Princeton University
Department of Economics

Wage Dispersion and Preferred Workers:

An Insider-Outsider Search Theoretic Model of the Labor Market

Submitted to Princeton University
Department of Economics
In Partial Fulfillment of the Requirements for the A.B. Degree

April 13, 2016

Acknowledgements

First, I would like to thank my advisor, Greg Kaplan, for his guidance throughout the process of writing this thesis. His dedication to teaching and his passion for economic research are abundantly clear in the way he engages with and listens to his students.

I would also like to thank the numerous professors who have motivated my interest in economics and economic research while at Princeton, including but not limited to Alan Blinder, Nobuhiro Kiyotaki and Dilip Abreu.

Michael Apicella, of the University of Iowa, thank you for your mentorship – your respect of my ideas and commitment to teaching cultivated within me five years ago an interest in research that still resonates today, even if the topic of research changed from biofilms to labor markets.

To all of my friends at Princeton, thank you for making this place and this experience what it was. A special thanks to James, Joe and Evan – without you I would still be a chemistry major.

Neal and Claire – it's immeasurably easier to walk a road already paved by others, and for that I am eternally grateful.

Finally, Mom and Dad, I could not imagine better exemplars of integrity and devotion, both of which you apply to all that you do. Thank you for your unconditional love and support.

Contents

Abstract

In this paper, we construct a search theoretic model that allows for heterogeneous attachment to the labor market in order to explain wage dispersion. We solve for the invariant wage distribution analytically. We also use a Markov chain to numerically solve a discretized version of the model by solving for the transition probability between unemployment and the different wage states. Finally, we apply this model to U.S. wage data, using a generalized method of moments approach to estimate the parameters. Compared to the data, our model overestimates wage dispersion, placing excess density at the lower end of the distribution. However, the addition of density in the upper tail relative to the traditional on-the-job search model matches the fat tails found in the data as shown by the ratio of wages at the upper end and lower end of the wage distribution, particularly for the 90-10 and 99-10 ratios.

1 Introduction

There is widespread empirical evidence of increasing wage inequality in the U.S. starting from the 1980s. Piketty and Saez (2001) find that the trajectory of income inequality is U-shaped since the beginning of the 20th century, first falling dramatically through the Great Depression and World War II and then increasing in the postwar period through the present. But it is not only the level of inequality that changes over time. Heathcote et al. (2010) find that the composition of wage inequality varies as well, primarily affecting the lower end of the wage distribution in the 1970s and moving to the upper tail through the 1980s and 1990s. Lee (1999) is able to partially explain the early wage dispersion, arguing that the real minimum wage decreases over this time, relaxing the floor at the bottom end of the wage distribution. But Autor et al. (2008) reject Lee's minimum wage argument as unable to sufficiently explain observed wage dispersion on the upper end of the wage distribution. Changes in minimum wage policy were one-time shocks to the labor market in the 1980s, so if such changes were truly the cause of the widening wage dispersion, then income inequality would be due to nonrecurring non-market factors. Autor et al. argue instead that observed wage inequality is due to secular changes in demand for workers and is rooted in the skill-based technical change hypothesis, where high-skilled workers benefit from technological change at the expense of moderately-skilled workers.

The skills gap alone does not explain observed wage dispersion, though admittedly it may play a critical role in its increase over the past decades. Skills-based technical change as an explanation for wage inequality is consistent with the theory of competitive markets, as workers are paid their marginal product. However, Krueger and Summers (1988) and Mortensen (2003), among others, provide empirical evidence of wage dispersion, even when controlling for ability. This finding implies that there are frictions in the labor market that prevent the efficient allocation of labor market resources.

Assuming in practice that rent-sharing occurs to some extent in wage negotiation, the frictions in the labor market can either come from firm-based hiring practices or worker-based acceptance decisions. Abowd et al. (1999) find that while heterogeneity of personal characteristics among workers, like in unemployment history, is important in determining wage differentials between workers, firm heterogeneity also plays a substantial role, so wage dispersion could come from different wage setting policies between firms. The ability of different firms to pay similar workers different wages could imply higher bargaining power on the part of the firm relative to the worker. If workers of similar skill had greater bargaining power, then one might expect to see relatively homogeneous wages as all workers bargain for a similar wage. On the other hand, if firms with greater bargaining power have different wage-setting policies, then one might expect to see wage dispersion.

This paper uses the insider-outsider model of labor markets to explore the role of firm-induced wage dispersion. The bases of the insider-outsider model, described by Lindbeck and Snower (2002), are turnover costs. The combination of costs associated with firing a current employee and finding, hiring and training a new employee can be non-negligible, where the current employee is the insider, while the potential new employee is the outsider. The existence of these costs creates a wedge in the hiring process, so that in order for outsider to replace an insider, his or her productivity must not only exceed the insider's, but it must also exceed the turnover costs. Generally, the insider-outsider model is useful in that it explains why certain workers or groups of workers are more attached to the labor market than others.

There are many applications of the insider-outsider model of the labor market. One of the most prominent examples is to describe the interaction between the employed and those looking for work. Blanchard and Summers (1986) and Ball (2009) show that hysteresis, or the degradation of valuable skills and a professional network, can occur when workers are long-term unemployed. The loss of labor market contacts makes re-entering the work force more difficult. Krueger et al. (2014) find evidence of an insider-outsider interaction in the U.S. during the Great Recession. They argue that the long-term unemployed from the Great Recession failed to exert significant downward pressure on wages, and were therefore only marginally attached to the labor market. The insider-outsider model can also be used to describe dynamics between unionized and non-unionized workers. Taschereau-Dumouchel (2015) develops a search model that describes the effect of unionization on firm hiring decisions. However, in the scope of this paper, insiders will represent those who are attached to the labor market, and outsiders represent those who are less attached.

