Job Dispersion and Compensating Wage Differentials

Paul Sullivan   Ted To

American University   *Bureau of Labor Statistics*

February 2019

Abstract

The empirical literature seeking for evidence of 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 this weak support is the result of bias due to dispersion in total job values, or "job dispersion." 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. The model incorporates three primary mechanisms that threaten the validity of traditional hedonic wage regressions: (1) search frictions, (2) dispersion in the value of job offers, (3) unobserved heterogeneity in worker ability. Estimating simple hedonic wage regressions using simulated data from the model reveals that estimates of the marginal willingness-to-pay (MWP) for non-wage job characteristics are severely attenuated. While worker heterogeneity and search frictions 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 competitive 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."

In models with frictions, identical workers can receive different compensation packages in equilibrium so that "jobs" (the total value of wage and non-wage amenities) are dispersed. In a labor market with frictions, observed combinations of wages and job amenities do not directly reveal worker's marginal-willingness-to-pay (MWP) for job amenities. As a result, traditional hedonic wage regression estimates of the MWP are biased. Not surprisingly, job dispersion is closely related to the extensive empirical literature on wage dispersion.[1] 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). In this broad class of models, dispersed wages imply that jobs are dispersed and as a result, hedonic wage regressions will yield biased estimates of the MWP.

To understand job dispersion and its role in biased hedonic wage estimates, it is natural to turn towards job search. While efficiency wages and oligopsony can also exhibit job dispersion, search models have both well-developed theoretical underpinnings and an extensive empirical literature.[2] A number of papers have studied non-wage job characteristics using a search framework. Gronberg and Reed (1994) develop a method for estimating the MWP that relies on job duration data. Hwang et al. (1998) use a theoretical equilibrium search model to demonstrate that the endogenous relationship between non-wage amenities and total job value can lead to compensating wage differential regressions that give misleading estimates of the MWP. Dey and Flinn (2005, 2008) estimate search models that include health insurance as well as wages. Bonhomme and Jolivet (2009) estimate a search model that includes a number of non-wage job characteristics. Sullivan and To (2014) demonstrate how the importance of non-wage utility can be identified through revealed preference using

[1]Some notable examples include Dunlop (1957), Brown and Medoff (1989) and Groshen (1991).

[2]See Eckstein and van den Berg (2007) for a survey.

data on accepted wages and job transitions, and focus on estimating the importance of wages relative to non-wage amenities. Hall and Mueller (2015) use the Krueger and Mueller (2011, 2016) survey data on reservation wages, job offers, acceptance decisions and prior wages to estimate the joint distribution of non-wage utility and wage offers in a search model. Sorkin (2015) uses matched firm-worker data (LEHD) to show that preferences for non-wage job characteristics are at least as important as wages for worker job choices. Taber and Vejlin (2016) use Danish matched worker-firm data to estimate a rich model that quantifies the contributions of comparative advantage (Roy model), human capital, non-wage utility, and search frictions to overall wage inequality.

Building on Bonhomme and Jolivet (2009) and Sullivan and To (2014), we quantify the magnitude of bias in MWP estimates and decompose the sources of bias. Following Sullivan and To, we estimate a structural, on-the-job search model where workers search across jobs that offer different wages and levels of nonwage utility. Then, as in Bonhomme and Jolivet, we use our estimated model to simulate a dataset that contains accepted wages and non-wage utilities, and then perform a detailed evaluation of the reduced-form hedonic wage regressions that are commonly used to estimate compensating wage differentials. Instead of focusing on search fictions,[3] as in Bonhomme and Jolivet, we decompose the bias in traditional hedonic wage estimates into three sources: (1) differences in worker ability, (2) search frictions, and (3) the dispersion of job offers (the variance of the job offer distribution).[4] Of particular importance for the empirical literature on compensating wage differentials, this is the first paper to clearly distinguish, and separately quantify, the bias resulting from these three channels.

We differ from Sullivan and To (2014) in two important aspects. First, this search model allows for the possibility that wages and non-wage utility from employment are correlated. By estimating this correlation, we can determine the extent to which workers who tend to receive high wage offers also tend to receive job offers that are desirable for other, non-monetary, reasons. This is a crucial empirical issue when attempting to evaluate the theory of equalizing differences. Second, while Sullivan and To briefly discuss biased compensating wage differential estimates in a search framework with non-wage utility, we elaborate considerably with a detailed analysis that breaks down these biases into several sources.

[3]It is important to note that search frictions are necessary but not sufficient for biased cross-sectional MWP estimates in a search model (Sullivan and To, 2018)

[4]The term "search frictions" typically refers to some amalgamation of job offer arrival rates and separation rates (Ridder and van den Berg, 2003) but could also include the dispersion of job offers. For our purposes we consider "search frictions" and "job offer dispersion" to be distinct aspects of a search model which together affect the degree to which there are frictions in the labor market.

The empirical literature on compensating wage differentials has focused on unobserved 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 job dispersion and hence the bias in compensating wage differential estimates can be attributed to unobserved worker ability. Nevertheless, although unobserved worker ability can increase job dispersion, it is not necessary for biased compensating wage differential estimates. Indeed, under ideal circumstances where the researcher can perfectly control for heterogeneity in worker ability, roughly 41 percent of the bias in cross-sectional MWP estimates based on our simulated dataset can only be due to the inherent dispersion of job offers.

