Job Dispersion and Compensating Wage Differentials

Paul Sullivan and Ted To

US Bureau of Labor Statistics

November 2013

Abstract

The empirical literature on compensating wage differentials has a mixed history. While there have been some successes, much of this research finds weak support for the theory of equalizing differences. We argue that these negative results can be explained by bias in estimated compensating wage differential regressions. The source of this bias is dispersion in total job values, or "job dispersion." We begin by using a simple theoretical model to demonstrate how dispersion in wages and non-wage values of jobs can lead to biased hedonic estimates of the marginal-willingness-to-pay (MWP) for non-wage job characteristics. Next, we quantify this bias by estimating a structural on-the-job search model that allows jobs to be differentiated by both wages and job-specific non-wage utility. Using simulated data from the model, we conduct a detailed analysis of the sources of bias in traditional hedonic wage estimates. Estimates of the MWP for non-wage job characteristics are severely attenuated. While worker heterogeneity and job dynamics are important sources of bias, a significant proportion can only be explained by randomness in job offers.

Keywords: compensating wage differentials, theory of equalizing differences, revealed preference, on-the-job search

JEL codes: J3, J42, J64

1 Introduction

Equally able workers in a frictionless labor market earn wages that equalize differences in the value of non-wage job characteristics between different jobs (Smith, 1776; Rosen, 1986). The argument is simple and compelling but empirical support for the theory is weak. For example, Brown (1980) summarized existing results at that time, and found that compensating wage differential estimates "provided rather limited support for the theory [of equalizing differences]" and that the most common explanation is inadequate control for unobserved worker ability. Using the National Longitudinal Survey Young Men's sample, he then showed that even after controlling for individual characteristics, compensating wage differential estimates are "often wrong signed or insignificant."

While unobserved heterogeneity in worker ability certainly contributes to biased compensating wage differential estimates, we argue that the often poor performance of hedonic estimates of the marginal-willingness-to-pay (MWP) for job characteristics is due to dispersion in equilibrium total job values, or "job dispersion" for short. Indeed, unobserved heterogeneity can increase job dispersion, increasing the bias from traditional compensating wage differential estimates. Our fundamental insight is that in models with frictions, identical workers can receive different equilibrium compensation packages (total value of wage and non-wage amenities) so that jobs are dispersed. These dispersed jobs can result in biased hedonic estimates of the MWP.

Not surprisingly, job dispersion is closely related to the extensive empirical literature on wage dispersion (some early examples include Dunlop, 1957; Brown and Medoff, 1989; Groshen, 1991). Theoretical explanations for wage dispersion have included on-the-job-search (Burdett and Mortensen, 1998), efficiency wages (Albrecht and Vroman, 1998) and oligopsony (Bhaskar and To, 2003). But in such models, dispersed wages imply that total job values or jobs are dispersed as well.

We begin by demonstrating how job dispersion can lead to inaccurate compensating wage differential estimates. To do so, we contrast the results from Rosen's (1986) competitive clean/dirty job model with a Hotelling model of the labor market where jobs are differentiated by commuting time (Bhaskar et al., 2002). While hedonic estimates based on the competitive model yield unbiased estimates of workers' willingness-to-pay, the bias from hedonic estimates based on the Hotelling model are in excess of 50%. While this simple example illustrates the basic intuition behind job dispersion and biased compensating wage differential estimates, it lacks the richness[1] necessary for an in-depth analysis.

[1]In particular, simple oligopsony models cannot generate unemployment-to-employment, job-to-job and

To develop a deeper understanding of the sources of job dispersion and bias we estimate a structural search model with wages and non-wage utility.[2] As in Sullivan and To (2013) and Taber and Vejlin (2013), the importance of non-wage utility to workers is identified through revealed preference. Using our estimated parameters, we simulate a dataset to perform a detailed evaluation of the reduced form hedonic wage regressions that are commonly used to estimate compensating wage differentials. Our compensating wage differential estimates have the correct sign but are moderately to severely biased vis the true MWP. This bias arises because for a worker moving out of unemployment, almost all acceptable jobs offer total utilities that are strictly greater than an unemployed worker's reservation utility. In other words, because accepted jobs are dispersed, accepted wages and non-wage amenities do not directly reveal workers' MWP.

On-the-job search creates further job dispersion as workers climb a "job ladder" by moving to jobs that offer higher utility. In our model, employed workers' reservation utilities are given by the utilities at their current jobs. As a result, variation across workers in the total utility of the first job out of unemployment translates directly into variation across workers in reservation utility levels. This dispersion in reservation utilities further increases the cross-sectional dispersion of job values through two channels. First, workers on the lower end of the job distribution are more likely receive a superior job offer, shifting the job distribution away from the lower end. Second, as these workers move to better jobs, the job distribution shifts towards the higher end.

Several papers have studied non-wage job characteristics using a search framework. Hwang et al. (1998) use a theoretical equilibrium search model to demonstrate that compensating wage differential regressions may give misleading estimates of the MWP. Gronberg and Reed (1994) develop a method for estimating the MWP that relies on job duration data. Dey and Flinn (2005, 2008) estimate search models that include health insurance as well as wages. Their first paper focuses on studying "job lock" using an equilibrium search model, and their second paper focuses on household job search and health insurance. Bonhomme and Jolivet (2009) estimate a search model that includes a number of non-wage job characteristics. Taber and Vejlin (2013) estimate a model that quantifies the contributions of comparative advantage (Roy model), human capital, and utility from non-wage job

employment-to-unemployment transitions observed in longitudinal data sets.

[2]In order to avoid potential problems with excluded job characteristics (e.g., correlation between job risk and health insurance would understate the value of health insurance), we bundle of all relevant job characteristics into "non-wage utility." Moreover, many job characteristics are intangible, unobserved or heterogeneously valued (Bhaskar and To, 1999; Bhaskar et al., 2002) and it would be impossible to accurately capture all of them.

attributes.[3]

The empirical literature on compensating wage differentials has focused on heterogeneous ability as the primary explanation for the weak support for the theory of equalizing differences. Our analysis suggests that this focus is not unwarranted since much of the bias in compensating wage differential estimates can be attributed to heterogeneous worker ability. Nevertheless, although worker heterogeneity can increase job dispersion, it is not necessary for biased compensating wage differential estimates. Indeed, under ideal circumstances, 33 percent of cross-sectional job dispersion in our simulated dataset can only be due to the inherent dispersion of job offers.

In the following section, we illustrate the basic intuition for why job dispersion results in inaccurate MWP estimates by contrasting the Rosen competitive clean/dirty job model with a simple Hotelling labor model in which job dispersion arises naturally. Next, we lay out a partial equilibrium model of on-the-job search with preferences for non-wage job characteristics in Section 3. Then in Section 4, we discuss the dataset used to estimate our partial equilibrium model and in Section 5 we discuss our econometric methodology and some important identification issues. In Section 6 we present our parameter estimates and in Section 7 we analyze the estimation of compensating differentials using simulated datasets. Section 8 concludes.

2 Job Dispersion

Consider a utility function over wages and non-wage job characteristics:

$$U_i = U(w, \boldsymbol{\xi}) \tag{1}$$

where w is the wage, $\boldsymbol{\xi}$ is a vector of non-wage job characteristics and the utility function, U, is increasing in w. As far as the workers' job choice decisions are concerned, U_i perfectly summarizes all of the information embodied in w and $\boldsymbol{\xi}$. Since total utility, not its individual components, determine job choice, for all intents and purposes, U_i is the job.

Consider the theory of equalizing differences using Rosen's (1986) "clean" and "dirty" job model, so that $\xi = 0$ for clean jobs and $\xi = 1$ for dirty jobs and for any w, $U(w, 0) > U(w, 1)$. In a competitive labor market, $U(w^0, 0) = U(w^1, 1) \equiv U^*$ so that the competitive wages offered for clean and dirty jobs are $w^0 < w^1$. Notice that every worker gets $U_i = U^*$ so there

[3] Similar to our approach, Taber and Vejlin (2013) incorporate a general, job-specific non-wage utility match effect rather than modeling specific non-wage job characteristics.

is no job dispersion. The standard wage hedonic is given by:

$$w_i = \alpha + \beta \xi_i + e_i.^4 \tag{2}$$

Since $w_i = w^0$ ($w_i = w^1$) when worker i works for a clean (dirty) employer, estimating this equation yields $\hat{\beta} = w^1 - w^0$ so that $\hat{\beta}$ represents a worker's willingness to pay for a "clean" job.[5]

However, if total utility, U_i, is not identical across jobs or workers then jobs are dispersed. But if jobs are dispersed, cross-sectional wage and amenity data do not directly reveal worker preferences over wages and amenities so that a regression equation such as (2) may lead to poor MWP estimates. The notion that total job values or jobs can be dispersed is quite general and arises naturally in a variety of theoretical models including oligopsony, job search and efficiency wage models.

