et al 2015). The oDesk platform allows employers to post jobs, hire from the online applicants, make payments to these globally distributed remote workers, and monitor workers with proprietary project management software.

The oDesk data allows a unique opportunity to study the role of luck in finding jobs because the transactions data used here contain records of employers' hiring along with the entire set of applications that employers receive. 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, their electronic resume as displayed on a profile maintained on oDesk, and, importantly, they also bid an hourly wage.[16] For workers who have worked on oDesk before, there is a public evaluation (1 to 5 score) of past performance done by

[15] There are a total of 384 unique 1990 codes throughout the dataset. 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 dataset for 1975, and a high of 373 occupational codes in 1994.

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

employers. Some workers are affiliated with agencies and others are unaffiliated. Stanton and Thomas (2015b), Ghani, Kerr, and Stanton (2014), Horton (2010), Pallais (2013) and Agrawal et al (2012) and Agrawal et al (2013) describe oDesk marketplace institutions in more detail.

Table 2 provides summary information about jobs posted on oDesk by different task category. There are 9 main categories of jobs posted on oDesk over the 2006 – 2010 data used here. The range of skills for these jobs is evident in the wages paid for the work. The two largest job categories are Administrative Support and Web Programming.

Empirical Results

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

Hypothesis 1: Bumping occurs. If two high ability workers show up at the firm, then the lowest ability of the two is bumped down to the easy job.

The oDesk are used to test for bumping. In this online job site, the employer posts a job opening, a multitude of applicants respond to that opening with their resumes, wage bids, and past evaluations. The employer hires one of these or conceivably, none at all.

To test for bumping, the first step is to build a model of the probability that each individual worker is hired within a given slot. Stanton and Thomas (2015a) provide this model for the oDesk data. The employer on job opening *i* hires the best worker available for the individual slot. For each worker *j*, the employer observes that part of worker quality that is commonly valued in all production tasks, q_j, and chooses the applicant in the choice set C_i that has the highest expected wage-adjusted quality given by

$\frac{q_j \exp(\varepsilon_{ij})}{w_{ij}}$, where $C_i = \{\{q_j, \varepsilon_{ij}, w_{ij},\}_{j=1}^{J_i}, \{0, \varepsilon_{i0}, 0\}\}$ includes a no hire option, $\{0, \varepsilon_{i0}, 0\}$. Adding a purely idiosyncratic Type-1 extreme value distributed error, ε_{ij}, implies that the probability of being hired takes a conditional logit form within job opening. Taking logs, the probability that worker j is hired is the probability that worker *j* is the highest order statistic in the choice set, or

$$\Pr\left(log(q_j) + \varepsilon_{ij} - \alpha log(w_{ij}) > log(q_k) + \varepsilon_{ik} - \alpha log(w_{ik})\right)$$

for all $k \in C_i$. Taking this to data, the commonly valued components of productivity are modeled as $\log(q_j) = X_{jt}\beta + v_{jt}$, where v_{jt} is an error component that is correlated with worker quality that may be observable to employers but is not observed in the data.[17] This creates the possibility that the bids submitted to jobs are endogenous.

[17] In the richer model in Stanton and Thomas (2015a), employers also have horizontal measures of preferences, but because the horizontal measures are orthogonal to worker skills, they are less important for quantifying the extent of luck due to slots.

The ability for bids to adjust to job characteristics, individual characteristics, or, most importantly, the extent of competition does not mean that luck is absent from labor market allocations. However, if wages adjust to compensate for different probabilities of being hired, endogenous bids are likely to make detection of luck more difficult. Therefore, a technique is needed to incorporate endogeneity of bids.

Wage bid endogeneity is accounted for using variation in exchange rates as an instrument for workers' bids (see Stanton and Thomas (2015a)). Workers' local labor market opportunities are denominated in their own currency, while contracts on oDesk are all denominated in dollars. With any friction that limits immediate adjustment of local prices to exchange-rate parity, appreciation of the local currency relative to the dollar (one dollar provides more local currency units) results in a reduction in the dollar-denominated value of workers' outside option. This shift in exchange rates is expected to change equilibrium bidding behavior online, but it is independent of unobserved worker quality, v_{jt}.

With this instrument, the control function approach of Petrin and Train (2009) is used to estimate the parameters governing the probability that an individual worker is hired. This approach uses the residuals from the first stage regression as a regressor in the conditional logit model to control for that part of a workers' bid that is orthogonal to observed worker characteristics and the exchange rate variation. It is this orthogonal component that is potentially observed by employers and correlated with wages, so the control function approach directly includes the endogenous portion of log wages.

One difficulty is that there may be sorting by workers on the instrument. A second instrument within a selection model permits the participation decision to be correlated with exchange rate movements. Details about this second instrument, the selection model, and instrument strength are in the Appendix.

Letting CF denote the residuals from this first stage regression, choice probabilities are

$$p_{j|i} = \exp\left(X_{jt}\beta - \alpha log\left(w_{ij}\right) + CF_{jt}\psi\right) / (1 + \Sigma_J^{J_i} exp\left(X_{jt}\beta - \alpha log\left(w_{ij}\right) + CF_{jt}\psi\right),$$

which is estimated by maximum likelihood. Table 3 displays the parameter estimates for the two largest samples that span low skill and high skill, data entry and web development, respectively. When the model uses generated regressors, block-bootstrapped standard errors are reported with each block corresponding to a job opening. Otherwise, robust standard errors are reported.

The basic results are sensible: the estimated parameter values show that employers value applicants with better feedback scores and more past experience. In most cases, the parameters on worker characteristics are larger when wages and the control function enter the model, suggesting that wage bids are positively correlated with characteristics that employers value. Also consistent with Stanton and Thomas (2015b), employers more highly value novice applicants with agency affiliation, but this effect fades out for more experienced applicants.

With estimates of the parameters in hand, it is possible to assess how sensitive the hiring probability of one worker is with respect to the characteristics of other workers who apply for a

job. To capture differences in quality as valued by employers through revealed preferences, the parameters $\beta_{JobCategory}$ are used as weights on an index of worker quality. An individual worker's quality index within each job category is $X_{jt}\beta_{JobCategory}$. The first six applicants to a job are used to assess luck; this is to abstract away from different number of job applicants to different vacancies and to ensure that hiring probabilities are roughly comparable across applicant order.[18]

Given this setup, it is now possible to test Hypothesis 1, that bumping occurs. Using the index of applicant quality, a worker is said to be lucky if the next applicant to arrive has a lower index of applicant quality. A worker is unlucky if the next applicant to arrive has a higher index of quality. Additionally, workers themselves can be ranked overall relative to the distribution of quality, and 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.

Panel B of Table 3 displays the results. Good applicants are uniformly more likely to get a job when the next applicant to arrive is of lower quality. For both data entry and web programming, the change in hiring probability for good applicants due to luck is about 30% (Columns 2 and 4). In levels, however, individual applicants in web programming are less likely to be hired regardless of luck, consistent with the predictions of the model that not all workers are qualified for more skill-intensive tasks. For bad applicants, those below the median of the quality index, luck also plays a role but the change in hiring probability due to luck is much smaller; these workers are very unlikely to be hired, and with multiple applicants to a job, conditioning only on the identity of the next applicant when computing luck has very little 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.

