Wage Dispersion and Search Behavior: The Importance of Non-Wage Job Values *

Robert E. Hall
Hoover Institution and Department of Economics
Stanford University
National Bureau of Economic Research
rehall@stanford.edu

Andreas I. Mueller
Columbia Business School
National Bureau of Economic Research and IZA
amueller@columbia.edu

September 15, 2016

Abstract

We use a rich new body of data on the experiences of unemployed jobseekers to determine the sources of wage dispersion and to create a search model consistent with the acceptance decisions the jobseekers made. Heterogeneity in non-wage job values or amenities among jobseekers and jobs is a central feature of our model. From the data and the model, we identify the distributions of four key variables: offered wages, offered non-wage job values, the value of the jobseeker's non-work alternative, and the jobseeker's personal productivity. We find that, conditional on personal productivity, the standard deviation of offered log-wages is moderate, at 0.24, whereas the dispersion of the non-wage component of offered job values is substantially larger, at 0.34. The resulting dispersion of offered job values is 0.38. We also find high dispersion of personal productivity, at 0.43.

JEL J31, J32, J64

Contents

List of Tables

List of Figures

We take a close look at the economics of job search in the modern U.S. labor market, using a new body of data from a survey of unemployed jobseekers. We develop a model with rich heterogeneity among jobseekers in multiple dimensions—(1) unobserved personal productivity, (2) the value they achieve when not working, and (3) the random flow of offers of wages and non-wage job characteristics from potential employers. We build on a large body of prior research on job search. The new survey permits some important advances. We infer personal productivity from the offers that jobseekers receive and from their previously reported reservation wages. We infer the dispersion of the non-wage values of jobs that jobseekers consider in their offer-acceptance decisions. And we infer the extent that offered wages embody compensating wage differentials that offset non-wage amenities of offered jobs.

It is well known that fitting search models to data on wages and labor flows is a challenge. Wages have huge dispersion among workers with similar observed characteristics. Traditional search theory hypothesized that jobseekers would keep considering wage offers until they found one high in the upper tail of the distribution of available opportunities. But the high rate of jobseekers' acceptance of job offers suggests that searchers are leaving money on the table by taking jobs long before it is likely that they have adequately sampled the upper tail. The addition of on-the-job search to the model relieves some of this tension, because jobseekers departing unemployment may do so by taking an interim job and continue to search for the dream job in the upper tail while employed in the interim job. Even then, models that calibrate the offer distribution to the distribution of wages across workers find that the exit rate from unemployment to jobs makes sense only if workers find unemployment virtually intolerable, else they would be more picky in their acceptance decisions. These points reflect the suspicions that search economists have harbored for some time. We advance the resolution of the puzzle by finding that the dispersion of personal productivity is quite high. Our measure of dispersion is the standard deviation of the log of a variable. We find that the dispersion of productivity is 0.43. The implication is that the dispersion of offered wages for a particular jobseeker is only 0.24.

A second major issue is that jobseekers care about more than wages. We find that the dispersion of jobseekers' perceived non-wage values of jobs is higher than the dispersion of their offered wages. This finding cuts in two directions. To the extent that jobs with high non-wage values pay correspondingly low wages (so wages embody compensating differences), the finding helps understand search behavior by showing that the dispersion of the total job value is lower than the dispersion of wages. If the wage does not fully offset the non-wage value, the dispersion of the total job value will be closer to the dispersion of wages. If the dispersion of the non-wage value is substantial and its correlation with the wage is less negative than minus one, it is entirely possible that the dispersion of total job values exceeds

the dispersion of the wage. In that case, bringing non-wage values into the picture further complicates the interpretation of search behavior.

Our results fall into the more complicated territory. We find that the dispersion of non-wage job value is 0.34, well above the dispersion of the offered wage of 0.24. Although the offered wage is negatively correlated with the non-wage value, suggesting some role for compensating wage differentials, the sheer size of the dispersion of the non-wage value makes it unlikely that the dispersion of the job value could be less than the dispersion of the wage. Based on our estimate of the extent of the compensating differential, we find that the dispersion of the job value is 0.38, well above the dispersion of the offered wage of 0.24. Our estimates of the dispersions have small standard errors, but our estimate of the compensating differential has a fair amount of sampling variation.

The survey data permit more refined measures of dispersion than do the data sources in earlier work. The data include prior wages collected from administrative sources and survey responses about reservation wages each week during a spell of unemployment, and the wages of job offers and of newly accepted jobs. Although earlier surveys have collected cross-section data on reservation wages, the survey is the first, as far as we know, that collects panel data on reservation wages. It is also the first U.S. source of data on reservation wages to match survey data and administrative data, we believe. The survey data, together with a reasonable set of assumptions, permit a solution to a problem that has significantly impeded research on labor search behavior. The problem is the lack of information on individual wages relative to personal productivity. Conditional on measures of personal characteristics available to the econometrician, wages have high dispersion. The dispersion implied by our work is quite a bit smaller, because we infer that the dispersion of personal productivity is 0.43. Our identifying assumption is that the wage-related variables measured in the survey are proportional to personal productivity. Under the proportionality assumption for the reservation wage, both the offered wage and the wage in the prior job are proportional to personal productivity. The covariance of their logs reveals the dispersion of personal productivity. Then the difference between the log variances of the offered wage and the prior wage is the log variance of the offers facing a jobseeker with a given level of productivity.

The survey allows the measurement of the non-wage job value in a new way, from the finding that many jobseekers accept jobs that pay less than the jobseekers' previously stated reservation wages, and a fair number reject jobs that pay more. From the acceptance data, we infer that the dispersion of the non-wage value is 0.34. Direct questions in the survey reveal that two-thirds of rejections of job offers are motivated by reasons unrelated to the offered wage. We use the observed relation between the acceptance probability and the

7

difference between the offered and reservation wages to infer the correlation between the wage and non-wage value of job offers as well as the overall dispersion of the non-wage value.

In our model, as in most models in the labor search literature, jobseekers assign a flow value to the time they spend not working. This non-work value reflects the value of alternative non-market activity made possible by not working, net of the disamenity of job search. We use it in dollars without taking logs, because we consider it entirely possible that it is negative for some jobseekers. In the simplest model, jobseekers accept any job offered with a job value (sum of log wage and log non-wage value) at least as high as the the flow value of non-work. In models that treat acceptance as the exercise of a real option, the non-wage value is influential on the reservation value, but the reservation value exceeds the flow value because of the option value of remaining an unemployed jobseeker. In our model, the offer flow to employed jobseekers is half the flow to unemployed jobseekers, so there is an option value. We find that the mean of the non-work value is \$2.41 per hour and the standard standard deviation is \$3.36.

The paper and accompanying appendixes contain numerous investigations of the robustness of our results. We believe that our main conclusions about the relative contributions of personal productivity, wages, and non-wage job values are quite robust. Our conclusions about the extent that compensating variations in wages offset the values of non-wage job characteristics is limited by moderately high sampling variation, but we believe that our conclusion is robust that the offset is well under a hundred percent.

1 Related Literature

Alan Krueger and Andreas Mueller (KM) carried out the survey that underlies this paper—see Krueger and Mueller (2011) and Krueger and Mueller (2016).

The challenge of reconciling the wide dispersion of offered wages to the limited number of job offers considered by most jobseekers came into sharp focus in an influential article, Hornstein, Krusell and Violante (2011) (HKV). HKV, Section II, discuss the challenges in detail. They note that most empirical search models that appear to rationalize observed unemployment-to-employment flows imply an implausibly low flow value of unemployment. It is frequently negative. These models generally infer the value of job search from estimates of the dispersion of wage offers derived from cross-sectional data, where dispersion is high. Sampling from that distribution is highly valuable activity, which implies that people must truly hate unemployment to take the first job that comes along as frequently as they do in practice. HKV has an extensive discussion of the literature on wage dispersion, with many cites, notably Mortensen (2003), Rogerson, Shimer and Wright (2005), Bontemps, Robin

and Berg (2000), Postel-Vinay and Robin (2002), Jolivet, Postel-Vinay and Robin (2006), and Jolivet (2009).

Abowd, Kramarz and Margolis (1999) (AKM) introduced the use of matched employee-employer data to study dispersion. In an equation with the log of the wage of a worker as the left-hand variable, they estimated fixed effects for workers and for firms. A reasonably consistent finding in the resulting line of research has been that the firm effects account for a little over 20 percent of the dispersion of the log wage. Although non-wage job values may be one of the determinants of the firm effect, rents from search frictions or other sources may be another, so the dispersion of the firm effects cannot be taken as a measure of the dispersion of non-wage component of job values. Further, the dispersion of worker effects includes any persistent tendency for a worker to pick jobs with high non-wage values, and presumably somewhat lower wages. Thus AKM does not provide a direct measure of the dispersion of non-wage job values. Rather, the line of research it inspired is an advance in the topics of how much wage dispersion arises from employers, with a full adjustment for worker heterogeneity, and how much from workers, with a full adjustment for employer heterogeneity.

Our use of data on rejection of offers with wages above the previously measured reservation wage and acceptance of those paying less than the reservation wage to infer the role of non-wage job values is a cousin of research that infers an improvement in the non-wage job value when a worker moves voluntarily to a lower-wage job from a higher-wage one. The papers in this literature closest to ours are Becker (2011) and Sullivan and To (2014), who infer the dispersion of non-wage amenities from the fraction of job-to-job transitions that result in a wage decrease. Both of these papers assume that wage and non-wage values are independent of each other, which precludes investigation of an important strand of the wage-dispersion literature, compensating variation in wages. Their estimates of the dispersion of the non-wage value are similar to ours.

Sorkin (2015) is an ambitious application of the idea that voluntary job-to-job transitions reveal information about non-wage values. Sorkin's extraction of information about non-wage job values has three phases. The first considers only job values. He uses a gigantic longitudinal body of data on the identity of the employers of many millions of workers. Flows of workers making voluntary changes to work for other firms identify a value for each firm. He uses a relative of the Google page-rank algorithm to solve the system of 1.5 million simultaneous equations defining the firm values. The second phase uses an AKM equation to assign a firm effect to wages. The firm effect accounts for 22 or 23 percent of the variance of earnings of workers. The third phase involves combining the values of firms with their wage effects. Amenities are a residual in an equation with value on the left and earnings on

9

the right. It would be interesting to divide total amenities into a compensating difference (which he associates with Sherwin Rosen) and a rent or other component not correlated -1 with earnings (associated with Dale Mortensen). But, while the variance of the Rosen component is identified (by its known relation to the wage), the variance of the Mortensen component is not, because its correlations with the Rosen component and the AKM firm effect are unknown. Thus Sorkin does not answer the question considered in this paper, of the dispersion of the non-wage value irrespective of its accompaniment by a compensating wage difference. His contribution is to show that the dispersion of non-wage job values that *are* accompanied by offsetting wage differences is 15 percent of the total dispersion of wages. The reason Sorkin cannot identify the total residual in his equation with his inferred total job value as left-hand variable is that he does not know the value of the multiplier ω relating that value to the AKM wage value. He cannot recover that parameter by regression or any other statistical technique because the right-hand variable, the wage value, is correlated with the disturbance in the equation. That disturbance is interpreted as the entire non-wage job value. Our approach differs from Sorkin's in a key respect. Because we use data on reservation wages, stated in dollars, our basic relation does not have an unknown slope parameter, and we can obtain values of the entire residual by subtraction.

Jarosch (2015) builds a model in which job security is a non-wage job value. The frictional Mortensen component of the wage distribution is substantial. Workers suffering involuntary job loss face large and persistent earnings losses, consistent with evidence about displaced workers in U.S. and German data. The paper has a thorough treatment of wage determination with two-dimensional job values, a topic we sidestep by an assumption that employers post wages and non-wage job characteristics.

Hagedorn and Manovskii (2010), Low, Meghir and Pistaferri (2010), and Tjaden and Wellschmied (2014) have estimated the extent of wage dispersion arising through search frictions—the Mortensen component, in Sorkin's lingo. These papers infer the extent of wage dispersion arising from differences in match quality from the higher volatility of wage growth of those who switch jobs compared to those who stay on their jobs. With this approach, estimates of wage dispersion depend critically on how the process of on-the-job search is modeled. If the efficiency of on-the-job search is high, workers move up the job ladder relatively fast, and most job-to-job transitions are associated with small wage gains as workers continue to search for new jobs even when they are far up on the ladder. This process implies that, for a given observed variance of wage changes, the inferred dispersion in offered wages is increasing in the search efficiency of on-the-job search—see Tjaden and Wellschmied. We estimate the dispersion in wages arising from search frictions with a different identification strategy from these papers. Our estimates of the dispersion in wage

offers are closest to those of Low and co-authors, who find a standard deviation of match-specific wage shocks of 0.23, but are substantially larger than the estimates in Tjaden and Wellschmied and Hagedorn and Manovskii.

An important challenge for many of the papers discussed in this section is to distinguish job-to-job transitions that are value increasing—movements up the job ladder—from transitions that arise from layoffs or other involuntary separations. The conclusions emerging from this literature depend on whether one interprets wage decreases as compensated for by higher non-wage characteristics or as falling off the job ladder. We use direct information on job acceptance decisions of unemployed workers, so the main parameters in our approach do not rest on properties about the process of on-the-job search, notably the relative probabilities of receiving offers while working and while unemployed. But our estimate of the flow value of non-work does rest on those properties, as we discuss later in the paper.

2 Model

Our model focuses on the behavior of a worker in an environment with the following key variables:

- Randomly arriving job offers, each with log flow value $\hat{v}$, which is the sum of a log wage $\hat{y}$ and a non-wage log value n

- Personal productivity, x, in logs

- A personal non-work value $\hat{h}$, the opportunity cost of employment, not in logs

We will explain the role of the hat, $\hat{\ }$, shortly. Jobseekers, who may be employed or unemployed, form a reservation job value $\hat{r}_v$ and accept the first job offer with a value at least as high as $\hat{r}_v$. Offers arrive at rate λ_u for the unemployed and λ_e for the employed jobseekers. This model is an extension of the job-ladder model of a large recent literature. The model generates equilibrium distributions of wages w and non-wage values n among workers. Dispersion of those variables across workers arises from (1) dispersion of productivity, (2) dispersion of non-work value, (3) dispersion in the position of workers on the job ladder, arising from their histories of random job offers and separations.

The KM survey asks a respondent for her reservation *wage* $\hat{r}$, not her reservation job value. It also asks if her rejection of a job offer was for non-wage reasons. Our model makes assumptions that enable identification of its parameters by making use of these responses.

2.1 Job offers and the interpretation of the distribution of job value

We use the term *offer* to describe a jobseeker's encounter with a definite opportunity to take a job. Nothing in this paper requires that employers make firm job offers and that jobseekers then make up-or-down decisions. The jobseeker's acceptance problem, upon finding a job opportunity, is the same whether the employer is making a single firm offer, or they engage in full-information alternating-offer bargaining. In the latter case, the jobseeker will participate in the bargaining process only if she anticipates that the ultimate job value will meet the reservation value. That said, the survey included a question about the nature of the job offer, and in the majority of cases, the employer did make a firm offer. The distribution of offers that we consider is the actual probability distribution of the job value $\hat{v}$ of a definite employment opportunity. It is specific to a jobseeker and reflects all of the selection of jobs that the jobseeker investigates and all the consideration of a jobseeker's qualification by the employer prior to the jobseeker understanding that the opportunity is definite. We do not model the distribution of offered job values as the censored version of an underlying general distribution of job values. As we describe in a later section, if a respondent receives more than one offer in a week, the survey gathers information about the best offer. As a result, our distribution of job values reflects the improvement that is available from running an auction, in effect, when a jobseeker can choose among competing offers.

We observe the wage component of an offer, $\hat{y}$, and infer a non-wage log value, n, as a residual, defined by other observable variables. Our conclusion about the importance of the non-wage value derives from the notion that the offered wage is a good indicator of the overall wage value of a job. If all jobs paid the same wage (the offered wage) for a week, and then adjusted the wage to reflect the new worker's skills, the non-wage value would be the whole story. There would be no dispersion in the distribution of offered wages. In support of the hypothesis that the offered wage is a good indicator of earnings in the future, Kudlyak (2014) shows that initial wages are highly persistent within the first years of a job, up to at least seven years, even in the face of substantial changes in the wages of more recently hired workers.

2.2 The key role of the acceptance function

The *acceptance function* is a central empirical object in our model. It is the probability that a job is accepted, as a function of $d = \hat{y} - \hat{r}$, the difference between the offered wage and the reported reservation wage. To demonstrate the value of the acceptance function, we consider a special case here. Suppose that jobseekers report their reservation job values

as their reservation wages—that is, when asked for a reservation wage, they give the wage that would be just enough to be acceptable for a job with a log non-wage value of zero. And suppose that $\hat{y}$ and n are uncorrelated. The acceptance function then satisfies:

$$A(d) = \text{Prob}[\hat{v} \geq \hat{r}_v] = \text{Prob}[\hat{y} + n \geq \hat{r}_v] = \text{Prob}[\hat{y} - \hat{r} \geq -n] = \text{Prob}[n \geq -d] \quad (1)$$

Thus we can write

$$A(d) = 1 - F_n(-d), \quad (2)$$

and we can calculate F_n directly from the acceptance function:

$$F_n(n) = 1 - A(-n). \quad (3)$$

We conclude that, in this special case, the acceptance function reveals the distribution of non-wage job values directly. If all jobs had the same non-wage value, the reservation wage would be a perfect predictor of acceptance. The incidence of acceptances of jobs whose wages are below the reservation wage reveals the frequency of high non-wage values and the frequency of rejection of jobs whose wages exceed the reservation wage reveals the frequency of low non-wage values.

2.3   Assumptions

We make three general assumptions to support identification of the model's parameters:

Assumption 1, *Observable and private values*: Jobseekers and prospective employers know the jobseeker's personal productivity x at the time they meet each other, but the non-work value $\hat{h}$ and the reservation job value $\hat{r}_v$ are private to the jobseeker.

Assumption 2, *Proportionality-to-productivity*: The distributions of $y = \hat{y} - x$, $v = \hat{v} - x$, $n = \hat{n} - x$, $r = \hat{r} - x$ and $h = \hat{h}/\exp(x)$ in the population with personal log productivity x are the same as the distributions of $\hat{y}$, $\hat{v}$, $\hat{n}$, $\hat{r}$ and $\hat{h}$ in the sub-population with $x = 0$.

Assumption 3, *Joint distribution*: Let

$$\eta = n - \kappa(y - \mu_y), \quad (4)$$

n adjusted for its correlation with y, where μ_y is the mean of y. The variables y, η, r, and x are jointly normally and independently distributed.

We make additional assumptions that support identification from specific features of the KM survey:

Assumption 4, *Reference non-wage value.* The survey asks for a reservation *wage*, not the *reservation job value* of the model. We assume that the reported reservation wage r is the reservation wage applicable to an offer with a zero value of the non-wage value, n, so $r_v = r$. The mean, μ_n, of the non-wage value is not necessarily zero.

Assumption 5, *Measurement errors.* We assume that the observed values of the offered wage and reservation wage contain measurement errors:

$$\tilde{y} = y + x + \epsilon_y \tag{5}$$

$$\tilde{r} = r + x + \epsilon_r \tag{6}$$

The measurement errors are normally distributed with mean zero and are independent of each other and the other variables of the model.

Assumption 6, *Preponderant reason for rejection*: Respondents report that they rejected a job offer for a non-wage reason if the deviation from the mean is more negative for the non-wage value than for the wage value: $n - \mu_n < y - \mu_y$.

2.4   Discussion of the assumptions

2.4.1   Observable and private values

With respect to personal productivity, mutual observability seems the natural starting point for modeling employment, though we recognize that the information is not perfect on either side. Both parties have strong incentives to track down information about the job and the jobseeker's fit to the job. Random future changes in productivity are consistent with the model.

We also assume that h and r are not observed by the employer. Unlike productivity, the jobseeker's work history is not very informative about flow values of non-work and of unemployment. Even if past acceptance decisions were observed, for most jobseekers there are too few data points to infer their reservation wages with useful precision. In the final section of the paper, we perform a robustness check where flow values are fully observed and find that it does not materially change the results.

2.4.2   Proportionality-to-productivity

The most controversial aspect of this hypothesis is that non-market productivity is higher by the entire amount of market productivity in the population with higher values of x. Low-x populations are not systematically more choosy about taking jobs than are high-x populations. While this assumption obviously fails if applied across the entire population

including those out of the labor force, it appears reasonable in a sample of workers eligible for unemployment compensation. Moreover, we find that the average acceptance rates do not differ systematically across different levels of educational attainment. Unemployment rates decline with productivity, but this is largely because separation rates decline, not because of heterogeneity in acceptance. Toward the end of the paper, we report robustness checks that suggest that the non-proportionality in our sample is unimportant.

2.4.3 Joint distribution

The principle of compensating wage differentials suggests that the correlation between wage offers y and non-wage values n should be negative—employers offer lower wages for jobs with favorable non-wage values. The correlation is not perfect, however, because there is a personal dimension to the non-wage value that the firm may ignore, under a posted-wage policy, or respond to only partially, in a bargained-wage policy. For example, commuting cost varies across individual workers. For this reason, we assume that the non-wage value n comprises (1) a component η that is uncorrelated with the other fundamentals and (2) a component that is the negative of a fraction κ of the offered wage minus its mean:

$$n = \eta - \kappa(y - \mu_y). \tag{7}$$

With respect to the independence or non-correlation assumption, the variables y, η, and r are subject to the proportionality-to-productivity assumption and thus are uncorrelated with x by definition. The correlation of η and y is zero by construction. The support for the assumption that the correlation of y and r is zero involves issues that we take up in the next section on the determination of the reservation wage and again in our discussion of the results.

2.4.4 Reference non-wage value

An unemployed jobseeker decides about accepting a job offer by comparing the *job value* $v = y + n$ to a reservation value, r_v. The KM survey asks about a reservation wage, not a reservation job value. We take the reference non-wage value to be zero. This choice is only a normalization, because we estimate the mean of the distribution of n, μ_n. Acceptance choices conditional on reservation wages are the only evidence we have about non-wage values, so we cannot distinguish between the mean of non-wage values and the reference level that respondents use in answering the question about the reservation wage. The fact that it is more common for an unemployed jobseeker to accept an offer below the reservation wage than reject one above the reservation wage is equally well explained by two views: (1) the distribution of non-wage values has a positive mean, or (2) the respondents use a

high reservation wage on account of answering the question with respect to a hypothetical offer with a job value well below average. Our assumption that the reservation job value is observed as the reservation wage corresponds to the first view.

2.4.5 Measurement errors

Many jobseekers accept wage offers below the reservation wage and a smaller fraction reject offers above the reservation wage. The main way we account for the first group is that the distribution of non-wage values has a positive mean but respondents use zero job value for the hypothetical job that lies behind the survey's question about the reservation wage. Our acceptance model accounts for the acceptances and rejections that appear contrary to the reservation wage in two ways. First, we invoke a non-wage value that is imperfectly correlated with the offered wage. Second, we attribute measurement errors to the reported values of the offered wage and the reservation wage. We assume that the observed values are:

$$\tilde{y} = y + x + \epsilon_y \tag{8}$$
$$\tilde{r} = r + x + \epsilon_r \tag{9}$$

where the measurement errors $\epsilon_y \sim N(0, \sigma_{\epsilon_y})$ and $\epsilon_r \sim N(0, \sigma_{\epsilon_r})$, and are independent.

Let $d = \tilde{y} - \tilde{r}$ be the difference between the observed offered wage and the reservation wage. Also let $m = v - r_v$ (recall that $v = y + n$, the job value). As above, we write the acceptance probability A as a function of d:

$$A(d) = \text{Prob}[m \geq 0|d] = 1 - \text{Prob}[0 \geq m|d] = 1 - F_m(0|d), \tag{10}$$

which differs from equation (1) because of the presence of measurement error and non-zero correlation between y and n.

2.4.6 Preponderant reason for rejection

The shape of the acceptance function does not separately identify the dispersion of the idiosyncratic part of non-wage values, σ_η, and the compensating differential parameter κ, as higher values of either parameter imply a flatter acceptance function. The survey includes a question for respondents who rejected a job offer if the rejection was because of a non-wage reason. We assume that respondents report that they rejected a job offer for a non-wage reason if the deviation from the mean is more negative for the non-wage value than for the wage value: $n - \mu_n < y - \mu_y$.

16

Let $p = (\eta - \mu_\eta) - (y - \mu_y)(1 + \kappa)$. The fraction of rejections for non-wage reasons for a person with reservation wage r, denoted J_r, is:

$$
\begin{aligned}
J_r &= P(\text{non-wage preponderates} \mid \text{offer rejected}) \\
&= P(n - \mu_n < y - \mu_y \mid v < r) \\
&= \frac{P(p < 0 \text{ and } v < r)}{P(v < r)} \\
&= \frac{\int_{-\infty}^{v=r} P(p < 0|v) dF_v(v)}{P(v < r)} \\
&= \int_{-\infty}^{v=r} \frac{F_p(0|v)}{F_v(r)} dF_v(v),
\end{aligned}
$$

and integrating over the distribution of r, we get:

$$
J = \int_{-\infty}^{\infty} \int_{-\infty}^{v=r} \frac{F_p(0|v)}{F_v(r)} dF_v(v) dF_r(r). \tag{11}
$$

3   Determination of the Reservation Job Value

In this section, we consider how a jobseeker sets her reservation job value while unemployed or employed. If search on the job is less effective than while unemployed, the decision to take a job offer while unemployed includes a real-option element because it involves a sacrifice of the superior flow of job offers. Absent that option value, the reservation job value is simply the opportunity cost, so $r_v = \log h$ if $h > 0$ and acceptance is automatic if $h \leq 0$.