Finally, we identify the manner in which search frictions add to job dispersion – the variance in the value of accepted offers increases as workers climb the "job ladder." This increased dispersion due to job dynamics works through three 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. Third, exogenous separations kick workers back to the bottom of the job ladder. Controlling for ability and using position in job cycle to control for job dynamics, MWP estimates remain significantly attenuated (specifications 5 and 6 from Table 5). To the best of our knowledge, we are the first to identify job dynamics as the mechanism by which search frictions add to the bias in hedonic estimates of the MWP.

In the following section, we lay out a partial equilibrium model of on-the-job search with preferences for non-wage job characteristics. Then in Section 3, we discuss the dataset used to estimate our partial equilibrium model and in Section 4 we discuss our econometric methodology and some important identification issues. In Section 5 we present our parameter estimates and in Section 6 we analyze the estimation of compensating differentials using simulated datasets. Section 7 concludes.

2 The Search Model: Wages and Non-Wage Utility

We now describe the discrete time, on-the-job search model used to study compensating wage differentials.[5] In each period, workers are either employed or unemployed and search for job offers to maximize their discounted expected utility. Workers have discount factor δ.

Unemployed workers get per-period log unemployment benefits, b. Employed workers

[5] Note that while there are some differences in the details of our implementation, the structure of the basic model is the same as in Sullivan and To (2014).

get utility from the wage and non-wage amenities. In particular, assume that an employed worker's one period utility is:

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

where w is the log wage and ξ is the match-specific, per-period non-wage utility.[6]

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.

When a worker receives a job offer, it is drawn from the cumulative distribution function $F(w, \xi)$ where w and ξ are fixed for the duration of the match. Given $F(w, \xi)$ and $U(w, \xi)$, it will simplify the description of the optimization problem to define the workers' problem in terms of total utility, $U \equiv U(w, \xi)$ where U is distributed according to $H(U)$.[7]

A worker who receives a job offer, U', compares the discounted expected value of the offer with the value of their fallback alternative – if the value of the offer exceeds the value of the fallback, the worker accepts the offer and otherwise rejects.

For our purposes, there are a number of important advantage to aggregating the value of all non-wage job characteristics into a single index, as opposed to attempting to estimate the value of particular job characteristics. First, Section 3.2 presents empirical evidence on the relationship between the non-wage job characteristics that are available in the NLSY97 and job mobility. In particular, the five tangible non-wage job characteristics in the data are only weakly related to wage changes at job transitions, and are unable to account for the prevalence of voluntary wage declines at job-to-job transitions.[8] Second, this approach avoids the potential bias 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

[6]Our additively separable utility function encompasses the class of two-factor Cobb-Douglas utility functions (Appendix A).

[7]For our additively separable utility function, $H(U) = \int_{-\infty}^{\infty} F_{w|\xi}(U - \xi|\xi) f_{\xi}(\xi) d\xi$ where $F_{w|\xi}$ is the conditional cumulative wage distribution and f_{ξ} is the unconditional probability density function for ξ.

[8]See Bonhomme and Jolivet (2009) for a paper that estimates the value of a large number of job characteristics using a structural search model.

excluded. This problem with excluded characteristics is exacerbated by the likelihood that workers have heterogeneous preferences over the employer provided benefits and tangible and intangible job attributes that differentiate jobs.

2.1 Unemployed Search

An unemployed worker's discounted, expected utility is given by

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

where $V^e(U')$ represents the discounted expected utility for a worker employed in a job with utility level U'.

With probability λ_u, an unemployed worker receives an offer with utility U'; the worker accepts the offer if $V^e(U') \geq V^u$ and rejects it otherwise, remaining unemployed. With probability $1-\lambda_u$, an unemployed worker who does not receive an offer remains unemployed getting V^u.

2.2 On-the-job Search

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

$$\begin{aligned} V^e(U) &= U + \delta[\lambda_e E \max\{V^e(U), V^e(U')\} + \lambda_l V^u \\ &\quad + \lambda_{le} E \max\{V^u, V^e(U')\} + (1 - \lambda_e - \lambda_l - \lambda_{le})V^e(U)]. \end{aligned} \tag{2}$$

With probability λ_e, an employed worker receives an offer with utility U'; the worker accepts the offer if $V^e(U') \geq V^e(U)$ and rejects it otherwise. With probability λ_l, the worker is separated and becomes unemployed getting utility V^u. With probability λ_{le}, the worker is separated and receives an offer with utility U'; the worker accepts the offer if $V^e(U') \geq V^u$ and rejects it otherwise. Finally, with probability $1 - \lambda_e - \lambda_l - \lambda_{le}$, the worker does not receive an offer and is not separated.

2.3 The Job Offer Distribution

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. Following this literature, we control for unobserved heterogeneity in worker ability when estimating the model.

Hedonic studies tend to control for as many observable worker characteristics as possible. Nevertheless, it is well known that wage regressions leave a large fraction of variation in wages unexplained. We minimize observable worker heterogeneity by estimating the model using a relatively homogeneous sample. Given the focus of our paper, we are not interested in estimating the effect of observable demographic characteristics on wages. Our only concern is controlling for worker heterogeneity, which we do with a flexible discrete mixture distribution for unobserved heterogeneity.

Another important concern when estimating the MWP for non-wage job characteristics is that wage offers and match-specific non-wage utility flows may be correlated. To the extent that this is the case, it will be reflected along with worker willingness-to-pay in the pairs of wages and amenities accepted by workers. This correlation enters our model through the job offer distribution $F(w, \xi)$.