In this paper, we use a search theoretic model with on-the-job search and endogenous search effort, in the spirit of Burdett (1978), Christensen et al. (2005), Rogerson et al. (2005), and Lise (2012) to test the effect that an insider-outsider model of the labor market has on wage dispersion. We insert the insider-outsider model into the traditional search theoretic model by allowing workers to decide to accept tenure, which decreases their job separation rate. However, because this model includes on-the-job search, accepting tenure also precludes the acceptance of a higher paying job. In my model, workers, upon receipt of a wage offer, must make two decisions. The first, similar to the models upon which this one is built, is whether or not to accept employment at the given wage. The second is whether or not to pay a set cost and accept tenure. Intuitively, workers who receive a low wage offer will value highly the possibility of receiving further wage increases and are likely to reject tenure, while those who receive a high wage offer will have smaller expected wage increases, and will therefore be willing to sacrifice further search in favor of a lower job separation rate.

The remainder of this paper is organized as follows. Section 2 outlines the search model of the labor market with tenure. Section 3 describes the steady state distribution that the model produces. Section 4 uses a generalized method of moments approach to estimate the parameters of the model, using household wage data from the Current Population Survey. Section 5 discusses the implications of the model for income inequality, and section 6 concludes and lays out paths for future research.

2 An Insider-Outsider Search Theoretic Model

Workers are infinitely-lived risk averse agents and have identical preferences and productivities. Workers gain utility from the wages that they accept, and disutility from searching for work and transitioning from being an outsider to an insider. The transition cost can be interpreted as either a loss of leisure or a monetary cost that is associated with the learning the additional skills that are necessary for insider employment. Utility functions follow Constant Relative Risk Aversion preferences. For the outsiders, preferences are as follows:

$$u_o(c) = \frac{c^{1-\zeta} - 1}{1-\zeta}$$

Insiders must face the cost of transitioning from outsider to insider, and their preferences are therefore the following:

$$u_i(c) = \frac{c^{1-\zeta} - 1}{1-\zeta} - \frac{E^{1-\zeta} - 1}{1-\zeta}$$

where E is the cost of becoming an insider, $\zeta > 0$ is the measure of risk aversion, with an increasing value corresponding to more risk averse preferences. For values of $\zeta = 1$, let workers have log-utility.

Let w be the wage that a worker accepts, drawn from a distribution $F(w)$, the cumulative distribution of wage offers. For outsiders, job offers follow a Poisson process and are received at an exogenous rate λ, weighted by the endogenous search intensity $s(w)$. The cost of search has the power form, as follows:

$$c(s) = \frac{\eta s^\gamma}{\gamma}$$

where η is a scalar, and $\gamma > 1$ denotes the elasticity of search, or how responsive the cost of search is to the effort put in. Outsiders lose their job, and transition to unemployment, with exogenous rate δ_o. Because we allow for on-the-job search in this model, outsiders can transition out of employment if they find employment at a higher wage or if they lose employment and become unemployed. Insiders have greater job security, but they lose on-the-job search. This means that insiders can only transition out of employment at rate δ_i, where $\delta_i < \delta_o$, because $\lambda = 0$.

Assuming that workers maximize their lifetime wealth, the continuous-time Bellman equations for unemployed, outsiders and insiders are as follows:

$$\begin{aligned} V_o^u &= u_o(b) - c(s) \\ &\quad + \beta\lambda s(b) \max\Bigg\{ \int_W \Big[\max\{V_o^u, V_o^e(x)\} - V_o^u\Big] dF(x), \\ &\int_W \Big[\max\{V_i(x), V_o^u\} - V_o^u\Big] dF(x)\Bigg\}, \end{aligned} \tag{1}$$

$$\begin{aligned} V_o^e(w) &= u_o(w) - c(s) \\ &\quad + \beta \max\Bigg\{\lambda s(w) \int_W \Big[\max\{V_o^e(w), V_o^e(x)\} - V_o^e(w)\Big] dF(x) \\ &\quad + \delta_o\big[V_o^u - V_o^e(w)\big], \\ &\quad \lambda s(w) \int_W \Big[\max\{V_i(x), V_o^e(w)\} - V_o^e(w)\Big] dF(x)\Bigg\}, \end{aligned} \tag{2}$$

$$V_i(w) = u_i(w) + \beta\Big[V_i(w) + \delta_i\big[V_o^u - V_i(w)\big]\Big] \tag{3}$$

The value functions $V_o^u, V_u^e(w), V_i(w)$ represent, respectively, the expected lifetime value of being unemployed, holding an outsider job at wage w, and holding an insider job at wage w, for wage $w \in W$, the set of all possible wages.

Starting with the value of the unemployed worker, the first two terms represent net current income from being unemployed. Unemployed workers get utility from unemployment insurance b but face a cost of searching for work, $c(s)$. The term on the second line represents the discounted value of finding work, where β is the continuous discount factor and λ is the exogenous rate at which workers receive job offers, weighted by search effort $s(w)$. The integral is the expected value of the new job offer when it is received, which is the expected value of receiving a wage above what the worker already has, weighted by the probability of receiving a wage above what the worker already has. If the worker receives a wage lower than the current unemployment benefits, it is preferable to reject the wage offer and continue being unemployed, and in such cases the value of the integral is equal to zero. The final term is the expected value of receiving a wage w and deciding to take it with insider status.

The value function for the employed outsider is analogous to the value function for the unemployed worker. The only conceptual difference between the two is the addition of the term on the third line, which represents the loss in value from losing employment and joining the ranks of unemployed workers. Employed outsiders lose work at the exogenous rate δ_o, and the magnitude of the loss in value is given by the difference between outsider employment and unemployment.