Under the assumption of proportionality, the value functions of employed workers are proportional to personal productivity. Our next step is to derive the Bellman equations and associated reservation job value for an individual with $x = 0$. As before, those for individuals with other values of x scale in proportion. The Bellman equation for an unemployed person with non-work value h and offer rate λ_u adjusts the reservation job value r_v to include the lost option value associated with accepting a job offer while unemployed:

$$
U(h) = h + \frac{1}{1+\rho} \max_{r_v} \left((1-s)\lambda_u \int_{r_v} W(h, \tilde{v}) dF_v(\tilde{v}) + (1 - (1-s)\lambda_u(1 - F_v(r_v))) U(h) \right). \tag{12}
$$

On the left is the value of being unemployed, $U(h)$. On the right, the individual receives the non-work flow value h and finds the best reservation job value to maximize the discounted asset value arising from the optimal choice of the reservation value, r_v. A higher r_v raises the capital gain upon re-employment but lowers the probability of receiving it.

The Bellman equation for an employee with non-work value h and offer rate λ_e is

$$W(h,v) = e^v + \frac{1}{1+\rho}\left\{(1-s)\left[\int_v \lambda_e W(h,\tilde{v})dF_v(\tilde{v}) + (1 - \lambda_e + \lambda_e F_v(v))W(h,v)\right] + sU(h)\right\}$$
(13)

The worker automatically accepts any job with a value greater than the current job value, v, because there is no loss of option value. There is a flow value from the probability of finding a better job with capital gain $W(h,\tilde{v}) - W(h,v)$. There is also a flow probability s, the separation rate, of suffering the capital loss $W(h,v) - U(h)$.

If employed jobseeking is just as effective as unemployed jobseeking, the reservation job value for the unemployed is the non-work value h. If there is an option value, it remains the case that unemployed jobseekers with higher non-work values have higher reservation job values. Our assumption of zero correlation of the reservation value and the offered value will fail if the jobseeker knows something about the possible job offer before contracting an employer, because the jobseeker will contact only the more promising employers. Choosier jobseekers with higher non-work values will get better job offers, though less often than other jobseekers. The correlation between the reservation value and the offered value will be positive, not zero. The issue of how much a jobseeker knows about job prospects is important in search theory. Models of the search process range on a spectrum from directed search to random search. With strictly directed search, the jobseeker knows the terms of a job prior to contacting an employer. The jobseeker visits only one employer and automatically accepts the job. With strictly random search, the jobseeker meets employers at random and lacks any ability to target a favorable employer. In reality, the jobseeking environment is somewhere in between. In Appendix D.1, we consider a model of partially directed search as an alternative to our main specification of random search.

4 The KM Survey

The KM survey enrolled roughly 6,000 jobseekers in New Jersey who were receiving unemployment insurance benefits in September 2009. The survey collected weekly data from them for several months up to April 2010. The sampling frame of the survey was based on a stratified random sample of *all* unemployment insurance recipients in NJ. The survey was conducted online and was administered by the Cornell Survey Research Institute in collaboration with the Princeton Survey Research Center. Individuals were initially invited to participate in the survey for 12 consecutive weeks, but the survey was extended for an additional 12 weeks for the very long-term unemployed—those with a duration of unemployment of 60 weeks or more at the start of the survey.

The KM survey is a novel data source on unemployed workers' search behavior and outcomes. It is unique in several dimensions: First, the survey provides a unique combination of information on reservation wages, job offers, and job acceptance decisions. Second, the data were collected for a large cross-section of *unemployed* workers, representative of the population of unemployment insurance recipients in New Jersey. Data sets that have some of the same information usually have substantially smaller samples and often are focused on particular segments of the population, such as the youth in the NLYS79. Third, the data have a weekly panel dimension, which is unprecedented. This feature is important for the research in this paper, because it allows us to relate the acceptance decisions to the reservation wage prior to the receipt of the job offer. Finally, the survey data can be matched to administrative records for the respondents, notably their wages on the jobs they held just prior to becoming unemployed.

The overall response rate in the survey was 9.7 percent and respondents completed on average about 5 interviews over the first 12 weeks of the survey. While the relatively low response rate may be a concern, there are several reasons to believe that the non-response should not lead to a major bias in the results in this paper: First, the public-use survey data include survey weights, which adjust both for sampling probability and non-response. Krueger and Mueller (2011) provide a detailed analysis of non-response and show that respondents were more likely to be female, white, and older and have a college degree, compared to the sample frame. After adjusting for survey weights, however, the characteristics of the sample of respondents closely matches the characteristics of the sample frame. Second, Krueger and Mueller (2011) provide additional evidence based on updated UI records that the weekly hazards of UI exit do not differ significantly between respondents and non-respondents during and after the survey. This finding suggests that search behavior of respondents and non-respondents did not differ markedly over the period of the survey. There is a significant difference in the first week of the unemployment spell, which is probably because some unemployed found a job by the time they were invited to the survey two weeks after the date when they were sampled. Finally, Krueger and Mueller (2016) provide evidence that the ratio of weekly re-employment wages to weekly prior wages in NJ administrative wage data were very similar between respondents and non-respondents. This finding is particularly relevant for this paper, as it shows that the unemployed workers in our sample did not differ significantly from non-respondents in their accepted wage distributions, and thus are unlikely to differ in their reservation wage choices and job offer distributions.

We follow Krueger and Mueller (2011) by restricting the sample to survey participants of ages 20 to 65, and exclude outlier observations of reservation and offered wages. Outliers are defined as observations where the wage expressed in weekly terms exceeded \$8000 or

was below \$100 or where the wage in hourly terms was greater than \$100 or below \$5. In addition, following Feldstein and Poterba (1984), we trimmed reservation wages if the ratio of the reservation wage over the prior wage exceeded 3 or was below one third. All major results in the paper are robust to not trimming outlier observations of reservation wages and offered wages (see Appendix Table 7).

4.1 Job offers

The KM survey asked respondents each week: "In the last 7 days, did you receive any job offers? If yes, how many?" The respondents in our sample received a total of 2,174 job offers in 37,609 reported weeks of job search. The ratio of the two, 0.058, is a reasonable estimate of the overall weekly rate of receipt of job offers.

For respondents who indicated that they received at least one job offer, the KM survey asked respondents: "What was the wage or salary offered (before deductions)? Is that per year, per month, bi-weekly, weekly or per hour?" In cases where respondents reported that they received more than one offer in a given week, the survey asked the offered wage only for the best offer. Among the individuals who reported at least one job offer, 86.3 percent reported that they received one offer in the last 7 days, 8.6 percent reported receiving two offers in the last 7 days, 2.4 percent received 3 offers and the remaining 2.7 percent received between 4 and 10 offers in the last 7 days.

Figure 1 reports the kernel density of the hourly offered wage for our sample of 1,153 job offers. In cases, where the wage was not reported on an hourly basis, to measure of the hourly offered wage, we divided the salary by the number of weeks in the reference period (if yearly, 52, and if monthly, 4.33) times the hours on the job. The sample is restricted to cases where details of the offer (including the wage) and a reservation wage from a previous interview were available. We use the same sample below when we compute the acceptance frequency conditional on the difference between the log of the offered wage and the log of the reservation wage from a previous interview.

The model interprets this distribution as the mixture of the distribution of wage offers for a worker with standardized personal productivity and the distribution of productivity across workers—by mixture, we mean the weighted average of the offer distribution for given productivity, with the weights taken as the distribution of productivity.

4.2 Reservation wage

Each week, the respondents in the KM survey answered a question about their reservation wages: "Suppose someone offered you a job today. What is the lowest wage or salary you
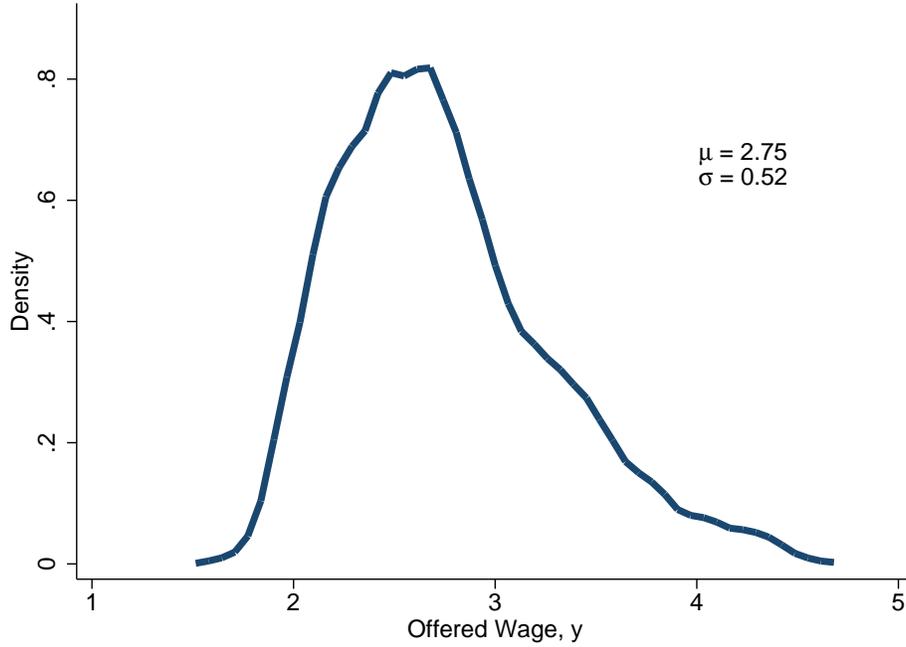
Figure 1: Kernel Density of the Log Hourly Offered Wage, y

would accept (before deductions) for the type of work you are looking for?" We only use the first reservation wage observation available for each person in the survey so that the sample is representative of the cross-section of unemployed workers. We apply the same sample restrictions as Krueger and Mueller (2011)—we exclude survey participants who reported working in the last seven days or already accepted a job offer at the time of the interview. Figure 2 shows the kernel density of the hourly reservation wage for our sample of 4,138 unemployed workers. We calculated the reservation in the same way as we calculated the offered wage. Not all unemployed workers in our sample received job offers during the survey period, but the mean and the standard deviations of the reservation wage are nearly identical for the full sample and the sample restricted to those who received job offers. The mean of the log reservation wage is 2.83 in the restricted sample compared to 2.82 in the full sample, and the standard deviation is 0.47 in both samples. In the estimation of our model, we rely on the restricted sample to estimate the acceptance function and the covariance of the wage offer and the reservation wage.

The model infers the value of the non-work option from the reservation wage of a job-seeker. The survey reveals the distribution of the reservation wage among all respondents. The model interprets this distribution as the mixture of the distribution of the reservation wage for a worker with standardized personal productivity and the distribution of productivity across workers.
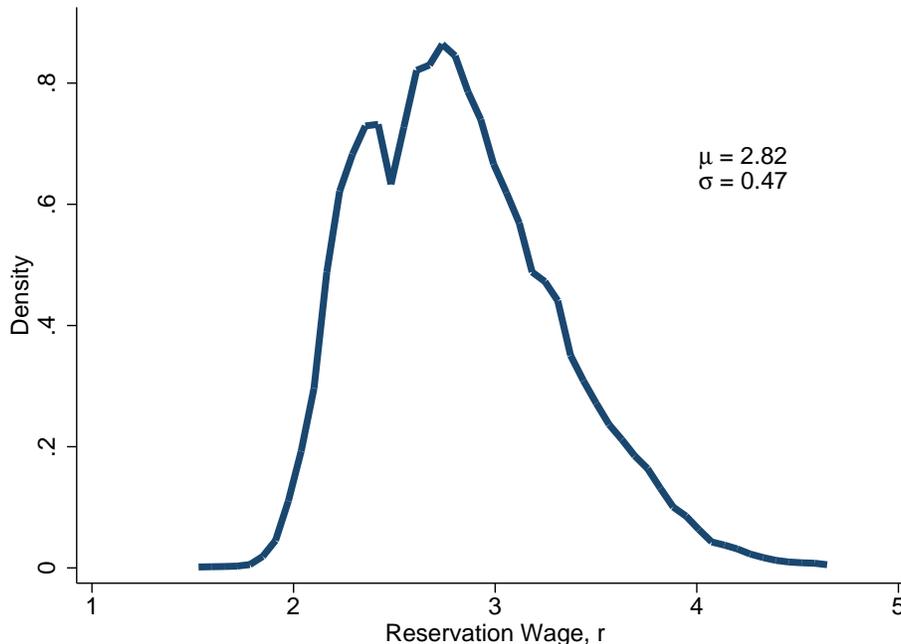
Figure 2: Kernel Density of the Log Hourly Reservation Wage, r

4.3 Acceptance

Many respondents accept job offers that pay less than the respondent's previously reported reservation wage. Some do the reverse, rejecting an offer that pays more than the reservation wage. Our model posits that jobs have non-wage values, to explain why the offered wage does not control the acceptance decision—jobseekers accept jobs paying less than the reservation wage because these jobs have positive non-wage values that offset the low wage. The model accounts for the bias toward acceptance by treating the reported reservation wage as referring to a job with below-normal non-wage value.

We study the acceptance probability as a function of the difference between the log of the offered wage and the log of the reservation wage. To avoid possible bias from cognitive dissonance among the respondents, we exploit the longitudinal structure of the survey and use the reservation wage value reported in the week *prior* to the receipt of a job offer. Krueger and Mueller (2016) give a detailed analysis of the acceptance frequency in the survey. The job acceptance frequency rises with $d = y - r$. The average frequency of job acceptance in our sample is 71.9 percent. In 20.9 percent of the cases, respondents indicated that they had not yet decided whether to accept the job offer or not.

To deal with the problem of missing data for acceptance of some job offers, we make use of administrative data on exit from unemployment insurance. UI exit is a potentially
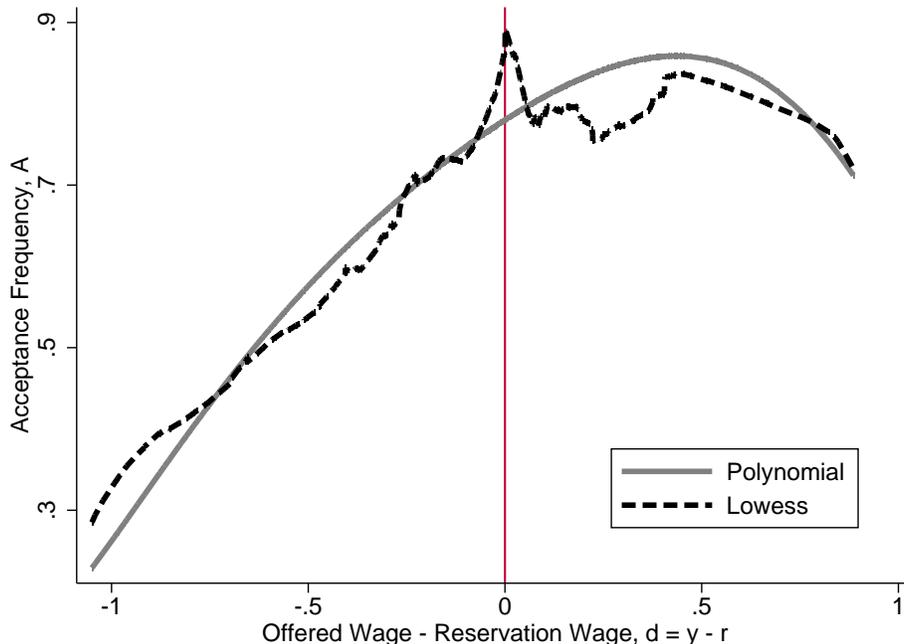
Figure 3: Smoothed Acceptance Frequency, A, as a Function of the Difference between the Log Hourly Offered Wage and the Log Hourly Reservation Wage, $d = y - r$

useful but imperfect indicator of acceptance, for four reasons: (1) A delay occurs between job acceptance and UI exit. (2) An exit from the UI system may relate to a different offer from the one reported in the survey. (3) UI exit data are censored at the point of UI exhaustion, as the data do not track recipients after they exhaust benefits. (4) An unemployed worker may perform limited part-time work while receiving benefits and thus acceptances of such offers will not be reflected in an exit from the UI system. Krueger and Mueller (2016) show that the rate of UI exit for those who were undecided was almost exactly halfway between the rate of UI exit for those who accepted the offer and the rate of UI exit for those who rejected the offer. We believe that this estimate is the best available. Notwithstanding the imperfect relation between exits and acceptances of offers, we believe it is the best way to handle the problem of missing data, so we create an indicator variable A that takes on the value zero for a rejected offer, 0.5 for an offer for which the respondent was undecided, and 1 for an accepted offer.

Figure 3 shows the acceptance frequency smoothed in two ways: (1) as the fitted values from a regression of A on a 6th-order polynomial in $y - r$ and (2) as the fitted values from a locally weighted regression (LOWESS) with bandwidth 0.3. The figure runs from first-percentile value of d to the 99th percentile value. Values outside that range are inherently unreliable for any smoothing method.

23

The survey also asked a question about reasons for rejecting a job offer: 32.3 percent indicated that they rejected because of "inadequate pay/benefits" and the remaining 67.7 percent indicated another reason for rejecting such as unsuitable working conditions, insufficient hours/too many hours, transportation issues, insufficient use of skills/experience. Consistent with our principle that the offer distribution includes the advantage of multiple competing offers, we exclude from the sample the 5.0 percent of offers that respondents rejected because they accepted another job offer. Unfortunately, the survey did not distinguish between inadequate pay and inadequate benefits, but in response to a similar question in the National Longitudinal Survey of Youth (NLSY) in 1986-87, 36.8 percent of respondents mentioned "inadequate pay" as the reason for rejecting a job offer, indicating that the inadequate pay is the most common reason for rejecting the job offer, not inadequate benefits. Moreover, as reported in Krueger and Mueller (2016), 40 percent of offers below the reservation wage were rejected for inadequate pay or benefits, whereas only 1 percent of offers above the reservation wage was rejected for the same reason. This evidence suggests that either benefits are not an important factor in the acceptance-rejection decision or that benefits are quite positively correlated with the offered wage, as otherwise we would expect at least some rejections for the reason of inadequate benefits for job offers with wages above the reservation wage. As explained further later in the paper, our model allows for correlation between wage offers and non-wage amenities.

In our approach to estimation, the shape of the acceptance function and the fraction of rejections for non-wage reasons together identify the dispersion of the non-wage value and the correlation of wages and non-wage values. The fact that many jobs are accepted that pay well below the reported reservation shows that fairly large positive non-wage values are common. We characterize the function by the acceptance rate at five values of d. Together with the fraction of offers rejected for non-wage reasons, these moments identify the mean and standard deviation of the log of the non-wage job value, as well as the correlation of wages and non-wage values in job offers.

4.4 Prior wage

Our model views the prior wage as based on the job-ladder model. A respondent searched during an earlier spell of unemployment and accepted the first job offered that exceeded the reservation job value (combining wage and non-wage components). While employed, the worker received offers, and accepted the ones that exceeded the job value of the prior job. The distribution of the observed wage on the job the respondent held just before the current spell of unemployment is the stationary distribution of the process of starting up the job
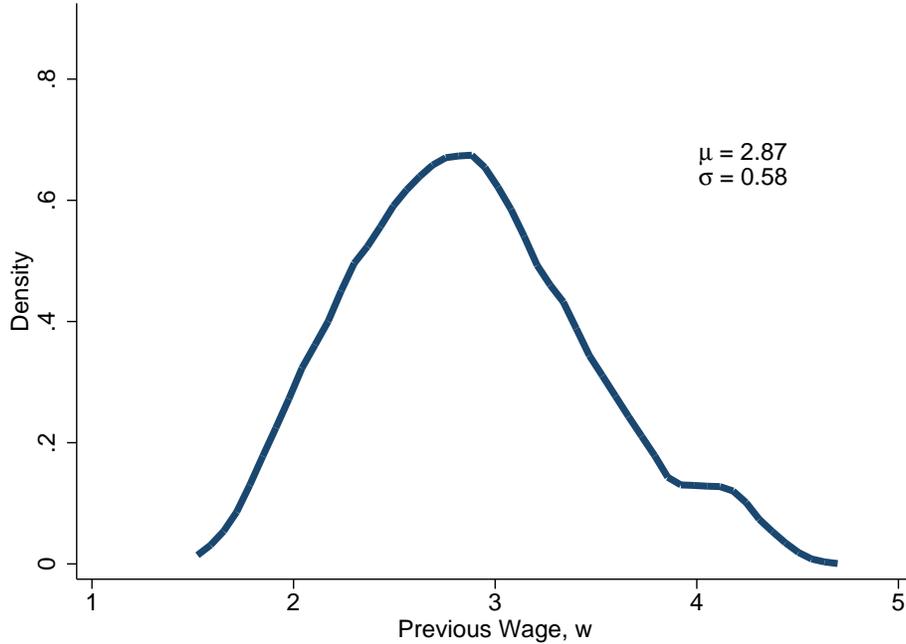
Figure 4: Kernel Density of the Log Hourly Wage at the Prior Job, w

ladder, making successive improvements, and occasionally suffering job loss and dropping back to the bottom of the ladder.

Figure 4 shows the kernel density of the hourly wage on the prior job. The wage is computed from administrative data on weekly earnings during the base year, which typically consist of the first four of the five quarters before the date of the UI claim, and from survey data on weekly hours for the previous employment. Hours on the previous job might not perfectly overlap with the period of the base year. Roughly 15 percent of the respondents answered that hours varied on their previous jobs. We imputed their hours based on demographic characteristics as in Krueger and Mueller (2011). For these reasons, the hourly previous wage includes some measurement error despite the fact that weekly earnings are taken from administrative data.

In the model, the distribution of the prior wage depends on all four unobserved distributions. We carry out a rather complicated calculation of the distribution and match it to the observed one. We update the wage by 2.75 percent to adjust for the time elapsed between the measurement of the respondents' earnings in March 2008 to the median survey month, November 2009, based on the Bureau of Labor Statistics' Employment Cost Index for the metro area including New Jersey, which is adjusted for changes in the composition of employment.

5   Estimation

The model is overidentified. We estimate its parameters from a sub-model that is conditional on the reservation wage, rather than incorporating the part of the model dealing with the optimal reservation wage. The moments we omit from estimation are the means and standard deviation of the log of the wage earned on the job prior to the current spell of unemployment and the covariances of the prior wage and the offered and reservation wages (m_w, s_x, $c_{y,w}$, and $c_{r,w}$). The reasons for not matching those moments are: (1) no parameter values can actually match the standard deviation of the prior log-wage, though the parameter values from the sub-model estimation come quite close, as we show in a later section; (2) we lack evidence about the process of on-the-job search, where the KM survey is silent because all of its respondents are unemployed; and (3) the KM survey sampled unemployed jobseekers, so its distribution of prior wages is not directly comparable to the distribution of wages among the employed. Nonetheless, we believe it is useful to calculate the implied distribution among the employed and compare it to the distribution of past wages among the survey respondents.

5.1   Moments

Table 1 shows the moments of the data that are the targets for matching with the sub-model. The moments for the acceptance frequency are taken from the predicted values of the polynomial of degree 6 evaluated at five values of d.

5.2   Matching the model's moments to the observed moments

We estimate the parameters of the distributions of the four variables y, r, η, and x, and the compensating-difference parameter κ. As described earlier, we take the distributions of the variables to be log-normal and independent. We normalize the mean of x to zero. The other three means, μ_y, μ_r and μ_η; the standard deviations, σ_y, σ_r, σ_η, and σ_x; and κ (the relation of the non-wage value n to the offered wage y,), are parameters to estimate, for a total of 8. We target the following 11 data moments: the means $m_{\hat{y}}$ and $m_{\hat{r}}$, standard deviations $s_{\hat{y}}$ and $s_{\hat{r}}$ of the two directly observed variables, the covariance $c_{\hat{y},\hat{r}}$, the five values $\hat{A}_1$-$\hat{A}_5$ of the acceptance frequency, and the fraction of rejections for non-wage reasons, $\hat{J}$. We minimize the sum of squares of the deviation of the model from the data moments, with appropriate weights for each moment. The weights correspond to the inverse of the variance of each moment, bootstrapped with 2000 repetitions.