Hypothesis 2: Low ability workers are more likely to be unemployed than high ability ones. High ability workers can work both the easy and difficult jobs, which makes them more employable. The model predicts that sufficiently high ability workers can never be unemployed.[19]

[18] A spline with applicant order is included in the characteristics, X, but to be included in the sample, job openings must have at least 6 applicants, making the assessment balanced across job openings. Only worker-initiated applicants are considered for the purposes of these calculations. 92 percent of the first 6 applications to web programming jobs are initiated by workers.

[19] 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

Figure 3 shows the unemployment rate by education using the 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-prices, 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.

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 sufficiently high ability worker 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 total job postings. The hypothesis is that this rate rises with skill level. The Conference Board data include the SOC 3-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: highly skilled occupations are those occupied by highly educated people.

The first test of Hypothesis 2 is in Table 4. 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. The unit of analysis is an occupation-year. The dependent variable is the number of unfilled postings in a month divided by total postings averaged over all the months for that year. Similarly, the independent variable is a year dummy interacted with average education in the occupation. Thus, there are 94 occupations times 9 years, or 846 observations. Each year-occupation variable is an independent test of the hypothesis so the fact that all nine 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 increases the percent unfilled by about .15, over a mean of .47.

The second test of Hypothesis 2 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 skill used are the mean wage in the job category and the 90^{th} percentile of wages in the job category. The probability of filling a job is negatively related to both measures. Table 5 also displays point estimates for fill probabilities for each of the main job categories in the data.

---

certain that both workers are employed. Thus, any worker with ability greater than **Error! Main Document Only.**-$\gamma/(\delta-\beta)$ can be certain that he will be at a firm that employs both applicants and can never suffer unemployment.

By way of background, note that in Table 2 applications per-vacancy are highest in low-skill tasks (as measured by mean 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 development or software development.

In sum, using measures of job vacancies for jobs posted online, vacancies rise with skill level. Although this may not be a completely novel finding, it 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 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 set so it is reasonable to conclude that highly educated workers not only have low unemployment rates, but that they are also found in jobs with high vacancy rates.

Hypothesis 4: The over-qualified and under-qualified will 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.

Two data sets are used to test this, the CPS and the PSID. The CPS is the much larger data set, but the PSID follows respondents over time.

For the CPS data, two definitions of over-qualification are used, for a broad and narrow definition of over-qualification..

1. Define "over-qualification 1" as the condition in which person i's educational level exceeds the average education in his occupation j by more than one standard deviation:
   (12) Individual i in occupation j is over-qualified for that occupation if
   $\text{Education}_{ij} > \text{MeanOccupationalEducation}_{ij} + \text{Standard DeviationOccupationalEducation}_{ij}$

2. Define "over-qualification 2" as the condition in which person i's educational level exceeds the average education in his occupation j by more than two standard deviations:
   (13) Individual i in occupation j is over-qualified for that occupation if
   $\text{Education}_{ij} > \text{MeanOccupationalEducation}_{ij} + 2\ (\text{Standard DeviationOccupationalEducation}_{ij})$

Two definitions of under-qualification are also used, for broad and narrow definitions.

1. Define "under-qualification 1" as the condition in which a person i's educational level is less than the average education in his occupation j by more than one standard deviations:
   (14) Individual i in occupation j is under-qualified for that occupation if

$Education_{ij} < MeanOccupationalEducation_{ij} + Standard\ DeviationOccupationalEducation_{ij}$

2. Define "under-qualification 2" as the condition in which a person i's educational level is less than the average education in his occupation j by more than two standard deviations: (15) Individual i in occupation j is under-qualified for that occupation if $Education_{ij} < MeanOccupationalEducation_{ij}$ j +2(Standard $DeviationOccupationalEducation_{ij}$)

It is now necessary to define the wage gap, $\hat{W}$ -*W*. This comes directly from the theory. The wage, W, is simply the wage that an individual currently has. For example, for over-qualified individuals, this would be the wage that they currently earn in the job for which their ability levels are too high. In the context of Figure 1, an individual with A>A* who is working in the easy job should be receiving wage $W = \alpha + \beta A$. However, that individual should be in the difficult job. Wages can be estimated for those who are in the difficult job where the wage depends on ability level, which is proxied by the given individual's education. For example, for engineers, among those qualified to do their job, there are some highly able and some less able and wages will rise with ability. Therefore, the predicted wage, W^, is the wage he would receive if he were appropriately assigned, as opposed to over-qualified for his job, given his education and age. Then define

(16) $\hat{W}_{it} = b_0 + b_1 f(Education_{it}) + b_2 g(Age_{it}) + b_3 h(Education_{it} * Age_{it}) + e_{it}$

The predicted wage is a function of quadratics in Education and Age. The coefficients for (16) are obtained by estimating the regression for the sample of qualified workers (who are neither over- nor under-qualified for their jobs as defined above) and those coefficients are then applied to the sample of over- and under-qualified workers.

The resulting wage gap, $\hat{W}$ -W, should be positive for both over- and under-qualified workers. For example, a worker who is over-qualified has A>A* in Figure 1, but is working in the easy job, earns only $\alpha + \beta A$ instead of the appropriate and higher $\gamma + \delta A$, as predicted by (16).

The empirical results are subject to a number of problems, some of which are data related. The major data-related issue is that some people who work in low-skilled jobs have very high pay. This may be informative, but more likely reflects measurement or reporting errors. Thus, outliers are dropped for the calculation of $\hat{W}$ -*W*, where an outlier is defined as the observed wage, *W*, that is more than three standard deviations from the average in the observed occupation.

Table 6 contains the results using CPS data that relate to over-qualification.[20] The columns of the table correspond to the two different definitions of over-qualification in (12) and

[20] Outliers, defined as observations for which the wage differs from the within-education group mean by more than three standard deviations are omitted.

(13). No matter which definition of over-qualification is used, the mean wage gap is significantly positive. For both measures, the wage gap is clearly positive for the over-qualified: a t-test confirms that the mean $\hat{W}$ -W is greater than zero.

There is another confounding interpretation of these results. Individuals who are deemed to be over-qualified for their jobs may be in those jobs because of some unobserved ability component that is not captured by measured education. For example, a Ph.D. who is working as an administrative assistant may be in that job not because someone else occupied the professorship so he took the clerical position, but perhaps because he is not qualified to be a professor. In the context of the model, this would be a person whose actual ability is A < A*, but whose measured ability based on education places him at A>A*. It would not be surprising that a highly educated but low ability individual would earn less than predicted, even if he were assigned properly, in this case to the easy job. It is possible to treat this using panel data, where an individual's fixed effect can be estimated and over-qualification is defined as being in a job that is low relative to that individual's lifetime mean. Panel data results will be discussed shortly after the results for the under-qualified are presented.