An Example

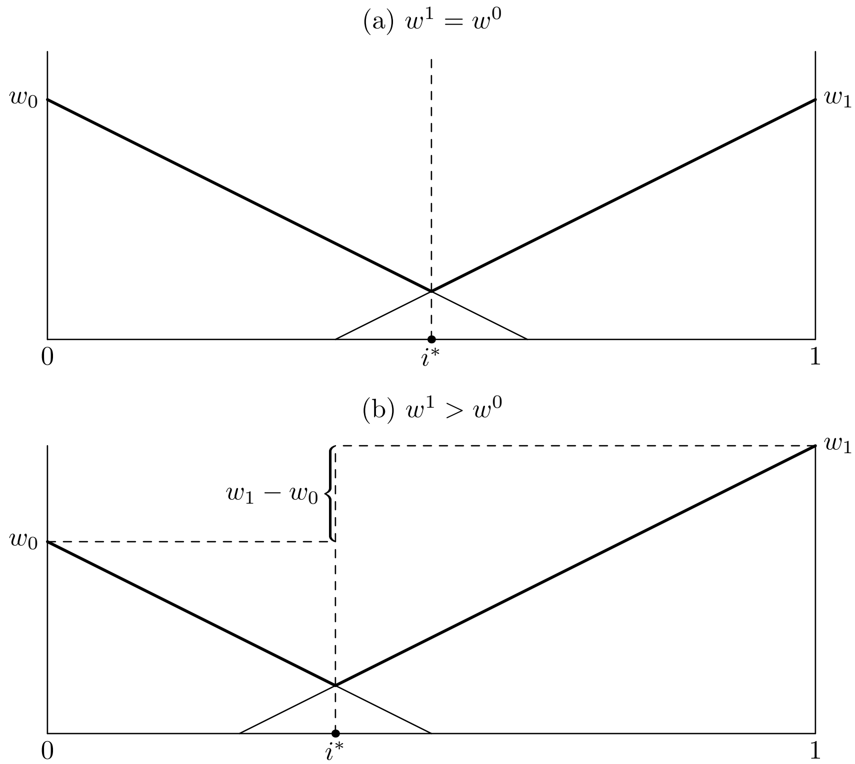
Consider a simple Hotelling labor market as described in Bhaskar et al. (2002). Equally able workers are uniformly distributed along the $[0, 1]$ interval and two employers, 0 and 1, who pay wages w^0 and w^1 are located at either end. Workers pay a "transportation cost" of t per unit of distance to go to work so that a worker located at i pays ti to work for employer 0 and $t(1 - i)$ to work for employer 1. In this case the distance to work, $\xi_i = i$ or $\xi_i = 1 - i$, is a negative job characteristic and the known MWP for ξ is positive and equal to t. A worker who works for employer $k = 0, 1$ gets utility that is the wage net of transportation costs or $U(w_i, \xi_i) = w_i - t\xi_i$. Figure 1 shows two examples where the downward and upward sloping line segments represent the worker's total utility. The worker who is indifferent between working for employers 0 and 1 is located at i^*. Workers located at 0 and 1 work for employers 0 and 1 and pay no transportation costs. As workers travel farther to work, they pay higher transportation costs and get lower utility.

If employers are equally productive then they pay the same wage and evenly split the labor market (Figure 1a). Even though wages are identical, jobs ($U_i = w_i - t\xi_i$) are dispersed. A standard compensating wage differential estimate (equation (2)) of the marginal-willingness-to-pay for commuting distance yields $\hat{\beta} = 0$, even though the true MWP in the model is

[4] Here and in the following, our notation reflects the fact that we assume hedonic regressions are estimated using individual level data.

[5] In Rosen's (1986) original treatment, he allowed for heterogeneous utility functions and $U_i = U^*$ for marginal workers so that $\hat{\beta}$ represents the marginal worker's willingness-to-pay for a clean job.

Figure 1: Hotelling job values



equal to t.

The simple hotelling model also generates biased hedonic estimates of the MWP in the more general case where employers differ in productivity. For example, if employer 1 is more productive than employer 0, it offers a higher wage rate ($w^1 > w^0$) and employs a larger share of the labor market (Figure 1b). On average, workers employed at firm 0 travel shorter distances than those working for firm 1, so wages are positively correlated with commuting distance. As a result of this positive correlation, estimating a hedonic wage equation by OLS yields an estimated MWP with the correct sign (i.e., $\hat{\beta} > 0$). However, this estimate is strictly bounded above by the true MWP of t.[6] This simple duopsony analysis extends to the case of more than two employers (Bhaskar and To, 2003).

While this example succinctly illustrates the basic intuition of how job dispersion can lead to attenuated MWP estimates, for the remainder of this paper, we elaborate on com-

[6] In fact, the upper-bound on $\hat{\beta}$ is $\frac{t}{2}$: $\hat{\beta}$ assumes is maximum value of $\frac{t}{2}$ when $i^* = \frac{3\pm\sqrt{3}}{6}$ and as $i^* \to 0$ or $i^* \to \frac{1}{2}$, $\hat{\beta} \to 0$.

pensating wage differentials and job dispersion using a partial equilibrium search framework. A search model's explicit dynamics play a key role in understanding the various sources of job dispersion observed in cross-section.

3 The Search Model: Wages and Non-Wage Utility

This section presents the search model used to study compensating differentials in a labor market with frictions. The model is set in discrete time. Agents maximize the discounted sum of expected utility over an infinite time horizon in a stationary environment. In each time period, individuals occupy one of two states: employment or unemployment.[7]

3.1 Preferences and the Total Value of a Job

The utility received by an employed agent is determined by the log-wage, w, and the match-specific non-wage utility flow, ξ. The one-period utility from employment is

$$U(w, \xi) = w + \xi$$

where both w and ξ are specific to a particular match between a worker and employer, and are constant for the duration of the match. A job offer consists of a random draw from the distribution $F(w, \xi)$, which is a primitive of the model. Since w and ξ are additively separable in the utility function, it is convenient to define the agent's decision problem in terms of total utility, $w + \xi$, where $U(w, \xi) \equiv U$ and U is distributed as $H(U)$.[8]

The non-wage match value, ξ, captures the net value of all the non-wage job characteristics associated with a particular job to a specific worker. These characteristics include employer provided benefits (health insurance), tangible job characteristics (commuting time), and intangible job characteristics (friendliness of co-workers). The non-wage match value represents the worker's personal valuation of a job, so in addition to capturing variation in non-wage characteristics across jobs, it also reflects heterogeneity in preferences for job characteristics across workers.

For our purposes, there are two primary advantages to aggregating the value of all non-wage job characteristics into a single index. First, this approach avoids the potential bias

[7]Following the majority of the search literature, the model does not distinguish between unemployment and non-participation in the labor market.

[8]In particular, $H(U) = \int_{-\infty}^{\infty} F_{w|\xi}(U - \xi \mid \xi) f_\xi(\xi) d\xi$ where $F_{w|\xi}$ is the conditional cumulative wage distribution and f_ξ is the unconditional probability density function for ξ.

that could result from focusing on a small number of observable characteristics while ignoring other relevant, but unobserved, job characteristics. As discussed in Rosen (1986), the theory of equalizing differences applies to the wage and the total non-wage value of a job. However, it will not necessarily apply if some job characteristics are excluded. Second, it is likely that workers have heterogeneous preferences over the employer provided benefits and tangible and intangible job characteristics that differentiate jobs.

3.2 Unemployed Search

Unemployed agents search for jobs, which arrive randomly with probability λ_u. The discounted expected value of lifetime utility for an unemployed agent is

$$V^u \quad = \quad b + \delta[\lambda_u E \max\{V^u, V^e(U')\} + (1 - \lambda_u)V^u], \tag{3}$$

where b represents log-unemployment-benefits, and δ is the discount factor. The term $V^e(U')$ represents the expected discounted value of lifetime utility for an agent employed in a job with utility level U'.

3.3 On-the-job Search

In each time period, with probability λ_e an employed agent receives a job offer from an outside firm. The worker may accept the job offer, or reject it and continue working for his current employer. Job matches end with exogenous probability λ_l. When a job ends for this reason, the worker becomes unemployed. With probability λ_{le}, a worker's current job exogenously ends and he receives a job offer from a new employer in the same time period. When this happens, the worker can accept the new offer, or become unemployed. Finally, with probability $(1 - \lambda_e - \lambda_l - \lambda_{le})$ the job does not end exogenously and no new offers are received, so the worker remains in his current job.

The discounted expected value of lifetime utility for a worker who is currently employed in a job with utility level U is

$$\begin{aligned} V^e(U) \quad = \quad & U + \delta[\lambda_e E \max\{V^e(U), V^e(U')\} + \lambda_l V^u \\ & + \lambda_{le} E \max\{V^u, V^e(U')\} + (1 - \lambda_e - \lambda_l - \lambda_{le})V^e(U)]. \end{aligned} \tag{4}$$

The first bracketed term in equation (4), $\lambda_e E \max\{V^e(U), V^e(U')\}$, represents the expected value of the best option available in the next time period for an employed individual who receives a job offer from a new employer.[9] The second bracketed term, $\lambda_l V^u$, corresponds to the case where a job exogenously ends and the worker is forced to enter unemployment. The third bracketed term, $\lambda_{le} E \max\{V^u, V^e(U')\}$, represents the case where the worker is laid-off but also receives a job offer from a new employer. The final bracketed term represents the case where the worker is neither laid-off nor receives an outside job offer.

3.4 Optimal Job Search and Job Dispersion

The optimal search strategies for unemployed and employed workers can be expressed in terms of reservation utilities, which are analogous to reservation wages in a standard income maximizing search model. A utility maximizing unemployed worker will accept any job offer with a one-period utility flow greater than the reservation level, U^* (Appendix A presents the formal derivation of U^*). The reservation utility search strategy implies that the distribution of accepted job offers generated by the model is truncated from below at U^*. Since unemployed agents in the model will choose to work in any job that offers utility level $U > U^*$, subject to the constraints imposed by search frictions, pairs of accepted job offers (w, ξ) do not directly reveal the marginal willingness to pay for non-wage job characteristics. Section 7 expands on this point in considerable detail by using simulated data from the estimated model to examine the performance of a standard compensating wage differential regression in our dynamic model of the labor market.

Employed agents in the model also adopt a reservation utility rule when evaluating outside job offers. In this stationary search environment, optimal decisions for employed agents are based on comparisons of one-period utility flows. When an employed worker receives an offer from an outside firm but does not experience an exogenous job ending, a simple reservation utility strategy is optimal. Since $V^e(U)$ is increasing in U, the rule is to accept the offer if it provides greater utility than the current job ($U' > U$), and reject the offer otherwise ($U' \leq U$). As a result, workers climb a "utility ladder" as they voluntarily move between jobs. Workers make optimal tradeoffs between wages and non-wage utility, but these decisions are constrained by search frictions.