The workers have two decisions to make in this model, denoted by the two *max* expressions, as agents in this model maximize their lifetime expected value. The first, within the integral, is the decision about whether to accept the job offer at the given wage or remain in his or her current state. If the worker decides to accept the job, he or she must then decide whether to accept the job as an outsider, or to pay a cost and accept the job as an insider.

Let us assume a wage reservation strategy for the worker. That is, there is a wage range $[w_{min}, w_R]$ where workers will reject all wage offers, $[w_R, w_T]$ where they accept outsider jobs, and $[w_T, w_{max}]$ where they accept insider jobs. With the introduction of reservation wages, we can further simplify Equation (2) to

$$V_o^e(w) = \frac{u_o(w) - c(s) + \beta\lambda s(w)\left[\int_w^{w_T} V_o^e(x)dF(x) + \int_{w_T}^{w_{max}} V_i(x)dF(x)\right] + \beta\delta_o V_o^u}{1 + \beta\lambda s(w)[1 - F(w)] + \beta\delta_o} \tag{4}$$

Written in this form, it becomes evident that the value function for outsiders satisfies both monotonicity and discounting, and therefore satisfies Blackwell's sufficient conditions for a contraction mapping. By the Contraction Mapping Theorem, we find existence and uniqueness of the value function, following Stokey and Lucas (1989). The Bellman equation above, in which current value is written in terms of current utility and future expected value, can be solved using recursive dynamic programming, following Sargent and Ljungqvist (2004). We solve for the value function by starting with an initial value for V_o^u, $V_o^e(w)$ and $V_i(w)$, and iterating until convergence. We choose an initial value of $V_*(w) = \frac{w}{1-\beta}$, where V_* is one of the three value functions, and $w = b$ when $V_* = V_o^u$. The initial function V_* is equivalent to the lifetime utility if a worker accepts a job permanently, without further search or a positive probability of transitioning to unemployment.

However, given the complexity of the Bellman equations, particularly with endogenous search as outlined below, it is not possible to analytically solve for each value function. We therefore solve the value functions numerically, as shown in Figure 1, where the wage range is divided evenly into 1000 steps. In terms of wage decisions, this approximation does not seem problematic, because empirically, wages are discretized, so this modification is not likely to have a substantial impact on workers' wage acceptance decisions.

At each wage offer, workers make their decision about whether to accept the wage offer and whether to become an insider such that their lifetime expected utility is maximized. The worker's value function is therefore piecewise, and is calculated according to the maximum value – unemployed, outsider, insider – at each wage. The existence of the insider status creates an additional kink in

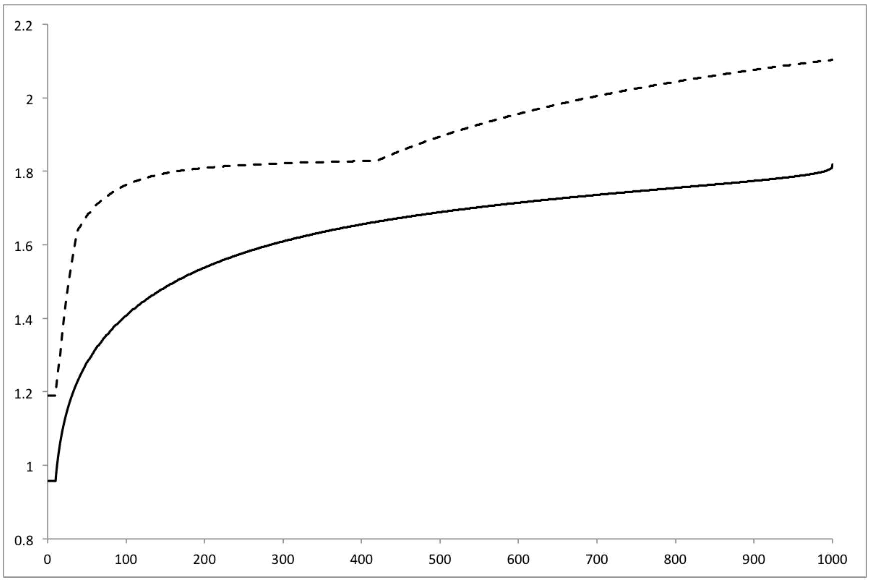


Figure 1: The filled line shows the value function for on-the-job search without tenure, while the dashed line shows for with tenure. The plot shows the value of being employed at a given wage w. The wages are normalized from 1 to 1000.

the value function, relative to the traditional on-the-job search model. Figure 1 shows the value function for on-the-job search without tenure as outlined in Christensen et al. (2005), in black, along with the value function with tenure as characterized by Equations (1)-(3), shown in the dashed line. In Christensen's model, the value function is the maximum of the value of being unemployed and employed. The value function with tenure adds in insider workers, which accounts for the additional kink.

For wages below the reservation wage, it is preferable to be unemployed and receive b, which can be interpreted as unemployment insurance, dividends, or some other source of income unrelated to employment. In this portion of the function, value is independent of wage, and is derived from the potential of future employment at a wage greater than w_R.

There are two interesting differences between the two value functions. The most obvious difference is for wages $w \in [w_T, w_{max}]$. In this range, workers in my model have tenure, so their job separation rate is significantly lower than in Christensen et al. A job with tenure is more valuable, because the prospect of becoming unemployed is much lower. However, we also notice that the value function for the insider-outsider model is higher in the range $w = [w_R, w_T]$ when compared with the model of on-the-job search without tenure. The difference lies in the gap in potential earnings. Workers in both models have the same utility functions, so the wage itself adds the same value to both. However, when workers have the ability to gain tenure, the value of jobs without tenure still increases because they incorporate the potential for becoming an insider and gaining a tenured position in the future.