Table 7 reports the results for the under-qualified. The wage gap is clearly negative for both definitions of underqualified even though it was predicted to be positive. As was the case for over-qualified individuals, the results are likely to be confounded by unobserved ability variation. An under-qualified worker is one with low levels of education who works in an occupation that generally requires more educated workers. The reason that she is in that occupation may be that her true ability may be higher than A*, even though it is measured to be below A*. As such, she is appropriately suited to the difficult job in which she is employed. The predicted wage is based on her being in the easy job and therefore understates her true earning capacity. Panel data mitigates these unobserved ability problems that surface in the context of testing Hypothesis 4 using the cross-section time-series CPS data. Recall that the concern was that the higher-than-predicted wage for the over-qualified might simply reflect unobserved ability.

The use of the Panel Study of Income Dynamics reduces these concerns because a given individual's job mobility over time can be used. In particular, ability can be taken out in the standard way by estimating worker fixed effects. More to the point, panel data allow for a more refined definition of over- and under-qualification.

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 wage should be higher in her usual job than in the unusual one for which she is over-qualified.

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, they hire the former machinist. His productivity as machinist, $\alpha+\beta A$, exceeds his productivity in the clerical position, $\gamma+\delta A$, so he is under-qualified 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 under-qualified 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.

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 under-qualified: 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 over-qualified. She would have earned the amount that corresponds to point 1 were she in the appropriate job (the difficult one), but earns only the amount that corresponds to point 2. Second, and important, even though over-qualified workers receive less in the job for which they are over-qualified 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, under-qualified workers not only receive less in the job for which they are under-qualified than in their usual job, they receive less than the typical worker in the difficult job. In Figure 4, the under-qualified 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 under-qualified worker also receives less than the typical worker in the job for which he is under-qualified 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, shown by point 4, where the wage that he receives is lower than the wage that the typical worker in the difficult job receives.

It is first necessary to define over- and under-qualified in the PSID data. This was done 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.[14] About one-fifth of the observations fit this definition of being unusual. The worker in an unusual occupation was defined as over-qualified 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 under-qualified 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 under-qualified and 10% being classified as over-qualified.

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

As was done with the CPS data, it was necessary to estimate the predicted wage for each worker-year in the sample. There were 19,397 person-year observations for individuals in their usual occupation used to estimate the wage regression across 4385 individuals. In this estimation, worker fixed effects are included to mitigate, if not remove, any ability bias that was a concern in the CPS estimation.s[215] The wage regression is estimated using only observations that are in the usual occupation, which implies that no under- or over-qualified job spells influence the results. The predicted log wage can be computed for the unusual job spells, using the coefficients and person fixed effects that are estimated in the usual occupation regression. This provides an estimate of what the worker would receive were he in his normal job.

Summarizing, there are four patterns that should appear in the results. First, those who are over-qualified for their jobs should 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. Second, those who are under-qualified 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 over-qualified earn less than in they would in their proper jobs, they should earn more than the typical worker in the job for which they are over-qualified. Fourth, not only do the under-qualified 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 under-qualified.

The results are reported in Table 8. Column 1 reports the results for those job spells that correspond to under-qualification and Column 2 for those that correspond to over-qualification.[21] Row A reports that the average log wage gap between the predicted wage and the actual wage is positive as predicted for both groups. 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 number reported is the average log wage gap across all person-years that fit the definition of under- or over-qualification. The average log wage gap is positive and statistically significant in both cases.

Row B reports the average difference between the received log wage and the mean log wage of workers who are in the occupation, most of whom are there appropriately. The prediction is that this gap should be negative for the under-qualified and positive for the over-qualified. Both predictions are borne out, again with numbers that are statistically significant.

The panel data approach is consistent with the theory and treats the issue of unobservables. Because person fixed effects are removed, the bias that plagued the CPS cross-sectional estimates appears to be removed here, at least insofar as producing the expected results.

---

[15]The regressions were done two ways. The first specification included age, education, quadratic terms and interactions as well as person fixed effects. The second included only year and person fixed effects. The R-squared was almost identical and above .8 in both specifications because once fixed effects are included, only time varying education and aging contributes to the regression, the latter being captured mostly by year effects. Results here are based on the simple specification.

[21] Individuals are dropped from the sample if their wage is an outlier, measured as their wage being greater than three standard deviations from the occupational average wage. Comparing Table 8 to Tables 6 and 7, Table 8 follows the rows that drop wages greater than three standard deviations, and there are no columns in Table 8 because the over- and under-qualified are those that are outside their modal occupation.

Finally, the PSID data allows an assessment of which workers are in jobs for which they are over and under-qualified. Figure 5 presents the results of local polynomial regressions that flexibly characterize the probability of over and under qualification as a function of the log wage in the usual occupation. Consistent with the theory, those in the middle of the distribution are those who are likely to be in unusual jobs. The probability of over qualification is increasing with skill up to a point and then declines. Underqualification is declining with the usual log wage. The previous results say nothing about who is in an unusual job; these results suggest that those who have high wages in their usual occupation are most at risk for overqualification up to a point. Interestingly, those with the highest usual log wages are less likely to be overqualified than those who earn slightly less in their usual occupations. That is, those at the very top are unlikely to be bumped down, and this inverted u-shape is exactly what comes out of the theory. Those with low wages in their usual occupation are most at risk for underqualification, again consistent with the theory.

Hypothesis 5: When the return to skills rises over time, there are increasingly adverse wage consequences of mismatch. This result depends on the assumption, made earlier, that the nature of skill-biased technical change steepens the relation of wages to ability more in difficult jobs than in easy jobs.

It is well known that over the last thirty 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. As discussed earlier, this implies that the difference between δ' and δ,. is greater than the difference between $\beta'$and β, as shown in Figure 2. The implication is that the variance of pay is greater today for the highly able than it was in the past: 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:

$$(17) \quad \sigma_{it} = b_1 Year_{it} + b_2 Skill_{it} + b_3 Skill_{it} * Year_{it} + e_{it}$$

where the variance of pay is calculated for each education group i and for each year t, resulting in a data set of 672 observations.[22] There are two measures of variance. The first, is the ratio of wages in the 90th percentile for that education group over wages in the 10th percentile for the education group. The second is the standard deviation of wages by educational group. The measure of Skill is the mean education in the occupation. The first implication is that $b_1>0$ because the rising return to skills increases the variance of earnings over 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.

[22] There are 40 years of data and an average of 17 educational groups, where the educational group varies by year as the CPS changes these educational categories over time.

Regression results in Table 9 are consistent with both implications. Column 1 shows the shows the 90/10 ratio of wages and column 2 shows the standard deviation of wages. The variance of pay has risen over time, but it has risen most for the highly skilled.

Conclusion

A worker's skills alone cannot determine the job in which he or she is hired or indeed, hired at all. The existence of slots or job positions 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 that position.

Although this idea is intuitive, it has not been modeled or nor have its implications been explored. The model and analysis herein not only provides many specific predictions on what should be observed in hiring and job assignment, but tests and validates those predictions using four different data sets.

First, the job to which one is assigned and whether hired at all depends not only on own skills but also on the skills of the competition. This is verified using oDesk data.