More specifically, we allow for unobserved variation in worker ability and for correlation between wage offers and non-wage utility flows by permitting the mean wage offer, μ_w, and the mean match-specific utility flow, μ_ξ, to vary across the population.[9] Following Keane and Wolpin (1997) and a large subsequent literature, we assume that the joint distribution of unobserved heterogeneity is a mixture of discrete types. Assume that there are K types of people in the economy, and let π_k represent the proportion of type k in the population. The parameters of the distribution of unobserved heterogeneity, $\{\mu_w(k), \mu_\xi(k), \pi_k\}_{k=1}^{K}$, are estimated jointly along with the other parameters of the model.

The job offer distribution faced by workers in the model is:

$$\begin{aligned} F(w,\xi) &= \sum_{k=1}^{K} \pi_k F(w, \xi \mid k) \\ F(w, \xi \mid k) &= N(\mu_w(k), \sigma_w) N(\mu_\xi(k), \sigma_\xi). \end{aligned}$$

The correlation between wage offers and non-wage utility offers is $\rho_{w\xi} = \text{cov}(w, \xi)/\sqrt{\text{var}(w)\text{var}(\xi)}$.

[9] 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.

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 $\mu_w(k)$ and $\mu_\xi(k)$ 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.

2.4 Optimal Job Search and the Job Ladder

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 B 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. Indeed, the accepted distribution of (w, ξ) is determined by the job offer distribution, parameters that govern search frictions, and worker preferences. Section 6 expands on this point in considerable detail by using simulated data from the estimated dynamic model of the labor market to examine the performance of a standard hedonic wage regression.

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. These voluntary moves result in further dispersion since voluntary job-to-job moves are truncated at reservation utilities that are themselves dispersed and truncated at U^*.

[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 $\{\mu_w(k), \mu_\xi(k), \pi_k\}_{k=1}^K$ determine whether or not the correlation is positive or negative.

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 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 Data

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 young, unmarried, low-skilled men who are at the beginning of their careers 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. Following the structural 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 this early phase of their career.

We aggregate the weekly employment record for each NLSY97 respondent up to monthly data.[11] 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, and the associated wage, based on the employer that the worker spent the most weeks working for during the month. The aggregated monthly employment record contains employment durations, direct transitions between employers that occur without an intervening spell of unemployment, transitions into unemployment, and wage changes at transitions.

In the NLSY data, wage changes are observed for jobs that persist across survey years. Our estimation approach does not make use of these repeated wage measures for two reasons. First, we only observe a single wage for 72 percent of all jobs in our data. Second, we are unable to reject the null hypothesis that mean wage growth is zero within job spells with multiple wage measurements (at the 5% significance level). Since there appears to be little chance of precisely estimating an on-the-job wage growth process, we restrict wages to be constant within job spells.[13]

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.

3.1 Descriptive Statistics

This section highlights the key characteristics of the data used to estimate the structural model. As is standard in the search literature, we describe labor market histories in terms of employment cycles. An employment cycle begins and ends with unemployment, and includes

[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.

[13] we use the first reported wage for each job as the wage for the entire job spell.

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	
	All Jobs		
Mean unemployment spell duration	5.908		
Mean number of cycles per person‡	2.878		
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.

all job spells that occur without an intervening unemployment spell. 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

[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 2: Changes in Observable Job Characteristics at Voluntary Job Changes

	Obtained	Lost	Kept	Never had
Health insurance	0.289	0.136	0.205	0.368
Life insurance	0.209	0.079	0.075	0.636
Retirement plan	0.217	0.085	0.079	0.617
Flexible hours	0.151	0.118	0.060	0.669
Stock options	0.096	0.057	0.022	0.823
Observations	759			

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.

3.2 Observable Job Characteristics in the NLSY97

The prevalence of wage declines at voluntary job-to-job transitions suggests the possibility that, consistent with the theory of compensating wage differentials, job-changers frequently trade off lower wages in exchange for improved non-wage job characteristics. In this section, we look for evidence of this behavior in the NLSY97 data. For each job, the NLSY97 contains information about whether or not the employer offers each of the following non-wage amenities: health insurance, life insurance, retirement plans, flexible hours, and stock options.[16]

Table 2 shows the patterns of changes in amenities at voluntary job transitions. Let $d_a(j)$ be a dummy variable indicating that amenity a is offered by job j. For each amenity, there are four possible outcomes that could occur when a worker voluntarily moves to a new job, which is indexed by j'. The worker could obtain the amenity ($d_a(j) = 0$, $d_a(j') = 1$), lose it ($d_a(j) = 1$, $d_a(j') = 0$), keep it ($d_a(j) = 1$, $d_a(j') = 1$), or have never had access to it ($d_a(j) = 0$, $d_a(j') = 0$). The first row of Table 2 shows that workers obtain health insurance at 29 percent of voluntary job transitions, while they lose it at 14 percent of these transitions. Interestingly, the majority of job changes (57%) involve no change in health coverage ($0.205 + 0.368 = 0.57$). Similar patterns hold for the other non-wage job

[16] As is the case in the majority of longitudinal data sets, the NLSY97 contains information about whether or not benefits are offered on jobs, but it unfortunately does not include information about take-up of benefits.

characteristics. Workers are more likely to obtain amenities than lose them, but in all cases a sizable fraction of workers chose to give up amenities at voluntary transitions. Moreover, for each amenity, the most likely outcome at a voluntary transition is no change.