The set up in Equation (4) is useful because it allows us to easily take the first-order conditions with respect to w. Using the Benveniste-Scheinkman theorem, we find that

$$V_o^{e'}(w) = \frac{u_o'(w)}{1 + \beta\lambda s(w)\left[1 - F(w)\right] + \beta\delta_o} > 0 \tag{5}$$

where $s(w)$ is the optimized search effort with respect to the given wage. Note that the value of employment $V_o^e(w)$ is increasing in w, so outsiders will always accept jobs with higher wage offers. Note from Equation (1) that there exists a wage at which point workers are indifferent between outsider employment and unemployment, and we denote this wage w_R. It follows that $w_R = b$.

Suppose that $w_R < b$. Then workers would accept a job offer that paid wage $w = [w_R, b)$. However, they receive such a wage w, which leaves the worker strictly worse off than accepting b, with probability $\int_w^b dF(x) > 0$, while they could receive unemployment insurance b with probability 1. For $w < b$, it is preferable to be unemployed, contradicting the assumption that $w_R < b$. A similar argument follows if we suppose that $w_R > b$. Then the worker rejects jobs with wage offers $w = [b, w_R)$ that make the worker strictly better off, even though such job offers are received with probability $\int_b^{w_R} dF(x) > 0$. There is a positive expected value of wages $w = [b, w_R)$, so it follows that $w_R = b$ in order to prevent the loss of surplus. The finding that $w_R = b$ is a result of the fact that there is on-the-job search in this model, so there is no trade-off between accepting a job and receiving a job offer that pays a higher wage.

In this model, search is costly, so workers optimize their effort so that the expected marginal benefit of continuing to search for another job is equal to the marginal cost of searching. Taking the first-order conditions of Equation (2) with respect to s allows us to solve for the optimal search effort:

$$c'(s) = \beta\lambda \left[\int_{w}^{w_T} V_o^e(x) - V_o^e(w) dF(x) + \int_{w_T}^{w_{max}} V_i(x) - V_o^e(w) dF(x) \right]$$

Making use of integration by parts, we can simplify $c'(s)$ further:

$$c'(s) = \beta\lambda \left[\int_{w}^{w_T} \left[1 + F(w_T) - F(x)\right] V_o^{e'}(x) dx + \int_{w_T}^{w_{max}} \left[1 - F(x)\right] V_i'(x) dx \right]$$

Using Equation (5) above and the fact that $V_i'(w) = \frac{u'(w)}{1-\beta+\beta\delta}$, $c'(s)$ becomes

$$c'(s) = \beta\lambda \left[\int_{w}^{w_T} \frac{\left[1 + F(w_T) - F(x)\right] u_o'(w)}{1 + \beta\lambda s(w) \left[1 - F(w)\right] + \beta\delta_o} dx + \int_{w_T}^{w_{max}} \frac{\left[1 - F(x)\right] u'(w)}{1 - \beta + \beta\delta_o} dx \right] \quad (6)$$

Equation (6) implicitly solves for the optimal search effort. Solving for $s(w)$, we find that search effort is decreasing in w. This is intuitively satisfying because as the wage offer increases, the value of outsider employment increases but the cost of search does as well, so there is less incentive to work to find a higher paying job.

Figure 2 shows how search effort depends on wage for the model with and without tenure. An interesting difference between the two functions is the level that effort converges to as wage increases. For the on-the-job search model with tenure, search effort eventually goes to zero. The intuition behind this is relatively straightforward – as the wage increases, the probability that the worker receives a job offer with a higher wage goes to zero. The marginal benefit of searching thus decreases dramatically. Furthermore, the marginal cost of searching is increasing in effort, so in order for the marginal cost of searching to equal its marginal benefit, search effort goes to zero.

Interestingly, search effort for the insider-outsider model asymptotically ap-

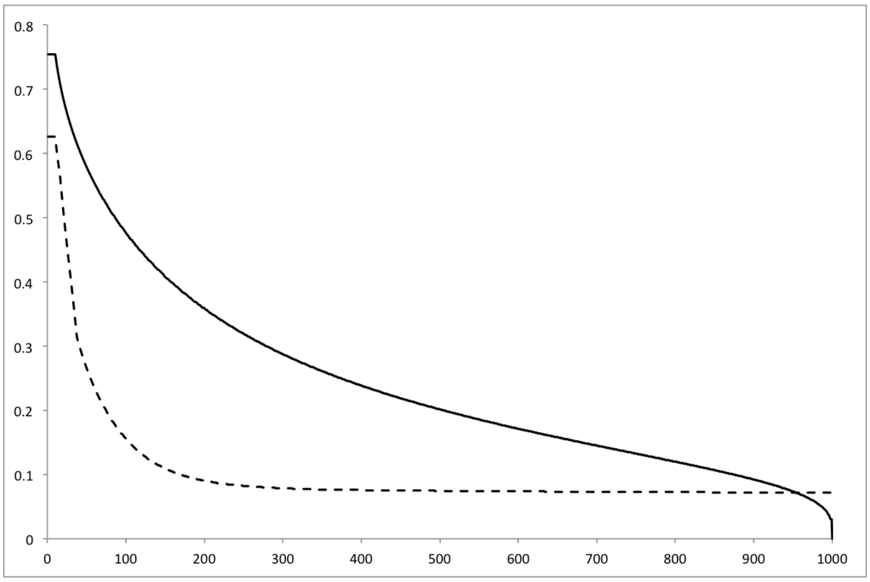


Figure 2: The filled line shows the search effort for on-the-job search without tenure but with endogenous search, while the dashed line show the effort when tenure is added to the model. The plot shows the search effort given wage w. The wages are normalized from 1 to 1000.

proaches a value greater than zero. This is again because of the ability to gain tenure. Only outsiders have the ability to search, and they have a large incentive to continue searching until they become an insider, as shown in Figure 1. In my model, workers reduce their search effort as wages increase up until w_T, at which point they decide to become an insider and give up search all together. This differs from Christensen's model, where workers search up until w_{max}. The existence of the potential for becoming an insider means that the marginal benefit of searching does not go to zero as the current wage increases. It is worth noting in Figure 2 that because insiders do not search, search effort should be zero for $w > w_T$. However, we leave the search effort over the full wage range in order to demonstrate its convergence to a value greater than zero.