Because the minimization is computationally demanding, we also used a different and much easier approach, by setting the weights for the moment conditions apart from the acceptance function and the rejection frequency to infinity—that is, we required that the

Moment	Symbol	Value
Mean offered wage	$m_{\hat{y}}$	2.75
Mean reservation wage	$m_{\hat{r}}$	2.82
Standard deviation of offered wage	$s_{\hat{y}}$	0.525
Standard deviation of reservation wage	$s_{\hat{r}}$	0.474
Covariance of offered wage and reservation wage	$c_{\hat{y},\hat{r}}$	0.183
Acceptance frequency at $d_1 = -1.0$	$\hat{A}_1$	0.262
Acceptance frequency at $d_2 = -0.5$	$\hat{A}_2$	0.576
Acceptance frequency at $d_3 = 0.0$	$\hat{A}_3$	0.780
Acceptance frequency at $d_4 = 0.5$	$\hat{A}_4$	0.856
Acceptance frequency at $d_5 = 1.0$	$\hat{A}_5$	0.618
Fraction of rejections for non-wage reasons	$\hat{J}$	0.677

Note : d_i refers to the difference between the log offered wage $\hat{y}$ and the log reservation wage $\hat{r}$.

Table 1: Target Moments

estimates solve the first five moment conditions exactly. For our baseline specification, we found that the results of this approach were identical to those for the estimation using weights derived from the sampling variances of the the means, standard deviations, and covariance of $\hat{y}$ and $\hat{r}$. Accordingly, we used the streamlined approach for estimating the alternative specifications later in the paper. The reasons that the streamlined approach gives identical results are that the sampling weights for the moments related to the acceptance function are smaller than the other weights, and that the parameters related to the distribution of y and r yield little or no gain in improving the fit of the acceptance function.

We allow for measurement error in the reservation wage and the offered wage, by using the finding of Bound and Krueger (1991) that 13 percent of the total variation in offered wages and reservation wages is due to measurement error. They obtained the estimate by comparing survey data to administrative data.

To sum up, the moment-matching conditions are:

$$m_{\hat{y}} = \mu_y \tag{14}$$

$$m_{\hat{r}} = \mu_r \tag{15}$$

$$s_{\hat{y}} = \sqrt{\sigma_y^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{y}}}^2} \tag{16}$$

$$s_{\hat{r}} = \sqrt{\sigma_r^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{r}}}^2} \tag{17}$$

$$c_{\hat{y},\hat{r}} = \sigma_x^2 \tag{18}$$

$$\hat{A}_i = 1 - \Phi(0, \mu_{m|d_i}, \sigma_{m|d_i}), i = 1, 2, 3, 4, 5 \tag{19}$$

$$\hat{J} = \int_{-\infty}^{\infty} \int_{-\infty}^{r} \frac{\Phi(0, \mu_{p|v}, \sigma_{p|v})}{\Phi(r, \mu_v, \sigma_v)} \phi(v, \mu_v, \sigma_v) dv \phi(r, \mu_r, \sigma_r) dr. \tag{20}$$

Here $\Phi(x, \mu, \sigma)$ is the normal cdf and $\phi(x, \mu, \sigma)$ is the normal pdf. Note that the functions $\mu_{m|d}$, $\sigma_{m|d}$, $\mu_{p|v}$, and $\sigma_{p|v}$; and the values μ_v, and σ_v are functions of the 8 parameters to be estimated—see Appendix A for details.

To measure sampling variation, we calculate the bootstrap distribution of the estimates. In our actual estimation procedure, we compute our moments from two different samples: We take the moments $m_{\hat{r}}$ and $s_{\hat{r}}$ from the first interview for all unemployed workers in the survey who were not working or had not yet accepted a job offer, whereas we take $m_{\hat{y}}$, $s_{\hat{y}}$, $c_{\hat{y},\hat{r}}$ and $\hat{A}_1$-$\hat{A}_5$ from the sample of 1,153 job offers with information on the offered wage and on the lagged reservation wage. The standard bootstrap strategy applies to single samples. Accordingly, we use only the smaller sample. This smaller sample appears not to be biased, as $m_{\hat{r}} = 2.83$ and $s_{\hat{r}} = 0.47$, which are almost identical to the estimates in the bigger sample. For the bootstrap, we thus sample with replacement from the 1,153 job offers, and compute the moments in the data and in the model for 100 draws. The resulting bootstrap distribution provides an upper bound on the dispersion of our actual sampling distribution.

5.3 Estimation results

Table 2 shows the estimation results. Our main findings are:

1. The dispersion in the offered wage among people with the same personal productivity is moderate but not small: $\sigma_y = 0.24$.

2. The dispersion in the reservation wage among people with the same personal productivity is small: $\sigma_r = 0.11$.

3. The dispersion of the independent component of the non-wage job value is substantial: $\sigma_\eta = 0.34$.

Parameter	Explanation	Estimate	(s.e.)
μ_r	Mean of reservation wages	2.82	(0.02)
μ_y	Mean of wage offers	2.75	(0.02)
μ_η	Mean of the independent component of non-wage value of wage offer	0.31	(0.06)
σ_x	Standard deviation of personal productivity	0.43	(0.02)
σ_r	Standard deviation of the reservation wage	0.11	(0.05)
σ_y	Standard deviation of the offered wage	0.24	(0.02)
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.34	(0.07)
κ	Compensating differential	0.25	(0.30)
σ_v	Standard deviation of offered job values $(v = y + n)$	0.38	(0.09)

Table 2: Parameter Estimates

4. The dispersion of personal productivity is substantial: $\sigma_x = 0.43$.

5. There is a moderate amount of compensating wage differentials: $\kappa = 0.25$.

6. The mean value of the non-wage value of a job offer is positive: $\mu_n = \mu_\eta = 0.31$.

The variance of observed offered wages decomposes as

$$s_{\hat{y}}^2 = 0.28 = \sigma_y^2 + \sigma_x^2 + 0.13 s_{\hat{y}}^2 = 0.06 + 0.18 + 0.04. \tag{21}$$

Thus $0.18/0.28 = 66$ percent of the cross-sectional variance in offered wages is explained by dispersion in personal productivity x, and only $0.06/0.28 = 21$ percent is explained by differences in wage offers among workers with the same productivity, y. The remaining 13 percent is explained by measurement error. Our results, however, also show that there is substantial dispersion in the non-wage job values, with the dispersion of non-wage job values being larger than the dispersion in offered wages. Our estimates imply that the standard deviation of job values $v = y + n$ is 0.38, which is much larger than the standard deviation for offered wages y alone.

Our data do not identify the amount of measurement error—we rely on extrinsic evidence from Bound and Krueger (1991) about measurement errors in actual wages. Measurement error in reservation wages is potentially higher than measurement error in actual wages,

if unemployed workers do not understand the intended meaning of the reservation-wage question or have different reference levels in mind when they express the reservation wage.

Another source of information about measurement errors in the reservation wage is based on the fact that many respondents in the survey reported their reservation wages more than once. The within-person variance of reservation wages among these respondents has a standard deviation of 0.0974 (if controlling for duration) and 0.0975 (if not controlling for duration). The tiny difference between these two figures is consistent with the results in Krueger and Mueller (2016) who find only a modest negative relationship between the reservation wage and unemployment duration.

The low within-person variance of reservation wages supports our conclusion about the extent of measurement error in that variable. The within-person variance of reservation wages is 4.2 percent of the total variance of reservation wages, which is below our baseline calibration (13 percent). Of course, some part of the measurement error in the reservation wage is persistent across interviews and does not show up in the within-person measure. Accordingly, we believe that our calibration with 13 percent of measurement error is a reasonable baseline.

Our moment conditions for $s_{\hat{r}}$ and $c_{\hat{y},\hat{r}}$ imply an upper bound on measurement error. According to the model, the difference between these two moments is:

$$s_{\hat{r}}^2 - c_{\hat{y},\hat{r}} = \sigma_r^2 + \sigma_{\epsilon_{\hat{r}}}^2. \tag{22}$$

Because $\sigma_r^2 \geq 0$, $\frac{\sigma_{\epsilon_{\hat{r}}}^2}{s_{\hat{r}}^2} \leq \frac{s_{\hat{r}}^2 - c_{\hat{y},\hat{r}}}{s_{\hat{r}}^2} = 18.4$ percent, from Table 1. This bound is relatively tight. Note that it arises from the high correlation of $\hat{y}$ and $\hat{r}$.

To illustrate the sensitivity of our main results to the extent of measurement error, Table 3 shows the share of the variance in the offered job value v accounted for by dispersion in non-wage values. Panel A reports the share for the idiosyncratic part η and panel B reports the share for the entire non-wage value n. The rows are for alternative values of the variance of the measurement error in the reservation wage as a ratio to the variance of the true reservation wage. The columns are for alternative values of the measurement error in the offered wage as a ratio to the variance of the true offered wage. The highest value for both rows and columns is 0.184, the upper bound discussed above. For the baseline calibration (in bold), more than three-quarters of the dispersion in job values is accounted for by dispersion in non-wage values. In general, this share is declining in the extent of measurement error, but still more than one half for the maximum degree of measurement error at 0.184 of the total variation in offered wage and the reservation wage. For lower amounts of measurement error, the dispersion of non-wage values is even more important, as the model can only generate the shape of the acceptance function with higher values of κ, that is, a higher correlation between

A: $\sigma_\eta^2 / \sigma_v^2$		$\sigma_{\varepsilon y}^2 / \sigma_y^2$			
		0.000	0.065	0.130	0.184
	0.000	1.00	1.00	1.00	0.99
$\sigma_{\varepsilon r}^2 / \sigma_r^2$	0.065	0.93	0.93	0.92	0.92
	0.130	0.86	0.83	**0.78**	0.71
	0.184	0.80	0.75	0.67	0.55
B: σ_n^2 / σ_v^2		$\sigma_{\varepsilon y}^2 / \sigma_y^2$			
		0.000	0.065	0.130	0.184
	0.000	2.55	2.26	1.96	1.70
$\sigma_{\varepsilon r}^2 / \sigma_r^2$	0.065	1.23	1.20	1.17	1.16
	0.130	0.95	0.89	**0.81**	0.72
	0.184	0.83	0.76	0.67	0.57

Table 3: The Contribution of Non-Wage Values to the Variance of Job Value for Alternative Amounts of Measurement Error

offered wages y and offered non-wage values n. Appendix Table 8 gives the estimates of κ and other parameters. Higher values of κ are also the reason that the variance of n is larger than the variance of v for low amounts of measurement error. This finding suggests that one should be careful in interpreting the estimate of the parameter κ as evidence of compensating differentials, as it appears to be a substitute for measurement error in explaining the shape of the acceptance function. However, the substantial contribution of non-wage values to the dispersion in job values v remains a strong result for all the calibrations of measurement error considered here.

Figure 5 shows the smooth acceptance frequency from the data (solid line) with a boot-strapped confidence interval and the acceptance frequency implied by the estimated param-eter. The range of the x-axis is restricted to the 1st to the 99th percentile of d. The fit of the model to the data appears to be quite good, except towards the extreme values of d, especially for values of $d > 0.5$. Note our model imposes that the the acceptance frequency converges to 1 as d increases, whereas the data shows a decline. Non-classical measurement error due to outliers could account for the apparent decline the acceptance frequency in the data. Less than 5 percent of our sample of offers has $d > 0.5$ and less than 1 percent of offers has $d > 0.9$. This sparsity accounts for the widening of the confidence interval for high and low values of d.

Figure 6 shows the kernel density of the log of the offered wage and the reservation wage, along with the normal distributions with the same mean and standard deviation. Both plots
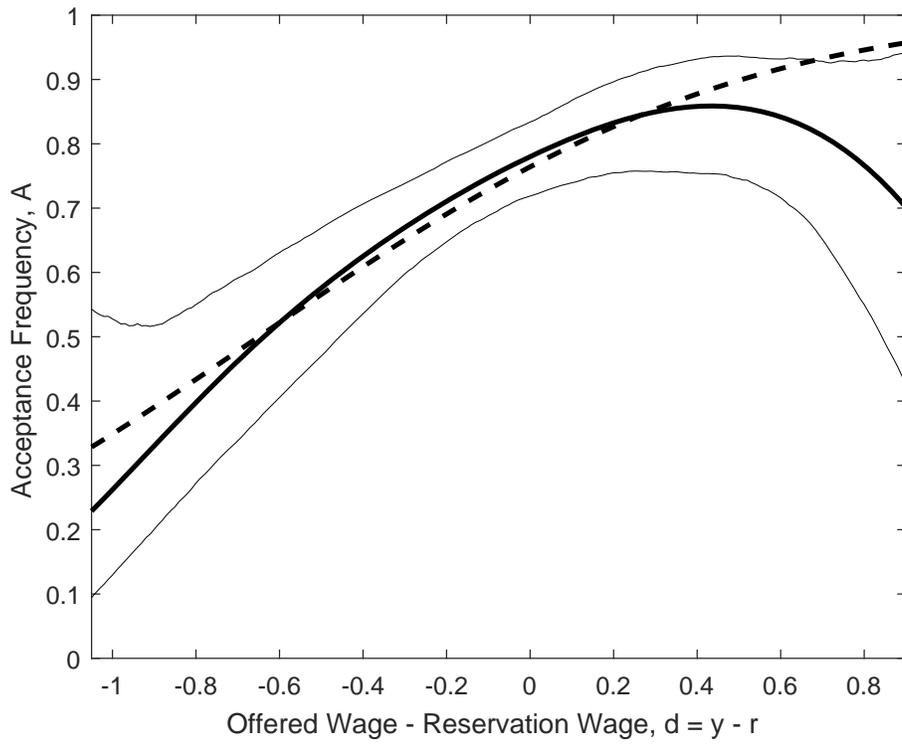
Figure 5: Acceptance Function: Model (Dashed Line) and Data (Solid Line, with 95 Percent Confidence Interval)
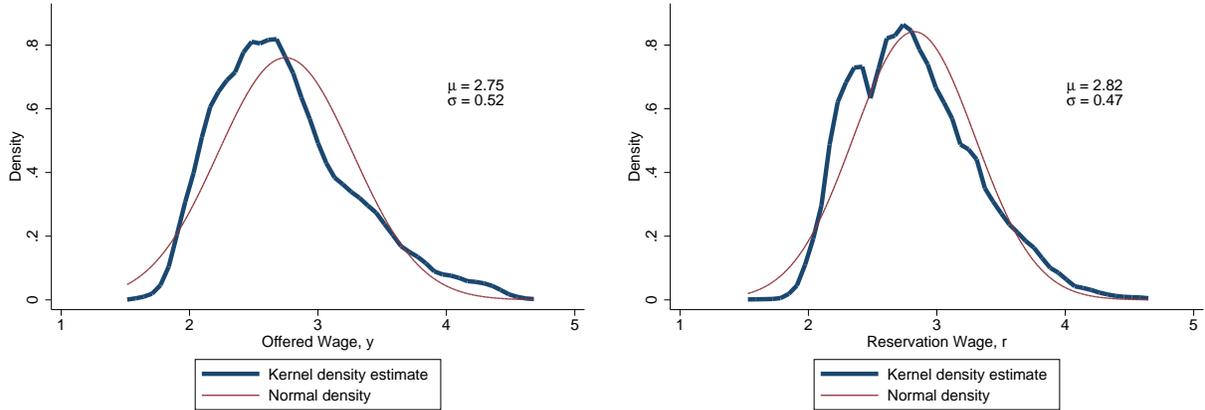
Figure 6: Kernel Density of the Log Hourly Offered Wage, y, and the Log Hourly Reservation Wage, r, Compared to Normal Distribution

show departures from the normal density, mostly in the form of right-skewness. The offered wage distribution has a skewness of 0.79 and an excess kurtosis of 0.30. The reservation wage distribution has a skewness of 0.59 and an excess kurtosis of 0.06. Skewness-kurtosis and Shapiro-Wilk tests rejected normality with a p-values of less than 0.001.

The fact that both distributions are right-skewed in a similar way suggests that it is underlying distribution of x rather than the distributions of y and r that are right-skewed. The two figures show that the log-normal framework of this paper is not completely successful at matching the two observed distributions. More flexible functional forms could improve the fit, at a considerable cost in complexity. See Appendix D.9 for results based on distributions with non-zero skewness and excess kurtosis. These results give improved fits to the distributions but do not change our main conclusions about the dispersions of the key variables.

6 The Model's Implications for the Distribution of Wages among Workers

We now turn to the implications of jobseekers' choices of reservation job values and the stochastic equilibrium of the job-ladder process. In this section, we consider the optimal reservation wage for unemployed jobseekers as derived from the system of Bellman equations. Recall that employed jobseekers' reservation job value is just the value of the current job and the reservation wage of a jobseeker whose job-finding efficiency is at least as high while working as while unemployed is just the opportunity cost—the value of non-market activities.

The hard part is finding the elevated reservation wage for unemployed jobseekers who sacrifice option value by taking a job.

6.1 The distribution of values in non-market activities

The reservation value condition $U(h) = W(h, r)$, defines a function $h = H(r)$ that relates the value of non-market activities h to the reported reservation wage r—see Appendix E for details. The cdf of the distribution of values in non-market activities, $F_h(h)$, satisfies

$$F_r(r) = F_h(H(r)), \tag{23}$$

so, from the estimated parameters of the distribution of reported reservation wage values, $F(r)$, and the function $H(r)$, we can compute the implied distribution of values in non-market activities, $F_h(h)$. Note that in the case where search on the job is equally effective as when unemployed, $\lambda_e = \lambda_u$, the model simplifies to $H(r) = e^r$ and thus $F_r(r) = F_h(e^r)$.

6.2 The stationary distribution of wages

We let $F_w(w)$ be the cdf of wages among workers with $x = 0$. An individual draws a non-work value h at the outset, associated with a reservation wage r through $h = H(r)$. A personal state variable records whether the individual is unemployed or employed. The flow value of the current job, $v = w + n$, is a second personal state variable for the employed. Jobs end because of the arrival of a better offer or through exogenous separation and a drop to the bottom of the ladder. The latter occurs with fixed probability s and sends the worker into unemployment at the bottom of the ladder.

Define

$$F_v(v) = \int f_{y,\eta}\left(\frac{v - \eta - \kappa\mu_y}{1 - \kappa}, \eta\right) d\eta, \tag{24}$$

the cdf of a job offer with value v. Here $f_{y,\eta}(y, \eta)$ is the joint density of y and η. The probability in one week that an unemployed worker with a reservation value r will remain unemployed in the next week is

$$T_{uu}(r) = 1 - (1 - s)\lambda_u(1 - F_v(r)). \tag{25}$$

The probability that an unemployed individual will be at work in the succeeding week with a job value not greater than v' is

$$T_{ue}(v'|r) = (1 - s)\lambda_u(F_v(v') - F_v(r)). \tag{26}$$

The probability that an employed worker will be unemployed in the next week is

$$T_{eu} = s. \tag{27}$$

The probability that an employed individual will remain employed at the same job value with value v is

$$T_{ee}(v|v) = (1-s)[1 - \lambda_e(1 - F_v(v))].\tag{28}$$

The probability that an employed individual will move to a better job with value $v' > v$ is

$$T_{ee}(v'|v) = (1-s)\lambda_e(F_v(v') - F_v(v)).\tag{29}$$

Let q be the compound state variable combining a binary indicator for unemployment/employment and the job value v and let $T(q'|q,r)$ be its transition cdf derived above. The stationary distribution of q, $F_q(q|r)$ satisfies the invariance condition,

$$F_q(q'|r) = \int T(q'|q,r)dF_q(q|r).\tag{30}$$

Throughout, an integral without limits of integration is over the support of the integrand. The ergodic distribution of the job value for employed workers, $F_v(v|r)$, is the conditional distribution of v for values of q for employed workers.

The cdf of the wage, w, conditional on the job value v, is

$$F_w(w|v) = \frac{\int^w f_{y,\eta}(y, v - y(1-\kappa) - \kappa\mu_y)dy}{\int f_{y,\eta}(y, v - y(1-\kappa) - \kappa\mu_y)dy}.\tag{31}$$

The implied ergodic distribution for the wage is

$$F_w(w|r) = \int F_w(w|v)dF_v(v|r).\tag{32}$$

Finally, the distribution in the population with $x = 0$ is the mixture,

$$F_w(w) = \int F_w(w|r)dF_r(r)\tag{33}$$

and the distribution in the overall population is the mixture,

$$F_{\hat{w}}(\hat{w}) = \int F_w(\hat{w} - x)dF_x(x).\tag{34}$$

6.3 Parameter values

The weekly offer arrival rate in the survey is $\lambda_u = 0.058$ and the average acceptance rate is $a = 0.72$. We calculate the entry rate to unemployment, s, as

$$s = \frac{u}{1 - u(1 - \lambda_u a)}\lambda_u a = 0.0041 \text{ per week,}\tag{35}$$

Parameter	Explanation	Actual Values	Estimate	(s.e.)
μ_h	Mean of non-work values		$2.41	($5.23)
σ_h	Standard deviation of non-work values		$3.36	($1.56)
$m_{\hat{w}}$	Mean previous wage, adjusted for intervening wage growth	2.90	2.91	(0.05)
$\sigma_{\hat{w}}$	Standard deviation of previous wage	0.58	0.52	(0.02)
T_{ee}	Monthly job-to-job transition rate (adjusted for time aggregation)	0.019	0.022	(0.00)

Table 4: Actual and Fitted Values of the Job-Ladder Model

the weekly rate consistent in stationary stochastic equilibrium with an unemployment rate of $u = 0.09$ and the observed job-finding rate. This calculation omits job-finding from out-of-the-labor force and exits from unemployment and employment to out-of-the-labor force.

We calibrate the offer rate for employed jobseekers, λ_e, as half the rate, λ_u, found in our survey. While we do not have a direct estimate of the job offer rate while employed, this calibration matches the rate of job-to-job transitions in the data. We compute the monthly job-to-job transition rate from the CPS monthly files for the years 2009 and 2010. Following Fallick and Fleischman (2004), we measure job-to-job transitions in the CPS using information from a question that asked whether a person worked at the same employer as in the previous month and compute the job-to-job transition *rate* as the fraction of workers changing employers between two consecutive monthly CPS interviews.

We adjust the moments from the model for time aggregation. To make the weekly job-to-job transition rates in the model comparable to the monthly job-to-job transition rates in the CPS data, we aggregate the weekly job-to-job transition rates to monthly rates, taking into account that short unemployment spells of duration less than a month may be misleadingly counted as job-to-job transitions. See the Appendix G for details.

We set the weekly discount rate $\rho = 0.001$, equivalent to an annual discount factor of 0.949.

6.4 Results

The full model including the distribution of the actual wage has no new estimated parameters. We solve it with the estimated parameters reported in Table 2 and the calibrated values of λ_u, λ_e, s and ρ. We ask, what are the estimates of the distribution of the value of non-

market activities h, and how well does the calibrated model match the additional moments not included in Table 1 such as the prior wage? Recall that we do not expect a perfect match for the reasons we listed earlier. Table 4 describes the match:

1. The model nearly matches the mean of the wage on the previous job, $m_{\hat{w}}$, in the case of the moderate amount of measurement error.

2. The model is not capable of matching the standard deviation of the prior wage, $s_{\hat{w}}$. The fitted value is about 0.06 log points below the actual value of the moment for both values of measurement error. The job-ladder model implies that the dispersion of offered wages is larger than the dispersion of wages on the prior job, which is violated in the data. We abstract here from any other sources of wage dispersion that may arise during an employment spell, such as heterogeneous job tenure effects or variation in wages due to changes in job- and firm-specific productivity, which may account for the shortfall.

3. The model does well in matching the job-to-job transition rates in the CPS data in 2009 and 2010.

4. The mean of non-work values is positive but relatively small. Recall that it is stated in dollars per hour, not log points. Figure 7 shows the pdf of h implied by our calibrated job ladder model for our baseline calibration. While the dispersion in h is rather small, there is a substantial fraction of h's with negative values, supporting our choice to express the non-work values in dollars rather than logs.

5. The bootstrap dispersion of the fitted values is quite small in all cases.

We also computed the covariance of $\hat{y}$ and $\hat{w}$ in the model and found $\text{cov}(\hat{y}, \hat{w}) = 0.183$, which matches exactly the covariance of $\hat{y}$ and $\hat{w}$ in the data.

7 The Flow Value of Non-Work

Hornstein et al. (2011) take earlier authors to task for failing to observe that search models imply an extremely low, even negative, value of non-work. The essential point is that the dispersion of offered wages is high enough to justify sampling a large number of offers before picking the best, so that the observed time to acceptance only makes sense if waiting to go to work is painful. They note that the problem remains, though less acute, with on-the-job search.
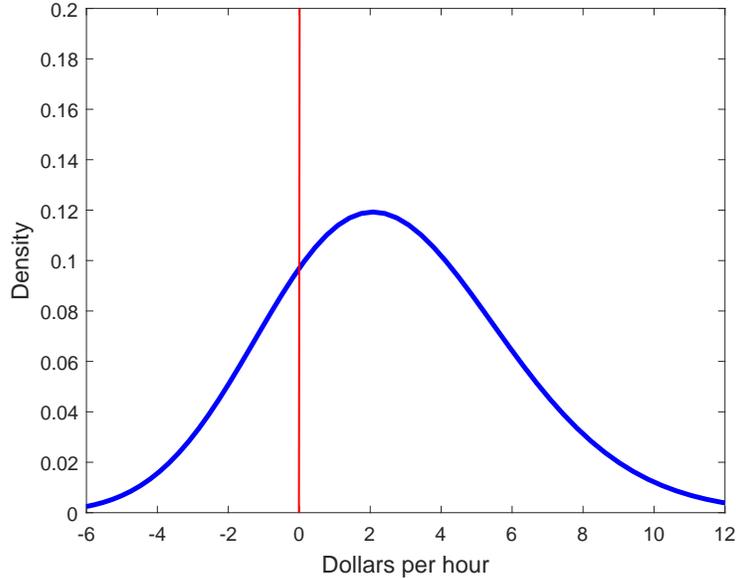
Figure 7: The Density of the Value of Non-Work Time, h

In the search-and-matching literature, whose canon is Mortensen and Pissarides (1994), a variable often called z describes the relation between the flow value of remaining out of the labor market and the flow value of participating in the market. z is often taken as a parameter in these models. It is the ratio of the flow value of non-work to the mean of the marginal product of labor.