Table 3: Log-Wage Changes and Changes in Job Characteristics

	(1) Change in Log-Wage OLS	(2) $1\{\triangle w < 0\}$ Probit
Obtained health insurance	0.029	−0.073
	(0.056)	(0.048)
Obtained life insurance	0.069	−0.082
	(0.057)	(0.054)
Obtained retirement plan	0.055	−0.055
	(0.049)	(0.052)
Obtained flexible hours	0.012	−0.011
	(0.065)	(0.049)
Obtained stock options	0.079	−0.029
	(0.074)	(0.067)
Lost health insurance	−0.064	0.139
	(0.092)	(0.061)
Lost life insurance	−0.005	−0.015
	(0.096)	(0.074)
Lost retirement plan	−0.096	0.007
	(0.103)	(0.065)
Lost flexible hours	0.053	−0.005
	(0.071)	(0.054)
Lost stock options	−0.092	0.094
	(0.088)	(0.070)
Constant	0.098	—
	(0.026)	
R^2	0.025	0.043
N	759	759

Notes: Estimates in Column (2) are marginal effects.

From the perspective of attempting to estimate the marginal willingness to pay for non-wage job characteristics, the empirical relationship between changes in amenities at job transitions and changes in wages is of central importance. Column (1) of Table 3 shows a regression of the change in the log wage on changes in amenities at voluntary job-to-job transitions. The explanatory variables are a set of dummy variables indicating whether a worker "obtained" or "lost" each amenity, and all coefficients are measured relative to the base outcome of no change in the amenity. The point estimates for each of the "obtain" variables are all positive, but none are statistically different from zero at conventional levels. This is weak evidence that workers tend to experience wage increases when they move to jobs that offer amenities. Turning to the point estimates for the "lost" variables, four out of the five coefficients are negative, but again, the parameters are imprecisely estimated. Together, changes in benefits account for only 2.5 percent of the observed variation in log-wage changes.

Column (2) of Table 3 shows the estimated marginal effects from a probit model of wage declines at voluntary job-to-job transitions. The dependent variable in this model is equal to one if a worker chooses to move directly to a lower paying job. As in the OLS regression, many of the parameters are estimated imprecisely, and there is no clear evidence that workers are frequently accepting wage cuts in exchange for access to observable job amenities. For example, the marginal effect for "obtaining health insurance" is -0.073, which indicates that workers are *less* likely to experience a wage decline when they move to a job that offers health insurance. The pseudo R^2 for the wage decrease probit model is only 0.043, so changes in observable non-wage job characteristics fail to explain the vast majority of voluntary wage declines observed in the data.

Broadly speaking, these reduced form empirical results provide further justification for our decision to model the total value of non-wage amenities (ξ) as a random effect, above and beyond the fact that this practice has become widely adopted in the literature (Sullivan and To, 2014; Sorkin, 2015; Hall and Mueller, 2015). In addition, as our structural estimates will show, we find strong evidence that non-wage amenities are quantitatively important, despite the weak reduced form relationships found for observable amenities shown in Table 3.

4 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.

4.1 Distributional Assumptions and Exogenous Parameters

Measurement Error in Wages

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.[17] 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, $b(k) = \ln(0.35 \times e^{\mu_w(k)})$. This specification allows unemployment benefits to vary across types, so that agents with higher expected wages receive higher unemployment benefits.

4.2 Data Simulation

As discussed in Section 2, 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 K discrete types that make up the population distribution of unobserved

[17] 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).

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, $\{U^*(k)\}_{k=1}^K$, requires numerically solving Equation (B3). 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.

4.3 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).[18] 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. The complete list of moments is shown in Appendix C, Table 4.

Let $\boldsymbol{\theta} = \{\sigma_w, \sigma_\xi, \lambda_u, \lambda_l, \lambda_e, \lambda_{le}\} \cup \{\mu_w(k), \mu_\xi(k), \pi_k\}_{k=1}^K$ 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 Q moments. The SMD estimate of the structural parameters minimizes the weighted difference between the simulated and sample moments. Let m_q represent the qth moment in the data, and let $m_q^S(\boldsymbol{\theta})$ represent the qth simulated moment, where the superscript S denotes averaging across the S artificial datasets. The vector of differences between the simulated and actual moments is $g(\boldsymbol{\theta})' = [m_1 - m_1^S(\boldsymbol{\theta}), \ldots, m_Q - m_Q^S(\boldsymbol{\theta})]$, and the simulated minimum distance estimate of $\boldsymbol{\theta}$ minimizes the following objective function,

$$\Phi(\boldsymbol{\theta}) = g(\boldsymbol{\theta})' W g(\boldsymbol{\theta}), \tag{3}$$

[18] 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).

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 (3) using simulated annealing. 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.

4.4 Identification

This section discusses identification of a number of important model parameters. In the interest of brevity, we omit a detailed discussion the transition parameters (λ's), because the identification of these parameters using data on accepted wages and employment transitions is well established in the existing literature.[19] The model developed in this paper generalizes the standard search model by allowing for non-wage utility. Sullivan and To (2014) discuss in detail how job-specific non-wage utility can be identified using worker-level data on accepted wages and job transitions. The intuition is that the importance of non-wage utility is identified by revealed preference. When workers make job mobility decisions that appear inconsistent with pure income maximization, such as voluntarily moving to lower wage jobs, it provides information about non-wage utility.

Two important features of the model are unobserved worker heterogeneity and correlation between wage offers and non-wage utility flows. In many cases, the intuition behind identification of the parameters $\{\mu_w(k), \mu_\xi(k), \pi_k\}_{k=1}^{2}$ 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 jobs 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 $\mu_\xi(k)$ 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)

[19]See, for example, French and Taber (2011) for a thorough discussion of identification in search models.