If a worker's job exogenously ends and he receives a new job offer at the same time, which occurs with probability λ_{le}, the situation is identical to the one faced by an unemployed agent who receives a new job offer. As a result, he will choose to accept or reject the offer based on the unemployed reservation utility level U^*. In the remainder of the paper, we

[9] The value function reflects the fact that in this model it is never optimal for a worker to quit a job and enter unemployment.

will refer to direct job-to-job transitions that occur as the result of a simultaneous layoff and job offer as "involuntary" transitions between employers. This terminology reflects the fact that although a direct job-to-job transition occurs, the worker's previous job ended involuntarily (exogenously). For agents in the model, voluntary and involuntary transitions are fundamentally different types of job mobility. When a voluntary job-to-job transition occurs, utility increases ($U' > U$). In contrast, when an involuntary transition occurs, the new job offer is preferable to unemployment ($U' > U^*$), but it may offer lower utility than the previous job which exogenously ended ($U' < U$).

3.5 Heterogeneity in Worker Ability

A primary concern of the existing empirical hedonic wage literature is the effect of heterogeneity in worker ability on estimates of compensating wage differentials (Brown, 1980). For example, Rosen (1986) contends that unobserved worker ability is the primary reason that low paying jobs tend to be the "worst" jobs. Using a competitive framework, Hwang et al. (1992) and Han and Yamaguchi (2012) show that unobserved worker productivity can significantly bias compensating wage differential estimates.

Hedonic studies tend to include as many worker characteristics as possible but nevertheless it is well known that wage regressions leave a large fraction of variation in wages unexplained. Instead, we minimize worker heterogeneity by selecting a relatively homogeneous sample and in addition allowing for unobserved worker heterogeneity.

We control for heterogeneity in worker ability by allowing the mean of the wage offer distribution, μ_w, to vary across the population. In addition, we also allow the mean of the distribution of match specific utility flows, μ_ξ, to vary across the population. To the best of our knowledge, this is the first study of hedonic wages that allows for the possibility that workers search for jobs in an environment where the mean quality of non-wage job offers varies across workers.

To model worker heterogeneity, we follow Keane and Wolpin (1997) and a large subsequent literature, and assume that the joint distribution of unobserved heterogeneity is a mixture of discrete types. Assume that there are J types of people in the economy, and let π_j represent the proportion of type j in the population. The parameters of the distribution of unobserved heterogeneity, $\{\mu_w(j), \mu_\xi(j), \pi_j\}_{j=1}^{J}$, are estimated jointly along with the other parameters of the model.

One important feature of the discrete mixture distribution is that it allows for a wide range of possible correlations between the mean wage offer and non-wage utility offer faced by workers.[10] For example, if $\mu_w(j)$ and $\mu_\xi(j)$ are positively correlated, then high ability workers tend to receive good (high w and high ξ) job offers. Arguments can be made for either positive (health insurance) or negative (risk of injury) correlation between w and ξ and our discrete mixture distribution provides a great deal of flexibility to match the correlation across our relatively homogeneous population.

4 Data

We use the 1997 rather than the venerable 1979 cohort of the NLSY to estimate our model for two reasons. First, the NLSY97 is more representative of current labor market conditions. Second, the NLSY97 design team incorporated lessons from the NLSY79 and has a more consistent methodology (Pergamit et al., 2001).

The NLSY97 is a nationally representative sample of 8,984 individuals who were between the ages of 12 and 16 on December 31, 1996. Interviews have been conducted annually since 1997. The NLSY97 collects extensive information about labor market behavior and educational experiences which provide the information needed to study the transition from schooling to employment, early career mobility between employers, and the associated dynamics of wages. Individuals enter the estimation sample when they stop attending high school. The information from the annual interviews is used to construct a weekly employment record for each respondent.

We select a particular subset of the NLSY97 in order to minimize unnecessary complications in estimating our model. Women are excluded for the usual reason of avoiding the difficulties associated with modeling female labor force participation. Similarly, in order to avoid issues relating to household search, men who are ever married during the sample period are excluded. Moreover, we use data from interviews up to the 2006 interview and we select workers who have never attended college because low-skilled workers with little work experience can be expected to have little or no bargaining power and hence conform best to our wage-posting model. Thus we focus on young, unmarried, low-skilled men who are at the beginning of their careers. As is standard in the empirical search literature, individuals who ever serve in the military or are self employed are excluded from the sample. Since the maximum age that an individual could reach during the sample period is only 26 years, our results should be viewed as applying to young workers who tend to be quite mobile during

[10] However, it is important to note that we do not impose any particular correlation between μ_w and μ_ξ. The estimated values of the parameters $\{\mu_w(j), \mu_\xi(j), \pi_j\}_{j=1}^J$ determine whether or not the correlation is positive or negative.

this early phase of their career. Whether the results generalize to older workers, or different cohorts of workers, is an open question.

The NLSY97 provides a weekly employment record for each respondent which is aggregated into a monthly[11] labor force history for the purposes of estimation. First, each individual is classified as unemployed or employed full time for each month depending on whether more weeks were spent employed or unemployed during the month.[12] Next, employed individuals are assigned a monthly employer based on the employer that the worker spent the most weeks working for during the month. The monthly wage is the one associated with the monthly employer. The monthly employment record contains a complete record of employment durations, direct transitions between employers that occur without an intervening spell of unemployment, transitions into unemployment, and the growth in wages resulting from mobility between employers.

Since the importance of non-wage job characteristics is identified in part by job-to-job transitions, we are careful to differentiate between those that are voluntary and those that are not. To identify involuntary job-to-job transitions we use the stated reason that a worker left their job. We consider "layoffs," "plant closings," "end of a temporary or seasonal job," "discharged or fired" or "program ended" to be involuntary. While these data may be somewhat noisy, we are reassured by the summary statistics which show that direct transitions we classify as strictly involuntary are more likely to result in a wage decline (Table 1). In addition, on average, workers who make involuntary transitions between employers experience nearly a 2 percent decline in wages. In contrast, wages increase on average by 8 percent at all direct transitions between employers.

The final issue worthy of discussion regarding the data is the treatment of within-job variation in wages. In the NLSY97, when a job persists across survey interviews, which occur approximately one year apart, a new measurement of the wage is taken. If a job does not last across interview years, only the initial measurement of the wage is available. In principle, it would be possible to allow for within-job variation in wages using these data. However, as discussed by Flinn (2002), jobs with observed wage changes are not a random sample from the population, so there are difficult selection issues which must be confronted when estimating an on-the-job wage process using these data. Even more importantly for our purposes, since the NLSY97 is still a relatively short panel, the majority of jobs do not persist across survey years. For these jobs, it is impossible to observe on-the-job wage

[11] For tie-breaking purposes, we use a 5-week month.

[12] Non-participation and unemployment are considered to be the same state for the purposes of aggregating the data. Full time employment is considered to be jobs that involve at least twenty hours of work per week.

growth. More specifically, we only observe a single wage for 72 percent of all jobs in our data. In addition, for our estimation sample we are unable to reject the null hypothesis that mean wage growth is zero within job spells.[13] Given these features of the data, there is little hope of precisely estimating an on-the-job wage growth process. As a result, we restrict wages to be constant within job spells for the purposes of estimation. When multiple wages are reported for a particular job, we use the first reported wage as the wage for the entire job spell. Moreover, for our application, with our focus on young, unskilled workers during a highly mobile stage of their career, constant wages within jobs does not seem unrealistic.

4.1 Descriptive Statistics

This section highlights the key characteristics of the data used to estimate the importance of non-wage job characteristics in determining employment outcomes. It is convenient to describe the labor market histories in the data and the data generated by the search model in terms of employment cycles, as in Wolpin (1992). An employment cycle begins with unemployment and includes all of the following employment spells that occur without an intervening unemployment spell. When an individual enters unemployment, a new cycle begins. In the remainder of the paper, whenever a job is referred to by number, it represents the position of the job within an employment cycle.

Table 1 shows the means and standard deviations of key variables from the sample of the NLSY97 used in this analysis. There are 980 individuals in the data who remain in the sample for an average of 54.2 months, and these people experience an average of 2.88 employment cycles. The top section of the table shows that as individuals move between employers within an employment cycle, the average wage and employment duration increase.[14] The middle section of the table shows that although mean wages increase as individuals move directly between jobs, conditional on switching employers without an intervening unemployment spell there is a 36 percent chance that an individual reports a lower wage at his new job.[15] For individuals who report that the direct transition between employers was involuntary, the mean wage change is negative and the probability of a wage decrease rises

[13]More specifically, we are unable to reject the null hypothesis that the mean of wage growth equals zero at the 5% level. Mean wage growth is computed using the first and last wage present for each job in the NLSY estimation sample.

[14]Statistics are not reported for more than three jobs within a cycle because only a very small number of people have four or more consecutive jobs without entering unemployment.

[15]This number is consistent with existing estimates of the fraction of direct employer-to-employer transitions that involve a wage decrease. Bowlus and Neumann (2006) report that 40 percent of direct transitions involve a wage decrease in the NLSY79.