One final difference between the search effort in the two models is the overall level. Over the whole wage range, except for at the very high end, search effort is lower when the possibility of tenure exists. At first glance, this may appear to be counterintuitive – if workers are unemployed or have a low-paying job as an outsider, then it seems plausible that they will be incentivized to search harder, because if they find a high enough paying job, then they can become an insider. As shown in Figure 1, the value of being an insider is larger than that of an outsider, so utility maximizing workers would want to put in more effort to find such an insider position.

However, we note that this cannot be the case, as the search effort is lower when tenure is allowed. The reason for the lower search effort is that the amount of effort is based on the marginal benefit of search. Note in Figure 1 that the value function with tenure strictly dominates the value function when tenure is not allowed, and while it increases quickly for low wages, it flattens out rather quickly, so the additional value gained from a higher wage is smaller than it is without tenure.

3 The Invariant Wage Distribution

To model the flows between unemployment and employment at wage w, we discretize the wage range $[w_{min}, w_{max}]$ and use a Markov chain with $n+1$ states, where n is the number of discretizations of the wage range, and the additional state represents unemployment. The resulting transition probability matrix P is aperiodic and positive recurrent, and therefore converges to a steady state distribution of accepted job offers, $G(w)$. In this paper, we only make use of observed wage data, and do not rely on observed wage offers, so in order to estimate the parameters of this model, $G(w)$ is necessary.

Workers flow from unemployment to employment at rate $\lambda s(w_R)[1-F(w_R)]$, and flow from employment to unemployment at rate δ_o for outsiders and δ_i for insiders. For workers in the unemployed state, the probability that they are unemployed the following period is the same as the probability that they do not receive a job offer above their reservation wage, or $1 - \lambda s(w_R) \int_{w_R}^{w_{max}} dF(x) = 1 - \lambda s(w_R)[1 - F(w_R)]$. The probability that an unemployed worker receives wage w is $\lambda s(w_R) f(w)$, where $f(w)$ is the density of wage offer distribution.

For workers employed at wage $w \in [w_R, w_T)$, the probability that they are unemployed in the next state is δ_o. Because this model allows for on-the-job search, outsiders will either move from employment to unemployment, remain employed at wage w, or accept a new job with wage $w^* > w$. This means that the probability of moving to a new state with a lower wage is zero. The probability of remaining employed at wage w is the same as the probability that the worker receives no new offers that pay less than the current wage, or $1 - \delta_o - \lambda s(w) \int_w^{w_{max}} dF(x) = 1 - \delta_o - \lambda s(w)[1 - F(w)]$. Finally, the probability of moving to a state with wage w^* is $\lambda s(w) f(w^*)$, similar to the unemployed worker.

For insiders, the dynamics are less complex, because they give up on-the-job search in favor of higher job security. Therefore, an insider employed at wage w will move to an unemployed state with probability δ_i, and remain in their current state with probability $1 - \delta_i$.

The dynamics described above are sufficient to create the transition probability matrix P, outlined more completely in Figure 7 in the Appendix. Iterating P on itself, for sufficiently large n, $G(w) = P^n$ converges to a steady state distribution of accepted wages, as shown in Figure 3. On-the-job search, both with and without tenure, increases wage dispersion when compared to the distribution of wage offers. The addition of tenure into the model further increases wage dispersion, as shown in the dashed line. The existence of insider positions means that it is more likely for a worker to hold a higher paying job, as high paying jobs come more frequently because search effort is greater in the insider-outsider model. And once a worker receives a high paying job, the probability of losing work and transitioning to unemployment is lower, because $\delta_i < \delta_o$. The extent to which wage dispersion occurs in the insider-outsider model depends on the differential between the insiders' and outsiders' respective rate of job separation, and the cost of becoming an insider, represented by E in the insider's utility function. Workers flow out of unemployment according

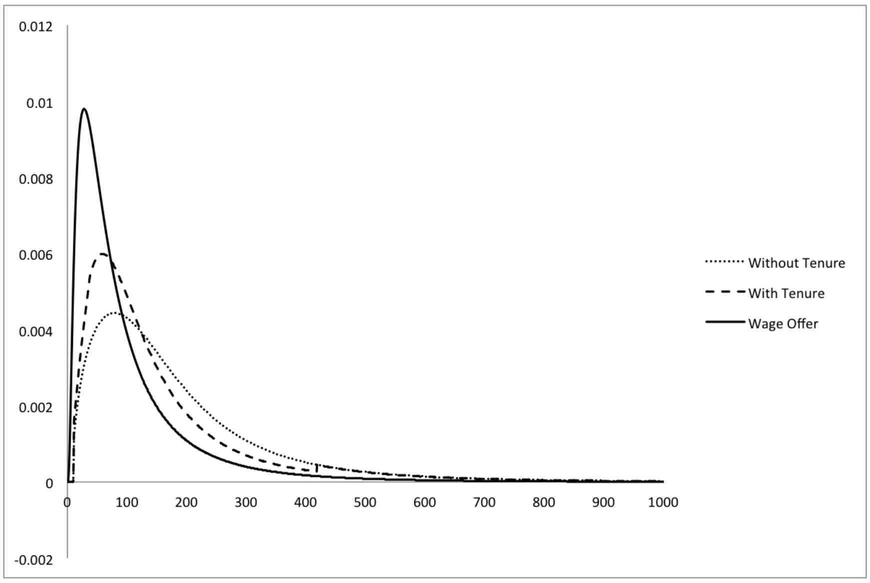


Figure 3: The filled line shows the original density of wage offers. The dotted line shows the distribution of accepted wage offers for the on-the-job search model without tenure, and the dashed line shows the accepted distribution when tenure is introduced. We use the estimated parameters for the model with tenure to calculate the distribution without tenure to allow for an appropriate comparison. The wages are normalized from 1 to 1000.