Finally, it remains to discuss the identification of the correlation between wage offers and non-wage utility flows, $\rho_{w\xi}$. This object 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). Intuitively, if high wage workers also tend to have short unemployment durations, this feature of the data suggests that $\rho_{w\xi}$ is positive.

5 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 (Appendix C, Table C1) 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.

5.1 Job Offers and Labor Market Frictions

The parameter estimates are shown in Table 4. 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 (Figure 1). In these data, as workers move between jobs 1–3, mean wages, non-wage utilities and total utilities increase.

The estimate of the population correlation between wages and match-specific non-wage utility flows, $\rho_{w\xi} = 0.3451$, indicates that the two components of the value of a job in the model are positively correlated. Since it can be argued that some characteristics are positively correlated (health insurance benefits) while others are negatively correlated (risk of injury), this correlation does not seem unreasonable. Moreover the correlation is statistically different from zero at conventional significance levels. From the perspective of the literature on compensating wage differentials, $\rho_{w\xi}$ is an important parameter because it will tend to

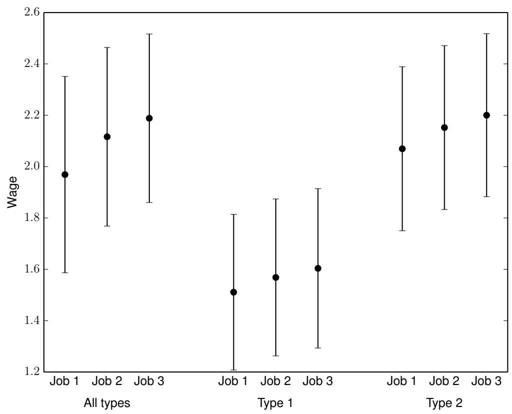
Table 4: Parameter Estimates

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)
Correlation(w, ξ)	$\rho_{w\xi}$	0.3451
		(0.0257)
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)
Type 1		
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)
Type 2		
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)

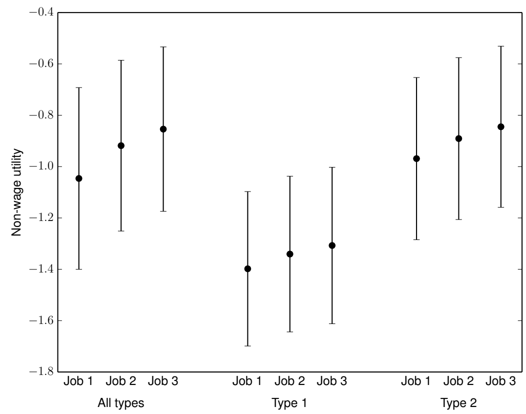
*The reservation utility levels are computed by solving Equation (B3) at the estimated parameters.

Figure 1: Mean simulated wages, non-wage utility and total utility

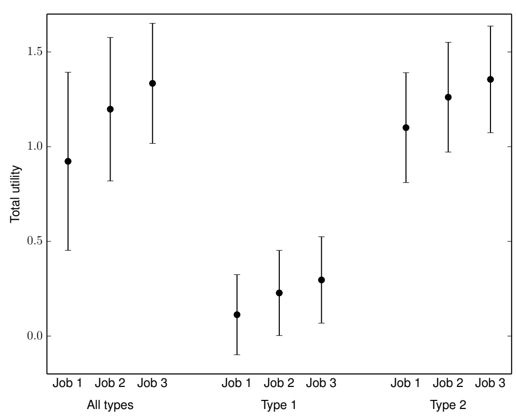
(a) Wages



(b) Non-wage utility



(c) Total utility

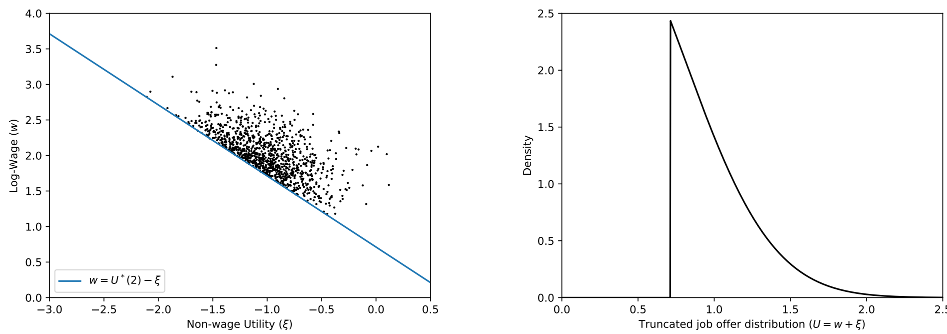


Intervals are the mean of the simulated values +/- one standard deviation.

Figure 2: First jobs – Type 2 workers

(a) Scatter plot

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



create a positive correlation between accepted pairs of wages and non-wage amenities in the cross section. Later in the section, we will examine how the job offer distribution, search frictions, and optimal worker search behavior jointly impact standard hedonic regressions.

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.[20] Not only are wages and non-wage utilities dispersed (Figure 2a) but total job values or "jobs" are also widely dispersed (Figure 2b). As discussed in Sullivan and To (2018) 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 35 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 ($\lambda_{le} = 0.0427$).

[20]Section 5.2 describes the differences in labor market outcomes between Type 1 and Type 2 workers. Section 6 provides a detailed analysis of how job-to-job mobility impacts accepted job offers.