Second, bumping occurs where workers take jobs for which they are not well-suited, but receive the offer because their skills are superior to those of other applicants but inferior to those applying to the job that they prefer. The model provides a clear definition of "overqualification" and of "underqualification" that has specific empirical meaning. Using these definitions, the CPS and PSID data provide evidence that over- and under-qualification occurs and that the wages that are received in those jobs are exactly as predicted by the model. Namely, both over- and under-qualified workers receive less in those jobs than they would in their appropriate positions, but over-qualified workers receive more than the average worker in that job. Conversely, underqualified workers not only receive less than they would in their appropriate job, but also less than the average worker in the job for which they are underqualified.

Third, less able workers are more likely to be unemployed because the more able workers are capable of doing a wider variety of jobs. This implication is not as obvious as it seems. The model provides this as an implication, and not surprisingly, the implication is found to hold using the CPS data.

Fourth, vacancy rates are higher in jobs that require higher 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 do the high skilled jobs. The Conference Board data on vacancy rates confirms this prediction.

References

Abowd, J. M., Haltiwanger, J., Lane, J., McKinsey, K. L., and K. Sandusky, 2007, "Technology and the demand for skill: an analysis of within and between firm differences," *NBER Working Paper 13043*.

Agrawal, Ajay, Lacetera, Nicola, and Elizabeth Lyons, 2012, "How Do Online Platforms Flatten Markets for Contract Labor?" *Working Paper*.

Agrawal, Ajay, Horton, John, Lacetera, Nicola, and Elizabeth Lyons, 2013, "Digitization and the Contract Labor Market: A Research Agenda," *NBER Working Paper 19525*.

Altonji, Joseph G., and Charles R. Pierret, 2001, "Employer Learning and Statistical Determination," *Quarterly Journal of Economics*, 116 (1): 313-350.

Andrews, M. J., Bradley, S., Stott, D., and R. Upward, 2008, "Successful employer search? An empirical analysis of vacancy duration using micro data," *Economica,* 75: 455–480.

Autor, D. H., 2001, "Wiring the labor market," *Journal of Economic Perspectives,* 15: 25–40.

Autor, D. H., Katz, L. F., and A.B. Krueger, 1998, "Computing inequality: Have computers changed the labor market?" *Quarterly Journal of Economics,* 113: 1169–1213.

Autor, Levy, and Murnane, 2003, "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118 (4): 1279-1334.

Baker, George P., Gibbs, Michael, and Bengt Holmstrom, 1994, "The Wage Policy of a Firm," *Quarterly Journal of Economics*, 109 (4): 921-955.

Brencic, V., 2009, "Do employers respond to the costs of continued search?" *Oxford Bulletin of Economics and Statistics,* 71: 1–25.

Bresnahan, T. F., Brynjolfsson, E., and L. M. Hitt, 2002, "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence," *Quarterly Journal of Economics,* 117: 339–376.

Burks, Stephen V., Cowgill, Bo, Hoffman, Mitchell, and Michael Housman, 2015, "The Value of Hiring Through Employee Referrals," *The Quarterly Journal of Economics,* 130 (2): 805-839.

Diamond, Peter, 1982, "Wage Determination and Efficiency in Search Equilibrium," *The Review of Economic Studies,* 49 (2): 217-227.

Gee, Laura, 2015, "The More You Know: Information Effects on Job Application Rates by Gender in a Large Field Experiment," *Working Paper*.

Ghani, Ejaz, Kerr, William, and Christopher Stanton, 2014, "Diasporas and outsourcing: Evidence from oDesk and India." *Management Science*, 60 (7): 1677-1697.

Gibbons, Robert, Katz, Lawrence, Lemieux, Thomas, and Daniel Parent, 2005, "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, 23 (4): 681-724.

Mitchell Hoffman, Kahn, Lisa, and Danielle Li, 2015, "Discretion in Hiring," *Working Paper*.

Horton, John, 2010, "Online Labor Markets," *Working Paper*.

Katz, Lawrence F., and Kevin M. Murphy, 1992, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107 (1): 35-78.

Kudlyak, Marianna, Lkhagvasuren, Damba, and Roman Sysuyev, 2012, "Systematic Job Search: New Evidence from Individual Job Application Data," *FRB Richmond WP #12-03*.

Kuhn, Peter, and Hani Mansour, 2014, "Is Internet Job Search Still Ineffective?" *Economic Journal,* 124 (581): 1213–1233.

Kuhn, Peter, and Kailing Shen, 2013, "Do Chinese Employers Avoid Hiring Overqualified Workers? Evidence from an Internet Job Board," *Research in Labor Economics,* 37: 1-30.

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

Kuhn, Peter, and Mikal Skuterud, 2004, "Internet Job Search and Unemployment Durations," *American Economic Review,* 94 (1): 218-232.

Lazear, Edward P., 1992, "The Job as a Concept," in *Performance Measurement and Incentive Compensation*, ed. William J. Bruns, Jr. Cambridge: Harvard Business School Press.

Lazear, Edward P., 1986, "Retail Pricing and Clearance Sales," *American Economic Review, 76 (6): 14-32.*

Marinescu, Ioana, and Ronald Wolthoff, 2015, "Opening the Black Box of the Matching Function: The Power of Words," *Working Paper*.

Modestino, Alicia, Shoag, Daniel, and Joshua Balance, 2015, "Upskilling: Do Employers Demand Greater Skills When Workers are Plentiful?" *Working Paper*.

Pallais, Amanda, 2014, "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104 (11): 3565-3599.

Petrin, Amil, and Kenneth Train, 2009, "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research,* XLVI.

Rogerson, Richard, Shimer, Robert, and Randall Wright, 2005, "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43: 959–988.

Spence, Michael, 1973, "Job Market Signaling," *Quarterly Journal of Economics*, 87 (3): 355-374.

Stanton, Christopher, and Catherine Thomas, 2015a, "Information Frictions and Observable Experience," *Harvard Business School Working Paper*.

Stanton, Christopher, and Catherine Thomas, 2015b, "Landing the First Job: The Value of Intermediaries in Online Hiring," *Review of Economic Studies*, Forthcoming.

van Ours, J., and G. Ridder, 1991, "Cyclical variation in vacancy durations and vacancy flows*," European Economic Review,* 35: 1143–1155.

van Ours, J., and G. Ridder, 1992, "Vacancies and the recruitment of new employees," *Journal of Labor Economics,* 10: 138–155.

van Ours, J., and G. Ridder, 1993, “Vacancy duration: Search or selection?” *Oxford Bulletin of Economics and Statistics,* 55: 187–198.