to $\lambda s(w_R)\,[1 - F(w_R)]\,u$, which is the probability of receiving an offer above the reservation wage, scaled by the fraction of workers who are unemployed. Worker flows into unemployment can come from both insiders and outsiders that lose their jobs, and is given by $\delta_o(1-u-i)+\delta_i i$. Therefore, in equilibrium, it follows that

$$\lambda s(w_R)\,[1 - F(w_R)]\,u = \delta_o(1 - u - i) + \delta_i i \tag{7}$$

While interesting in that it allows us to determine the fraction of unemployed, outsiders and insiders, analysis of flows in and out of unemployment do not reveal a relationship between the distribution of wage offers, $F(w)$, and of accepted wages, $G(w)$. To do this we must analyze flows in and out of both outsider and insider employment.

To evaluate the flows in and out of outside employment, we focus on the wage range $w \in [w_R, w_T]$. Unemployed workers enter into outside employment with a wage of w or less according to $\lambda s(w_R)\,[F(w) - F(w_R)]\,u$. Workers can leave outside employment that pays $[w_R, w]$ by either losing employment or finding a higher paying job. Outsiders flow to unemployment according to $\delta_o(1-u)G(w)$, which is the probability of job separation weighted by the fraction of outsiders and the probability of accepting a job of w or less. Note that flows for employed workers depend on the distribution of accepted wages, $G(w)$, rather than the distribution of wage offers, $F(w)$. This is important because wage offers that are not accepted are not relevant to the distribution of wages.

Workers can also leave their current job by finding another job that pays a higher wage. This flow is given by $\lambda \int_{w_R}^{w} s(x)[1-F(w)](1-u)dG(x)$, which represents the probability of finding a wage greater than w for workers at each wage in the range $[w_R, w]$, weighted by the probability of accepting a job at that wage. Therefore, in equilibrium, we find that

$$\delta_o G(w) + \lambda \int_{w_R}^{w} s(x)[1-F(w)]dG(x) = \frac{\lambda s(w_R)\,[F(w)-F(w_R)]\,u}{1-u}$$

$$\int_{w_R}^{w} s(x)dG(x) = \frac{\lambda s(w_R)\,[F(w)-F(w_R)]\,u - \delta_o(1-u)G(w)}{[1-F(w)](1-u)} \tag{8}$$

While less than elegant, Equation (8) implicitly solves for the distribution of accepted wages as a function of the initial wage offer distribution. We note that the left-hand side of the equation and the denominator of the right-hand side are strictly greater than zero, so it follows that the numerator must be strictly positive, or

$$G(w) < \frac{\lambda s(w_R)}{\delta_o} \cdot \frac{u}{1-u} \cdot [F(w)-F(w_R)]$$

Christensen et al. find that $G(w)$ stochastically dominates $F(w)$, so in order for the insider-outsider model to align with the literature, it must hold true that u and $\lambda s(w_R)$ are sufficiently large and δ_o sufficiently small.

Figure 3 shows the disjointed nature of the wage distribution under the insider-outsider model. This is because different workers have different connectedness to the labor market and therefore experience different flows between work and unemployment. The flows into insider employment from the unemployed is given by $\lambda s(w_R)[F(w) - F(w_T)]u$. However, outsiders can also transition to insider employment because of on-the-job search. They enter inside employment at or below wage w according to $\lambda \int_{w_R}^{w_T} s(x)[F(w) - F(w_T)](1-u)dG(x)$. Because insiders lose on-the-job search, the only flow out of insider employment is from job separation, which occurs at an exogenous rate of δ_i, so the flow of workers out of insider employment is $\delta_i G(w)i$. In steady state, the flows in and out of insider employment are the same, so

$$\lambda s(w_R)[F(w) - F(w_T)]u + \lambda \int_{w_R}^{w_T} s(x)[F(w) - F(w_T)](1-u)dG(x) = \delta_i G(w)i$$

$$G(w) = \frac{\lambda}{\delta_i i} \cdot [F(w) - F(w_T)] \cdot \left[s(w_R)u + (1-u) \int_{w_R}^{w_T} s(x)dG(x) \right] \quad (9)$$

where $G(w_R) = 0$, as workers are indifferent between accepting unemployment income and working income. The accepted wage distribution $G(w)$ is solved according to Equation (8) for outsider employment, where wage $w \in [w_R, w_T)$, and is solved according to Equation (9) for insider employment, where wage $w \in [w_T, w_{max}]$.

4 Estimation of Initial Parameters

This model takes the wage offer distribution, job finding and job destruction rates as inputs, and outputs the distribution of accepted wages. However, we only have readily available data on the accepted wage distribution, so we will infer the wage offer distribution by fitting the model's output, the calculated distribution of accepted wages, to the data. This is done using a generalized method of moments approach, which will be explained later in the section.

In order to solve the model outlined in Section 2, it is necessary to know the parameters for the utility functions, ζ, the coefficient of risk aversion and E, the cost of becoming an insider, for the search cost function, η, the coefficient of search cost and γ, the elasticity of cost with respect to search effort, and for the value function, the discount rate β, the arrival rate of job offers λ, the job destruction rate for outsiders δ_o, and for insiders δ_i and the distribution of job offers $F(w)$. However, not all of these parameters, specifically those relating to utility functions and the cost of search, are readily derivable from available data. Therefore, for β, ζ, η and γ, we rely on estimates found in Lise (2012). In order to estimate the remaining parameters, we rely on data on the accepted wage distribution in the US, the estimated separation rate, and the share of the labor force with a bachelor's degree as a proxy for insiders.