7.1 The implied value of z

In the presence of non-wage job values, the calculation of z depends on how much of the benefit of an amenity is a cost to the employer. If the amenity is incidental to employment and comes at no cost to the employer, the marginal product of labor is the observed wage plus the part of the surplus accruing to the employer. For a typical calibration of a DMP-type model, as in Hall and Milgrom (2008), the ratio of the wage to the marginal product is 0.985, so the marginal product is the wage divided by 0.985. On the other hand, if the job value n generates an equal cost to the employer, the job value is effectively an element of the wage. The marginal product of labor is the wage plus the non-wage value, divided by 0.985. We find that the mean of the non-wage value, μ_n, is fairly large and positive, so the adjustment is materially upward.

Table 5 shows the calculation of z for the baseline calibration of the model. Line 1 shows the value of non-work as estimated in that table, expressed in dollars per hour at the median of the distribution of h. Line 2a shows the median wage, whereas line 2b shows the median flow value of work. Line 3 gives an estimate of the marginal product, which is computed by

Step	Explanation	Values	
1	Value of non-work at median for x=0, μ_h	2.41	2.41
2a	Earnings while employed, median for x=0, $\exp(m_w)$	17.90	
2b	Job value while employed, median for x=0, $\exp(m_v)$		34.34
3	Implied marginal product	18.17	34.86
4	Ratio of value of non-work to marginal product	0.13	0.07

Table 5: Ratio of the Flow Value of Non-Work to the Marginal Product of Labor

dividing the estimates in lines 2a and 2b by 0.985. Line 4 reports the resulting value of z, the ratio of the value of non-work to the marginal product. The values are robustly positive, but considerably smaller than in the Hall-Milgrom calibration.

Outside information about the value of z is scant. Chodorow-Reich and Karabarbounis (2016), a deep investigation of the time-series properties of z, is agnostic about its level. Hall and Milgrom (2008) finds a value of 0.71 based on an assumed functional form that satisfies certain elasticity conditions. If the Frisch constant-marginal-utility-of-consumption labor supply function is not a smooth curve in the hours-wage space, but has zero hours until the wage nears a reservation level and then shoots up, the value of z is much lower than Hall and Milgrom calculated.

Another important consideration is that the formula for z in Hall-Milgrom and Chodorow-Reich-Karabarbounis includes the replacement rate for unemployment insurance with a coefficient of one. Our sample is drawn from workers who receive benefits, so the replacement rate is likely to be higher than the 25 percent that Hall and Milgrom assume. The corresponding value of z is much higher— about equal to the median wage—with the 50-percent replacement rate we believe is more realistic. We do not believe that z could possibly be that high. Rather, it shows that the calibration does not give reasonable results with a higher replacement rate. This observation supports the proposition that Hall-Milgrom probably overstated z by choosing an unrealistic functional form for the Frisch supply function.

As discussed in detail in HKV, the crucial parameter for the estimate of z is the offer rate while employed, λ_e, as it determines the option value of remaining unemployed in the event of receiving a job offer. For example, if we calibrated $\lambda_e = 0.7\lambda_u$, our estimate of z lies in the range of 0.28 to 0.53 instead of 0.07 to 0.13, while yielding a job-to-job transition rate of 2.4 percent, which is somewhat larger than in the CPS data at the time of the survey.
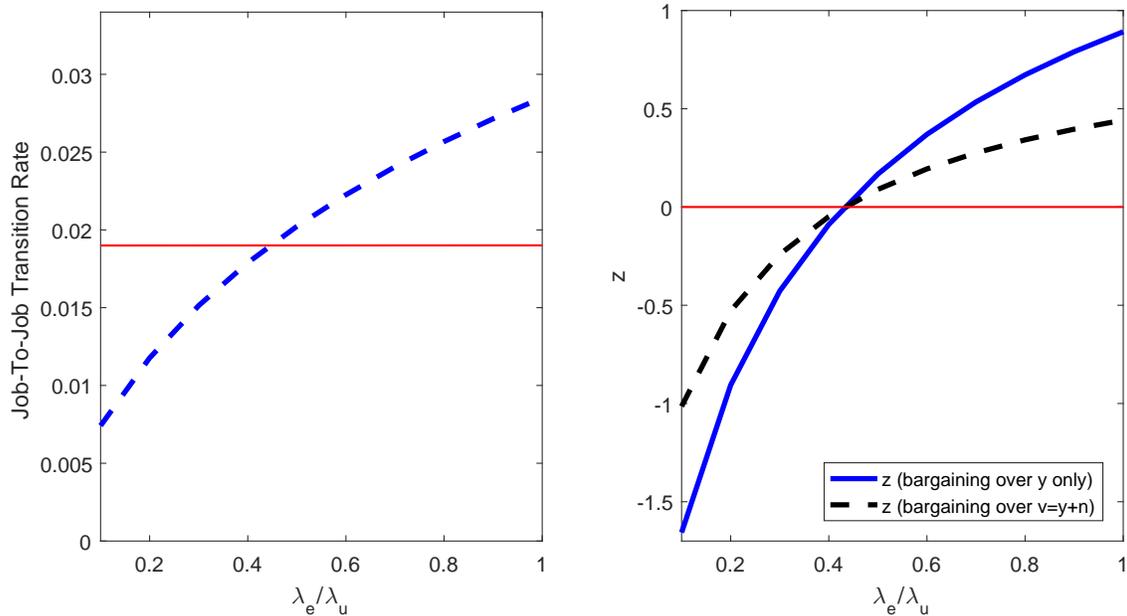
Figure 8: The Efficiency of Search on the Job and the Flow Value of Non-Market Time

See Figure 8, which shows the job-to-job transition rates and values of z for values of $\frac{\lambda_e}{\lambda_u}$ ranging from 0.1 to 1. Blau and Robins (1990) find that offer rates are, if anything, higher for employed jobseekers, suggesting a value of z closer to the one calculated by Hall and Milgrom.

7.2 Re-employment wages

Job-ladder models focus on employment spells—chains of jobs linked by job-to-job transitions. One feature that is common to most job-ladder models is that the combination of high wage dispersion and high offer rates while employed leads to substantial wage growth during an employment spell, as employed workers transition from low- to high-paying jobs. This feature implies a substantial drop in the wage when a worker falls off the job ladder and resumes employment at the bottom of the ladder after an unemployment spell. Our results suggest an important but not overwhelming drop in wages of 9 percent—the mean accepted log wage is 2.81 compared to the mean log wage on the prior job of 2.90, adjusted for wage growth as in Table 4. Research has demonstrated substantial earnings shortfalls occur after job loss. Reconciling the difference in detail is beyond the scope of this paper, but we are aware of a number of differences. First, a major component of the earnings loss is from unemployment rather than declines in wage rates. The KM survey, with a low jobfinding rate, confirms earnings loss from unemployment. Second, research on earnings losses from displacement usually focuses on the losses of higher-tenure (often three years or

more) workers, and these tend to be greater than the losses of low-tenure workers, who make up the great majority of jobseekers who have lost previous jobs.

Our model perfectly matches the mean wage on the prior job as wages do not grow much during a spell of employment despite the job-to-job transitions. The reason is that the dispersion in the idiosyncratic part of non-wage values is larger than the dispersion in offered wages alone, and thus non-wage values tend to dominate wages in the acceptance decision. In other words, employed workers in our model transition frequently from one job to the next, but mostly because new jobs offer higher non-wage values rather than higher wages, and while there is little growth in wages over the course of an employment spell, non-wage values grow substantially, as can be seen from comparing lines 2a and 2b in Table 5. As emphasized earlier in the paper, we think of non-wage values not only as comprising employee benefits such as health insurance, but also preferences over other characteristics of the job, such as commuting distance, relationships with co-workers, and the flexibility of the work schedule. What we label as *non-wage* values may also capture differences in the chances of promotion and pay raises at a future date within the same firm, as in the models of Cahuc, Postel-Vinay and Robin (2006) and Jarosch (2015).

8   Extensions, Robustness Checks, and Further Discussion

In this section, we test the sensitivity of our results to a number of alternative identification assumptions and estimation procedures. In particular, an important identification assumption is that y, r, and x are independently distributed. We investigate the plausibility of these assumptions and test the robustness of the main results to deviations from them.

8.1   Directed search

Unemployed jobseekers with higher non-work values have higher reservation job values. Our assumption of zero correlation of the reservation value and the offered value will fail if the jobseeker knows something about the possible job offer before contracting an employer, because the jobseeker will contact only the more promising employers. Choosier jobseekers with higher non-work values will get better job offers, though less often than other jobseekers. The correlation between the reservation value and the offered value will be positive, not zero.

To illustrate the importance of the issue, suppose that the jobseeking process works the way we describe, with one exception. Instead of seeing all the offers that jobseekers receive, there is a probability χ that the jobseeker knows the offer's terms without contacting the employer. If the job value falls short of the reservation wage, we never learn about the

offer, whereas if the offer is acceptable, it goes into our data. This setup induces a positive correlation between r and y because of the truncation of observations with low values of y.

Table 11 in Appendix D.1 shows the estimation results for different calibrations of χ in the range between 0 and 0.74. The results show that the mean of the wage offer distribution is somewhat smaller for higher values of χ, but the estimated dispersion of y, r and x remains unchanged. This finding may be somewhat surprising, as the censoring of offers should introduce a correlation between y and r and thus lower the estimated dispersion of x and increase the estimated dispersion of y and r. The main reason that the estimates of σ_y, σ_r, and σ_x remain unchanged is that the dispersion of non-wage amenities, σ_η, is more important than the dispersion in wages, σ_y, and thus most of the censoring of offers occurs due to low values of η rather than low values of y. Moreover, the estimated σ_η increases with higher values of χ, which implies that little censoring occurs based on low values of y at any reasonable value of χ.

8.2   Independence of y and r

As we noted earlier, one important assumption in our estimation strategy is that—conditional on personal productivity x—offered wages and reservation wages are uncorrelated, that is, $\text{cov}(\hat{y}, \hat{r}|x) = 0$, as it implies that $\text{cov}(\hat{y}, \hat{r}) = \sigma_x^2$. One possible concern with this assumption is that it may not hold if the employer knows the outside option of the jobseeker and thus tailors the job offer accordingly. Evidence against this is that 76 percent of the survey respondents indicated that the offer was a take-it-or leave-it offer as opposed to 24 percent who said that some bargaining was involved over pay. In any case, our estimate of σ_y changed little when we restricted the sample to take-it-or leave-it offers only— $\sigma_y = 0.21$ as opposed to 0.24 in the baseline case.

A model where the employer knows the reservation wage of the job applicant also implies that $\text{cov}(\hat{y}, \hat{r}) > \text{cov}(\hat{y}, \hat{w})$, as the correlation between wages and the values of non-market activities will be dissipated through the process of on-the-job search and job-to-job transitions. The reason is that, while for an unemployed jobseeker the value of non-market activities may, through bargaining, directly influence the final wage offered, for an employed jobseeker the value of non-market activities is less relevant for the bargaining outcome as the employed worker's outside option is the value of the current job (it still matters to the extent that the value of non-market activities affected the current wage, but less so). However, as mentioned in 6.4, in the data $\text{cov}(\hat{y}, \hat{w}) = \text{cov}(\hat{y}, \hat{r}) = 0.183$.

Finally, in Appendix D.2, we study a model with Nash bargaining and find that our main results do not change in this case. The main reason for this result is that the variance of r is small, so it would require a high correlation of y and r to have a meaningful impact on

the overall covariance of $\hat{y}$ and $\hat{r}$. In other words, as long as the worker's bargaining share α is not too close to 0, the estimate of σ_r will be small and thus the estimate of σ_x large, as in our baseline model.

A related concern with our estimation strategy may be that measurement error in y and r are correlated, which would also violate our assumption that $\text{cov}(\hat{y}, \hat{r}|x) = 0$. Recall that we exploit the longitudinal structure of the survey and use the reservation wage value reported in a week *prior* to the receipt of the job offer. In addition, in the presence of correlated measurement error, we would expect this correlation to be much larger for the pair (y, r) than for the pair (y, w). The reason is that the prior hourly wage is computed from administrative data on weekly wages and hours on last job reported in the first week of the survey. Thus, we gain confidence from the finding that $\text{cov}(\hat{y}, \hat{r}) = \text{cov}(\hat{y}, \hat{w}) = 0.183$.

8.3 Proportionality-to-productivity

As explained earlier, we make the assumption that the distributions of $\hat{y} - x$ and $\hat{r} - x$ in the population with personal productivity x are the same as the distributions of y and r. The most controversial aspect of this hypothesis is that non-market productivity is higher by the entire amount of market productivity in the population with higher values of x. One can test for the presence of non-proportionality in reservation wages by looking at the acceptance rates of job offers across different education levels. Under the proportionality-to-productivity assumption, the average acceptance rate should be the same across workers with characteristics associated with different market productivity x, as these workers should all be equally picky about accepting a job offer. We find that the average acceptance rates do not differ systematically across different levels of educational attainment: The acceptance rate for those with a high-school degree or less is 72.6 percent, for those with some college education is 67.4 percent and for those with a college degree is 74.9 percent, and the differences are not statistically significant. These results are not consistent with a major deviation from the proportionality-to-productivity assumption.

In addition, we estimated the model with a set of moments based on deviations from a model relating wages to their determinants instead of the moments reported in Table 1 based on the wages themselves. More precisely, we ran a Mincer-type regression of the log reservation and offered wage on years of schooling, potential experience, potential experience squared, and dummies for gender, marital status, race and ethnicity, and used the residuals of these regressions to compute the same moments as in Table 1 (except for the means, which we left unchanged from Table 1). One would expect the estimation results to change if the proportionality-to-productivity assumption does not hold in the data. To see this, consider the extreme case where the observable characteristics capture all the variance in

productivity x. In this case, the proportionality-to-productivity assumption is not necessary for identification as the residualized moments of $\hat{y}$ and $\hat{r}$ are independent of x and thus directly capture the moments of interest (plus some measurement error). The results in Appendix Table 7, however, show that all estimated parameters are similar to the results in Table 2 except for the variance of x, which, as expected, is estimated to be substantially smaller, and the compensating differential parameter κ. Appendix Table 7 also shows sub-sample results for those with some college education and less as well as those with a college degree. The mean of the job offer distribution is 38 log points higher for those with a college degree compared to those with some college education or less, whereas the mean of the reservation wage is 47 log points higher (the difference of 38 log points is within sampling variation). The standard deviation of offered wages y is also similar across the two groups, though there is a big difference in terms of the compensating differential parameters κ. The reason is that the sample used to estimate the shape of the acceptance function is quite small and thus, the estimated parameters κ, μ_η and σ_η, which are identified of the shape of the acceptance function, have to be taken with caution in the sub-sample analysis. Overall, these results suggest that proportionality-to-productivity is a reasonable assumption.

Finally, we extend the model by allowing for non-proportionality in the reservation wage variable. This enables us to analyse whether deviations from the assumption of proportionality have an impact on the estimation results. We assume that $\hat{r} = (1 + \kappa_r)x + r + \epsilon_{\hat{r}}$ and use the same moment conditions to re-estimate the model (see Appendix D.4 for details) for different values of κ_r. The sub-sample analysis by education group gives some indication of the potential magnitude of the non-proportionality parameter κ_r. The point estimates of μ_y and μ_r for the two education groups imply that $\kappa_r = 0.2$, because the difference in μ_r is 0.47, which is slightly larger than the difference in μ_y of 0.38. Results in the Appendix show that the non-proportionality tends to raise the dispersions of y and η slightly, but the differences from the estimates of the baseline model where $\kappa_r = 0$ are small.

8.4 Identification of σ_η and κ

As discussed earlier, the shape of the acceptance function, $A(\hat{y} - \hat{r})$, does not separately identify σ_η and κ. The reason is that both parameters increase the likelihood that a high-wage offer is associated with a low non-wage value and thus both parameters increase the probability that a high-wage offer is rejected. σ_η raises the probability of rejection of a high-wage offer because it increases the variance of the non-wage values n, whereas κ raises the probability of rejection of a high-wage offer mainly because positive values lead to a negative correlation between the wage value y and the non-wage value n. For these reasons, we use the fraction of rejections for non-wage reasons, $\hat{J}$, as an additional moment to estimate the

model in our base specification. To make sure that the model is identified, for a given σ_η, we estimated the 7 parameters μ_y, μ_r, σ_x, σ_y, σ_r, μ_η and κ by using the first 7 moment conditions above but not the moment condition for J. In Appendix Figure 10, we plot the fraction of rejections for non-wage reasons, J, for various values of the parameter σ_η. The figure shows that the value of J is strongly increasing in σ_η, demonstrating that the 8 parameters of the model are fully identified with this additional moment. The main reason that the fraction of rejections for non-wage reasons adds valuable information for separately identifying σ_η and κ is that, while higher values of both σ_η and κ make the acceptance functions flatter, the fraction of rejections for non-wage reasons depends mainly on σ_η, because it depends strongly on the relative importance of the idiosyncratic variance of y and n, but is not much affected by the correlation between y and n (and thus κ).

8.5   The acceptance function

For the baseline estimation of the model, we target the acceptance frequency at

$$d = \hat{y} - \hat{r} = [-1, -0.5, 0, 0.5, 1]. \tag{36}$$

In additional results reported in Table 6, we target the acceptance frequency at two, four, seven, eight and nine points at equidistant on the intervals $[-1, 0.5]$, $[-1, 0.75]$ or $[-1, 1]$. We minimize the weighted sum of squared differences of the acceptance frequency at these points, along with the fraction of rejections for non-wage reasons, where the weights correspond to the inverse of the variance of each moment, which was bootstrapped with 2000 repetitions. Table 6 shows that the estimated parameters are similar to the ones in the baseline estimation.

We also take a different estimation approach: Instead of the points on the acceptance function, we match the coefficients of a probit model that was estimated on the KM data. The specification of the probit model is: $A_i = \alpha + \beta d_i$. One can show that matching the two probit coefficients is equivalent to the maximum likelihood estimator of μ_η and σ_η given κ. Together with the fraction of rejections for non-wage reasons, the model is therefore identified. The advantage of this approach is that it takes into account the information contained in all observations in the sample, but we could not impose $A = 0.5$ in the probit estimation for the undecided and had to drop these observations. For this reason, we prefer our approach of matching points on the acceptance function. In any event, the estimated parameters are similar to the ones in the baseline estimation.

Figure 9 in the Appendix also shows the fit of the acceptance function for the baseline calibration and for alternative specification where match different points or the probit coefficients. The fit appears to be similar across all specifications and within the 95 percent

confidence interval for nearly the entire interval except at the very top near $d = 1$. As we noted earlier, it is possible or even likely that non-classical measurement error involving outliers explains the deviation of the model from the data for values of $d > 0.5$.

8.6 Non-stationarity

In our baseline model, we assume a stationary environment for the unemployed jobseeker and thus abstract from forces that lead to changes in the reservation wages over the spell of unemployment. The limited duration of unemployment benefits, declining savings, or changes in the wage offer distribution throughout the spell of unemployment could lead to declining reservation wages over the spell of unemployment. However, as documented in detail in Krueger and Mueller (2016), reservation wages for a given unemployed worker tend to decline little over the spell of unemployment, with point estimates ranging from 1.4 to 3.4 percent over a 25 week period. Moreover, a tendency for the flow value of non-work to change over the spell of unemployment should be reflected in the dispersion of non-work values, but our estimates show little dispersion in non-work values and thus are consistent with close to constant reservation wages over unemployment spells.

8.7 Flow versus stock sampling

Our sample is representative of the *stock* of unemployed workers in New Jersey in 2009, but it may be preferable to estimate the model on a sample representative of the *inflow* of unemployed individuals, as those with low reservation wages or characteristics associated with higher job-offer rates find jobs and thus leave the sample more quickly than those with high reservation wages and low job-offer rates. To assess this issue, we divided our sample into short- and long-term unemployed individuals, using a cutoff duration of unemployment of 26 weeks at the start of the survey. While the short-term unemployed tend to be individuals with higher personal productivity, we find that the point estimates of our main parameters of interest are similar across the two groups and the differences are statistically ambiguous. We find that σ_y is 0.23 for the short-term unemployed and 0.25 for the long-term unemployed, σ_r is 0.08 for the short-term unemployed and 0.20 for the long-term unemployed, and σ_η is 0.31 for the short-term unemployed and 0.32 for the long-term unemployed. Appendix Table 7 provides the details. An alternative way to investigate this issue would be to reweight the sample based on observable demographic characteristics, to make it representative of the inflow, but this would not account for the role of selection based on unobservable characteristics and, in any event, the sub-sample results provided here suggest that reweighting would make little difference.

Parameter	Explanation	Baseline	A(d_i) evaluated at 2 points on interval [-1.0, 0.5]	A(d_i) evaluated at 4 points on interval [-1.0, 0.5]	A(d_i) evaluated at 7 points on interval [-1.0, 0.5]	A(d_i) evaluated at 8 points on interval [-1.0, 0.75]	A(d_i) evaluated at 9 points on interval [-1.0, 1.0]	Matching probit coefficients	Matching probit coefficients (trimmed sample)
μ_r	Mean of reservation wages	2.82	2.82	2.82	2.82	2.82	2.82	2.82	2.82
μ_y	Mean of wage offers	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75
σ_x	Standard deviation of personal productivity	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43
σ_r	Standard deviation of the reservation wage	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
σ_y	Standard deviation of the offered wage	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24
μ_η	Mean of the independent component of non-wage value of wage offer	0.31	0.24	0.32	0.32	0.31	0.31	0.35	0.38
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.34	0.36	0.34	0.34	0.34	0.34	0.32	0.33
κ	Compensating differential	0.25	0.15	0.20	0.19	0.26	0.29	0.38	0.21
σ_v	Standard deviation of offered job values	0.38	0.42	0.39	0.39	0.38	0.38	0.35	0.38

Table 6: Robustness Checks for Acceptance Function

9 Concluding Remarks

The KM data provide a novel view of unemployed workers' search behavior and the dispersion in potential wage offers they face when looking for a job. The data are unique—they contain direct information on reservation wages, job offers, and job acceptance decisions. The data on reservation wages permit identification of the variation in job offers that is due to differences in personal productivity. We use the jobseeker's acceptance decisions to infer the dispersion in non-wage values and to account for the asymmetry in acceptance frequencies of offers above and below the previously reported reservation wage.

We find that the dispersion of the wage offer distribution is moderate, but larger than what HKV associate with the search model without on-the-job search. We find that the dispersion of the non-wage value in job offers is at least as large as the dispersion of wages. The implied overall dispersion in job values for a jobseeker relative to the jobseeker's productivity is substantial. A related finding is that the implied value of non-market time, though not negative, is quite low—around 10 percent of a worker's productivity. We believe that this finding does not contradict other evidence about labor supply. We study an alternative specification of the job ladder model with lower job-finding efficiency among employed searchers, but find that the specification implies even lower values of non-work. We think these findings point in the direction of equal job-finding efficiency for on-the-job search. The pronounced tendency for jobseekers to accept the first job offer they receive is inconsistent with the sacrifice of option value that occurs when a worker takes a job that interferes with subsequent on-the-job search.

Our model has the property that the offered wage remains in effect for the duration of a job. In fact, wage rates do adjust as a worker accumulates tenure. Kudlyak (2014) shows that initial wages are strongly persistent; her evidence supports our assumption. HKV noted that job-ladder models with sequential auctions, such as in Cahuc et al. (2006) weaken the link between the offer rate while employed and the estimate of z, as in these models firms may make counter-offers if a worker receives an outside offer. Outside offers lead to job-to-job transitions only if the outside offer comes from a more productive firm, which can outbid the employee's current firm. See also Papp (2013) who provides a detailed analysis of this issue. Similarly, Christensen, Lentz, Mortensen, Neumann and Werwatz's (2005) model with endogenous search effort implies that workers further up the wage ladder search less and thus transition less frequently to other jobs. Therefore, these models can accommodate larger dispersion in wage offers with higher values of z, as the data on job-to-job transitions do not imply a large option value of unemployment in these models.