Figure 1

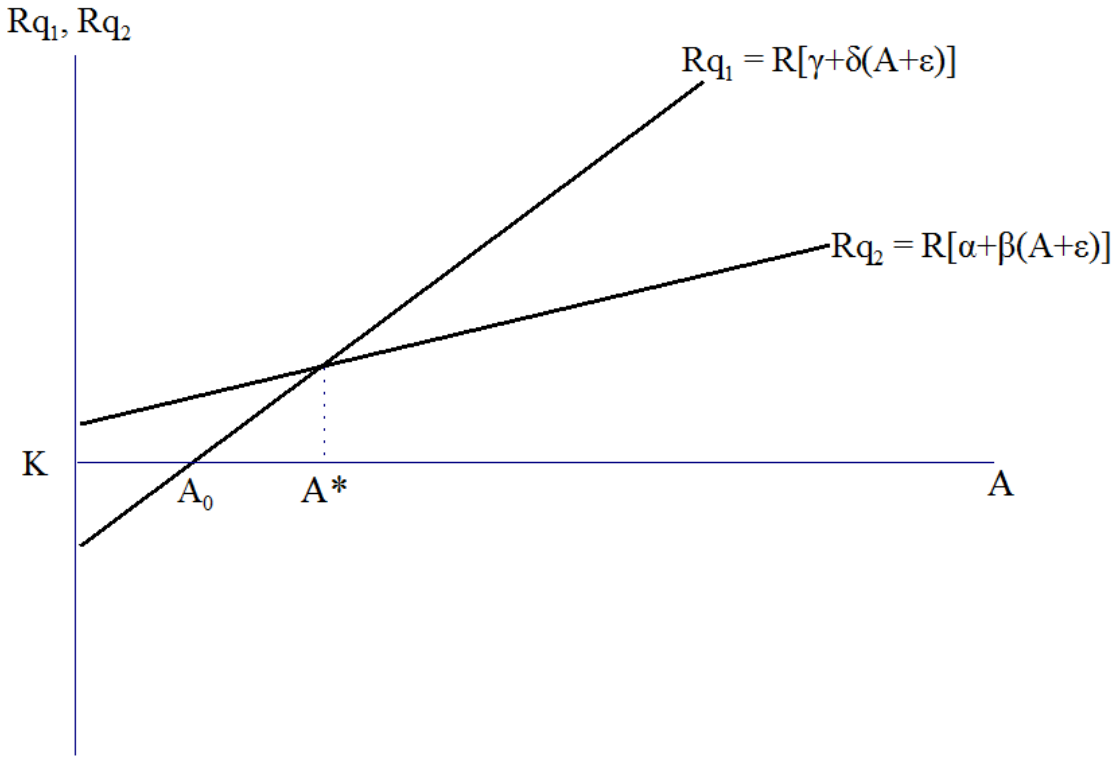

Rq_1, Rq_2
$Rq_1 = R[\gamma+\delta(A+\varepsilon)]$
$Rq_2 = R[\alpha+\beta(A+\varepsilon)]$
K
A_0
A*
A

Figure 2

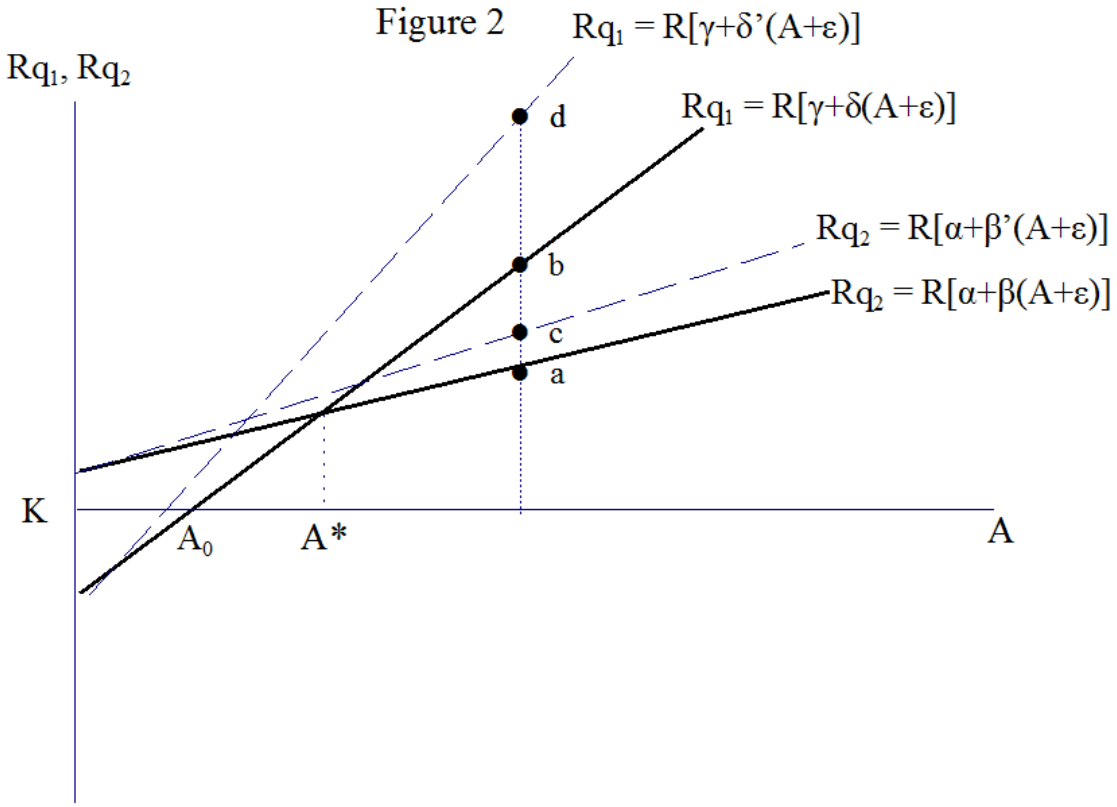
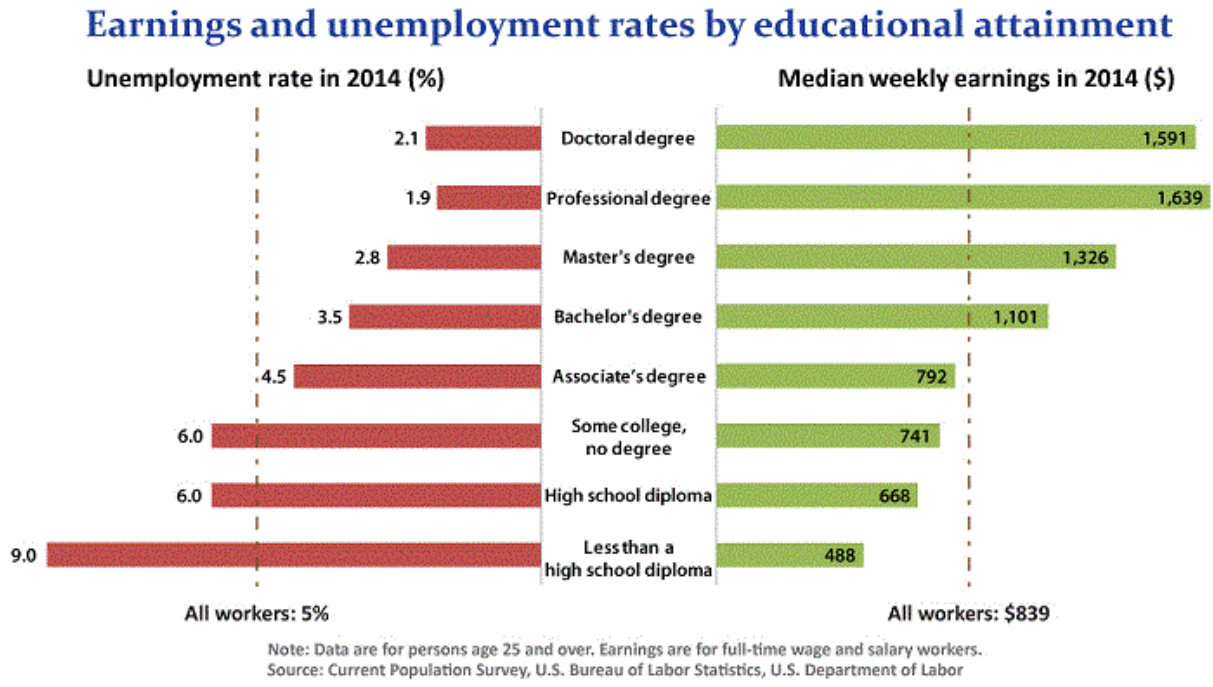