Table 1: Descriptive Statistics: NLSY97 Data

	Job Number within Cycle		
	Job 1	Job 2	Job 3
Mean log-wage	1.979	2.038	2.061
Standard deviation of log-wage	0.425	0.458	0.457
Mean employment spell duration*	8.939	9.271	9.738
Number of observations	2614	940	382
	Type of Employer Switch		
	All	Involuntary	
Pr(wage decrease) at job-to-job move	0.364	0.460	
Mean Δw at job-to-job switch†	0.081	-0.017	
Median Δw at job-to-job switch	0.074	0.00	
Mean Δw at job-to-job switch $\|\Delta w > 0$	0.359	0.322	
Median Δw at job-to-job switch $\|\Delta w > 0$	0.231	0.211	
Mean Δw at job-to-job switch $\|\Delta w < 0$	-0.327	-0.345	
Median Δw at job-to-job switch $\|\Delta w < 0$	-0.163	-0.206	
	All Jobs		
Mean unemployment spell duration	5.908		
Mean number of cycles per person‡	2.878		
Std. dev. of number of cycles per person	1.793		
Mean total work experience at end of sample period§	40.01		
Fraction of job-to-job transitions that are involuntary	0.151		
Number of people	980		
Mean number of months in sample per person	54.153		

Notes: *All durations are measured in months.
†Δw represents the change in the wage at a job-to-job transition.
‡An employment cycle begins with the first job after an unemployment spell, and includes all subsequent jobs that begin without an intervening unemployment spell.
§This is the across-person mean of total work experience in the final time period. The final time period is either the end of the sample time frame, or the final time period before an observation is truncated due to missing data.

to 46 percent. Measurement error in wages certainly accounts for some fraction of the observed wage decreases at voluntary transitions between employers. However, the prevalence of these wage decreases and the increased probability of observing a wage decline at an involuntary transition both suggest a role for non-wage job characteristics in determining mobility between jobs.

5 Estimation

The parameters of the model are estimated by simulated minimum distance (SMD). This section begins by specifying the distributional assumptions about the job offer distribution, measurement error in wages, unemployment benefits, and the discount factor needed to estimate the model. Then it explains how the simulated data is generated, describes the estimation algorithm and discusses identification.

5.1 Distributional Assumptions and Exogenous Parameters

The Job Offer Distribution and Measurement Error in Wages

Estimating the model requires specifying the distribution $F(w, \xi)$. As we noted at the end of Section 3.5, our discrete mixture distribution allows for correlation between log-wage offers and match-specific utility flows across the population. Thus we assume for each type that they are independent. Moreover, it is necessary to make parametric assumptions over $F(w, \xi)$; we assume that they are normally distributed so that,

$$\begin{aligned} F_j(w, \xi) &\sim \Omega_j(w)\Psi_j(\xi) \\ \Omega_j(w) &\sim N(\mu_w(j), \sigma_w) \\ \Psi_j(\xi) &\sim N(\mu_\xi(j), \sigma_\xi), \end{aligned}$$

where j indicates that these distributions vary over the J discrete types of agents in the model.

Wages in typical sources of microeconomic data are measured with error (see Bound et al. (2001) for a comprehensive survey). We account for measurement error by assuming that the relationship between the log-wage observed in the data and the true log-wage is $w^o = w + \varepsilon$, where w^o is the observed log-wage, w is the true log-wage, and $\varepsilon \sim N(0, \sigma_\varepsilon)$

represents measurement error in wages that is independent of the true wage.[16] Based on existing estimates of the extent of measurement error in wages, we set $\sigma_\varepsilon = 0.15$.

Unemployment Benefits and the Discount Factor

Many papers in the search and dynamic labor supply literature have found that the discount factor is either not identified, or is in practice very difficult to estimate. Following these papers, we set the monthly discount factor to $\delta = 0.998$. Finally, estimating the model requires choosing a value for b, the amount of unemployment benefits. The unemployment insurance system in the U.S. is quite complicated, and the details of the program such as eligibility requirements, maximum duration of benefits, and the generosity of benefits varies widely across States (see Kletzer and Rosen, 2006). Kletzer and Rosen also documents that the average replacement rate for UI benefits across the U.S. was 0.36 during the years 1975-2004. Given the complexity of the UI program, we adopt the following stylized model of unemployment benefits, $b(j) = \ln(0.35 \times e^{\mu_w(j)})$. This specification allows unemployment benefits to vary across types, so that agents with higher expected wages receive higher unemployment benefits.

5.2 Data Simulation

As discussed in Section 3, the optimal decision rules for the dynamic optimization problem can be described using simple static comparisons of one-period utility flows. It is straightforward to simulate data from the model using these optimal decision rules without numerically solving for the value functions that characterize the optimization problem.

The first step when simulating the model is to randomly assign each individual in the data to one of the J discrete types that make up the population distribution of unobserved heterogeneity. Next, a simulated career is formed for each individual in the NLSY97 estimation sample by randomly generating job offers and exogenous job endings, and then assigning simulated choices for each time period based on the reservation value decision rules. Computing the reservation utility levels for each type, $\{U^*(j)\}_{j=1}^{J}$, requires numerically solving Equation (A3). The number of time periods that each simulated person appears in the simulated data is censored to match the corresponding person in the NLSY97 data. Measurement error is added to the simulated accepted wage data based on the assumed measurement error process.

[16] Accounting for measurement error in this way is standard in the search literature. See, for example, Stern (1989), Wolpin (1992), and Eckstein et al. (2009).

5.3 Simulated Minimum Distance Estimation

Simulated minimum distance estimation finds the vector of structural parameters that minimizes the weighted difference between vectors of statistics estimated using two different data sets: the NLSY97 data, and simulated data from the model. We use the terminology simulated minimum distance to make it clear that during estimation we match moments from the data (as in the simulated method of moments) and the parameters of an auxiliary model (as in indirect inference).[17] In this application, the auxiliary parameters are the parameters of a reduced form wage regression. In the remainder of the paper, for brevity of notation we refer to all of the statistics from the data that are matched during estimation as moments.

Let $\boldsymbol{\theta} = \{\sigma_w, \sigma_\xi, \lambda_u, \lambda_l, \lambda_e, \lambda_{le}\} \cup \{\mu_w(j), \mu_\xi(j), \pi_j\}_{j=1}^{J}$ represent the parameter vector that must be estimated. The search model is used to simulate S artificial datasets, where each simulated dataset contains a randomly generated employment history for each individual in the sample. The simulated and actual data are each summarized by K moments. The SMD estimate of the structural parameters minimizes the weighted difference between the simulated and sample moments. Let m_k represent the kth moment in the data, and let $m_k^S(\boldsymbol{\theta})$ represent the kth simulated moment, where the superscript S denotes averaging across the S artificial datasets. The vector of differences between the simulated and actual moments is $g(\boldsymbol{\theta})' = [m_1 - m_1^S(\boldsymbol{\theta}), \ldots, m_K - m_K^S(\boldsymbol{\theta})]$, and the simulated minimum distance estimate of $\boldsymbol{\theta}$ minimizes the following objective function,

$$\Phi(\boldsymbol{\theta}) = g(\boldsymbol{\theta})' W g(\boldsymbol{\theta}), \tag{5}$$

where W is a weighting matrix. We use a diagonal weighting matrix during estimation, where each diagonal element is the inverse of the variance of the corresponding moment. We estimate W using a nonparametric bootstrap with 300,000 replications. Bootstrapping the matrix W is convenient because it is not necessary to update the weighting matrix during estimation. Parameter estimates are obtained by minimizing the objective function shown in equation (5) using simulated annealing.[18] Simulated moments are averaged over $S = 25$ simulated datasets. The standard errors are computed using a nonparametric bootstrap using 900 draws from the NLSY97 data.

[17] See Stern (1997) for a survey of simulation based estimation, and Smith (1993) for the development of indirect inference. Recent examples of papers that use this approach to estimating search models include Eckstein et al. (2009) and Yamaguchi (2010).

[18] See Goffe et al. (1994) for a discussion of the simulated annealing algorithm and FORTRAN source code to implement the algorithm. The primary advantage of this algorithm is that it is a global search algorithm that can escape local optima.

5.4 Choice of Moments and Identification

This section discusses the moments targeted during estimation and provides a discussion of how they identify the parameters of the structural model. Table B1 lists the 51 moments from the NLSY97 that are used to estimate the model. This section begins by describing how the wage offer distribution and non-wage utility offer distribution are identified. Next, it demonstrates how the distribution of person-specific unobserved heterogeneity is identified. The section concludes by turning to a discussion of how the transition parameters (λ's) are identified. Throughout this section, we provide examples of how different parameters impact the moments used in estimation to illustrate identification. We do not attempt to provide an exhaustive discussion of the effect of each parameter on every simulated moment.

Identifying the Wage and Non-Wage Offer Distributions

As is standard in the structural search literature, and described in Section 5.1, we must assume a parametric functional form for the job offer distribution, $F(w, \xi)$.[19] For clarity of exposition, we initially abstract away from person-specific unobserved heterogeneity when discussing identification of the job offer distribution. Following this discussion, we explain how the distribution of unobserved heterogeneity is identified.

The mean and standard deviation of the wage offer distribution are identified by moments that describe accepted wages and wage growth from mobility. More specifically, our simulated minimum distance estimation procedure attempts to match the mean and standard deviation of accepted wages, and the mean employment spell duration for the first three employment spells (Panel 1 of Table B1). Recall that employers within a cycle represent a sequence of direct transitions between employers that occur without an intervening spell of unemployment, so mean wages conditional on employer number also provide information about wage growth from job search. As discussed in Barlevy (2008), wage gains from mobility provide useful identifying information about the wage offer distribution. Also, the increase in the average employment duration as workers move between jobs reflects improvements in job match values from on-the-job search.

The extent of variation in non-wage utility, σ_ξ, is identified by moments which summarize the relative importance of wages and non-wage utility in determining job mobility. In particular, moments relating to wage changes in job-to-job transitions are particularly informative about the magnitude of σ_ξ (Panel 3 of Table B1). For example, the proportion of voluntary

[19] The major difference between our paper and existing work is that we must specify the distribution of ξ in addition to the wage distribution.

job-to-job transitions in the simulated data that involve a wage decrease responds strongly to changes in σ_ξ. When $\sigma_\xi = 0$, the only explanation for the relatively large number of observed wage decreases at job-to-job transitions in the NLSY97 data is measurement error in wages which for our application has been fixed to $\sigma_\varepsilon = 0.15$. As σ_ξ increases away from zero, the model generates an increasing number of voluntary transitions where workers are willing to accept lower wages in exchange for higher non-wage utility.