The wage distribution is taken from the Current Population Survey, collected by the U.S. Census Bureau and the Bureau of Labor Statistics. The distribution is derived from respondents' pre-tax total wage and salary income. We clean the data to remove outliers where income is reported as below \$0 or above \$1 million. There is no interpretation of negative incomes in the context of this model, so such values are ignored. In theory, there should be no accepted wage offers below the reservation wage, but in the data, accepted wages approach zero. These can be interpreted as part-time workers or contractors who work irregular jobs. It is also possible that there is an error in data collection and that respondents misreport their wages and salaries. However, in order to alleviate this potential discrepancy, we set the value of unemployment insurance b close to zero, to prevent a substantial disparity between the model and the data.

Similarly, above \$1 million, the income data ceases to be approximately continuous. That is, given my approximation of wages, there are gaps in between wages earned. While this would not be problematic given complete data on wages and salaries earned by U.S. workers, the CPS relies on a survey data, so the inclusion of a few high income earners in the sample has the potential to give differing estimates of mean and variance from one year to the next. Therefore, as a simplification, we do not use them when estimating the parameters of this model. This simplification modestly reduces the mean and variance of the distribution. When calculating skewness and kurtosis, these reductions are amplified, leaving estimates of the third and fourth moments rather lower. However, we contend that this model is not attempting to explain the existence of dramatic disparities in individual incomes in the top fraction of a percent of wage earners, but instead seeks to explain wage dispersion more generally throughout the whole wage distribution. When evaluating the estimated parameters though, it is important to note that they may understate the fatness of the right tail.

However, the distribution of accepted wages alone is insufficient to properly estimate the parameters of the model. We also introduce the separation rate as a parameter. There are numerous empirical studies that estimate flows in and out of unemployment, such as Hall (2006), Davis, Faberman and Haltiwanger (2006), Hobijn and Sahin (2009), and Shimer (2012). For the sake of this study, we will use Shimer's estimates. Shimer finds that the separation rate changes over time, so Shimer's estimates allow us to see what effects a changing separation rate has in explaining changing income inequality in the postwar period. The separation rate differs from the weighted average of job destruction rates δ_o and δ_i because of on-the-job search. Separation includes both job destruction and workers finding higher paying jobs. Therefore the separation rate d is calculated as follows:

$$d = \int_{w_R}^{w_T} \delta_o + \lambda s(w)\left[1 - F(w)\right] dG(w) + \int_{w_T}^{w_{max}} \delta_i dG(w)$$

By targeting the separation rate, we establish a relationship between δ_o, δ_i and λ. However, the separation rate and the accepted wage distribution together are insufficient to pin down three variables. We therefore target E, the cost of transitioning from an outsider to an insider. In practice, however, this cost is difficult to pin down – it accounts for both monetary costs associated with the transition and for non-monetary costs associated with the disutility of undertaking additional training. But it is still possible to estimate, because the cost of transitioning has a direct impact on the share of insiders and outsiders in the economy. A higher cost of transitioning means that workers need a higher wage for the transition to be worth it, so w_T is increasing in E. Given a higher tenure reservation wage w_T, it is less likely that workers receive a wage offer high enough to transition, so the share of the population that are insiders i is decreasing in w_T. By extension, it follows that i is decreasing in E, so either one can be used as an additional parameter to the same effect.

Up until now, we have left the concept of insiders rather vague, to avoid limiting the applications of this model. We now target the share of the population with a bachelor's degree as a proxy for insider employment. A college education is used as a proxy, because college graduates are more attached to the labor market, as shown in Figure 4. Workers with a bachelor's degree have lower unemployment rates, implying that they either find jobs more quickly when unemployed or transition to unemployment less frequently, or likely a combination of both. College graduates gain skills in college that make them more productive employees, and are therefore given a wage premium. The cost of replacing a skilled worker is higher than replacing an unskilled worker, so college graduates have more bargaining power in their wage contracts. Finally, if a college graduate does lose employment, he or she has the skill set and network of potential employers so that unemployment spells are relatively short-lived.

Admittedly, college graduates are an imperfect proxy for insiders. One of primary assumptions in this model is that all workers are identical and that wage dispersion is a product of search in the labor market, not of intrinsic differences between workers. This of course does not necessarily hold for tertiary education, as there is a correlation between ability and the decision to pursue higher education. This is potentially problematic because we are matching our model to the data without accounting for ability. The estimated parameters assume identical workers, while in the data, part of the dispersion is likely due to heterogeneity of workers.

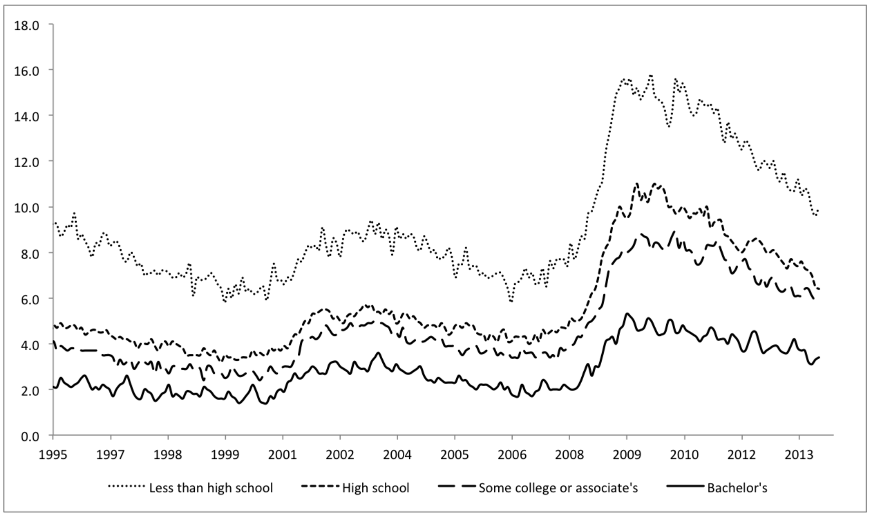


Figure 4: The plot above shows the unemployment rate from 1995 through the present, decomposed by level of education. Higher levels of education are associated with lower unemployment rates.