We believe that our assumption that the distributions of key observed and latent variables are log-normal or normal is reasonable as a starting point for research on the multiple dimensions of wage dispersion, but the methods of this paper could be extended to other more flexible parametric distributions, such as mixtures of log-normal distributions. We also believe that our finding of high dispersion in non-wage job values shows the potential value of new surveys that collect data on the non-wage characteristics of job offers such as benefits, commuting time, hours, flexibility, job security, firm size, and promotion prospects.

References

Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Becker, Dan, "Non-Wage Job Characteristics and the Case of the Missing Margin," April 2011. http://ssrn.com/abstract=1805761.

Blau, David M. and Philip K. Robins, "Job Search Outcomes for the Employed and Unemployed," *Journal of Political Economy*, 1990, *98* (3), 637–655.

Bontemps, Christian, Jean-Marc Robin, and Gerard J. van den Berg, "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, 2000, *41* (2), 305–358.

Bound, John and Alan B. Krueger, "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," *Journal of Labor Economics*, January 1991, *9* (1), 1–24.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin, "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, March 2006, *74* (2), 323–364.

Chodorow-Reich, Gabriel and Loukas Karabarbounis, "The Cyclicality of the Opportunity Cost of Employment," *Journal of Political Economy*, 2016. Forthcoming.

Christensen, Bent Jesper, Rasmus Lentz, Dale T. Mortensen, George R. Neumann, and Axel Werwatz, "On-the-Job Search and the Wage Distribution," *Journal of Labor Economics*, 2005, *23* (1), 31–58.

Fallick, Bruce and Charles A. Fleischman, "Employer-to-Employer-to-Flows in the U.S. Labor Market: the Complete Picture of Gross Worker Flow," 2004. Finance and Economics Discussion Series 2004-34.

Feldstein, Martin and James Poterba, "UnemploymenI Insurance and Reservation Wages," *Journal of Public Economics*, 1984, *23* (1-2), 141–167.

Hagedorn, Marcus and Iourii Manovskii, "Search Frictions and Wage Dispersion," Manuscript, University of Zurich and University of Pennsylvania, October 2010.

Hall, Robert E. and Paul R. Milgrom, "The Limited Influence of Unemployment on the Wage Bargain," *American Economic Review*, September 2008, *98* (4), 1653–1674.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 2011, *101* (7), 2873–98.

Jarosch, Gregor, "Searching for Job Security and the Consequences of Job Loss," December 2015. Department of Econmics, Stanford University.

Jolivet, Grégory, "A Longitudinal Analysis of Search Frictions and Matching in the U.S.Labor Market," *Labour Economics*, 2009, *16* (2), 121–134.

___ , Fabien Postel-Vinay, and Jean-Marc Robin, "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US," *European Economic Review*, 2006, *50* (4), 877–907.

Krueger, Alan B. and Andreas I. Mueller, "Job Search, Emotional Well-Being, and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data," *Brookings Papers on Economic Activity*, Spring 2011, (1), 1–70.

___ and ___ , "A Contribution to the Empirics of Reservation Wages," *American Economic Journal: Economic Policy*, February 2016, *8* (1), 142–179.

Kudlyak, Marianna, "The Cyclicality of the User Cost of Labor," *Journal of Monetary Economics*, 2014, *68*, 53 – 67.

Low, Hamish, Costas Meghir, and Luigi Pistaferri, "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 2010, *100* (4), 1432–67.

Mortensen, Dale T., *Wage Dispersion: Why Are Similar Workers Paid Differently?*, Cambridge, Massachusetts: MIT Press: Zeuthen Lecture Book Series, 2003.

___ and Christopher Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 1994, *61*, 397–415.

Mueller, Andreas I., "Separations, Sorting and Cyclical Unemployment," September 2015. Columbia University.

Papp, Tamas, "Frictional Wage Dispersion with Bertrand Competition: an Assessment," *Review of Economic Dynamics*, July 2013, *16* (3), 540–552.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

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

Sorkin, Isaac, "Ranking Firms Using Revealed Preference," October 2015. University of Michigan.

Sullivan, Paul and Ted To, "Search and Nonwage Job Characteristics," *The Journal of Human Resources*, 2014, *49* (2), 472–507.

Tjaden, Volker and Felix Wellschmied, "Quantifying the Contribution of Search to Wage Inequality," *American Economic Journal: Macroeconomics*, 2014, *6* (1), 134–161.

Appendixes

Note: The appendixes are intended for online distribution. They repeat some text from the body of the paper for clarity.

A   Details on the Moment Conditions

The observed moments and their counterparts in the model are:

$$m_{\hat{y}} = \mu_y \tag{37}$$

$$m_{\hat{r}} = \mu_r \tag{38}$$

$$s_{\hat{y}} = \sqrt{\sigma_y^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{y}}}^2} \tag{39}$$

$$s_{\hat{r}} = \sqrt{\sigma_r^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{r}}}^2} \tag{40}$$

$$c_{\hat{y},\hat{r}} = \sigma_x^2 \tag{41}$$

$$A(d_i) = 1 - \Phi(0, \mu_{m|d_i}, \sigma_{m|d_i}), i = 1, 2, 3, 4, 5 \tag{42}$$

$$J = \int \int^r \left[\Phi(0, \mu_{p|v}, \sigma_{p|v}) \frac{\phi(v, \mu_v, \sigma_v)}{\Phi(r, \mu_v, \sigma_v)} dv \right] \phi(r, \mu_r, \sigma_r) dr. \tag{43}$$

where $m = v - r$, $d = \hat{y} - \hat{r}$ and $p = (\eta - \mu_\eta) - (y - \mu_y)(1 + \kappa)$. The functions $\mu_{m|d}$, $\sigma_{m|d}$, $\mu_{p|v}$ and $\sigma_{p|v}$ are determined by the parameters μ_y, μ_r, σ_y, σ_r, μ_η, σ_η, κ, $\sigma_{\epsilon_{\hat{y}}}^2$ and $\sigma_{\epsilon_{\hat{r}}}^2$, as follows:

$$\mu_{m|d} = \mu_m + \frac{\sigma_{m,d}}{\sigma_d^2}(d - \mu_d) \tag{44}$$

$$\sigma_{m|d}^2 = \sigma_m^2 - \frac{\sigma_{m,d}^2}{\sigma_d^2} \tag{45}$$

$$\mu_{p|v} = \frac{\sigma_\eta^2 - (1 - \kappa)(1 + \kappa)\sigma_y^2}{\sigma_v^2}(v - \mu_v) \tag{46}$$

$$\sigma_{p|v}^2 = \sigma_y^2(1 + \kappa)^2 + \sigma_\eta^2 - \frac{(\sigma_\eta^2 - (1 - \kappa)(1 + \kappa)\sigma_y^2)^2}{\sigma_v^2}. \tag{47}$$

where

$$\mu_d = \mu_y - \mu_r \tag{48}$$

$$\mu_m = \mu_y + \mu_\eta - \mu_r \tag{49}$$

$$\sigma_d^2 = \sigma_y^2 + \sigma_{\epsilon_{\hat{y}}}^2 + \sigma_r^2 + \sigma_{\epsilon_{\hat{r}}}^2 \tag{50}$$

$$\sigma_m^2 = (1-\kappa)^2\sigma_y^2 + \sigma_\eta^2 + \sigma_r^2 \tag{51}$$

$$\sigma_{m,d} = (1-\kappa)\sigma_y^2 + \sigma_r^2 \tag{52}$$

$$\mu_v = \mu_y + \mu_\eta \tag{53}$$

$$\sigma_v^2 = (1-\kappa)^2\sigma_y^2 + \sigma_\eta^2. \tag{54}$$

B Additional Results

B.1 Moments calculated from regression residuals

Column 2 of Table 7 shows the estimation results using residualized data instead of the raw data for the moment conditions. To be precise, we ran Mincer-type wage regressions of the log hourly offered wage $\hat{y}$ and the log hourly reservation wage $\hat{r}$ on observable characteristics (years of schooling, potential experience, potential experience squared, and dummies for gender, marital status, race and ethinicity) and used the residuals of these regressions to compute the same moments as in Table 1 (except of course for the means, which we left unchanged from Table 1). One would expect the estimation results to change if the proportionality-to-productivity assumption does not hold in the data. To see this, consider the extreme case where the observable characteristics capture all the variance in productivity x. In this case, the proportionality-to-productivity assumption is not necessary for identification as the residualized moments of $\hat{y}$ and $\hat{r}$ are independent of x and thus directly capture the moments of interest (plus some measurement error). The results in Table 7, however, show that all estimated parameters are similar to the results in Table 2 except for the variance of x, which as expected is estimated to be substantially lower, and the compensating differential parameter κ, which is estimated with a lot of sampling error. Overall, these results suggest that proportionality-to-productivity is a reasonable assumption.

Parameter	Explanation	Baseline	Residualized data	Some college education or less	College degree or more	Short-term unemployed	Long-term unemployed	Excluding those with date to return to previous employer	No trimming	Only observations where reservation and offered wages reported in same units
μ_r	Mean of reservation wages	2.82	2.82	2.70	3.17	2.87	2.78	2.83	2.84	2.82
μ_y	Mean of wage offers	2.75	2.75	2.60	2.98	2.85	2.62	2.74	2.75	2.70
μ_η	Mean of the independent component of non-wage value of wage offer	0.31	0.26	0.32	0.38	0.25	0.35	0.31	0.33	0.31
σ_x	Standard deviation of personal productivity	0.43	0.32	0.33	0.46	0.44	0.38	0.43	0.47	0.45
σ_r	Standard deviation of the reservation wage	0.11	0.07	0.18	0.00	0.08	0.20	0.10	0.08	0.00
σ_y	Standard deviation of the offered wage	0.24	0.22	0.25	0.26	0.23	0.25	0.24	0.26	0.18
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.34	0.31	0.38	0.25	0.31	0.32	0.34	0.36	0.26
κ	Compensating differential	0.25	0.38	0.89	0.11	0.27	0.70	0.30	0.43	-0.14
σ_v	Standard deviation of offered job values ($v = y + n$)	0.38	0.34	0.38	0.34	0.35	0.33	0.38	0.39	0.33

Table 7: Additional Estimation Results

B.2 Results for education groups

The results in columns 3 and 4 of Table 7 show results for respondents with some college education or less and for those with a college degree or more. The standard deviation of offered wages is similar across the two samples, whereas the mean is about 39 log points higher (or 48 percent higher). The standard deviation of the reservation wage is somewhat higher in the sample of those with some college education or less, and the differences in the means of the reservation wage is 47 log points (or 60 percent). The remaining parameters differ more substantially across the two samples. In particular, the parameter κ differs between the two sub-samples, but sampling error was big even in the full sample and thus the sub-sample results with respect to this parameter are noisy.

B.3 Results for groups by unemployment duration

The results in columns 5 and 6 show the sub-sample results for short-term and long-term unemployed workers, where short-term unemployed was defined as any individual who was unemployed for less than 26 weeks at the start of the survey. The results show that the results are very similar across the two samples.

B.4 Sensitivity of additional parameter estimates to measurement error

Table 8 shows the parameter estimates underlying Table 3 in the paper. As discussed in the body of the paper, the parameter estimate that changes materially with different amounts of measurement error is κ. Note that this finding is asymmetric, and more important for measurement error in $\hat{r}$, because κ interacts with σ_y in the acceptance model, where we defined $n = \eta - \kappa(y - \mu_y)$.

B.5 Sensitivity to the choice of matching points for the acceptance function

Table 6 shows results with the acceptance frequency targeted at two points ($d = -1$ and $d = 0.5$), four points ($d = -1$, $d = -0.5$, $d = 0.0$, $d = 0.5$), seven points equidistant on the interval $[-1, 0.5]$, eight points equidistant on the interval $[-1, 0.75]$ and nine points equidistant on the interval $[-1, 1]$. The baseline specification targets 5 points ($d = -1$, $d = -0.5$, $d = 0.0$, $d = 0.5$, $d = 1$). The estimated parameters are similar to the ones in the baseline estimation for all of these and the fit of the acceptance function shown in Figure 9 appears to be about the same, except the estimation with two points, which tends to underpredict acceptance around $d = 0$.

A: σ_v		$\sigma_{\varepsilon y}^2 / \sigma_y^2$			
		0.000	0.065	0.130	0.184
	0.000	0.24	0.23	0.23	0.22
	0.065	0.38	0.34	0.30	0.26
$\sigma_{er}^2 / \sigma_r^2$	0.130	0.45	0.42	**0.38**	0.35
	0.184	0.49	0.46	0.43	0.39

B: σ_y		$\sigma_{\varepsilon y}^2 / \sigma_y^2$			
		0.000	0.065	0.130	0.184
	0.000	0.30	0.27	0.24	0.20
	0.065	0.30	0.27	0.24	0.20
$\sigma_{er}^2 / \sigma_r^2$	0.130	0.30	0.27	**0.24**	0.20
	0.184	0.30	0.27	0.24	0.20

C: σ_η		$\sigma_{\varepsilon y}^2 / \sigma_y^2$			
		0.000	0.065	0.130	0.184
	0.000	0.24	0.23	0.23	0.22
	0.065	0.37	0.33	0.29	0.25
$\sigma_{er}^2 / \sigma_r^2$	0.130	0.41	0.38	**0.34**	0.30
	0.184	0.43	0.40	0.35	0.29

D: κ		$\sigma_{\varepsilon y}^2 / \sigma_y^2$			
		0.000	0.065	0.130	0.184
	0.000	0.98	0.96	0.94	0.90
	0.065	0.68	0.66	0.64	0.63
$\sigma_{er}^2 / \sigma_r^2$	0.130	0.45	0.37	**0.25**	0.08
	0.184	0.28	0.15	-0.03	-0.30

Table 8: Estimates for Additional Parameters and Different Amounts of Measurement Error

B.6 Matching a probit specification for the acceptance function

In an alternative estimation approach, we match the coefficients of a probit model estimated on the KM data. The specification of the probit model is: $A = \alpha + \beta d$. Matching the two probit coefficient is equivalent to the maximum likelihood estimator of μ_η and σ_η given κ (or more generally, two of these parameters given the value of the third parameter). Together with the fraction of rejections for non-wage reasons, therefore the model is identified. See the next section for the details of the moment conditions. The advantage of this approach is that it takes into account the information contained in all observations in the sample, but we could not impose $A = 0.5$ in the probit estimation for the undecided and had to drop these observations. For this reason, we prefer our approach of matching points on the acceptance function. In any event, the estimated parameters in columns (7) and (8) are similar to the ones in the baseline. Note that the results in column (8) are based on probit model estimates where we trimmed the sample below -1.0 and above 0.89 of $d = \hat{y} - \hat{r}$, which correspond to the 1st respectively the 99th percentile of the distribution of d. As shown in Figure 9, the fit of the acceptance function is somewhat worse than for the other specifications where we match points on the acceptance function, because we had to drop the undecided from the sample for this estimation procedure.
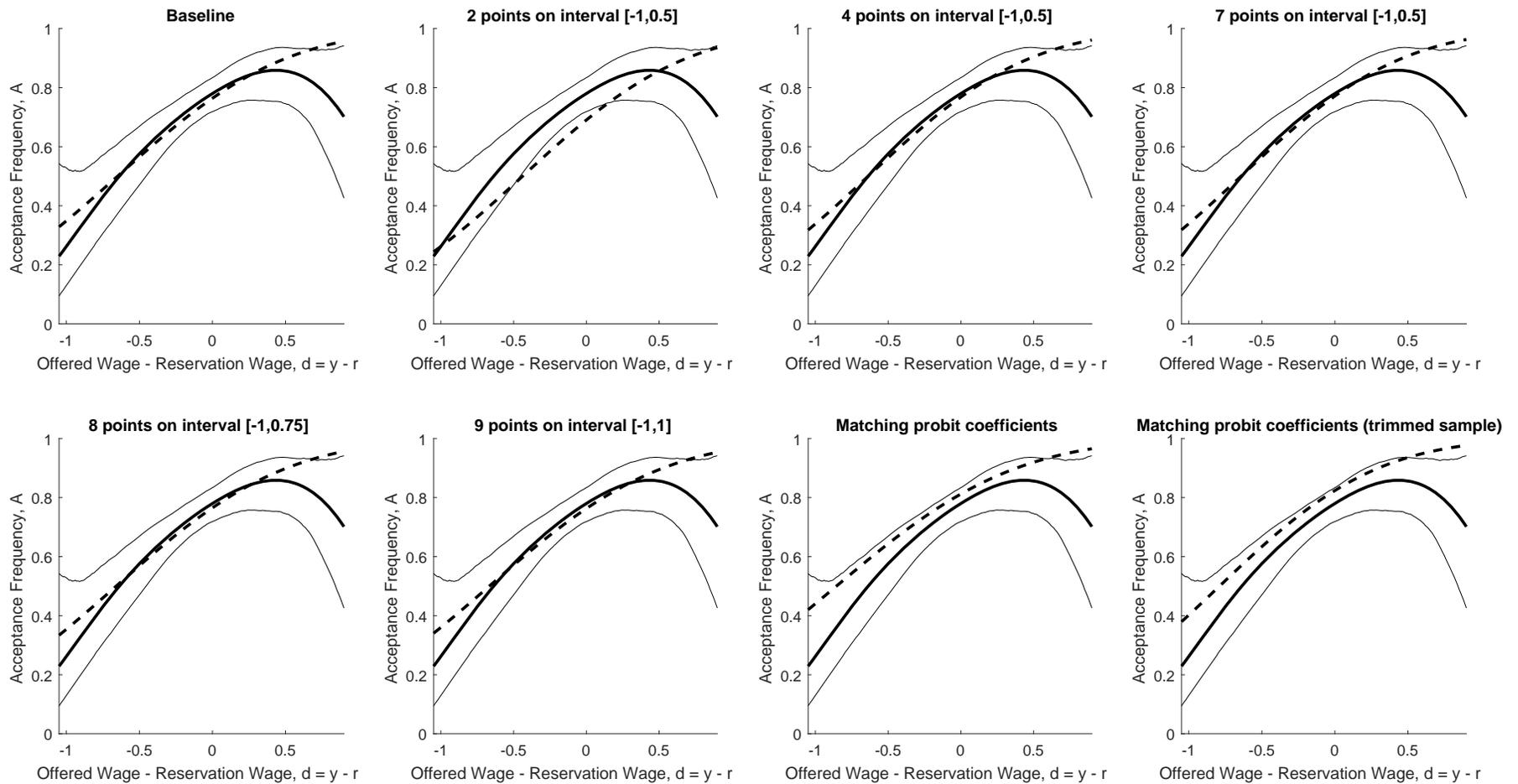
Figure 9: Alternative Specifications for the Acceptance Function: Model (Dashed Line) and Data (Solid Line, with 95 Percent Confidence Interval)

B.7 Identifying power of the frequency of rejection for non-wage reasons

As explained in the body of the paper, the shape of the acceptance function does not separately identify σ_η and κ. We achieve identification by matching the fraction of rejections for non-wage reasons, $\hat{J}$. To demonstrate how the model is identified by this moment, Figure 10 plots the model's calculated fraction of rejections for non-wage reasons, J, for various values of the parameter σ_η. To perform this calculation for a given σ_η, we estimate the 7 parameters μ_y, μ_r, σ_x, σ_y, σ_r, μ_η and κ by targeting the first 7 moments in the data (the moment conditions 37-42 above). The figure shows that the value of J is strongly increasing in σ_η, demonstrating that the 8 parameters of the model are fully identified with this additional moment. The main reason that the fraction of rejections for non-wage reasons adds valuable information for identifying σ_η and κ separately is the following: On the one hand, higher values of both σ_η and κ make the acceptance functions flatter, because they both increase the likelihood that a high-wage offer is associated with a low non-wage value. On the other hand, the fraction of rejections for non-wage reasons mainly depends on σ_η, because it strongly depends on the relative importance of the idiosyncratic variance of y and n, but is not much affected by the correlation between y and n (and thus κ).

B.8 An empirical analysis of the determinants of the job-offer rate and the acceptance probability

Table 9 reports the results of linear regressions for the probability of receiving a job offer (columns 1 to 3) and accepting a job offer (columns 4 to 6). Columns 1 to 3 show that the receipt of a job offer is not correlated with the reservation wage, which we lagged by one interview to address concerns of reverse causality. Columns 4 to 6 show regression results of job acceptance conditional on the receipt of a job offer. As expected, the likelihood of accepting a job offer is positively correlated with the offered wage and negatively correlated with the reservation wage. Observable characteristics have little influence on the likelihood of receipt or acceptance of a job offer. In particular, the likelihood of receipt or acceptance of a job offer does not appear to be correlated with years of school, indirectly supporting our assumption of proportionality in productivity. In other words, the results in this table suggest that (1) more highly educated unemployed workers are equally likely to be in our sample of job offers, and (2) more highly educated unemployed workers are equally picky when it comes to accepting these offers.
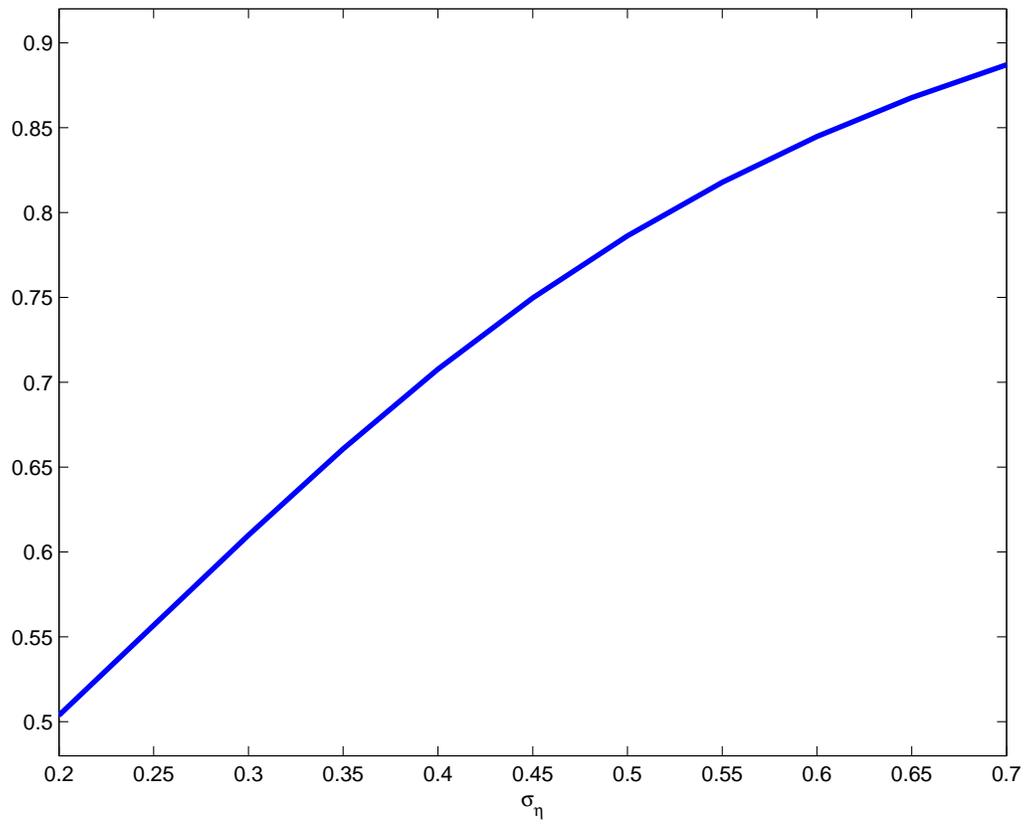
Figure 10: The Fraction of Rejections for Non-Wage Reasons for a Given σ_η

Variable	Dependent variable: Received at least one offer in last 7 days			Dependent variable: Accepted job offer		
log(lagged hourly reservation wage)	0.002		0.003	-0.187		-0.199
	(0.006)		(0.008)	(0.086)**		(0.084)**
log(hourly offered wage)				0.236		0.218
				(0.068)***		(0.067)***
Years of school		0.002	0.002		0.010	0.01
		(0.002)	(0.002)		(0.013)	(0.013)
Potential experience, in years		-0.001	-0.001		0.001	0.001
		(0.000)**	(0.000)**		(0.002)	(0.002)
Female		-0.003	-0.003		-0.015	-0.007
		(0.007)	(0.007)		(0.052)	(0.050)
Married		-0.002	-0.002		-0.070	-0.062
		(0.007)	(0.007)		(0.055)	(0.055)
Black		0.001	0.002		0.032	0.050
		(0.009)	(0.009)		(0.056)	(0.056)
Asian or other		0.004	0.004		0.107	0.073
		(0.016)	(0.016)		(0.076)	(0.061)
Race not available		0.015	0.015		-0.120	-0.146
		(0.012)	(0.012)		(0.118)	(0.117)
Hispanic		0.000	0.000		-0.034	-0.003
		(0.010)	(0.010)		(0.083)	(0.078)
Ethnicity not available		-0.021	-0.021		0.211	0.204
		(0.009)**	(0.009)**		(0.117)*	(0.120)*
Constant	0.035	0.027	0.023	0.627	0.627	0.594
	(0.019)*	(0.026)	(0.027)	(0.140)***	(0.204)***	(0.206)***
N	21,515	21,491	21,491	1,153	1,150	1,150
R-squared	0.000	0.005	0.005	0.050	0.052	0.092

Notes: Robust standard errors in parentheses (clustered at the inidividual level in columns 1-3); * p<0.1, ** p<0.05, *** p<0.01. The sample for the regressions in columns 1-3 consists of all interviews, where a reservation wage was available in the previous interview and the respondent was unemployed and had not accepted a job offer prior to the interview. The sample for columns 4-6 is the baseline sample of job offers in the paper.