Identifying the Distribution of Unobserved Heterogeneity

In many cases, the intuition behind identification of the parameters $\{\mu_w(j), \mu_\xi(j), \pi_j\}_{j=1}^2$ closely parallels simpler panel data models of wages and employment durations. For example, the within-person covariance in wages (moment 46) helps identify the person-specific component of wages, just as it would in a simpler panel data model of wages. When there is no heterogeneity in μ_w across people, the model generates a within-person covariance of zero between wages on employers that are separated by unemployment spells. The mean non-wage utility offer is identified by the combination of moments that summarize employment durations and unemployment durations. As the $\mu_\xi(j)$ parameters increase, jobs on average offer higher utility relative to employment, so unemployment spells tend to be shorter. The variation in μ_ξ across people is identified by moments that summarize the variation in unemployment durations across people (moments 38, 41)

Finally, it remains to discuss the identification of the covariance between the mean person-specific wage and non-wage offers. Given the assumed discrete distribution of unobserved heterogeneity, and allowing for $J = 2$ types of people, this covariance term is $\text{cov}(\mu_w, \mu_\xi) = \sum_{j=1}^2 \pi_j \mu_w(j)\mu_\xi(j) - [\sum_{j=1}^2 \pi_j \mu_w(j)][\sum_{j=1}^2 \pi_j \mu_\xi(j)]$. This covariance is identified by the covariance between the first wage observed after unemployment and the unemployment duration (moment 47), and the within-person covariance between the average wage and the fraction of months spent unemployed (moment 49).

Identifying Transition Parameters

We now discuss how the transition rate parameters $\{\lambda_u, \lambda_l, \lambda_e, \lambda_{le}\}$ are identified.

The layoff rate, λ_l, is identified non-parametrically by the empirical transition rate from employment into unemployment (moment 11 in Table B1). The job offer arrival rate, λ_e, is identified by moments that describe job-to-job transitions.[20] Within the model, the proba-

[20]We follow the search literature in assuming that job-to-job mobility is restricted by randomness in offer arrivals, but there is no direct monetary or non-monetary job switching cost.

bility of a job-to-job transition for a worker employed in a job with utility U is $\lambda_e \Pr(U' > U)$. Taking the parametric distribution $H(U)$ as given, λ_e is identified by moments that describe the frequency of job-to-job transitions, such as the empirical job-to-job transition rate (moment 12), and the average number of voluntary transitions that a worker makes over his career (moment 13).[21]

The probability that a layoff occurs in the same period is represented in the model by the parameter λ_{le}. As discussed in Section 4, we use data from the NLSY97 on the reason that jobs end to distinguish between voluntary and involuntary direct transitions between employers. If an individual reports that a job ends involuntarily, and he moves to a new job without experiencing an intervening spell of unemployment, then a simultaneous exogenous job ending and accepted outside offer has occurred. The probability that this type of transition occurs is $\lambda_{le} \Pr(U' > U^*)$. Taking $H(U)$ and U^* as given, the fraction of direct job-to-job transitions in the data that are involuntary (moment 33) identifies λ_{le}.

It remains to discuss identification of the arrival rate of job offers for the unemployed, λ_u. The reservation utility level for the unemployed is defined by Equation (A3) where the transition rate out of unemployment is $\lambda_u \Pr(U' > U^*)$. Note that U^* is not a primitive of the model – it is determined by optimal job search behavior. During estimation, as discussed in section 5.1, we fix the monthly discount rate to $\delta = .998$ and set $b = \ln(0.35 \times e^{\mu_w})$. It is clear from equation (A3) that λ_u will impact the unemployment durations generated by the model, so λ_u is identified by the moments that summarize the unemployment duration distribution (moments 10, 18-20).

6 Parameter Estimates

This section discusses the estimated parameters for the search model with non-wage job characteristics. In general, the model does a good job of fitting the data (Table B1) but in the interest of space we do not discuss this in further detail.

Our discussion begins with an examination of the importance of wages and non-wage utility and the magnitude of search frictions implied by the estimates. The discussion concludes by quantifying the importance of person-specific unobserved heterogeneity.

[21] As discussed in French and Taber (2011), non-parametric identification of λ_e requires exclusion restrictions in the form of observable variables that affect λ_e but do not affect $\Pr(U^{'} > U)$. In our model and with these data, there are no obvious candidates for exclusion restrictions. Unfortunately, information on rejected job offers, which would provide direct information about λ_e, is not available in the NLSY97.

Table 2: Parameter Estimates

Parameter	Notation	Estimate
Stand. dev. of wage offer	σ_w	0.4052
		(0.0066)
Stand. dev. of non-wage match	σ_ξ	0.3942
		(0.0113)
Pr(offer while unemployed)	λ_u	0.9655
		(0.0509)
Pr(layoff)	λ_l	0.0430
		(0.0072)
Pr(offer while employed)	λ_e	0.6295
		(0.0103)
Pr(offer and layoff)	λ_{le}	0.0427
		(0.0026)
<u>Type 1</u>		
Mean wage offer	$\mu_w(1)$	0.8875
		(0.0574)
Mean non-wage utility offer	$\mu_\xi(1)$	-1.9882
		(0.0430)
Reservation utility*	$U^*(1)$	-0.1333
		(0.0592)
Pr(type 1)	π_1	0.2497
		(0.0261)
<u>Type 2</u>		
Mean wage	$\mu_w(2)$	1.6328
		(0.0144)
Mean non-wage utility offer	$\mu_\xi(2)$	-1.3820
		(0.0154)
Reservation utility*	$U^*(2)$	0.7117
		(0.0230)
Pr(type 2)	π_2	0.7503
		(0.0261)

*The reservation utility levels are computed by solving Equation (A3) at the estimated parameters.

6.1 Job Offers and Labor Market Frictions

The parameter estimates are shown in Table 2. The estimate of the standard deviation of wage offers (σ_w) is 0.4052. Interestingly, the estimate of $\sigma_\xi = 0.3942$ indicates that a worker faces approximately the same amount of variation in non-wage utility across job matches as in wages. The relatively large amount of variation in non-wage utility across job matches indicates that non-wage considerations are an important factor as workers evaluate job offers. In other words, focusing only on wages, as is commonly done in the on-the-job search literature, misses a significant determinant of worker search behavior and total utility. This result is clearly demonstrated by examining simulated data generated from the estimated model (summary statistics are given in Table 3). In these data, as workers move between jobs 1–3, both wages and non-wage utilities increase.

Job dispersion due to search frictions is easily demonstrated by examining a scatterplot of accepted wages and non-wage utility and a histogram of total utility (Figure 2) where for clarity we focus on the first accepted job offer after unemployment for Type 2 workers.[22] Not only are wages and non-wage utilities dispersed (Figure 2a) but accepted total job values or "jobs" are also widely dispersed (Figure 2b). As we discussed in Section 2 and as we will illustrate in the following section, job dispersion results in severely biased estimates of the marginal-willingness-to-pay. Furthermore, as workers leave their current jobs to move up the "utility ladder," jobs become further dispersed.

The four transition parameters (λ's) determine the magnitude of frictions in the labor market. Recall that the model is estimated using monthly data, so all parameters are monthly arrival rates. The estimated offer arrival probability for unemployed agents is close to one ($\lambda_u = 0.9655$), and the estimated job offer arrival rate for employed workers is approximately 60 percent lower ($\lambda_e = 0.6295$). An employed agent faces approximately a 4 percent chance of exogenously losing his job in each month and being forced into unemployment ($\lambda_l = 0.0430$). Similarly, an employed worker has approximately a 4 percent chance of losing his job, but simultaneously receiving a new job offer that gives him the option of avoiding unemployment ($\lambda_{le} = 0.0427$).

6.2 Person-Specific Unobserved Heterogeneity

As discussed earlier in the paper, the extent of person-specific unobserved heterogeneity in the model is determined by the estimated values of the parameters $\{\mu_w(j), \mu_\xi(j), \pi_j\}_{j=1}^2$.

[22] Section 6.2 describes the differences in labor market outcomes between Type 1 and Type 2 workers. Section 7 provides a detailed analysis of how job-to-job mobility impacts accepted job offers.