$Rq_1 = R[\gamma+\delta'(A+\varepsilon)]$
Rq_1, Rq_2
d
$Rq_1 = R[\gamma+\delta(A+\varepsilon)]$
$Rq_2 = R[\alpha+\beta'(A+\varepsilon)]$
b
$Rq_2 = R[\alpha+\beta(A+\varepsilon)]$
c
a
K
A_0
A*
A

Figure 3:



From: Bureau of Labor Statistics, Employment Projections, April 2, 2015: http://www.bls.gov/emp/ep_chart_001.htm

Figure 4

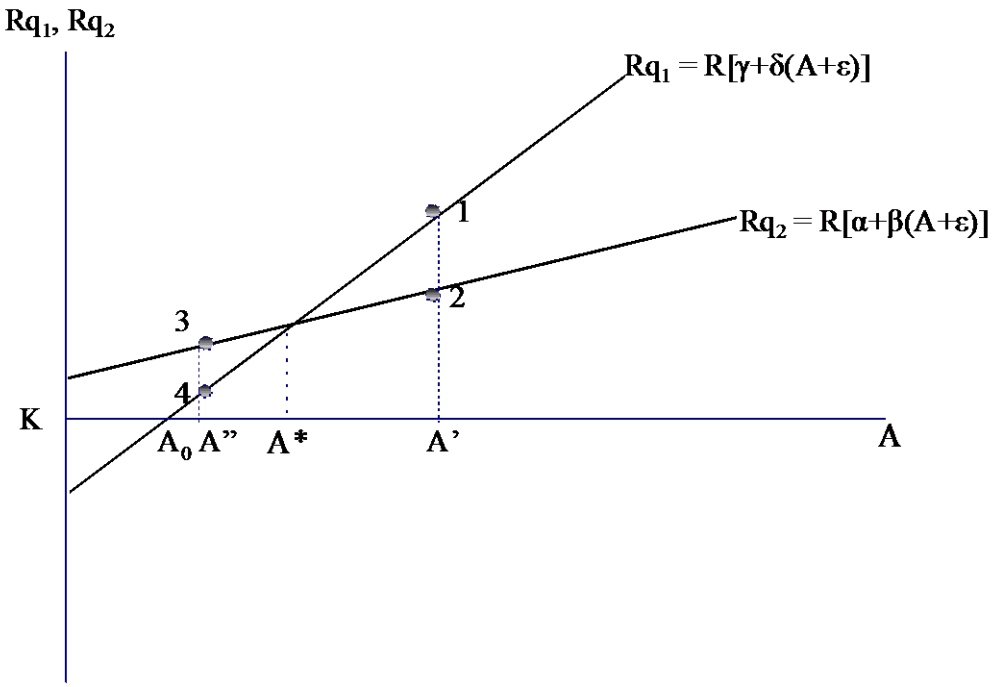


Figure 5: Probability of Over and Under Qualification by Log Wage in Usual Occupation