Table 9: Linear Regressions of Job Offer and Job Acceptance on Reservation Wage and Observed Characteristics

B.9 Unemployment, separation rate, and job-finding rate

Table 10 reports the average unemployment rate (u), monthly separation rate (s), and monthly job-finding rate (f), from matched Current Population Survey (CPS) micro data by educational attainment. To be more precise, s and f are measured as the rate of transitions between employment and unemployment and vice versa. The table shows that unemployment rates differ strongly by educational attainment. Most of the variation in unemployment is attributable to differences in the separation rate rather than differences in the job-finding rates. The job-finding rate hardly varies by educational attainment. To assess this more formally, we compute the flow steady-state unemployment rate, $u_{ss}(s, f) = \frac{s}{s+f}$, as well as two counterfactual unemployment rates: (1) The steady-state unemployment rate, holding constant the job-finding rate across groups, $u_{ss}(s, f_{avg}) = \frac{s}{s+f_{avg}}$. This counterfactual shows how much of the variation in unemployment rates across education groups is driven by the separation margin. (2) The steady-state unemployment rate, holding constant the separation rate across groups, $u_{ss}(s_{avg}, f) = \frac{s_{avg}}{s_{avg}+f}$. This counterfactual shows how much of the variation in unemployment rates across education groups is driven by the job-finding margin. Nearly all of the variation in unemployment rates across educational attainment is driven by the separation margin. This finding suggests that the fact that unemployment rates differ across education groups is *not* at odds with the assumption of proportionality-in-productivity. Instead, it supports the evidence from the KM survey in Table 9 that more highly educated unemployed workers are equally likely to generate and accept offers as unemployed workers with fewer years of school.

A. Sample: U.S. (2009-2010)	u	s	f	$u_{ss}(s,f)$	$u_{ss}(s,f_{avg})$	$u_{ss}(s_{avg},f)$
Less than high school	0.167	0.035	0.180	0.164	0.170	0.083
High school degree	0.113	0.021	0.164	0.112	0.106	0.090
Some college or associate degree	0.087	0.016	0.170	0.084	0.083	0.087
Undergraduate degree	0.056	0.010	0.190	0.050	0.054	0.079
Graduate degree	0.035	0.006	0.187	0.033	0.035	0.080
Total	0.090	0.016	0.173	0.086	0.086	0.086
Variance of group averages*100	0.269	0.013	0.012	0.273	0.275	0.002

B. Sample: New Jersey (2009-2010)	u	s	f	$u_{ss}(s,f)$	$u_{ss}(s,f_{avg})$	$u_{ss}(s_{avg},f)$
Less than high school	0.137	0.030	0.190	0.138	0.149	0.079
High school degree	0.112	0.019	0.147	0.115	0.099	0.100
Some college or associate degree	0.094	0.016	0.152	0.097	0.086	0.097
Undergraduate degree	0.074	0.011	0.166	0.064	0.062	0.090
Graduate degree	0.039	0.006	0.201	0.028	0.032	0.075
Total	0.089	0.015	0.160	0.086	0.080	0.092
Variance of group averages*100	0.140	0.009	0.056	0.186	0.192	0.012

Source: The authors' estimates with matched CPS monthly data for the years 2009 and 2010 for individuals of age 20 to 65.

Table 10: Unemployment, Separation and job-finding Rates, by Educational Attainment

B.10 Distributions of wages in the CPS and in the KM survey's data on prior wage

Figure 11 shows the kernel density of the log hourly wage for the currently employed and the log hourly prior wage for the currently unemployed. We use the CPS Outgoing Rotation Group data and follow Mueller (2015) to compute the distribution of prior wages for those currently unemployed in New Jersey in the years 2009 and 2010. The densities show that on average unemployed workers tend to be predominantly low-wage workers, which is the consequence of higher separation rates, as Mueller shows.

Figure 12 compares the distribution of the hourly *prior* wage in the CPS data and the KM administrative data. The mean in the CPS data is 2.88, which nearly matches the 2.87 in the KM data—see Figure 4. The standard deviation of prior wages in the CPS data is 0.68, which is quite a bit higher than the standard deviation in the KM administrative data, 0.58. The likely reason is that—while the KM data is representative of the UI recipient population in New Jersey in 2009—the CPS unemployed include a substantial fraction of UI-ineligibles (those who quit their jobs, new and re-entrants in the labor force and other unemployed not satisfying the earnings criteria for being eligible for UI). The UI recipient population is consequently likely to have lower dispersion of productivity x than the population of unemployed in general. Another consideration is that there are only 60 unemployed workers in the New Jersey CPS sample with information on prior wages and, thus, fairly high sampling error. The standard deviation of log hourly prior wages in the *national* sample is 0.55, close to the standard deviation of prior wages in the KM data (0.58).

B.11 Distributions of prior wages in the model and the data

Figure 13 compares the distribution of the log hourly prior wage in the KM data and the stationary distribution in the model, which was simulated for 50,000 individuals. We discussed in the main text why a perfect fit is not to be expected. As Figure 11 and Table 10 suggest, high-wage/high-skill workers have lower separation rates and thus are less likely to be unemployed.

C   The Likelihood Function

We can write the acceptance frequency A as a function of $d = \hat{y} - \hat{r}$:

$$\begin{aligned} A(d) &= \text{Prob}[m \geq 0|d] & (55) \\ &= 1 - \text{Prob}[0 \geq m|d] & (56) \\ &= 1 - F_m(0|d) & (57) \end{aligned}$$
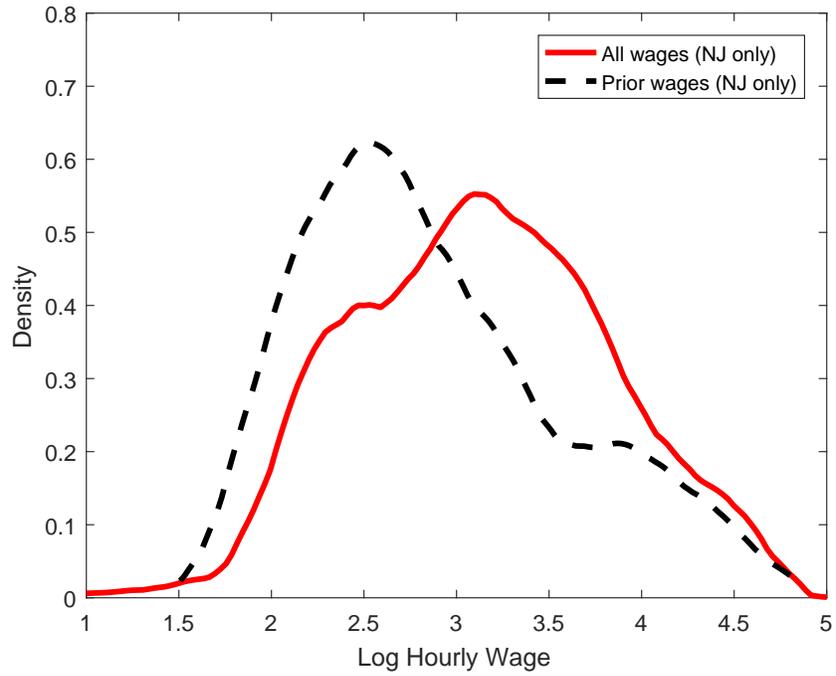
Figure 11: Kernel Density of the Log Hourly Wage in the CPS (2009-2010)
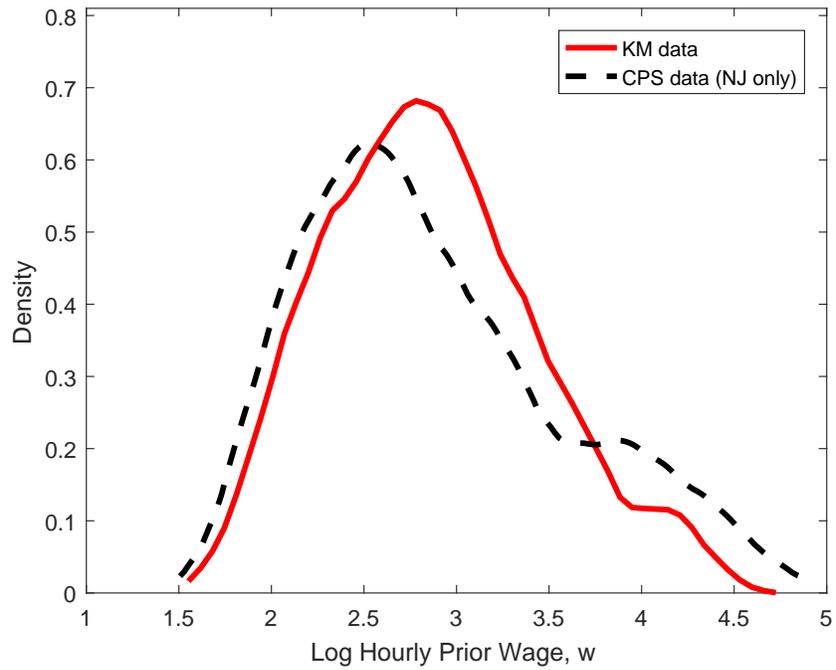
Figure 12: Kernel Density of the Log Hourly *Prior* Wage in the KM Data and the CPS Data
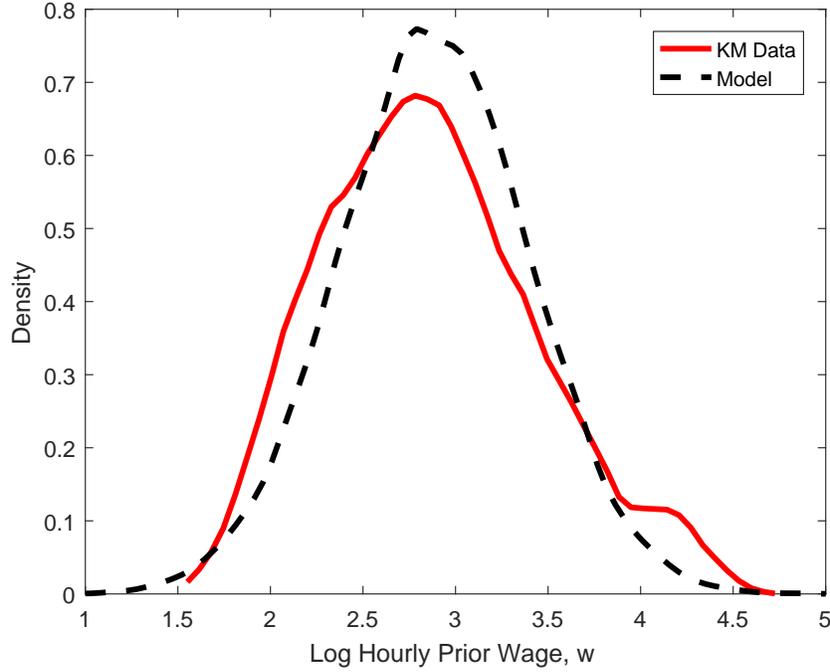
Figure 13: Kernel Density of the Log Hourly Prior Wage in the KM Data and the Model

where $m = v - r_v = \kappa\mu_y + (1-\kappa)y + \eta - r$.

If we assume that all variables are log normally distributed then $m|d$ is log normally distributed with mean and variance as follows:

$$\mu_{m|d} \;=\; \mu_m + \frac{\sigma_{m,d}}{\sigma_d^2}(d - \mu_d) \tag{58}$$

$$\sigma_{m|d}^2 \;=\; \sigma_m^2 - \frac{\sigma_{m,d}^2}{\sigma_d^2}, \tag{59}$$

where

$$\mu_d \;=\; \mu_y - \mu_r \tag{60}$$

$$\mu_m \;=\; \mu_y + \mu_\eta - \mu_r \tag{61}$$

$$\sigma_m^2 \;=\; (1-\kappa)^2\sigma_y^2 + \sigma_\eta^2 + \sigma_r^2 \tag{62}$$

$$\sigma_d^2 \;=\; \sigma_y^2 + \sigma_{\epsilon_{\hat{y}}}^2 + \sigma_r^2 + \sigma_{\epsilon_{\hat{r}}}^2 \tag{63}$$

$$\sigma_{m,d} \;=\; (1-\kappa)\sigma_y^2 + \sigma_r^2. \tag{64}$$

Given the estimates for μ_y, μ_r, σ_y, σ_r and σ_x and the calibrated values of $\sigma_{\epsilon_{\hat{y}}}$ and $\sigma_{\epsilon_{\hat{r}}}$, there are 3 unknowns: μ_η, σ_η, κ. The log likelihood function for the acceptance model can be written as:

$$\ln L(\mu_\eta, \sigma_\eta, \kappa) = \sum_{i=1}^{N} [(A_i \ln(1 - F_m(0|d_i))) + (1 - A_i) \ln F_m(0|d_i)]. \tag{65}$$

C.1   Equivalence to probit model

The log likelihood function of a Probit model for the latent variable $y^* = \alpha + \beta d$ is:

$$\ln L(\alpha, \beta) = \sum_{i=1}^{N} [(A_i \ln(1 - G(-\alpha - \beta d_i))) + (1 - A_i)G(-\alpha - \beta d_i)]. \tag{66}$$

From the formula of the cdf of the normal distribution it is apparent then, that this likelihood function is equal to the likelihood function in (65) if and only if:

$$\alpha + \beta d = \frac{\mu_{m|d}}{\sigma_{m|d}}. \tag{67}$$

$\frac{\mu_{m|d}}{\sigma_{m|d}}$ is a linear function of d and thus the likelihood functions (65) and (66) are identical for:

$$\alpha = \frac{1}{\sigma_{m|d}} \left[\mu_d + \mu_\eta - \frac{(1-\kappa)\sigma_y^2 + \sigma_r^2}{\sigma_d^2} \mu_d \right] \tag{68}$$

$$\beta = \frac{1}{\sigma_{m|d}} \left[\frac{(1-\kappa)\sigma_y^2 + \sigma_r^2}{\sigma_d^2} \right], \tag{69}$$

where

$$\sigma_{m|d}^2 = (1-\kappa)^2 \sigma_y^2 + \sigma_\eta^2 + \sigma_r^2 - \frac{((1-\kappa)\sigma_y^2 + \sigma_r^2)^2}{\sigma_d^2}. \tag{70}$$

Therefore, the moment conditions α and β together with the moment conditions (37)-(41) and the condition for the fraction of rejections for non-wage reasons (43) identify the 8 parameters of the model.

To summarize, the following are the 8 moment conditions that are used for the estimation results reported in Columns (7) and (8) in Table 6:

$$m_{\hat{y}} = \mu_y \tag{71}$$

$$m_{\hat{r}} = \mu_r \tag{72}$$

$$s_{\hat{y}} = \sqrt{\sigma_y^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{y}}}^2} \tag{73}$$

$$s_{\hat{r}} = \sqrt{\sigma_r^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{r}}}^2} \tag{74}$$

$$c_{\hat{y},\hat{r}} = \sigma_x^2 \tag{75}$$

$$\alpha = \frac{1}{\sigma_{m|d}}\left[\mu_d + \mu_\eta - \frac{(1-\kappa)\sigma_y^2 + \sigma_r^2}{\sigma_d^2}\mu_d\right] \tag{76}$$

$$\beta = \frac{1}{\sigma_{m|d}}\left[\frac{(1-\kappa)\sigma_y^2 + \sigma_r^2}{\sigma_d^2}\right] \tag{77}$$

$$J = \int\int^r \left[\Phi(0, \mu_{p|v}, \sigma_{p|v})\frac{\phi(v, \mu_v, \sigma_v)}{\Phi(r, \mu_v, \sigma_v)}dv\right]\phi(r, \mu_r, \sigma_r)dr. \tag{78}$$

D   Model Extensions

D.1   Model with directed search

Unemployed jobseekers with higher non-work values have higher reservation job values. Our assumption of zero correlation of the reservation value and the offered value will fail if the jobseeker knows something about the possible job offer before contracting an employer, because the jobseeker will contact only the more promising employers. Choosier jobseekers with higher non-work values will get better job offers, though less often than other jobseekers. The correlation between the reservation value and the offered value will be positive, not zero.

To illustrate the importance of the issue, suppose that the jobseeking process works the way we describe, with one exception. Instead of seeing all the offers that jobseekers receive, there is a probability χ that the jobseeker knows the offer's terms without contacting the employer. If the job value falls short of the reservation wage, we never learn about the offer, whereas if the offer is acceptable, it goes into our data. This setup induces a positive correlation between r and y because of the truncation of observations with low values of y.

The observed offer rate for unemployed workers with reservation wage r is:

$$\tilde{\lambda}(r) = \lambda(\chi A(r) + (1 - \chi))$$

where $A(r) = 1 - F_v(r)$ is the acceptance frequency given r and λ is the true underlying offer rate (or contact rate). The share of accepted offers where the unemployed worker knew the terms beforehand is $\pi(r) = \frac{\chi A(r)}{\chi A(r) + (1-\chi)}$.

The observed acceptance rate for unemployed workers with reservation wage r is:

$$
\begin{aligned}
\tilde{A}(r) &= \pi(r) + (1 - \pi(r))A(r) \\
&= \frac{\chi A(r)}{\chi A(r) + (1 - \chi)} + \frac{1 - \chi}{\chi A(r) + (1 - \chi)}A(r) \\
&= \frac{1}{\chi A(r) + (1 - \chi)}A(r) \\
&= \frac{1 - F_v(r)}{1 - \chi F_v(r)}.
\end{aligned}
$$

The observed acceptance rate for unemployed workers with $d = \hat{y} - \hat{r}$:

$$
\tilde{A}(d) = \pi(d) + (1 - \pi(d))A(d)
$$

where

$$
\begin{aligned}
\pi(d) &= \int \pi(r)dF(r|d) \\
&= \int \frac{\chi(1 - F_v(r))}{1 - \chi F_v(r)}dF(r|d) \\
A(d) &= \mathrm{P}(v \geq r|d) \\
&= 1 - \mathrm{P}(v < r|d) \\
&= 1 - F_m(0|d),
\end{aligned}
$$

where $m = v - r$. The key issue here is that $d = \hat{y} - \hat{r}$ is not normally distributed and thus it is not possible to derive an expression for $F_m(m|d)$.

D.1.1   Simulated moments

In view of the difficulty of deriving an expression for $F_m(m|d)$, we calculate moments from draws from that distribution, targeting the same moments as in the baseline estimation: $m_{\hat{y}}$, $m_{\hat{r}}$, $s_{\hat{y}}$, $s_{\hat{r}}$, $c_{\hat{y},\hat{r}}$, $\tilde{A}(d)$ for $d = [-1, -0.5, 0, 0.5, 1]$ and J. We use a weighting matrix with the inverse of the variance of each moment on the diagonal (bootstrapped with 2000 repetitions). We generate a million observations from the model and solve for the best fitting 8 parameters, conditional on a range of values of χ.

D.1.2   Results

Table 11 shows the estimation results for the baseline where $\chi = 0$ together with a range of positive values of χ. The mean of the wage-offer distribution is somewhat lower for higher values of χ, but the estimated dispersions of y, r, and x remain unchanged. This finding may be somewhat surprising, as the censoring of offers should introduce a correlation between y

and r, and thus lower the estimated dispersion of x and increase the estimated dispersion of y and r. The main reason that the estimates of σ_y, σ_r, and σ_x remain unchanged is that the dispersion of non-wage amenities, σ_η, exceeds the dispersion in wages, σ_y, and thus most of the censoring of offers occurs from low values of η rather than low values of y. Moreover, the estimated σ_η increases with higher values of χ, which implies that little censoring occurs based on low values of y at any reasonable value of χ.

Parameter	Explanation	Baseline	Value of parameter χ		
			$\chi=0.25$	$\chi=0.50$	$\chi=0.74$
χ	Fraction of offers where v is observed before offer stage	0.00	0.25	0.50	0.74
μ_r	Mean of reservation wages	2.82	2.82	2.83	2.82
μ_y	Mean of wage offers	2.75	2.74	2.72	2.69
μ_η	Mean of the independent component of non-wage value of wage offer	0.31	0.27	0.19	0.00
σ_x	Standard deviation of personal productivity	0.43	0.43	0.43	0.43
σ_r	Standard deviation of the reservation wage	0.11	0.12	0.12	0.11
σ_y	Standard deviation of the offered wage	0.24	0.24	0.25	0.24
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.34	0.38	0.43	0.50
κ	Compensating differential	0.25	0.14	0.02	-0.20
σ_v	Standard deviation of offered job values ($v = y + n$)	0.38	0.43	0.49	0.57

Table 11: Results for Model with Directed Search

The reason that σ_η increases with higher values of χ is that the censoring raises the observed acceptance rate, $\tilde{A}$, as offers where the unemployed worker observed the value of the offer before the offer stage, the acceptance rate of offers made is one by definition. The model also generates a natural asymmetry in the acceptance function $A(d)$ and thus relies less on positive values μ_η to account for it. In particular, for $\chi = 0.74$ the model returns an estimate of $\mu_\eta = 0$, which implies that the asymmetry in $A(d)$ is accounted for by the censoring of offers alone. Figure 14 shows that the fit of $A(d)$ is equally good for each value of χ.

D.2 Nash bargaining with observable values of non-market activities

Our main model assumes that the value of non-market activities h is not known to the firm when making the wage offer and thus wage offers are independent of the reservation wage r. This assumption does not hold in the standard search-and-matching framework with Nash bargaining *if* the value of non-market activities h is known to the employer. In that case, the Nash bargaining solution implies that:

$$e^y = \alpha e^{p_f} + (1 - \alpha)e^r, \tag{79}$$

where all variables are expressed in natural logarithms, p_f is the firm- or match-specific productivity, and α is the worker's bargaining share, taken to be 0.5. It is difficult to model rejections of offers in this environment, but we assume here that firms make wage offers even if $p_f < r$ and thus the offer is going to be rejected by the worker. Note also that we start here with a model where we assume that there are no non-wage amenities and thus only match the first five moments in equations 37-41 (see the next section for Nash bargaining with non-wage amenities). We also assume that there is no on-the-job search. In this model, the moment conditions are:

$$m_{\hat{y}} = E(y(p_f, r)) \tag{80}$$

$$m_{\hat{r}} = \mu_r \tag{81}$$

$$s_{\hat{y}} = \sqrt{Var(y(p_f, r)) + \sigma_x^2 + \sigma_{\epsilon_{\hat{y}}}^2} \tag{82}$$

$$s_{\hat{r}} = \sqrt{\sigma_r^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{r}}}^2} \tag{83}$$

$$c_{\hat{y},\hat{r}} = \sigma_x^2 + \text{cov}(y(p_f, r), r). \tag{84}$$

There are two new parameters to be estimated in this model (μ_{p_f} and σ_{p_f}, instead of μ_y and σ_y in the baseline model). The most important change relative to the estimation of the baseline model is that now the covariance of $\hat{y}$ and $\hat{r}$ not only depends on the variance
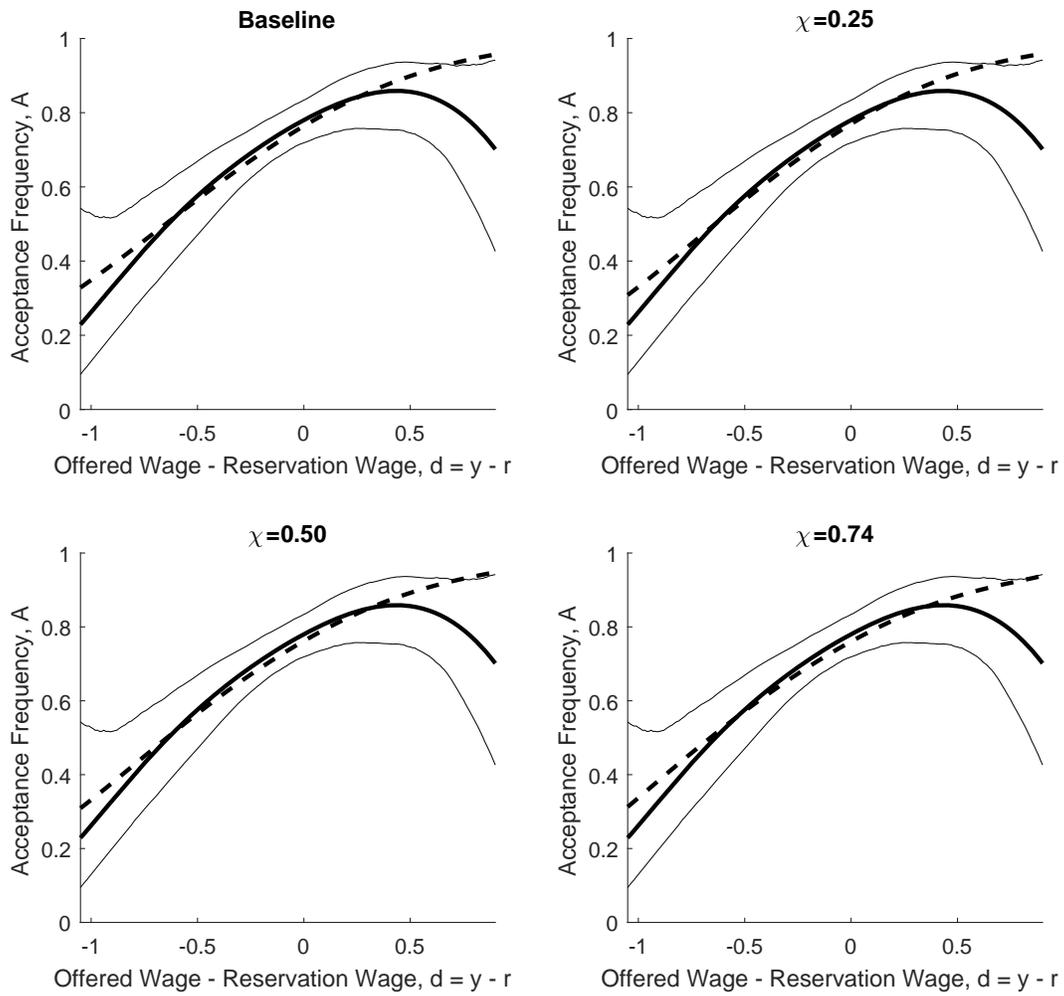
Figure 14: Acceptance Functions for Alternative Values of χ: Model (Dashed Line) and Data (Solid Line, with 95 Percent Confidence Interval)

of x but also on the covariance of the bargained wage y and the reservation wage r. The estimation of the model yields a value of $\sigma_x = 0.42$, which is only slightly below the baseline estimate, and thus the remaining parameter estimates of the model are affected only to a minor degree. The main reason for this result is that the variance of r is small, thus it would require a high correlation of y and r to have a meaningful impact on the overall covariance of $\hat{y}$ and $\hat{r}$. More precisely, one can reformulate the moment conditions such that

$$s_{\hat{r}}^2 - c_{\hat{y},\hat{r}} - \sigma_{\epsilon_{\hat{r}}}^2 \quad = \quad \sigma_r^2 - \text{cov}(y(p_f, r), r) \tag{85}$$

$$= \quad \sigma_r^2(1 - \rho_{y,r}\frac{\sigma_y}{\sigma_r}). \tag{86}$$

Given that the right hand side is relatively small and as long as the correlation coefficient $\rho_{y,r}$ (which is determined mainly by the worker's bargaining share α) is not too close to 0, the estimate of σ_r will be small and thus the estimate of σ_x large, as in our baseline model.