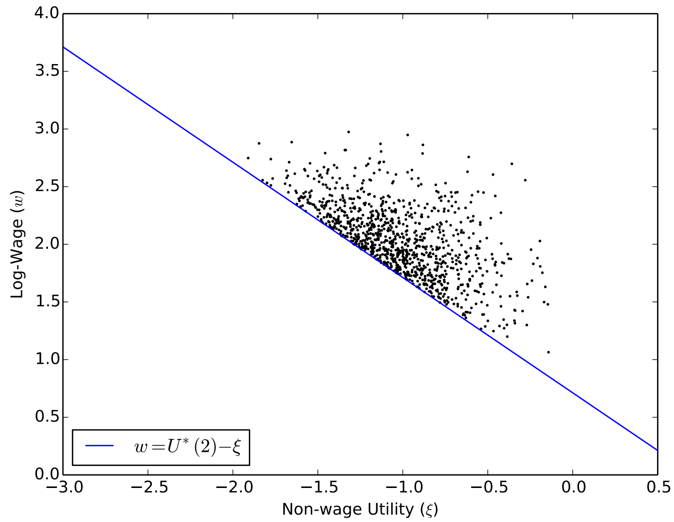
Table 3: Summary statistics for simulated data

	Variable	All types	Type 1	Type 2
All jobs	w	2.0297	1.5188	2.1082
		(0.3781)	(0.3044)	(0.3232)
	ξ	-0.9931	-1.3900	-0.9322
		(0.3531)	(0.3019)	(0.3192)
	$w+\xi$	1.0366	0.1288	1.1760
		(0.4624)	(0.2174)	(0.3056)
Job 1	w	1.9689	1.5108	2.0694
		(0.3825)	(.3033)	(0.3195)
	ξ	-1.0461	-1.3980	-0.9689
		(0.3540)	(0.3008)	(0.3159)
	$w+\xi$	0.9227	0.1128	1.1005
		(0.4702)	(0.2115)	(0.2900)
Job 2	w	2.1162	1.5683	2.1520
		(0.3478)	(0.3057)	(0.3192)
	ξ	-0.9184	-1.3407	-0.8908
		(0.3326)	(0.3032)	(0.3156)
	$w+\xi$	1.1978	0.2277	1.2612
		(0.3787)	(0.2248)	(0.2897)
Job 3	w	2.1883	1.6037	2.2001
		(0.3282)	(0.3104)	(0.3176)
	ξ	-0.8540	-1.3073	-0.8449
		(0.3202)	(0.3047)	(0.3139)
	$w+\xi$	1.3343	0.2963	1.3553
		(0.3169)	(0.2281)	(0.2814)
Mean job #		1.4817	1.1451	1.5334
		(0.7476)	(0.3899)	(0.7755)
Emp. dur.		12.8688	11.8966	13.0182
		(12.3730)	(11.4008)	(12.5090)
Enemp. dur.		13.3564	23.8201	4.9887
		(18.4453)	(23.3195)	(4.4567)
Disc. utility		167.7080	-14.6839	228.4081
		(106.4703)	(10.4777)	(17.7844)
Employment		0.6255	0.3335	0.7226

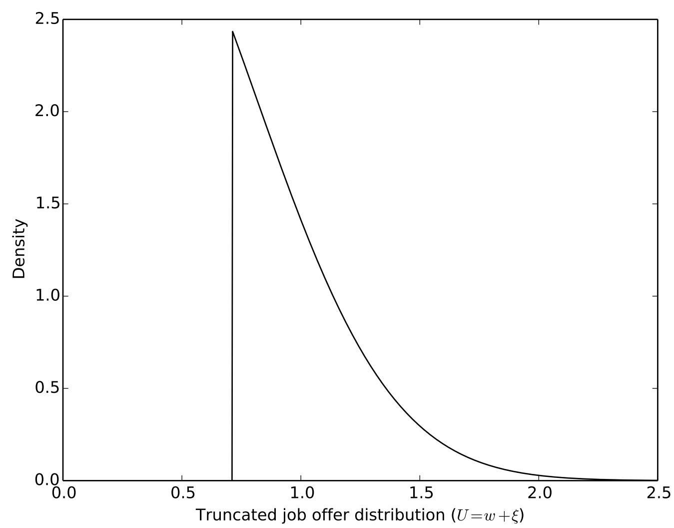
Notes: 5 million individuals simulated over 3 thousand periods.
Numbers in parentheses are standard deviations.

Figure 2: Accepted first jobs
Type 2 workers

(a) Scatter plot



(b) Density of $U = w + \xi$



The most common type of person in the economy makes up three-quarters of the population ($\pi_2 = 0.7503$), and has an expected wage offer of $\mu_w(2) = 1.6392$ and expect to receive a non-wage utility offer of $\mu_\xi(2) = -1.3820$. Recall that the non-wage utility flow from employment is measured relative to the value of unemployment, so the fact that this parameter estimate is negative simply indicates that on average these workers get disutility from working. The remaining one-quarter of the population consists of Type 1 workers. Relative to Type 2 workers, this segment of the population has lower labor market ability, and receives worse job offers (both wage and non-wage). More specifically, the expected wage offer for a Type 1 worker is approximately half as large as that of a Type 2 worker ($\mu_w(1) = 0.8875$ vs $\mu_w(2) = 1.6392$). Clearly, the estimates indicate that there is substantial unobserved heterogeneity in this sub-sample from the NLSY97.

The most straightforward way to quantify the importance of person-specific unobserved heterogeneity is by comparing simulated outcomes for the two types of workers. It is apparent from Table 3 that unobserved heterogeneity results in large differences in outcomes between Type 1 and 2 workers, not only for wages and non-wage utility but also employment and unemployment durations and for lifetime utility. As we will show, although controlling for worker ability reduces the bias observed in MWP estimates, it will not completely eliminate it. Indeed, regressing $w + \xi$ on a type dummy reveals that worker type can only explain about 59.2 percent of the variation in job utility. That is, even controlling for worker ability, 40.8 percent of the variation in total utility remains unexplained.

Finally, as discussed in Section 3.5, worker heterogeneity in mean wage and non-wage match values can give rise to correlation across the population. Given our estimates, this correlation is 0.339. Since it can be argued that some characteristics are positively correlated (health insurance benefits) while others are negatively correlated (risk of injury), this does not seem unreasonable.

7 Estimating Compensating Wage Differentials

In our model, the marginal-willingness-to-pay for ξ is known and fixed at -1. With this in mind, we can use our model to better understand the sources of job dispersion in a search framework and to illustrate how the various sources of job dispersion lead to biased hedonic compensating wage differential estimates. To do so, we estimate several variants of the

Table 4: Hedonic wage regressions

Type Dummies	Job Dummies	Type/Job Interaction	$\hat{\beta}$	R^2
N	N	N	-0.2160	0.0407
Y	N	N	-0.5780	0.5151
N	Y	N	-0.2784	0.1147
Y	Y	N	-0.6127	0.5548
Y	Y	Y	-0.6128	0.5549

following hedonic wage equation:

$$w_i = \alpha + \beta\xi_i + \boldsymbol{\gamma} \times \text{type dummies} + \boldsymbol{\kappa} \times \text{job dummies} + \boldsymbol{\tau} \times \text{type/job dummies} + e$$

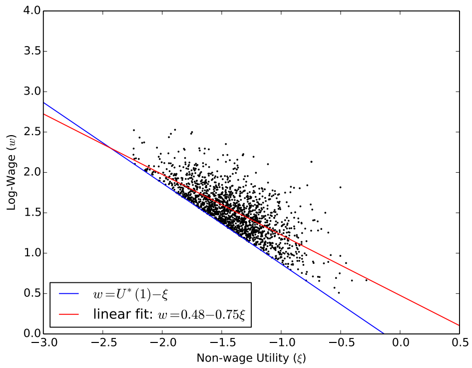
where the compensating wage differential literature interprets an estimate of β as the marginal-willingness-to-pay for ξ. At its most inclusive, this specification controls for worker heterogeneity and job dynamics through worker job numbers within employment cycles and their interactions. Table 4 presents $\hat{\beta}$ estimates and R^2 coefficients.

The standard, naïve wage hedonic with no controls for ability or job dynamics yields an extremely biased MWP of $\hat{\beta} = -0.22$ (row 1, Table 4). Controlling for worker type using dummy variables yields a greatly improved but still significantly biased MWP estimate of -0.58 (row 2, Table 4). Fully controlling for worker type by allowing MWP estimates to vary by type yields estimates of -0.75 for Type 1 workers and -0.55 for Type 2 workers (not shown in table).

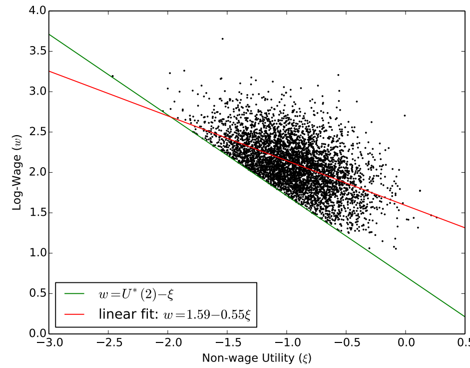
Figure 3 decomposes the effects of job dispersion and worker heterogeneity on hedonic MWP estimates. This figure plots steady-state cross-section wages and non-wage utilities in the simulated data separately for each type (Figures 3a and 3b), and pooled across both types (Figure 3c). Figures 3a and 3b show that conditional on type, there is a considerable amount of job dispersion in the simulated data because accepted jobs are truncated from below at the reservation utility level ($w + \xi > U^*$). As a result, hedonic estimates of the MWP are attenuated from the true value of -1 that was used to generate the simulated data. It is important to note that the estimates shown in this figure represent, in some respects, a best case scenario for the simple hedonic regression approach. These regressions allow the econometrician to have perfect information about worker type, so he is able to fully control for heterogeneity by estimating separate regressions for each type. In contrast, empirical applications typically rely on imperfectly measured, and undoubtedly incomplete, proxies

Figure 3: Worker heterogeneity and bias

(a) Type 1



(b) Type 2



(c) Both types

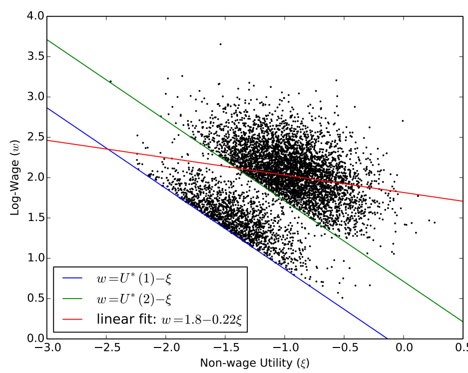
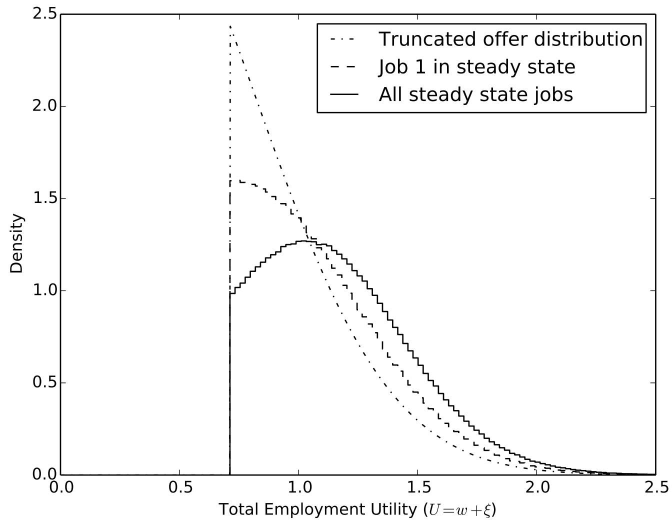


Figure 4: Job dispersion
Type 2 workers



for worker heterogeneity. These results show that even in this optimistic scenario where the econometrician is able to eliminate bias due to worker heterogeneity, job dispersion leads to seriously biased estimates of the MWP.