5.2 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 $\{\mu_w(k), \mu_\xi(k), \pi_k\}_{k=1}^2$. The most common type of person in the economy makes up three-quarters of the population ($\pi_2 = 0.7503$), has an expected wage offer of $\mu_w(2) = 1.6392$, and expects 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 indicates that these workers receive 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). For instance, 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 Figure 1 that unobserved heterogeneity results in large differences in outcomes between Type 1 and 2 workers and between jobs within an employment cycle. As we will show, although controlling for worker ability reduces the bias observed in MWP estimates, it does 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.

6 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 following hedonic wage equation:

$$w_i = \alpha + \beta\xi_i + \beta_1 T_i \xi_i + \gamma T_i + \sum_{j=1}^{J-1} \zeta_j D_{ij} + e_i \tag{4}$$

Table 5: Hedonic wage regressions

Specification	Type Dummy (T_i)	Interaction ($T_i\xi_i$)	Job Dummies (D_{ij})	$\hat{\beta}$	$\hat{\beta}+\hat{\beta}_1$	R^2
1	N	N	N	-0.2160		0.0407
2	Y	N	N	-0.5780		0.5151
3	Y	Y	N	-0.7491	-0.5545	0.5179
4	N	N	Y	-0.2784		0.1147
5	Y	N	Y	-0.6127		0.5548
6	Y	Y	Y	-0.7597	-0.5921	0.5568

Regressions estimated using simulated data from the estimated model. "Y" indicates that the variable was included in the regression, "N" indicates that the variable was not included.

where the compensating wage differential literature interprets an estimate of β (or β and $\beta + \beta_1$) as the marginal-willingness-to-pay for ξ. At its most inclusive, this specification controls for the non-wage amenity, worker ability, job dynamics, and the interaction between the non-wage amenity and worker ability. Specifically, T_i is a dummy variable equal to 1 if worker i is of type 2. The variable D_{ij} is a dummy variable equal to 1 if worker i is employed in job number j. Job number refers to the consecutive job number within a job cycle. Table 5 presents $\hat{\beta}$ estimates and R^2 coefficients.

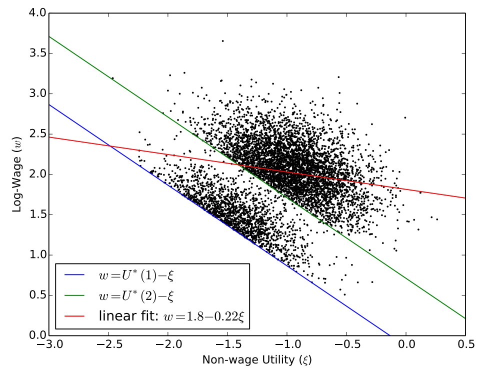
6.1 Wage Hedonics and Worker Ability

The standard, naïve wage hedonic with no additional controls ($w = \alpha + \beta\xi_i + e_i$) yields an extremely biased MWP estimate of $\hat{\beta} = -0.22$ and explains only 4 percent of the variation in wages (specification 1). Controlling for worker type using dummy variables yields a greatly improved but still significantly biased MWP estimate of -0.58, explaining 52 percent of wage variation (specification 2). 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, still explaining only 52 percent of wage variation (specification 3).

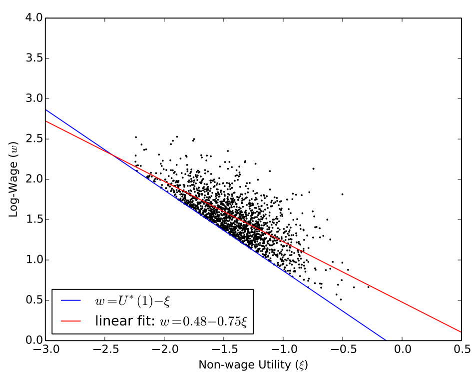
Figure 3 decomposes the effects of job dispersion and worker heterogeneity on hedonic MWP estimates. This figure plots steady-state cross-section wages and non-wage utilities in the simulated data pooled across both types (Figure 3a), and separately for each type (Figures 3b and 3c). Figures 3b and 3c shows conditional on type, there is a considerable amount of job dispersion in the simulated data. As a result, hedonic estimates of the MWP

Figure 3: Worker heterogeneity and bias

(a) Both types



(b) Type 1



(c) Type 2

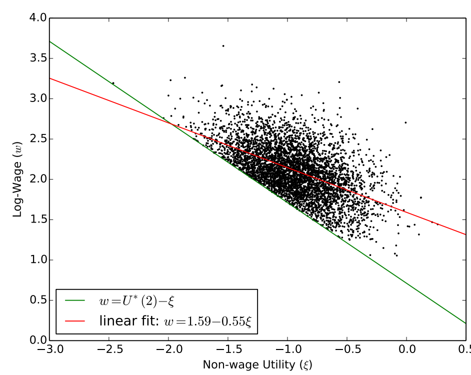
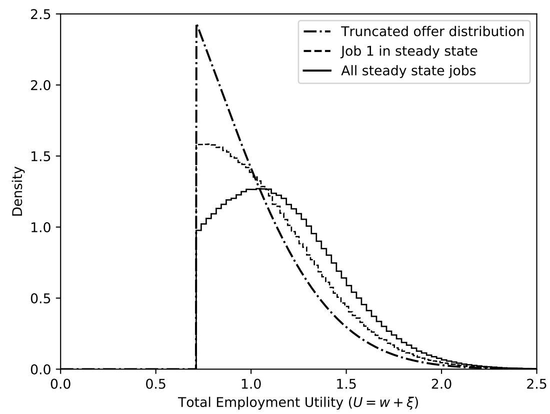


Figure 4: Job dispersion
Type 2 workers



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 Figures 3b and 3c 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 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 3a 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.