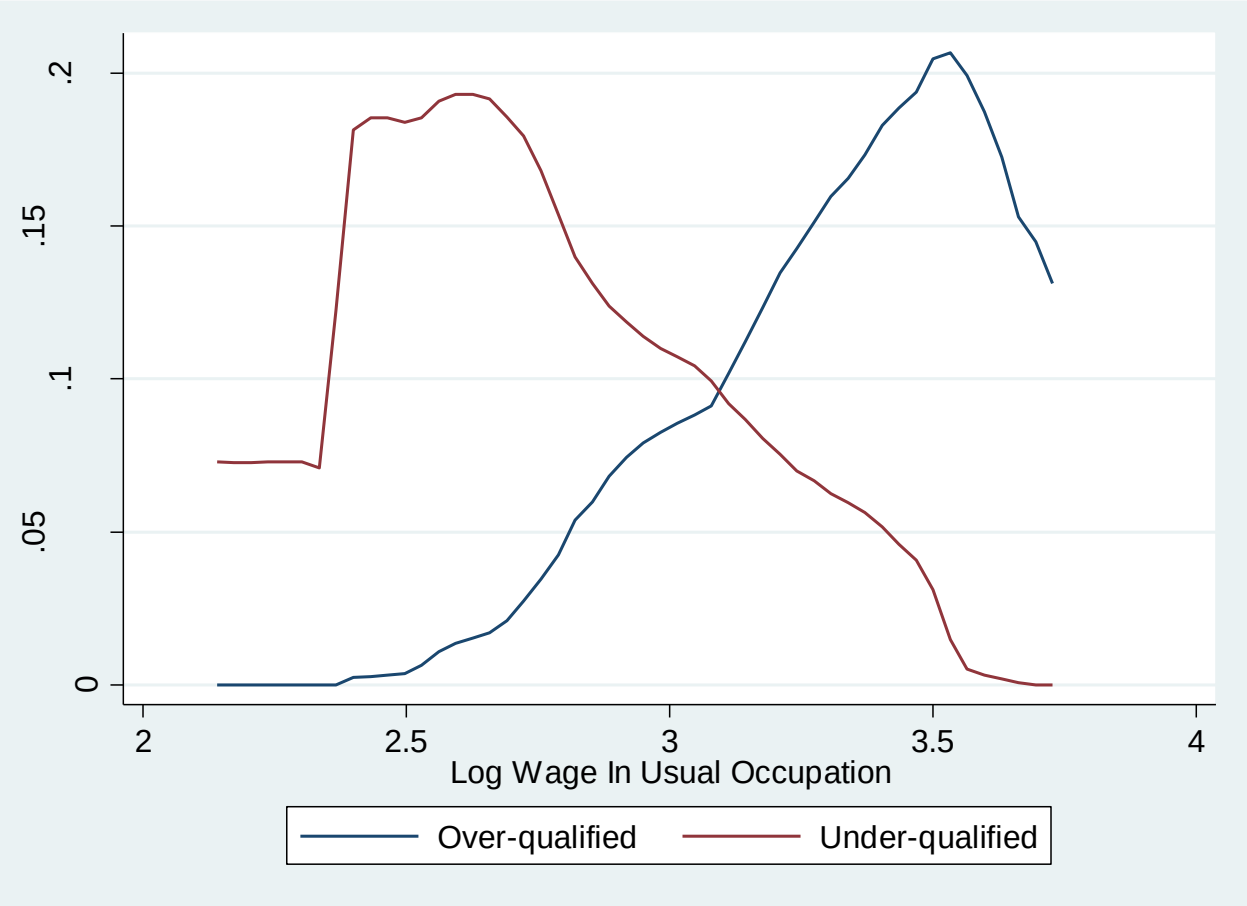


Table 1:
Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Conference Board Data:					
Mean Education by Occupation	846	14.02793	1.615229	9.166667	18.13372
Unfilled Jobs Ratio	846	.4704046	.0970282	.0761404	.7067247
CPS Data:					
Age	870665	39.04593	8.267556	25	54
Education	870665	13.61305	2.763408	0	19
Yearly Earnings	866432	66849.35	58876.55	1.085666	1845631
PSID Data:					
Age	28255	40.76602	10.73039	25	65
Education	28106	13.2164	2.697894	0	17
Tenure (weeks)	27010	49.25442	82.5911	0	780
Hourly Wage	28255	25.01306	13.7365	.0004947	154.7392

Note: The sample size for the Conference Board data is the 9 years (2006-2014) for 94 occupations. The sample size for the CPS data is the number of men age 25-54 for March years 1975-2013. The sample size for the PSID data is men age 25-65 from 1968 to 2010.

Table 2:
Summary Statistics by Job Category, oDesk Data

Job Category	Vacancies	Applications Per Vacancy	Mean Probability that Vacancy Is Filled	90th Percentile of Log Hourly Wages (Filled Vacancies)	Mean Log Hourly Wage	St. Dev. Of Log Hourly Wage
Administrative Support	32,175	26.15	0.49	2.08	0.97	0.91
Business Services	3,129	14.79	0.33	3.00	1.64	1.01
Customer Service	1,565	28.85	0.29	2.57	1.41	1.07
Design & Multimedia	18,737	15.14	0.40	3.00	2.35	0.70
Networking & Information Systems	4,905	11.35	0.34	3.51	2.67	0.77
Sales & Marketing	14,349	11.32	0.37	2.66	1.56	0.87
Software Development	21,818	10.21	0.33	3.22	2.57	0.96
Web Development	100,051	14.33	0.40	3.09	2.49	0.65
Writing & Translation	20,766	13.02	0.43	2.81	1.75	0.88

Table 3
oDesk Conditional Logit Parameter Estimates and Assessment of Luck

	(1)	(2)	(3)	(4)
Sample:	Data Entry	Data Entry, With Control Function	Web Programming	Web Programming, With Control Function
Panel A: Parameter Estimates and Standard Errors				
Log Hourly Bid		-9.124		-8.706
		(2.291)		(2.886)
Control Function		8.401		8.246
		(2.266)		(2.889)
Feedback Score Out of 5	0.242	1.062	0.237	0.838
	(0.0905)	(0.295)	(0.0410)	(0.218)
Agency Affiliate Indicator	-0.181	0.166	-0.459	-0.107
	(0.127)	(0.198)	(0.0511)	(0.141)
Agency Affiliate x Inexperienced Worker	0.498	0.955	0.235	0.244
	(0.243)	(0.289)	(0.122)	(0.123)
Indicator for Prior Experience on oDesk	1.981	1.419	1.771	2.794
	(0.144)	(0.22)	(0.100)	(0.345)
Mean Own-Bid Elasticity		-9.036		-8.627
		(2.65)		(2.869)
Number of Job Openings	1781	1781	8230	8230

Table 3, continued
oDesk Conditional Logit Parameter Estimates and Assessment of Luck

	(1)	(2)	(3)	(4)
Panel B: Hiring Probabilities by Own and Future Applicant Type for the First 6 Applicants				
Good Applicant and Better Next Applicant	0.0148	0.0152	0.0121	0.0135
Good Applicant and Worse Next Applicant	0.0235	0.0199	0.0194	0.0176
(Standard Error on Difference)	(0.0011)	(0.0012)	(0.0003)	(0.0006)
Bad Applicant and Better Next Applicant	0.004	0.0101	0.004	0.007
Bad Applicant and Worse Next Applicant	0.005	0.0118	0.005	0.008
(Standard Error on Difference)	(0.0002)	(0.0007)	(0.0001)	(0.0002)

The sample is described in the text. Job openings must have had more than six applicants, with at least three worker-initiated applicants. Estimates come from a conditional logit model that includes an option not to hire an applicant. In models with the control function, standard errors come from using 50 block bootstrap replications of the entire procedure in which each job opening forms a block. Other controls are included for English score, an indicator for no feedback, an indicator for being experienced without having feedback, as well as the country dummies and applicant order spline as described in the table notes for the first stage regression. In Panel B, a good or bad applicant is defined as one above or below the median for the set of worker characteristics, X times coefficients, β, excluding the control function and the log bid. Whether the next applicant is better or worse, ordering applicants by application time, is coded using the same underlying measure of applicant quality. All standard errors in Panel B are calculated using 50 block bootstrap replications to account for parameter uncertainty in forming $X\beta$.

Table 4
Conference Board Online Job Postings Regression Results

	(1) Unfilled Jobs Ratio
Occupational Education * Year 2006	0.040***
	(0.0070)
Occupational Education * Year 2007	0.038***
	(0.0067)
Occupational Education * Year 2008	0.037***
	(0.0065)
Occupational Education * Year 2009	0.036***
	(0.0065)
Occupational Education * Year 2010	0.035***
	(0.0063)
Occupational Education * Year 2011	0.034***
	(0.0062)
Occupational Education * Year 2012	0.036***
	(0.0063)
Occupational Education * Year 2013	0.037***
	(0.0062)
Occupational Education * Year 2014	0.036***
	(0.0062)
Constant	-0.044
	(0.094)
N	846
R^2	0.4144

The dependent variable is the percent of online job postings that are unfilled in an average month by year. The Occupational Education level is the mean education level from CPS 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 clustered by occupation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5
oDesk Probability of Filling a Job

Dependent Variable: Indicator for Filling Vacancy

	All Job Categories			2 Largest
Mean Wage in Job Category	-0.0683***			
	(0.00538)			
90th Percentile of Wages in Category		-0.105***		
		(0.00631)		
Admin Support (Baseline)				
Design and Multimedia			-0.101***	
			(0.00805)	
Networking and IS			-0.145***	
			(0.0116)	
Sales and Marketing			-0.127***	
			(0.00846)	
Software Development			-0.171***	
			(0.0132)	
Web Development			-0.100***	-0.104***
			(0.00675)	(0.00818)
Writing and Translation			-0.0804***	
			(0.00772)	
Firm Effects	Y	Y	Y	Y
Time Effects	Y	Y	Y	Y
N Job Openings	217,753	217,753	217,496	132,226
N Firms	60199	60199	60130	41402

Standard errors clustered by employer. Customer and Business Service estimates not displayed. 64 percent of employers post vacancies in different categories in the same month.