Figure 3c plots accepted jobs for both types of worker, the reservation utility frontiers ($U^* = w + \xi$) for each type, and a fitted hedonic regression line which assumes that worker type is not observed by the econometrician. In this scenario, MWP estimates will suffer from bias due to both dispersion in job offers and unobserved worker heterogeneity. The plotted regression line in this figure illustrates how failing to control for unobserved differences between workers leads to a severe downward bias in the estimated MWP. Simply put, because Type 2 workers tend to accept jobs that offer higher total utility than Type 1 workers, the relationship between accepted values of w and ξ uncovered by a naïve specification of a hedonic regression bears little resemblance to the true MWP of workers in the model.

7.1 Job dynamics

The dynamics of our search model also contributes to the dispersion in jobs. In particular, dynamics increase job dispersion through two channels. First, workers on the lower end of the job distribution are more likely receive a superior job offer, shifting the job distribution away from the lower end. Second, as these workers move to better jobs, the job distribution shifts towards the higher end. For example, the density of the total utility of Type 2 job offers truncated at $U^*(2)$ (i.e., utility of jobs accepted out of unemployment) is given by the dash-dotted line in Figure 4. In a steady state cross-section, as workers in bad first jobs accept better jobs, the histogram of workers' first jobs shows greater dispersion as workers on the lower end of the job distribution accept better offers (the dashed line). Finally, adding all subsequent jobs, the histogram over jobs illustrates the dispersion in cross-sectional steady state jobs (the solid line).

In terms of our hedonic compensating wage differential estimates, if we control for job number within an employment cycle but exclude ability controls the MWP estimate moderately improves, rising to -0.28 from -0.22 (row 3 vs row 1, Table 4). Even controlling for both type and job number, the MWP estimate is still significantly biased at -0.61 (row 4, Table 4). This specification is able to explain only about 55.5 percent of the variation in wages.[23]

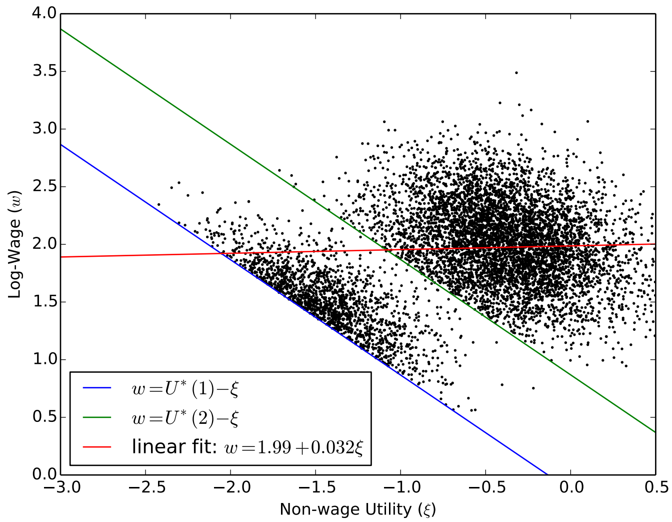
7.2 Worker heterogeneity

Unlike some of the compensating wage differential literature, our marginal willingness to pay estimates using our simulated dataset all have the correct sign. Using a counterfactual experiment, we examine how increasing worker heterogeneity can yield MWP estimates with the wrong sign.

We now show that a perturbation of our estimated parameters is sufficient to generate wrong-signed MWP estimates. Consider a model where the mean non-wage utility for Type 2 workers increases from $\mu_\xi(2) = -1.382$ to $\mu_\xi(2) = -0.75$ but all other parameters remain the same. As a result of the increase in mean non-wage utility offers Type 2 workers optimally increase their reservation utility from $U^*(2) = 0.7117$ to $U^*(2) = 0.8682$. Wages and non-wage utility for Type 2 workers shift to the northeast resulting in a positive estimated willingness-to-pay (see Figure 5). The hedonic wage regression without ability controls yields a simulated dataset with an estimated MWP of 0.0320 and a standard error of 0.0004.

[23]This can only be marginally improved by allowing for interactions in explanatory variables (row 5, Table 4).

Figure 5: Worker heterogeneity and wrong signed MWP



7.3 Discussion

With a few exceptions (Gronberg and Reed, 1994; Hwang et al., 1998; Bonhomme and Jolivet, 2009), the literature on compensating wage differentials has focused on unobserved worker ability as the reason for weak support for the theory (Brown, 1980; Hwang et al., 1992; Han and Yamaguchi, 2012). To some extent, this focus on unobserved ability is warranted but nevertheless, there are further, important sources of job dispersion due to job dynamics and search frictions.

In general, when frictions are an important feature of the labor market, compensating wage differential estimates will be biased. Our findings are similar to those of Bonhomme and Jolivet (2009) where they estimate a partial equilibrium version of Hwang et al. (1998) with several non-wage job characteristics. They find strong preferences for amenities but little evidence of compensating differentials in their simulated data. Our aggregate approach with choice over just two dimensions draws a clear picture of precisely how search frictions bias compensating wage differential estimates. Our biased compensating wage differential estimates arise because frictions imply that acceptable jobs typically provide utility greater than the reservation level. This job dispersion does not depend on but is exacerbated by heterogeneity in worker ability and job dynamics.

Since these biases are due to job dispersion (i.e., $\text{var}(U_i) > 0$), it is useful to examine the sources of job dispersion. In our discussion thus far, we have discussed worker ability and job-to-job mobility as sources of job dispersion. But job-to-job mobility, while observable to the practitioner, is an imperfect proxy for the dispersion of worker reservation utilities due to job dynamics. As the modelers, we perfectly observe reservation utilities, U_i^{**}: $U_i^{**} = U^*$ for the first job out of unemployment and $U_i^{**} = w_i(-1) + \xi_i(-1)$ for all other jobs where $w_i(-1)$ and $\xi_i(-1)$ are the worker's prior wage and non-wage utility. Regressing U_i on type dummies and U_i^{**}, we find that only about 67 percent of the job dispersion observable to us as modelers can be explained by ability differences and dispersion in reservation utilities. Thus, the remaining 33 percent can only be due to frictions resulting from the inherent dispersion of job offers.

8 Concluding remarks

In a frictionless and competitive labor market, equally able workers must receive the same total compensation and the estimated wage differential for a job attribute will equal the workers' willingness-to-pay for that attribute. Unfortunately, evidence in support of the theory is weak (Brown, 1980). In contrast, in labor markets with frictions, total job values or "jobs" are dispersed and total utility will in general exceed a worker's reservation utility so that different, equally-able workers will receive different compensation packages, biasing estimates of compensating wage differentials.

In this paper we explore the links between job dispersion and the often weak evidence for compensating wage differentials. We begin by estimating an on-the-job search model which allows workers to search across jobs based on both wages and job-specific non-wage utility flows. The importance of non-wage utility is revealed through voluntary job-to-job moves, wage changes at transitions, and job durations. Since not accounting for worker ability is a common explanation for the frequent failure of compensating wage differential estimates, we control for worker heterogeneity.

Using a simulated data set based on our model and parameter estimates, we show that job dispersion leads to severely biased compensating wage differential estimates. Job dispersion is exacerbated by differences in worker ability and in an on-the-job search framework, by dispersion in reservation wages and controlling for these sources of job dispersion ameliorates the bias. Nevertheless, MWP estimates still have a downward bias of nearly 40 percent. Indeed, estimating total utility on worker type and the reservation utility reveals that 33 percent of job dispersion must be due to frictions inherent in the dispersion of job offers.

Appendix

A Derivation of Reservation Utility

The reservation utility level for unemployed agents, U^*, solves $V^e(U) = V^u$. To derive U^*, we must first rearrange (4) and (3) so that common terms can be collected when evaluated at $U = U^*$. Subtracting $\delta V^e(U)$ from both sides of (4):

$$\begin{aligned}(1-\delta)V^e(U) = U + \delta[\lambda_e E \max\{0, V^e(U') - V^e(U)\} + \lambda_l(V^u - V^e(U)) \\ + \lambda_{le} E \max\{V^u - V^e(U), V^e(U') - V^e(U)\}.\end{aligned}$$

Evaluating this at $U = U^*$:

$$\begin{aligned}(1-\delta)V^e(U^*) &= U^* + \delta\left[\lambda_e \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U') + \lambda_{le} \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U')\right] \\ &= U^* + \delta(\lambda_e + \lambda_{le}) \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U')\end{aligned} \tag{A1}$$

Similarly, subtracting δV^u from both sides of (3),

$$\begin{aligned}(1-\delta)V^u &= b + \delta\lambda_u E \max\{0, V^e(U') - V^u\} \\ &= b + \delta\lambda_u \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U').\end{aligned} \tag{A2}$$

Evaluating at $U = U^*$, we can equate (A1) and (A2), integrate by parts and solve to get:

$$\begin{aligned}U^* &= b + \delta[\lambda_u - (\lambda_e + \lambda_{le})] \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U') \\ &= b + \delta[\lambda_u - (\lambda_e + \lambda_{le})] \int_{U^*}^{\infty} V^{e\prime}(U')[1 - H(U)]dU' \\ &= b + \delta[\lambda_u - (\lambda_e + \lambda_{le})] \int_{U^*}^{\infty} \frac{1 - H(U')}{(1-\delta) + \delta\{\lambda_e[1 - H(U')] + \lambda_l + \lambda_{le}\}} dU'.\end{aligned} \tag{A3}$$

When $\lambda_u > \lambda_e + \lambda_{le}$ (the probability of receiving an offer while unemployed is greater than that when employed), an unemployed worker's reservation wage exceeds the one-period utility flow from unemployment.