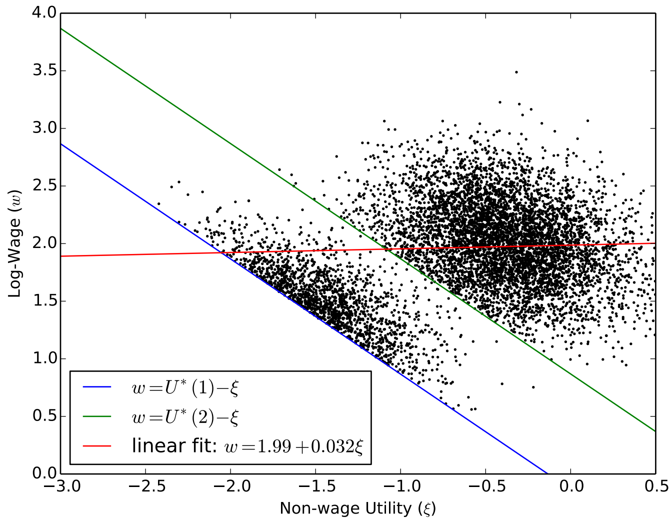
6.2 Search Frictions and Job Dynamics

The search frictions in our model contribute to job dispersion through the dynamics of job-to-job and job-to-unemployment transitions. In particular, these dynamics increase job dispersion through three 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. Third, some employed workers are involuntarily removed from the job ladder, reentering the job ladder at the bottom rung. Without involuntary transitions of employed workers, the former two channels could reduce job dispersion as workers are eventually all in "good" jobs. To illustrate, 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 of workers' first jobs, as workers in bad first jobs accept better jobs, the histogram 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 a job cycle but exclude ability controls the MWP estimate moderately improves, falling from -0.22 to -0.28, explaining roughly 11 percent of wage variation (specification 4, Table 5). Even controlling for both type and job number, the MWP estimate is still significantly biased at -0.61 (specification 5, Table 5). This specification is able to explain only about 55 percent of the variation in wages. Our most inclusive specification (6, Table 5) yields MWP estimates of -0.76 and -0.59, explaining 56 percent of the variation in wages.

Interestingly, job dynamics, while observable to the practitioner, are an imperfect proxy for the dispersion of worker reservation utilities due to job dynamics. However, as the modelers, we 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. If we substitute these reservation utilities, U_i^{**}, for the job dummies from for specification 5 (or 6) of Table 5, the estimated MWP improves to -0.64 (or -0.76 and -0.62) and is able to explain 59 percent of the variation in wages. But even in this unrealistic, best case, MWP estimates are significantly biased and 41 percent of the variation in wages remains unexplained.

Figure 5: Worker heterogeneity and wrong signed MWP



6.3 Worker Heterogeneity and Wrong Signed MWP

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.

6.4 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

Table 6: Bias $(1+\hat{\beta})$ for counterfactual job offer dispersions and search friction indices

		Type 1			Type 2			Population		
	σ_u	0.283	*0.565*	0.848	0.283	*0.565*	0.848	0.283	*0.565*	0.848
κ	3.673	0.113	0.242	0.318	0.308	0.400	0.432	0.338	0.718	0.658
	7.345	0.113	**0.254**	0.344	0.333	**0.445**	0.484	0.368	**0.784**	0.714
	11.018	0.117	0.267	0.371	0.359	0.486	0.527	0.397	0.834	0.758

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 search frictions and dispersion in job offers. In particular, in our simulated labor market where it is possible to perfectly control for ability differences across workers, MWP estimates are seriously biased.

Our results suggest that worker heterogeneity and dispersion in job offers appear to be of greater significance than search frictions as contributors to biased cross-sectional MWP estimates. To further evaluate this observation, we simulate counterfactual datasets where we vary the dispersion in total utility ($\sigma_u = \sqrt{{\sigma_w}^2 + {\sigma_\xi}^2}$) and in the level of search frictions[21] ($\kappa = \lambda_e/(\lambda_l + \lambda_{le})$) and estimate hedonic wage regressions. Table 6 shows MWP biases for increasing and decreasing offer dispersions and search frictions by 50 percent.[22] The italics numbers are baseline parameters, the bold numbers are the bias at the baseline parameters and the other numbers are those for counterfactual simulations. It can be seen that, disaggregated by type, frictions and offer dispersion are complementary[23] but offer dispersion tends to be a more significant determinant of bias than the level search frictions. These relations are even more apparent graphically as shown in Figure 6 where frictions (κ) and offer dispersion (σ_u) are varied on the horizontal axes and the magnitude of the bias is on the vertical axis. These findings are in accordance with our prior analysis showing that while search-frictions/job-dynamics are significant contributers to the bias in MWP estimates, they are less important than worker heterogeneity and job offer dispersion.

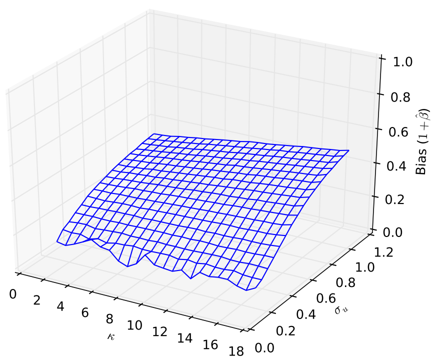
[21] This expression for κ follows from the frequently used Ridder and van den Berg (2003) definition.