Table 6
CPS Mean Annual Wage Gap
$\widehat{W} - W$, for Workers Who are Over-qualified for Their Job

	Over-qualified Defined as Those With Education Greater Than 1 SD of Mean Occupational Education		Over-qualified Defined as Those With Education Greater Than 2 SD of Mean Occupational Education	
Variable	**Obs**	**Mean**	**Obs**	**Mean**
Earnings Level	853572	63306.59	853572	63306.59
Over-qualified Wage Gap[17]	**110297**	**5395.926 (136.70)**	**8193**	**17504.48 (499.75)**

[17] The wage gap for individuals who are over-qualified for their job and whose earnings are within 3 standard deviations of the mean earnings for their occupation.

Table 7
CPS Mean Annual Wage Gap
$\widehat{W} - W$ for Workers Who are Under-qualified for Their Job

	Under-qualified Defined as Those with Education Less Than 1 SD of Mean Occupational Education		Under-qualified Defined as Those with Education Less Than 2 SD of Mean Occupational Education	
Variable	**Obs**	**Mean**	**Obs**	**Mean**
Earnings	853572	63306.59	853572	63306.59
Under-qualified Wage Gap [18]	**119770**	**-15157.9 (95.14)**	**30227**	**-14735.2 (203.15)**

[18] The wage gap for individuals who are under-qualified for their job and whose earnings are within 3 standard deviations of the mean earnings for their occupation.

Table 8
PSID Average Hourly Log Wage Gap and Average Difference Between Actual Log Wage and Mean Occupation Log Wage for Workers In Job Spells for Which they are Over- and Under-Qualified for Their Job

	1 Work Spells Corresponding to a Worker Being in a Job for Which He is Under-Qualified	2 Work Spells Corresp0onding to a Worker Being in a Job for Which She is Over-Qualified
A: Predicted Log Wage - Actual Log Wage (Std Error)	.048 (.011) (0.011)	.06 (.011)
B: Actual Log Wage - Mean Log Wage in Inappropriate Occupation (Std Error)	-0.175 (0.010)(0.010	0.058 (0.010)(0.21)
Number of Observations	2629	2803

Table 9
Rising Mismatch Over Time

	(1) Mismatch Variance	(2) Mismatch Variance
Year	770.4***	-273.5
	(73.6)	(200.0)
Occupational Skill	0.86***	-34.5***
	(0.071)	(7.54)
Skill*Year		0.018***
		(0.0038)
Constant	-1543801.5***	541478.8
	(146782.0)	(399728.7)
N	12733	12733
R^2	0.5209	0.5415

Weighted by number of observations per occupation
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The sample size for the CPS data by Occupation is 39 years (1975-2013) times an average of 343occupational 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 to 373 in 1994.The dependent variable is the variance of income within occupational group by year.. Regression observations are weighted by number of observations per occupation. In some years, there is only 1 observation for some occupations; these observations are dropped.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix; Endogeneity of Wage Bid in oDesk Data

There may be sorting on the instrument, and this section details an econometric correction for that sorting, along with diagnostics for the first stage regression.

Because workers observe the distribution of bids with high frequency, changes in bids by other workers are likely to be observed and to affect individual decisions to use the platform. To capture this, other workers' average bids (aggregate) excluding own country are used as an instrument for selection. A monthly time fixed effect is then removed. This instrument picks up cross-country differences in the intensity of competition from other workers. The instrument is correlated with participation behavior on the platform as a whole, but other workers' bids are plausibly uncorrelated with the error in the individual participation decision after controlling for aggregate sources of time effects. With this additional instrument, an aggregate participation equation at the monthly level using a probit model is used to correct for aggregate sorting based on the instrument.[19] The inverse mills ratio estimated from this probit model is included when estimating the control function.

To test the strength of the instrument, the first stage regression is

$$log(w_{ij}) = a_1 + Z_{1jt}\gamma_1 + X_{jt}\gamma_2 + \rho\frac{\phi(Z_{2jt})}{1-\Phi(Z_{2jt})} + u_{jt}$$

where the instruments are the z-scores of the local to dollar exchange rates for worker j and the inverse mills ratio from a monthly participation probit that has workers' average bids (aggregate) excluding bids from the own country. Table X3 contains the results of the first stage regression and the probit model used to construct the inverse mills ratio. The instruments are estimated precisely and are quite strong.

[19] There is no attempt to correct for endogenous sorting to individual job openings. This sorting is not a problem for pairwise comparisons between workers, but it may be problematic in assessing an employers' comparison of a worker and the no-hire option.

Appendix Table 1
oDesk First Stage Regression of Log Hourly Bids and Participation on Exchange Rate Instruments

Dependent Variable	Log Hourly Wage Bid (Job Opening Level)				1+ application during the month	
Model	Linear Regression				Probit	
Sample	Data Entry	Data Entry	Web Programming	Web Programming	Data Entry	Web Programming
	(1)	(2)	(3)	(4)	(5)	(6)
Log Local Currency to Dollar (monthly; standardized)	-0.0524***	-0.0454***	-0.0168***	-0.0160***	0.0939***	0.051***
	(0.00523)	(0.00574)	(0.00156)	(0.00178)	(0.042)	(0.003)
Mean Bid from Workers in Other Countries (Aggregate)					-1.637***	-0.233***
					(0.0447)	(0.067)
Feedback Score (Out of 5)	0.0660***	0.155***	0.0658***	0.0759***	0.560***	0.523***
	(0.00419)	(0.00780)	(0.00120)	(0.00174)	(0.0027)	(0.002)
Agency Affiliate Indicator	0.0657***	0.0778***	0.0411***	0.0430***	0.100***	0.188***
	(0.00690)	(0.00704)	(0.00168)	(0.00171)	(0.0072)	(0.004)
Prior Work Experience	-0.0684***	0.00728	0.117***	0.128***	0.5837***	0.626***
	(0.00725)	(0.00940)	(0.00266)	(0.00333)	(0.0087)	(0.005)
Monthly Participation Inverse Mills Ratio		0.468***		0.151***		
		(0.0386)		(0.0230)		
Number of Observations	73,056	73,056	240,806	240,806	1,475,701	1,836,876
R-Squared	0.169	0.171	0.212	0.212	0.3068	0.3587
F Statistic on Excluded Instruments	100.4	62.74	115.9	81.04		

Notes: The sample is experienced employers who have hired 2 or more previous workers from any job category and have posted at least 2 previous jobs in the job category in question. Robust standard errors in parentheses. All models contain a fifth order polynomial in calendar time, fixed effects for 6 country groups, controls for English skills, and an indicator for having zero feedback. The last country group includes many countries with very small application shares. Models in columns 1 - 4 also include a piecewise-linear spline with 4 knots for the application number, an indicator for an employer-initiated application, and an indicator that the worker only applies to this job during the month. The Log Local Currency to Dollar exchange rate is calculated using monthly data and z-scores are used to make the measure comparable across countries. The inverse mills ratio in columns 2 and 4 is taken from columns 5 and 6. Other workers' average bids (aggregate) in the probit models in columns 5 and 6 are first calculated excluding own-country and then a monthly time fixed effect is removed. A separate interaction for workers in the United States and the basket of other currencies (not reported) is included because these workers do not have any variation in own exchange rates.