B Estimation Moments

Table B1: Moments of the NLSY97 Data and Simulated Data

Moment #	Description	Data	Simulated
	Cycle Moments (Panel 1)		
1	Mean log-wage (employer 1)	1.9791	1.9469
2	Std. dev. of log-wage (employer 1)	0.4249	0.3942
3	Mean employment spell duration (employer 1)	8.9392	8.5086
4	Mean log-wage (employer 2)	2.0377	2.0690
5	Std. dev. of log-wage (employer 2)	0.4582	0.3767
6	Mean employment spell duration (employer 2)	9.2713	9.3163
7	Mean log-wage (employer 3)	2.0608	2.1142
8	Std. dev. of log-wage (employer 3)	0.4572	0.3553
9	Mean employment spell duration (employer 3)	9.7382	8.9228
	Transition and Duration Moments (Panel 2)		
10	mean unemp. spell duration	5.9087	5.5116
11	Pr(transition into unemp.)	0.0469	0.0563
12	Pr(job-to-job transition)	0.0364	0.0426
13	mean total number of voluntary job-to-job transitions	1.4510	1.4497
14	mean total number of involuntary job-to-job transitions	0.2571	0.2704
15	mean total number of transitions into unemployment	1.8786	1.6663
16	mean # of firms per cycle	1.6983	1.6515
17	mean total # of employers over entire career	4.3755	4.3597
18	Pr(unempdur = 1)	0.2375	0.3287
19	Pr(unempdur = 2)	0.1697	0.1272
20	Pr(unempdur = 3)	0.1092	0.1007
21	Pr(empdur = 1)	0.1423	0.1294
22	Pr(empdur = 2)	0.1412	0.1148
23	Pr(empdur = 3)	0.1209	0.0951
24	across-person mean fraction of months unemployed	0.2745	0.2907
	Wage Change Moments (Panel 3)		
25	Mean Δw at job-to-job switch	0.0812	0.1002
26	Mean Δw at job-to-job switch $\|\Delta w > 0$	0.3592	0.4100
27	Mean Δw at job-to-job switch $\|\Delta w < 0$	−0.3273	−0.3438
28	Pr(wage decrease at job-to-job transition)	0.3640	0.4091
29	Pr(wage decrease at involuntary job-to-job transition)	0.4601	0.5607
30	Mean Δw at involuntary job-to-job switch	−0.0168	−0.0791
31	Mean Δw at involuntary job-to-job switch $\|\Delta w > 0$	0.3224	0.3489
32	Mean Δw at involuntary job-to-job switch $\|\Delta w < 0$	−0.3454	−0.4152
33	Fraction of job-to-job transitions that are involuntary	0.1505	0.1569

Continued on next page

Table B1 – contintued from previous page

Moment #	Description	Data	Simulated
	Wage Regression (Panel 4)		
34	Constant	1.9311	1.9389
35	Experience	0.0058	0.0057
36	Experience^2/100	-0.0021	-0.0061
	Variance and Covariance Moments (Panel 5)		
37	across-person std. dev. of wages	0.3131	0.2774
38	across-person std. dev. of unemp. duration	5.9004	4.9495
39	across-person std. dev. of fraction of months unemp.	0.2587	0.2317
40	across-person std. dev. total number of firms	2.9437	2.7312
41	std. dev. of unemp. duration	7.7319	7.6788
42	std. dev. of # of firms per cycle	1.1513	0.9566
43	by person: std. dev. of total # of vol. job-to-job trans.	1.7091	1.4688
44	by person: std. dev. of total # of invol. job-to-job trans.	0.6253	0.5566
45	by person: std. dev. of total # of transitions into unemp.	1.7930	1.4824
46	within-person cov. in wages	0.0448	0.0421
47	cov(1st wage, 1st unemp. duration)	-0.1439	-0.5042
48	cov(1st unemp. duration, 1st emp. duration)	-1.4050	0.1323
49	within-person cov(ave. wage, fraction of months unemp.)	-0.0332	-0.0638
50	cov(wage, employment duration)	0.9138	0.1541
51	cov(Δw, $\Delta empdur$) at vol. job-to-job switch	0.7491	0.3222

References

Albrecht, J. and S. Vroman, 1998, "Nash Equilibrium Efficiency Wage Distributions," *International Economic Review*, 39(1):183–203.

Barlevy, G., 2008, "Identification of Search Models using Record Statistics," *Review of Economic Studies*, 75(1):29–64.

Bhaskar, V., A. Manning and T. To, 2002, "Oligopsony and Monopsonistic Competition in Labor Markets," *Journal of Economic Perspectives*, 16(2):155–174.

Bhaskar, V. and T. To, 1999, "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition," *The Economic Journal*, 109:190–203.

—, 2003, "Oligopsony and the Distribution of Wages," *European Economic Review*, 47(2):371–399.

Bonhomme, S. and G. Jolivet, 2009, "The Pervasive Absence of Compensating Differentials," *Journal of Applied Econometrics*, 24(5):763–795.

Bound, J., C. C. Brown and N. Mathiowetz, 2001, "Measurement Error in Survey Data," in E. E. Learner and J. J. Heckman, eds., "Handbook of Econometrics," pp. 3705–3843, North Holland Publishing, New York.

Bowlus, A. J. and G. R. Neumann, 2006, "The Job Ladder," in H. Bunzel, B. J. Christensen, G. R. Neumann and J.-M. Robin, eds., "Structural Models of Wage and Employment Dynamics," Elsevier.

Brown, C., 1980, "Equalizing Differences in the Labor Market," *Quarterly Journal of Economics*, 94(1):113–134.

Brown, C. and J. Medoff, 1989, "The Employer Size-Wage Effect," *Journal of Political Economy*, 97:1027–1059.

Burdett, K. and D. T. Mortensen, 1998, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2):257–273.

Dey, M. and C. Flinn, 2008, "Household Search and Health Insurance Coverage," *Journal of Econometrics*, 145(1–2):43–63.

Dey, M. S. and C. J. Flinn, 2005, "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 73(2):571–627.

Dunlop, J. T., 1957, "The Task of Contemporary Wage Theory," in G. W. Taylor and F. C. Pierson, eds., "New Concepts in Wage Determination," pp. 117–139, McGraw Hill, New York.

Eckstein, Z., S. Ge and B. Petrongolo, 2009, "Job and Wage Mobility with Minimum Wages and Imperfect Compliance," *Journal of Applied Econometrics*, forthcoming.

Flinn, C., 2002, "Labour market structure and inequality: A comparison of Italy and the US," *Review of Economic Studies*, 69(3):611–645.

French, E. and C. Taber, 2011, "Identification of models of the labor market," *Handbook of Labor Economics*, 4:537–617.

Goffe, W. L., G. D. Ferrier and J. Rogers, 1994, "Global optimization of statistical functions with simulated annealing," *Journal of Econometrics*, 60(1-2):65 – 99.

Gronberg, T. J. and W. R. Reed, 1994, "Estimating Workers' Marginal Willingness to Pay for Job Attributes Using Duration Data," *The Journal of Human Resources*, 29(3):911–931.

Groshen, E., 1991, "Sources of Intra-Industry Wage Dispersion: How Much Do Employers Matter?" *Quarterly Journal of Economics*, 106:869–884.

Han, S. and S. Yamaguchi, 2012, "Compensating Wage Differentials in Stable Job Matching Equilibrium," Working paper.

Hwang, H.-S., D. T. Mortensen and W. R. Reed, 1998, "Hedonic Wages and Labor Market Search," *Journal of Labor Economics*, 16(4):815–847.

Hwang, H.-s., W. R. Reed and C. Hubbard, 1992, "Compensating Wage Differentials and Unobserved Productivity," *Journal of Political Economy*, 100(4):835–858.

Keane, M. P. and K. I. Wolpin, 1997, "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3):473–522.

Kletzer, L. G. and H. F. Rosen, 2006, "Reforming Unemployment Insurance for the Twenty-First Century Workforce," Hamilton Project working paper, 2006-06.

Pergamit, M. R., C. R. Pierret, D. S. Rothstein and J. R. Veum, 2001, "The National Longitudinal Surveys," *Journal of Economic Perspectives*, 15(2):239–253.

Rosen, S., 1986, "The Theory of Equalizing Differences," in O. C. Ashenfelter and R. Layard, eds., "Handbook of Labor Economics," vol. 1, chap. 12, pp. 641–692, Elsevier, New York.

Smith, A., 1776, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Clarendon Press (1976), Oxford.

Smith, A. A., Jr., 1993, "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8(S):S63–84.

Stern, S., 1989, "Estimating a Simultaneous Search Model," *Journal of Labor Economics*, 7(3):348–369.

—, 1997, "Simulation-based estimation," *Journal of Economic Literature*, 35(4):2006–2039.

Sullivan, P. and T. To, 2013, "Search and Non-wage Job Characteristics," *Journal of Human Resources*, forthcoming.

Taber, C. and R. Vejlin, 2013, "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," Working paper.

Wolpin, K. I., 1992, "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth," *The Journal of Political Economy*, 100(3):535–560.

Yamaguchi, S., 2010, "Job Search, Bargaining, and Wage Dynamics," *Journal of Labor Economics*, forthcoming.