[22] To vary σ_u by factor ϕ, we multiply σ_w and σ_ξ by ϕ and to vary κ we multiply and λ_e by $\sqrt{\phi}$ and divide λ_l and λ_{le} by $\sqrt{\phi}$.

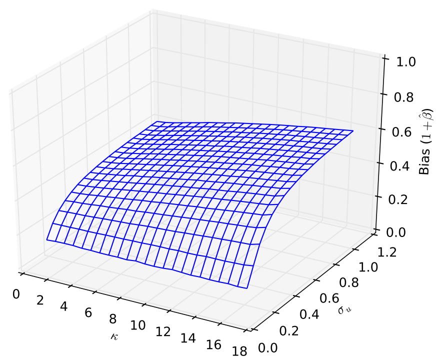
[23] Aggregated, frictions and offer dispersion are complementary until dispersion is relatively large at which point further increases in offer dispersion actually reduce bias as it becomes easier for Type 1 workers to eventually get good jobs.

Figure 6: Bias by κ and σ_u

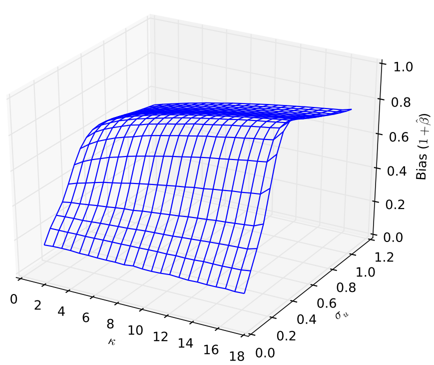
(a) Type 1



(b) Type 2



(c) Both types



7 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 select a relatively uniform sample and control for unobserved 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 job dynamics ("utility ladder") 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 Additive separability

Given the stationarity of the worker's problem, our additively separable utility function is quite general, encompassing any Cobb-Douglas function over wages and non-wage utility. Take $U(w, \xi) = Aw^{\alpha}\xi^{\beta}$.

Taking logs and dividing by α,

$$\frac{\ln(U(w, \xi))}{\alpha} = \ln w + \frac{\ln(A\xi^{\beta})}{\alpha}.$$

Defining $\tilde{w} = \ln w$ and $\tilde{\xi} = \ln(A\xi^{\beta})/\alpha$, call the functionally equivalent, transformed utility function, $\tilde{U}(\tilde{w}, \tilde{\xi}) = \tilde{w} + \tilde{\xi}$. This is precisely our assumed functional form.

B 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 (2) and (1) so that common terms can be collected when evaluated at $U = U^*$. Subtracting $\delta V^e(U)$ from both sides of (2):

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

Evaluating this at $U = U^*$:

$$\begin{aligned}(1-\delta)V^e(U^*) &= U^* + \delta\left[\lambda_e \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U') + \lambda_{le} \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U')\right] \\ &= U^* + \delta(\lambda_e + \lambda_{le}) \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U')\end{aligned} \tag{B1}$$

Similarly, subtracting δV^u from both sides of (1),

$$\begin{aligned}(1-\delta)V^u &= b + \delta\lambda_u E \max\{0, V^e(U') - V^u\} \\ &= b + \delta\lambda_u \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U').\end{aligned} \tag{B2}$$

Evaluating at $U = U^*$, we can equate (B1) and (B2), integrate by parts and solve to get:

$$\begin{aligned}U^* &= b + \delta[\lambda_u - (\lambda_e + \lambda_{le})] \int_{U^*}^{\infty} [V^e(U') - V^u]dH(U') \\ &= b + \delta[\lambda_u - (\lambda_e + \lambda_{le})] \int_{U^*}^{\infty} V^{e\prime}(U')[1 - H(U)]dU' \\ &= b + \delta[\lambda_u - (\lambda_e + \lambda_{le})] \int_{U^*}^{\infty} \frac{1 - H(U')}{(1-\delta) + \delta\{\lambda_e[1 - H(U')] + \lambda_l + \lambda_{le}\}} dU'.\end{aligned} \tag{B3}$$

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.

C Estimation Moments

Table C1: 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
	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

Continued on next page

Table C1 – continued from previous page

Moment #	Description	Data	Simulated
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.

Bhaskar, V. and T. To, 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, New York: North Holland Publishing.

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, New York: McGraw Hill.

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

Eckstein, Z. and G. J. van den Berg, 2007, “Empirical Labor Search: A Survey,” *Journal of Econometrics*, 136: 531–564.

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

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.

Hall, R. E. and A. I. Mueller, 2015, "Wage Dispersion and Search Behavior," NBER Working Paper 21764.

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.

Krueger, A. B. and A. I. Mueller, 2011, "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.*

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

Ridder, G. and G. J. van den Berg, 2003, "Measuring Labor Market Frictions: A Cross-Country Comparison," *Journal of the European Economic Association*, 1: 224–244.

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, New York: Elsevier.

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

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

Sorkin, I., 2015, "Ranking Firms Using Revealed Preference," Working paper.

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

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

Sullivan, P. and T. To, 2014, "Search and Non-Wage Job Characteristics," *Journal of Human Resources*, 49(2)(Spring 2014): 472–507.

Sullivan, P. and T. To, 2018, "When are Wage Differentials Compensating?" Working paper.

Taber, C. and R. Vejlin, 2016, "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*, 28(3): 595–631.