D.3 Nash bargaining with non-wage amenities

This sub-section extends the Nash-bargaining to a model with non-wage amenities in the total compensation package v. The Nash-bargained compensation package satisfies the following equation:

$$e^v = \alpha e^{p_f} + (1 - \alpha)e^r. \tag{87}$$

Note that the Nash-bargain outcome v does not provide any guidance into whether y and n are positively or negative correlated. On the one hand, predetermined aspects of n would lead y and n to be negatively correlated (as the offered wage should compensate for non-wage values), whereas more productive employers may offer more of both and thus y and n may be positively correlated. We let χ be the predetermined part of the non-wage value and ψ be the part of the non-wage value that is determined in the Nash bargain (note that the notation here deviates from the main text; the variance of χ and ψ is captured in our baseline model by the parameters κ and σ_η). We further assume that employers use the following simple rule $y = \gamma_y(v - \psi)$ and $\chi = (1 - \gamma_y)(v - \psi)$ such that $v = y + \chi + \psi$. In this case:

$$y = \gamma_y(ln(\alpha e^{p_f} + (1 - \alpha)e^r) - \psi). \tag{88}$$

As mentioned, the model's parameters γ_y and ψ do not directly map into the parameters of the baseline model and thus we would have to derive slightly different moment conditions for the moments involving acceptance and rejection. As a short cut, we calibrate the dispersion of non-wage values and investigate how the remaining parameters of the model are affected by the presence of non-wage amenities. To do this, we estimate the five parameters of the

75

model μ_{p_f}, σ_{p_f}, μ_r, σ_r and σ_x, for a given γ_y and a given σ_ψ, matching the moment conditions 80-84 above. Our main estimation results are:

1. For $\gamma_y = 1$ and $\sigma_\psi = 0$ (i.e., the baseline from Section D.2 above), we obtain the values $\mu_{p_f} = 2.56$, $\sigma_{p_f} = 0.58$, $\mu_r = 2.83$, $\sigma_r = 0.13$ and $\sigma_x = 0.42$.

2. For $\gamma_y = 0.5$ and $\sigma_\psi = 0$, we obtain the values $\mu_{p_f} = 6.15$, $\sigma_{p_f} = 0.53$, $\mu_r = 2.83$, $\sigma_r = 0.09$ and $\sigma_x = 0.43$.

3. For $\gamma_y = 1$ and $\sigma_\psi = 0.2$, we obtain the values $\mu_{p_f} = 2.61$, $\sigma_{p_f} = 0.37$, $\mu_r = 2.83$, $\sigma_r = 0.13$ and $\sigma_x = 0.42$.

4. For $\gamma_y = 0.5$ and $\sigma_\psi = 0.2$, we obtain the values $\mu_{p_f} = 6.15$, $\sigma_{p_f} = 0.48$, $\mu_r = 2.83$, $\sigma_r = 0.09$ and $\sigma_x = 0.43$.

These results indicate that adding non-wage amenities to the model in Section D.2 leaves our conclusion unchanged that Nash bargaining has little effect on the estimated level of σ_x. The reason is that a higher variance of non-wage amenities will lead to a lower estimated variance of p_f but the total variance of the offered wage $y(p_f, r, \psi)$ as well as the covariance of $y(p_f, r, \psi)$ and r is hardly affected.

D.4   Non-proportionality

In this exercise, we drop the assumption that reservation wages are fully proportional to personal productivity x. Instead, we assume that

$$\hat{r} = x + r(x) + \epsilon_{\hat{r}} \tag{89}$$

$$r(x) = \kappa_r x + r. \tag{90}$$

The moment conditions are:

$$m_{\hat{y}} = \mu_y \tag{91}$$

$$m_{\hat{r}} = \mu_r \tag{92}$$

$$s_{\hat{y}} = \sqrt{\sigma_y^2 + \sigma_x^2 + \sigma_{\epsilon_{\hat{y}}}^2} \tag{93}$$

$$s_{\hat{r}} = \sqrt{\sigma_r^2 + (1 + \kappa_r)^2 \sigma_x^2 + \sigma_{\epsilon_{\hat{r}}}^2} \tag{94}$$

$$c_{\hat{y}, \hat{r}} = (1 + \kappa_r) \sigma_x^2 \tag{95}$$

$$A(d_i) = 1 - \Phi(0, \mu_{m|d_i}, \sigma_{m|d_i}), i = 1, 2, 3, 4, 5 \tag{96}$$

$$J = \int \int^r \left[\Phi(0, \mu_{p|v}, \sigma_{p|v}) \frac{\phi(v, \mu_v, \sigma_v)}{\Phi(r, \mu_v, \sigma_v)} dv \right] \phi(r, \mu_r, \sqrt{\sigma_r^2 + \kappa_r^2 \sigma_x^2}) dr, \tag{97}$$

76

where $\Phi(x,\mu,\sigma)$ and $\phi(x,\mu,\sigma)$ are the cdf and the pdf of the normal distribution with mean μ and standard deviation σ, evaluated at x, and where $m = v - r(x)$, $d = \hat{y} - \hat{r} = y - r(x) + \epsilon_{\hat{y}} - \epsilon_{\hat{r}}$ and $p = (\eta - \mu_\eta) - (y - \mu_y)(1 + \kappa)$. The parameters $\mu_{m|d}$, $\sigma_{m|d}$, $\mu_{p|v}$ and $\sigma_{p|v}$ are determined by the parameters μ_y, μ_r, σ_y, σ_r, μ_η, σ_η, κ, κ_r, $\sigma^2_{\epsilon_{\hat{y}}}$ and $\sigma^2_{\epsilon_{\hat{r}}}$, as follows:

$$\mu_{m|d} = \mu_m + \frac{\sigma_{m,d}}{\sigma^2_d}(d - \mu_d) \tag{98}$$

$$\sigma^2_{m|d} = \sigma^2_m - \frac{\sigma^2_{m,d}}{\sigma^2_d} \tag{99}$$

$$\mu_{p|v} = \frac{\sigma^2_\eta - (1 - \kappa)(1 + \kappa)\sigma^2_y}{\sigma^2_v}(v - \mu_v) \tag{100}$$

$$\sigma^2_{p|v} = \sigma^2_y(1 + \kappa)^2 + \sigma^2_\eta - \frac{(\sigma^2_\eta - (1 - \kappa)(1 + \kappa)\sigma^2_y)^2}{\sigma^2_v}, \tag{101}$$

where

$$\mu_d = \mu_y - \mu_r \tag{102}$$

$$\mu_m = \mu_y + \mu_\eta - \mu_r \tag{103}$$

$$\sigma^2_d = \sigma^2_y + \sigma^2_{\epsilon_{\hat{y}}} + \sigma^2_r + \sigma^2_{\epsilon_{\hat{r}}} + \kappa^2_r \sigma^2_x \tag{104}$$

$$\sigma^2_m = (1 - \kappa)^2 \sigma^2_y + \sigma^2_\eta + \sigma^2_r + \kappa^2_r \sigma^2_x \tag{105}$$

$$\sigma_{m,d} = (1 - \kappa)\sigma^2_y + \sigma^2_r + \kappa^2_r \sigma^2_x \tag{106}$$

$$\mu_v = \mu_y + \mu_\eta \tag{107}$$

$$\sigma^2_v = (1 - \kappa)^2 \sigma^2_y + \sigma^2_\eta. \tag{108}$$

We estimate the model for different values of κ_r. The sub-sample analysis by education group gives some indication of the potential magnitude of the non-proportionality parameter. The point estimates in Appendix Table 7 of μ_y and μ_r for the two education groups imply that $\kappa_r = 0.205$, since the difference in μ_r between the two education groups is 0.47, which is slightly larger than the difference in μ_y of 0.39. The results for the model with $\kappa_r = 0.2$ in Table 12 show that the standard deviation of y and η are somewhat larger than in the baseline model with $\kappa_r = 0$ but the differences in estimates are relatively small.

It is important to note that there is a natural upper bound on κ_r as the moment conditions imply that

$$\kappa_r = \frac{s^2_{\hat{r}} - \sigma^2_r - \sigma^2_{\epsilon_{\hat{r}}}}{c_{\hat{y},\hat{r}}} - 1.$$

The upper bound occurs at $\sigma_r = 0$ and, for $\sigma^2_{\epsilon_{\hat{r}}} = 0.13 s^2_{\hat{r}}$, we get an upper bound of

$$\kappa_r = \frac{0.87 s^2_{\hat{r}}}{c_{\hat{y},\hat{r}}} - 1 = 0.07.$$

		Value of non-proportionality parameter			
Parameter	Explanation	$\kappa_r = 0$	$\kappa_r = 0.1$	$\kappa_r = 0.2$	$\kappa_r = 0.3$
μ_r	Mean of reservation wages	2.82	2.82	2.82	2.82
μ_y	Mean of wage offers	2.75	2.75	2.75	2.75
μ_η	Mean of the independent component of non-wage value of wage offer	0.31	0.34	0.35	0.35
σ_x	Standard deviation of personal productivity	0.43	0.41	0.39	0.38
σ_r	Standard deviation of the reservation wage	0.11	0.00	0.00	0.00
σ_y	Standard deviation of the offered wage	0.24	0.27	0.30	0.31
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.34	0.39	0.42	0.42
κ	Compensating differential	0.25	0.17	0.34	0.48
σ_v	Standard deviation of offered job values $(v = y + n)$	0.38	0.45	0.46	0.45

Table 12: Estimation Results for the Model with Non-Proportionality

For the calibrations in Table 12 where we assumed that $\kappa_r > 0.07$, we assumed that the measurement error was smaller so as to meet the moment condition for $s_{\hat{r}}^2$.

D.5 A model with endogenous search effort

Here we develop a job-ladder model with endogenous job-search effort, similar to Christensen, Lentz, Mortensen, Neumann and Werwatz's (2005), but with the addition of non-wage characteristics of job offers. We generate simulated data from the model and re-estimate our baseline specification using the simulations to investigate the potential effects of endogenous effort on our parameter estimates. We use this approach because not all the parameters of the extended model are identified by our survey data.

The value functions are:

$$
\begin{aligned}
U(h,x) &= \max_q \Big\{ h e^x - c_u(x,q) + \frac{1}{1+\rho} \max_{r_v} \Big((1-s)\lambda_u(q) \int_{r_v} W(h,x,\tilde{v}) dF_v(\tilde{v}) \\
&+ \big((1 - \lambda_u(q) + \lambda_u(q) F_v(r_v) + s\lambda_u(q)(1 - F_v(r_v)) \big) U(h,x) \Big) \Big\} \\
W(h,x,v) &= \max_q \Big\{ e^{v+x} - c_e(x,q) + \frac{1}{1+\rho} \Big((1-s)\lambda_e(q) \int_v W(h,x,\tilde{v}) dF_v(\tilde{v}) \\
&+ (1-s)(1 - \lambda_e(q) + \lambda_e(q) F_v(v)) W(h,x,v) + s U(h,x) \Big) \Big\},
\end{aligned}
$$

where

$$
\begin{aligned}
c_i(q,x) &= c(x)q^{1+\frac{1}{\gamma}} \\
\lambda_i(q) &= b_i q,
\end{aligned}
$$

and the FOCs are:

$$
\begin{aligned}
c_{u_q}(x,q_u^*) &= b_u \frac{1-s}{1+\rho}\left(\int_{r_v} (W(h,x,\tilde{v}) - U(h,x))dF_v(\tilde{v}) \right) \\
c_{e_q}(x,q(v)) &= b_e \frac{1-s}{1+\rho}\left(\int_v (W(h,x,\tilde{v}) - W(h,x,v))dF_v(\tilde{v}) \right) \\
&\quad he^x - c_u(x,q_u^*) + \frac{1}{1+\rho}\max_{r_v}\left((1-s)\lambda_u(q_u^*)\int_{r_v} W(h,x,\tilde{v})dF_v(\tilde{v}) \right. \\
&\quad + \left. ((1-(1-s)\lambda_u(q_u^*)(1-F_v(r_v)))W(h,x,r_v) \right) \\
&= e^{r_v+x} - c_e(x,q_e^*(r_v)) + \frac{1}{1+\rho}\left((1-s)\lambda_e(q_e^*(r_v))\int_{r_v} W(h,x,\tilde{v})dF_v(\tilde{v}) \right. \\
&\quad + \left. (1-(1-s)\lambda_e(q_e^*(r_v))(1-F_v(r_v)))W(h,x,r_v) \right),
\end{aligned}
$$

Rearranging, using the functional form assumptions from above, and assuming proportionality in search costs ($c(x) = c_0 e^x$), gives

$$
\begin{aligned}
\frac{c_0(1+\frac{1}{\gamma})(q_u^*)^{\frac{1}{\gamma}}}{b_u} &= \frac{1-s}{1+\rho}\int_{r_v}(W(h,\tilde{v}) - W(h,r_v))dF_v(\tilde{v}) \\
\frac{c_0(1+\frac{1}{\gamma})q_e^*(v)^{\frac{1}{\gamma}}}{b_e} &= \frac{1-s}{1+\rho}\int_v(W(h,\tilde{v}) - W(h,v))dF_v(\tilde{v}) \\
\frac{e^{r_v} - h + c_0(q_u^*)^{1+\frac{1}{\gamma}} - c_0 q_e^*(r_v)^{1+\frac{1}{\gamma}}}{b_u q_u^* - b_e q_e^*(r_v)} &= \frac{1-s}{1+\rho}\int_{r_v}(W(h,\tilde{v}) - W(h,r_v))dF_v(\tilde{v})
\end{aligned}
$$

We follow HKV by assuming a finite upper bound $\bar{v}$ to the offer distribution. Integrating by parts, as in the online Appendix of HKV, and differentiating $W(h,v)$, we get

$$
W_v(h,\tilde{v}) = \frac{e^{\tilde{v}}}{1 - \frac{1-s}{1+\rho}(1 - b_e q_e^*(\tilde{v})(1-F_v(\tilde{v})))},
$$

and

$$
\begin{aligned}
q_u^*(r_v)^{\frac{1}{\gamma}} &= \frac{b_u}{c_0}\frac{\gamma}{1+\gamma}\int_{r_v}^{\bar{v}}\frac{(1-s)(1-F_v(\tilde{v}))}{\rho + s + (1-s)b_e q_e^*(\tilde{v})(1-F_v(\tilde{v}))}e^{\tilde{v}}dF_v(\tilde{v}) \quad (109) \\
q_e^*(v)^{\frac{1}{\gamma}} &= \frac{b_e}{c_0}\frac{\gamma}{1+\gamma}\int_v^{\bar{v}}\frac{(1-s)(1-F_v(\tilde{v}))}{\rho + s + (1-s)b_e q_e^*(\tilde{v})(1-F_v(\tilde{v}))}e^{\tilde{v}}dF_v(\tilde{v}) \quad (110) \\
\frac{e^{r_v} - h + c_0(q_u^*)^{1+\frac{1}{\gamma}} - c_0 q_e^*(r_v)^{1+\frac{1}{\gamma}}}{(1-s)(b_u q_u^* - b_e q_e^*(r_v))} &= \int_{r_v}^{\bar{v}}\frac{1-F_v(\tilde{v})}{\rho + s + (1-s)b_e q_e^*(\tilde{v})(1-F_v(\tilde{v}))}e^{\tilde{v}}dF_v(\tilde{v}). \quad (111)
\end{aligned}
$$

79

D.5.1 Solution

To solve for the search effort of the employed and unemployed, we start with $q_e^*(\bar{v}) = 0$ and then use equation 110 to solve backwards from the top to the bottom of the job-ladder. We solve for $q_u^*(r_v)$ by combining equations 110 and 109.

D.5.2 Calibration

We use the parameter estimates from Table 2 and calibrate the parameters specific to the model here the following way: Following Christensen, Lentz, Mortensen, Neumann and Werwatz (2005), we set $\gamma = 1.18$, $b_u = b_e$ and target a weekly offer rate for an unemployed worker at the median reservation wage of 0.058. Note that c_0 is not separately identified from b_u and b_e, so we normalize them to 1, and choose c_0 to match the offer rate of the media unemployed worker. We also explore a calibration with $b_u = 1$ and $b_e = 0.5$. The remaining parameters are calibrated as for the main model in the paper.

D.5.3 Results

We simulate the model based on the calibration above and generate the 11 moments that go into the estimation. Then we use the generated moments to estimate the 8 parameters of the baseline model *without* search intensity. To the extent that the estimated parameters differ from the true data generating process, this indicates that the results in Table 2 are biased in the presence of endogenous search intensity. However, as the results in Table 13 show, the estimated parameters are similar to the baseline estimates in the paper, suggesting that endogenous search intensity does not lead to meaningful biases in our estimates. The main reason for this result is that there is little dispersion in reservation wages and thus search intensity, so almost no endogenous selection of job offers occurs based on the reservation wage of the job seeker.

D.6 Multiple offers

We noted earlier that our concept of the offer distribution includes the effect of the auction that benefits a jobseeker who receives multiple simultaneous job offers. This appendix explores the differences in our results when we adopt the alternative definition of the offer distribution as describing all offers, including the ones that are excluded by the auction.

In the survey, among those who received at least one offer in a given week, 86.3 percent received exactly 1 offer, whereas 8.6 percent received 2 offers, 2.4 percent received 3 offers, 0.6 percent received 4 offers, and the remaining 2.1 percent received between 4 and 10 offers. Because the survey only asked about the terms of the best offer, the occurrence of multiple

| Parameter | Explanation | Baseline | Data generated by model with endogenous search intensity | |
			$b_e = b_u$	$b_e = 0.5b_u$
μ_r	Mean of reservation wages	2.825	2.825	2.825
μ_y	Mean of wage offers	2.747	2.747	2.747
μ_η	Mean of the independent component of non-wage value of wage offer	0.312	0.312	0.315
σ_x	Standard deviation of personal productivity	0.428	0.428	0.428
σ_r	Standard deviation of the reservation wage	0.110	0.108	0.107
σ_y	Standard deviation of the offered wage	0.238	0.237	0.237
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.339	0.330	0.331
κ	Compensating differential	0.247	0.269	0.270
σ_v	Standard deviation of offered job values $(v = y + n)$	0.384	0.373	0.374

Table 13: Estimates Using Simulated Data from Model with Endogenous Search Intensity

offers may bias the estimates of our baseline model, which allows only one offer at a time. To address this issue, we extend the baseline model and allow multiple offers in a given week and then use only the best offer, the one with the highest value of v, to compute the moment conditions. We use the probabilities of multiple offers from the survey (see above). Because it is not possible to solve the moment conditions analytically, we estimated the model by matching simulated moments.

The results in the second column of Table 14 show that the estimated parameters are similar to the ones in the baseline model. The dispersion of η is slightly higher, because of the censoring of offers with low values of v. The estimated κ is slightly lower, to compensate the effect of the the higher value of σ_η on the shape of the acceptance function.

Parameter	Explanation	Baseline	Multiple offers	$v=log(e^y+e^n)$	$\sigma_{\mu\eta} = \sigma_{\varepsilon r}$	$\sigma_{\mu\eta} = 2\sigma_{\varepsilon r}$	Skew-normal x, y and r
					Model extensions		
μ_r	Mean of reservation wages	2.82	2.82	2.82	2.82	2.82	2.82
μ_y	Mean of wage offers	2.75	2.75	2.74	2.75	2.75	2.75
μ_η	Mean of the independent component of non-wage value of wage offer	0.31	0.31	1.18	0.36	0.40	0.31
σ_x	Standard deviation of personal productivity	0.43	0.43	0.44	0.43	0.43	0.43
σ_r	Standard deviation of the reservation wage	0.11	0.11	0.07	0.11	0.11	0.11
σ_y	Standard deviation of the offered wage	0.24	0.24	0.15	0.24	0.24	0.24
σ_η	Standard deviation of the independent component of non-wage value of wage offer	0.34	0.37	1.44	0.38	0.40	0.32
κ	Compensating differential	0.25	-0.01	0.05	0.04	-0.12	0.17
σ_v	Standard deviation of offered job values ($v = y + n$)	0.38	0.44	0.25	0.44	0.48	0.37

Table 14: Results of Estimates of Further Extensions of the Baseline Model

We also simulated the job-ladder model with multiple offers and found that the moments of the prior wage distribution were similar to the ones in the baseline estimation: the mean of the prior wage distribution was 2.88, compared to 2.91 in the baseline, and the standard deviation remained unaffected at 0.52. The slightly lower mean of the prior wage distribution is due to the fact that in the face of multiple job offers, selection of job offers based on n rather than y becomes somewhat more pronounced.

D.7   Log-additivity of y and n

Our specification of multiplicative interaction of wage and non-wage job values, with $v = y + n$, implies that the non-wage value n is more important for high-y jobs. In this extension, we consider the alternative of additive interaction, with $v = \log(\exp(y) + \exp(n))$, and estimate the model by matching simulated moments. The results in the third column of Table 14 show that the estimates of μ_η and σ_η differ from our baseline values because of their different interpretations. The other parameters tend to be similar, though not exactly: σ_y and σ_r is smaller, whereas σ_x is slightly larger in the modified model, and κ is slightly smaller, too. The implied dispersion of v is quite a bit smaller, which is in part due to the lower dispersion in y.

The fit of the model is substantially worse compared to the baseline model. With the weighted squared sum of deviations of the model moments from the data moments as a measure of goodness of fit, the baseline model yields a value of 5.46, whereas this model extension a value of 11.06. The additive model does not only in worse in fitting the shape of the acceptance function, but also the moments related to the distributions of y and r. The reason is that it is more difficult to fit the acceptance function for high values of σ_y and thus the estimation trades off lower values of σ_y for a better fit of the moments related to the job acceptance decision. We conclude that log-additivity of y and n is a better specification than additivity.

D.8   Dispersion in μ_η

We extend our baseline model by allowing for heterogeneity among jobseekers in the mean of the non-wage values, μ_η. We model the dispersion by specifying that μ_η is normally distributed with mean μ_{μ_η} and standard deviation σ_{μ_η}:

$$\mu_\eta \sim \mathcal{N}(\mu_{\mu_\eta}, \sigma_{\mu_\eta}^2), \tag{112}$$

which implies that

$$\eta \sim \mathcal{N}(\mu_{\mu_\eta}, \sigma_\eta^2 + \sigma_{\mu_\eta}^2) \tag{113}$$

$$\eta - \mu_\eta \sim \mathcal{N}(0, \sigma_\eta^2). \tag{114}$$

D.8.1 Moment conditions

The first five moment conditions regarding the distribution of $\hat{y}$ and $\hat{r}$ remain the same, but the moment conditions regarding the acceptance decision change to:

$$A(d_i) = 1 - \Phi(0, \mu_{m|d_i}, \sigma_{m|d_i}), i = 1, 2, 3, 4, 5 \tag{115}$$

$$J = \int \int^r \left[\Phi(0, \mu_{p|v}, \sigma_{p|v}) \frac{\phi(v, \mu_v, \sigma_v)}{\Phi(r, \mu_v, \sigma_v)} dv \right] \phi(r, \mu_r, \sigma_r) dr, \tag{116}$$

where $m = v - r$, $d = \hat{y} - \hat{r}$ and $p = (\eta - \mu_\eta) - (y - \mu_y)(1 + \kappa)$. The functions $\mu_{m|d}$, $\sigma_{m|d}$, $\mu_{p|v}$ and $\sigma_{p|v}$ are determined by the parameters μ_y, μ_r, σ_y, σ_r, μ_{μ_η}, σ_η, σ_{μ_η}, κ, $\sigma_{\epsilon_{\hat{y}}}^2$ and $\sigma_{\epsilon_{\hat{r}}}^2$, as follows:

$$\mu_{m|d} = \mu_m + \frac{\sigma_{m,d}}{\sigma_d^2}(d - \mu_d) \tag{117}$$

$$\sigma_{m|d}^2 = \sigma_m^2 - \frac{\sigma_{m,d}^2}{\sigma_d^2} \tag{118}$$

$$\mu_{p|v} = \frac{\sigma_\eta^2 - (1 - \kappa)(1 + \kappa)\sigma_y^2}{\sigma_v^2}(v - \mu_v) \tag{119}$$

$$\sigma_{p|v}^2 = \sigma_y^2(1 + \kappa)^2 + \sigma_\eta^2 - \frac{(\sigma_\eta^2 - (1 - \kappa)(1 + \kappa)\sigma_y^2)^2}{\sigma_v^2}, \tag{120}$$

where

$$\mu_d = \mu_y - \mu_r \tag{121}$$

$$\mu_m = \mu_y + \mu_{\mu_\eta} - \mu_r \tag{122}$$

$$\sigma_d^2 = \sigma_y^2 + \sigma_{\epsilon_{\hat{y}}}^2 + \sigma_r^2 + \sigma_{\epsilon_{\hat{r}}}^2 \tag{123}$$

$$\sigma_m^2 = (1 - \kappa)^2 \sigma_y^2 + \sigma_\eta^2 + \sigma_{\mu_\eta}^2 + \sigma_r^2 \tag{124}$$

$$\sigma_{m,d} = (1 - \kappa)\sigma_y^2 + \sigma_r^2 \tag{125}$$

$$\mu_v = \mu_y + \mu_{\mu_\eta} \tag{126}$$

$$\sigma_v^2 = (1 - \kappa)^2 \sigma_y^2 + \sigma_\eta^2 + \sigma_{\mu_\eta}^2. \tag{127}$$

D.8.2 Estimation and results

The model does not identify the dispersion of μ_η, as this would require a large sample of multiple job offers per person, which the survey lacks. Instead, we perform a sensitivity check, where we calibrate σ_{μ_η} to (1) the same and (2) twice as large as the measurement error in $\hat{r}$. The results in the Table 14 show that similar to measurement error the dispersion μ_η tends to lead to lower values of κ. The estimated dispersion of η at the individual level, σ_η, is slightly larger. The reason is that the lower value of κ not only affects the shape of the acceptance function, but also tends to lower the fraction of rejections for non-wage reasons. Therefore, higher values of σ_η compensate for the latter in the estimation. All other parameters are essentially unaffected by the dispersion in μ_η.

D.9 Normality of x, y and r

As described in Section 5, the model misses the skewness (especially, of $\hat{y}$), and excess kurtosis, of $\hat{y}$ and $\hat{r}$. To achieve a better fit, we estimate a model whose distributions of x, y and r follow the skew-normal distribution, which has three parameters: ξ (location), ω (scale) and α (shape). The location parameter identifies the mean of the distribution, the scale parameter identifies the variance, and the shape parameter the skewness and kurtosis. Compared to the baseline model with normal distributions, we have three additional parameters in the estimation. We identify these additional parameters in the estimation, by targeting the skewness and the kurtosis for $\hat{r}$ and $\hat{y}$ as additional moments from the data. We estimate the model by matching simulated moments.

We impose an upper limit of 50 for the αs—the relevant parameter for skewness and kurtosis is $\tilde{\delta} = \frac{\alpha}{\sqrt{1+\alpha^2}}$, which converges to 1 for high values of α. A value of $\alpha = 50$ implies $\tilde{\delta} = 0.9998$, very close to the maximum skewness with the skew-normal distribution. Because of this property, the estimate and sampling distribution of α are meaningless. Instead, we report the value and standard error of δ and the corresponding values of the skewness and excess kurtosis, with standard errors.

As shown in Table 15, our estimates for the shape parameter δ are 0.969 for the distribution of x, 0.9998 for the distribution of y, and 0.922 for the distribution of r. Bootstrapped standard errors, based on 10 repetitions, for the δs are 0.016, 0.000, and 0.658, respectively. δ_r has a relatively high standard error, but with limited impact on the standard error of the skewness of $\hat{r}$, because the distribution of $\hat{r}$ is dominated by the moments of x rather than r due to x's much higher variance. The standard errors of the model's fitted skewness and kurtosis were in the range of 0.05-0.06 for the distributions of both $\hat{y}$ and $\hat{r}$, which is rather small.

Parameters	Estimate	S.e.
δ_x	0.969	0.016
δ_y	0.9998	0.000
δ_r	0.922	0.658

Fitted moments	Estimate	S.e.
Skewness of $\hat{y}$	0.51	0.06
Skweness of $\hat{r}$	0.58	0.05
Kurtosis of $\hat{y}$	3.30	0.05
Kurtosis of $\hat{r}$	3.41	0.05

Table 15: Parameter Estimates of δ and Additional Fitted Moments of $\hat{y}$ and $\hat{r}$

For consistency with the results for other extensions in Table 14, the column for the extension to skew-normal distributions does not show the estimated parameter values of the those distributions (ξ, ω, and α), but rather shows the fitted values of the moments that correspond to the parameters in the baseline estimation (means and standard deviations). These moments are μ_y, μ_r, σ_x, σ_y and σ_r, which are all the means and standard deviations of the skew-normal distributions. The results show that these 5 fitted moments as well as the parameters μ_η, σ_η and κ are similar to the baseline parameter estimates. Therefore, we conclude that assuming normality in x, y and r does not lead to meaningful biases in the estimated parameters. For example, our key conclusion that non-wage job values have greater dispersion than the offered wage, conditional on personal productivity, applies with equal force in the model extended to skew-normal distributions of x, y and r.

The skew-normal distributions do a substantially better job of fitting the skewness and kurtosis of the distributions of $\hat{y}$ and $\hat{r}$. The skewness of $\hat{y}$ is 0.51 and the skewness of $\hat{r}$ is 0.57 in the model, compared to 0.79 and 0.59 in the data. The excess kurtosis in the model is 0.30 for $\hat{y}$ and 0.41 for $\hat{r}$, compared to 0.30 and 0.06 in the data. In the baseline model, the skewness and excess kurtosis are zero for both variables, as they are taken to be normal. The extended model cannot match both the skewness and kurtosis of these distributions, as those moments are controlled by the single parameter α. A second reason is that the skew-normal distribution has an upper bound in skewness, which is binding for y. The model could fit the skewness in $\hat{y}$ better by allowing for more skewness in x but this would lead to too much skewness in $\hat{r}$.

E   The Distribution of Values in Non-Market Activities

This section presents the details of the job-ladder model where search on the job is less effective than search while unemployed, and derives the implied distribution of values in non-market activities h for a given distribution of reported reservation wages r. We start by defining the value functions of the unemployed and employed worker of type (h, x), where h stands for the flow value of non-work and x for personal productivity. The value functions in discrete time become:

$$
\begin{aligned}
U(h, x) &= \hat{h}(h, x) + \frac{1}{1+\rho} \max_{r_v} \left((1-s)\lambda_u \int_{r_v} W(h, x, \tilde{v}) dF_v(\tilde{v}|x) \right. \\
&+ \left. ((1 - \lambda_u + \lambda_u F_v(r_v) + s\lambda_u(1 - F_v(r_v|x)))U(h, x) \right) \\
W(h, x, v) &= e^{\hat{v}(v,x)} + \frac{1}{1+\rho} \left((1-s)\lambda_e \int_v W(h, x, \tilde{v}) dF_v(\tilde{v}|x) \right. \\
&+ \left. (1-s)(1 - \lambda_e + \lambda_e F_v(v|x))W(h, x, v) + sU(h, x) \right),
\end{aligned}
$$

where ρ is the discount rate, λ_u the offer rate while unemployed, λ_e the offer rate while employed, s the separation rate, $U(h, x)$ the value of being unemployed for type (h, x), $W(h, x, v)$ the value of being employed with flow value $e^{\hat{v}(v,x)}$ for type (h, x). Note that $\hat{v}(v, x)$ is the log of the flow value during employment, whereas $\hat{h}(h, x)$ is the flow value during unemployment, which is expressed in absolute values in order to allow for negative values. Under the assumption of proportionality, $\hat{h}(h, x) = he^x$, $\hat{v}(v, x) = v + x$ and $dF_v(v|x) = dF_v(v)$, and thus:

$$
\begin{aligned}
U(h, x) &= he^x + \frac{1}{1+\rho} \max_{r_v} \left((1-s)\lambda_u \int_{r_v} W(h, x, \tilde{v}) dF_v(\tilde{v}) \right. \\
&+ \left. ((1 - \lambda_u + \lambda_u F_v(r_v) + s\lambda_u(1 - F_v(r_v)))U(h, x) \right) \\
W(h, x, v) &= e^{v+x} + \frac{1}{1+\rho} \left((1-s)\lambda_e \int_v W(h, x, \tilde{v}) dF_v(\tilde{v}) \right. \\
&+ \left. (1-s)(1 - \lambda_e + \lambda_e F_v(v))W(h, x, v) + sU(h, x) \right),
\end{aligned}
$$

The reservation value r_v satisfies $U(h, x) = W(h, x, r_v(h, x))$ and thus:

$$
\begin{aligned}
U(h, x) & = he^x + \frac{1}{1+\rho}\Bigg((1-s)\lambda_u \int_{r_v(h,x)} W(h,x,\tilde{v})dF_v(\tilde{v}) \\
& \quad + \; ((1-\lambda_u + \lambda_u F_v(r_v(h,x))) + s\lambda_u(1-F_v(r_v(h,x))))W(h,x,r_v(h,x)) \Bigg) \\
& = e^{r_v(h,x)+x} + \frac{1}{1+\rho}\Bigg((1-s)\lambda_e \int_{r_v(h,x)} W(h,x,\tilde{v})dF_v(\tilde{v}) \\
& \quad + \; (1-s)(1-\lambda_e + \lambda_e F_v(r_v(h,x)))W(h,x,r_v(h,x)) + sW(h,x,r_v(h,x)) \Bigg),
\end{aligned}
$$

which can be simplified to

$$
\begin{aligned}
U(h, x) & = he^x + \frac{1}{1+\rho}\Bigg((1-s)\lambda_u \int_{r_v(h,x)} W(h,x,\tilde{v})dF_v(\tilde{v}) \\
& \quad + \; ((1-(1-s)\lambda_u(1-F_v(r_v(h,x))))W(h,x,r_v(h,x)) \Bigg) \\
& = e^{r_v(h,x)+x} + \frac{1}{1+\rho}\Bigg((1-s)\lambda_e \int_{r_v(h,x)} W(h,x,\tilde{v})dF_v(\tilde{v}) \\
& \quad + \; ((1-(1-s)\lambda_e(1-F_v(r_v(h,x))))W(h,x,r_v(h,x)) \Bigg)
\end{aligned}
$$

and solving for h, we get

$$
\begin{aligned}
h & = e^{r_v(h,x)} - e^{-x}(\lambda_u - \lambda_e)\frac{1-s}{1+\rho}\Bigg(\int_{r_v(h,x)} W(h,x,\tilde{v})dF_v(\tilde{v}) - (1-F_v(r_v(h,x)))W(h,x,r_v(h,x)) \Bigg) \\
& = e^{r_v(h,x)} - e^{-x}(\lambda_u - \lambda_e)\frac{1-s}{1+\rho}\Bigg(\int_{r_v(h,x)} (W(h,x,\tilde{v}) - W(h,x,r_v(h,x)))dF_v(\tilde{v}) \Bigg).
\end{aligned}
$$

We follow HKV and assume that there is a finite upper bound $\bar{v}$ to the offer distribution. Integrating by parts, as in the online Appendix of HKV, we get

$$h = e^{r_v(h,x)} - e^{-x}(\lambda_u - \lambda_e)\frac{1-s}{1+\rho}\int_{r_v(h,x)}^{\bar{v}}(W(h,x,\tilde{v}) - W(h,x,(r_v(h,x))))dF_v(\tilde{v})$$

$$= e^{r_v(h,x)} - e^{-x}(\lambda_u - \lambda_e)\frac{1-s}{1+\rho}[[(W(h,x,\tilde{v}) - W(h,x,r_v(h,x)))F_v(\tilde{v})]_{r_v(h,x)}^{\bar{v}}$$

$$- \int_{r_v(h,x)}^{\bar{v}}W_v(h,x,\tilde{v})dF_v(\tilde{v})]$$

$$= e^{r_v(h,x)} - e^{-x}(\lambda_u - \lambda_e)\frac{1-s}{1+\rho}\int_{r_v(h,x)}^{\bar{v}}W_v(h,x,\tilde{v})(1 - F_v(\tilde{v}))d\tilde{v},$$

and differentiating $W(h,x,v)$, we get

$$W_v(h,x,\tilde{v}) = \frac{e^{\tilde{v}+x}}{1 - \frac{1-s}{1+\rho}(1 - \lambda_e(1 - F_v(\tilde{v})))},$$

and, therefore,

$$h = e^{r_v(h,x)} - (\lambda_u - \lambda_e)\frac{1-s}{1+\rho}\int_{r_v(h,x)}^{\bar{v}}\left[\frac{1 - F_v(\tilde{v})}{1 - \frac{1-s}{1+\rho}(1 - \lambda_e(1 - F_v(\tilde{v})))}\right]e^{\tilde{v}}d\tilde{v},$$

which can be simplified to

$$h = e^{r_v(h,x)} - (\lambda_u - \lambda_e)(1 - s)\int_{r_v(h,x)}^{\bar{v}}\left[\frac{1 - F_v(\tilde{v})}{\rho + s + (1-s)\lambda_e(1 - F_v(\tilde{v}))}\right]e^{\tilde{v}}d\tilde{v}. \tag{128}$$

and thus $r_v(h,x) = r_v(h)$, which implies that one can define value functions $U(h)$ and $W(h,v)$ such that:

$$U(h,x) = U(h)e^x$$
$$W(h,x,v) = W(h,v)e^x.$$

As explained in the main text, we treat the reported reservation wage as the lowest wage a jobseeker will accept for a job with a reference level of its non-wage value of zero, and thus the reservation wage function is $r(h) = r_v(h)$. The inverse function $H(r) = r^{-1}(h)$, will give the value of non-market activities implied by the job-ladder model for a given observed reservation wage r, and is defined by:

$$H(r) = e^r - (\lambda_u - \lambda_e)(1 - s)\int_r^{\bar{v}}\left[\frac{1 - F_v(\tilde{v})}{\rho + \sigma + (1-s)\lambda_e(1 - F_v(\tilde{v}))}\right]e^{\tilde{v}}d\tilde{v}, \tag{129}$$

where

$$H'(r) = e^r\left(1 + (\lambda_u - \lambda_e)(1 - s)\frac{1 - F_v(r)}{\rho + \sigma + (1-s)\lambda_e(1 - F_v(r))}\right). \tag{130}$$

Therefore, given the distribution of reported reservation wages r, one can find the distribution of values of non-market activities h, by solving:

$$F_r(r) = F_h(H(r)), \tag{131}$$

$$f_r(r) = f_h(H(r))H'(r). \tag{132}$$

F   The Distribution of Wages in the Job-Ladder Model

Let u be the fraction of the labor force unemployed and let $F_e(v|h)$ be the fraction of the labor force employed at a job value not higher than v. Note that $F_e(v|h)$ is not a cdf; rather, $F_e(\infty|h) = 1 - u(h)$, the fraction employed among those with non-work value h. The transition equation for the unemployment rate is

$$u(h)' = s(1 - u(h)) + [1 - (1 - s)\lambda_u(1 - F_v(r_v(h)))]u(h), \tag{133}$$

so the ergodic unemployment rate is

$$u^*(h) = \frac{s}{s + (1 - s)\lambda_u(1 - F_v(r_v(h)))}. \tag{134}$$

The transition equation for the value distribution is

$$F_e(v'|h) = (1 - s)\lambda_u(F_v(v') - F_v(r_v(h)))u + (1 - s)\int_{r_v(h)}^{v'}(1 - \lambda_e + \lambda_e F_v(v'))dF_e(v|h), \tag{135}$$

The first term says that a fraction $(1 - s)\lambda_u(F_v(v') - F_v(r_v(h)))$ of the unemployed find jobs with values not greater than v'. The second term says that a fraction $1 - s$ of those currently employed at value no greater than v do not suffer an exogenous shock sending them into unemployment. Among the survivors, a fraction $1 - \lambda_e$ receive no offer and remain at value $v' = v$. A fraction $\lambda_e F_v(v)$ receive an offer no better than the current job and a fraction $\lambda_e(F_v(v') - F_v(v))$ take a better job with value no greater than v'. Then

$$F_e(v'|h) = (1 - s)\lambda_u(F_v(v') - F_v(h))u(h) + (1 - s)(1 - \lambda_e + \lambda_e F_v(v'))F_e(v'|h), \tag{136}$$

because $F_e(r_v(h)|h) = 0$.

The ergodic distribution F_e satisfies

$$F_e(v|h) = (1 - s)\lambda_u(F_v(v) - F_v(r_v(h)))u^*(h) + (1 - s)(1 - \lambda_e + \lambda_e F_v(v))F_e(v|h). \tag{137}$$

Integrating over h, we have

$$F_e(v) = \int^v (1 - s)\lambda_u(F_v(v) - F_v(r_v(h)))u^*(h)dF_h(h) + (1 - s)(1 - \lambda_e + \lambda_e F_v(v))F_e(v). \tag{138}$$

where

$$F_e(v) = \int^v F_e(v|h)dF_h(h).$$

(139)

Finally,

$$F_e(v) = \frac{\int^v (1-s)\lambda_u(F_v(v) - F_v(r_v(h)))u^*(h)dF_h(h)}{1 - (1-s)(1 - \lambda_e + \lambda_e F_v(v))}.$$

(140)

F.1 F_v and dF_v

The distribution of job values among job offers is

$$F_v(v) = \int F_y(v - n)dF_n(n).$$

(141)

and its differential is

$$dF_v(v) = \int dF_y(v - n)dF_n(n).$$

(142)

F.2 F_e

The distribution of job values among the employed is

$$F_e(v) = \frac{N(v)}{D(v)},$$

(143)

where

$$N(v) = \int^v (1-s)\lambda_u(F_v(v) - F_v(r_v(h)))u^*(h)dF_h(h),$$

(144)

$$D(v) = 1 - (1-s)(1 - \lambda_e + \lambda_e F_v(v)),$$

(145)

and

$$u^*(h) = \frac{s}{s + (1-s)\lambda_u(1 - F_v(r_v(h)))}.$$

(146)

Then

$$dN(v) = (1-s)\lambda_u dF_v(v) \int^v u^*(h)dF_h(h)$$

(147)

and

$$dD(v) = -(1-s)\lambda_e dF_v(v).$$

(148)

Finally,

$$dF_e(v) = \frac{dN(v)}{D(v)} - \frac{N(v)dD(v)}{D(v)^2}.$$

(149)

F.3 $F_w(w)$

$$F_w(w) = \frac{1}{1 - u^*} \int \frac{\int_{-\infty}^{w} f_n(v - y)dF_y(y)}{\int_{-\infty}^{\infty} f_n(v - y)dF_y(y)} dF_e(v), \tag{150}$$

where $u^* = \int u^*(h)dF_h(h)$. Note that

$$F_w(w) = \frac{1}{1 - u^*} \int F_y(w|v)dF_e(v), \tag{151}$$

and thus with log-normal distributed variables, $F_y(y|v)$ is the cdf of a normal distribution with the following mean and variance:

$$\mu_{y|v} = \mu_y + \frac{(1 - \kappa)\sigma_y^2}{(1 - \kappa)^2\sigma_y^2 + \sigma_\eta^2}(v - \mu_y - \mu_\eta) \tag{152}$$

$$\sigma_{y|v}^2 = \frac{\sigma_y^2\sigma_\eta^2}{(1 - \kappa)^2\sigma_y^2 + \sigma_\eta^2}. \tag{153}$$

F.4 $F_{\hat{r}}(\hat{r})$, $F_{\hat{y}}(\hat{y})$, $F_{\hat{w}}(\hat{w})$

$$F_{\hat{r}}(\hat{r}) = \int F_h(H(\hat{r} - x))dF_x(x), \tag{154}$$

$$F_{\hat{y}}(\hat{y}) = \int F_y(\hat{y} - x)dF_x(x), \tag{155}$$

and

$$F_{\hat{w}}(\hat{w}) = \int F_w(\hat{w} - x)dF_x(x). \tag{156}$$

G   Measuring the Job-to-Job Transition Rate

G.1   The rate in the CPS data

Following Fallick and Fleischman (2004), we measure job-to-job transitions in the CPS data using information from a question that asked whether a person worked at the same employer as last month for those who were employed both this and last month (the variable *puiodp1*). We compute the job-to-job transition rate as the fraction of workers changing employers between two consecutive monthly CPS interviews, excluding the 1st and 5th interview because no information on previous employer is available in these interviews. We restrict our sample to those of age 20 to 65 and compute the average monthly job-to-job transition rate in the years 2009 and 2010, which is 0.019.

G.2 Computing the monthly transition rate in the model

In the model, we compute the monthly job-to-job transition rate as measured in the CPS, T_{ee}^{CPS}, as:

$$T_{ee}^{CPS} = \frac{P_{ee}^{1m} + P_{eue}^{1m}}{P_{e}^{1m}} \tag{157}$$

where P_{ee}^{1m} is the probability of at least one job-to-job transition over the period of 1 month, P_{eue}^{1m} is the probability of at least one intermittent spell of unemployment over the period of 1 month, and P_{e}^{1m} is the probability of being employed in the next CPS interview in one month from now. We compute the monthly probabilities as a weighted average of a 4-week probability and a 5-week probability, given that the average month in a year has a duration of 4.33 weeks, that is,

$$
\begin{aligned}
P_{ee}^{1m} &= 0.67 P_{ee}^{4w} + 0.33 P_{ee}^{5w} \\
P_{ee}^{1m} &= 0.67 P_{eue}^{4w} + 0.33 P_{eue}^{5w} \\
P_{ee}^{1m} &= 0.67 P_{e}^{4w} + 0.33 P_{e}^{5w}.
\end{aligned}
$$

The probability of at least one job-to-job transition, P_{ee}^{4w}, is the probability of not separating in the next 4 weeks multiplied by one minus the probability of no job-to-job transition in the next 4 weeks:

$$P_{ee}^{4w} = (1-s)^4 (1 - (1-T_{ee})^4),$$

and analogously for the 5-week period.

The probability of at least one intermittent spell of unemployment, P_{eue}^{4w}, is

$$
\begin{aligned}
P_{eue}^{4w} =&(1-s)^2 sf + (1-s)s(f(1-s) \\
&+ f(1-f)) + sf((1-s)^2 + sf) + s(1-f)(f(1-s) + f(1-f)),
\end{aligned}
$$

and analogously for the 5-week period.

The probability of being employed in 4 weeks from now, P_{e}^{4w}, is the sum of the probability of not being separated in the next 4 weeks and the probability of at least one intermittent spell of unemployment in the next 4 weeks:

$$P_{e}^{4w} = (1-s)^4 + P_{eue}^{4w},$$

and analogously for the 5 week period.

We take the weekly separation rate, $s = 0.0041$, and compute the average weekly job-finding rate and job-to-job transition rate in the model,

$$f = \lambda_u(1-s) \int (1 - F_v(r))dF(r) \tag{158}$$

$$T_{ee} = \int T_{ee}(v)dF(v), \tag{159}$$

and then compute the monthly rate as explained above. For the baseline specification, we get $P_{ee}^{1m} = 0.0200$, $P_{eue}^{1m} = 0.0012$ and $P_e^{1m} = 0.9836$, and thus $T_{ee}^{CPS} = 0.0216$. This estimate implies that time aggregation bias accounts only for 0.12 percentage points (or 5.6 percent) of the job-to-job transition rate measured in the CPS data.