In order estimate the parameters for this model, we use generalized method of moments estimation outlined in Hall (2005). This method is useful in this model because it does not impose a form on the distribution of accepted wages, as a maximum likelihood estimator might. In this way, we are able to let the data speak. In this model, we are trying to find the wage offer distribution $F(w)$ and parameters λ, δ_o and δ_i such that the accepted wage distribution, separation rate and insider share match the data. We minimize the sum of squared differences between the observed data and the models estimates:

$$\min_{\theta} L(\theta) = \sum_{k \in K} a_k (\hat{\mu}_k - \mu_k)^2$$

where k is an individual moment to be matched, from the set K of all moments, μ is a vector of the first four moments – mean, variance, skewness and kurtosis – of the observed wage distribution, separation rate and share of insiders. The vector $\hat{\mu}$ is the corresponding values estimated by the model, given an input vector of θ. The vector a is a weighting on each parameter. Values for the moments in μ are shown in Table 1. The vector of parameters, $\hat{\theta}$, is therefore the solution to the following, written in matrix notation:

$$\hat{\theta} = \arg\min_{\theta \in \Theta} m(\theta)^T A \, m(\theta) \tag{10}$$

where $m(\theta)$ is a $k \times 1$ vector of the difference between θ_k and μ_k, as above. A is a $k \times k$ weighting matrix.

For a linear model, the first-order conditions of Equation (10) would yield a closed-form solution to $\hat{\theta}$ as a function of the parameters from the data. However, for a non-linear mapping such as this, it is necessary to solve for the optimal parameter vector $\hat{\theta}$ numerically. In order to find the vector $\hat{\theta}$, we use the unconstrained nonlinear optimization method `fminsearch` in MATLAB. This method uses the Nelder-Mead simplex algorithm, based on Lagarias et al. (1998). This minimization technique takes in an initial argument θ_0 and iteratively searches $\theta \in \Theta$ until convergence to $\hat{\theta}$ occurs. As with any search algorithm, there is the risk that optimization finds a local minimum rather than the global minimum that is the solution. While it is possible that the solved solutions, presented below, are local minima, the distribution $G(w)$, outlined in Equations (8) and (9), and the separation rate d and insider share i give little basis for believing the function L is multimodal, so we assume that the solution $L(\hat{\theta})$ is the global minimum.

5 Results and Policy Implications

Table 1 lists the moments from the data, the corresponding values of the moments in the data, and Table 2 lists the estimated parameters that describe the model. Figure 5, found in the Appendix, compares the estimated accepted wage distribution with a histogram of CPS wage data.

The introduction of a segregated labor market, where certain workers are more preferable, has the effect of increasing skewness and kurtosis relative to the model of on-the-job search without the insider-outsider choice, as shown

Moment	μ	$\hat{\mu}$
$E(W)$	157.7	157.5
$E(W^2)$	17579.1	21401.7
$E(W^3)$	2.01	2.22
$E(W^4)$	9.19	8.91
Separation rate	0.17	0.20
Labor force with Bachelor's	0.24	0.071

Table 1: The moments used to estimate the parameters and the corresponding values that minimize the loss function in the model. The values denoted by μ are from the data, while the values labeled $\hat{\mu}$ are calculated from the model. The first four moments in the data are calculated from the distribution of wages, W, in 2015. The separation rate is taken from Shimer (2012). The share of the labor force with a college education is available via the U.S. Bureau of Labor Statistics.

	Point estimate
υ_F	4.211
σ	0.938
λ	0.680
δ_o	0.146
δ_i	0.089
E	78.7628
β	0.96
ζ	1.455
η	1.0
γ	1.168

Table 2: The estimated parameters that most closely match the data. The parameters υ_F and σ, mean and standard deviation, respectively, define the distribution of the log-normal job offer distribution. λ is the job finding rate, δ_o and δ_i are the respective job destruction rates for outsiders and insiders, and E is the cost of becoming an insider. The estimates for β, ζ, η and γ are taken from Lise (2012).

	Percentile				Ratios			
	0.10	0.50	0.90	0.99	50-10	90-50	90-10	99-10
Data	35	126	316	660	3.60	2.51	9.03	18.86
Model:								
without tenure	38	121	319	676	3.18	2.64	8.39	17.78
with tenure	39	111	339	737	2.85	3.05	8.69	18.90

Table 3: On the left, the table shows the wage (normalized on a scale from 1 to 1000) that corresponds to the given percentile of the wage distribution. The right then compares the ratio of the specified high wage worker to the specified low wage worker.

in Figure 3. This model places greater mass lower in the income distribution, and increases the right tail. The addition of insiders and outsiders is useful to study income inequality, as Krueger and Perri (2005) find a divergence between the tails of the accepted wage distribution from 1980 through the early 2000s. Indeed, plotting the calculated distribution with a histogram of CPS wage data shows that the calibrated overestimates income inequality, placing too much density at the lower end of the wage distribution and then again too much in the tail at the upper end of the wage range. We can quantify this by examining the ratio of wages earned at the top and bottom of the wage distribution, shown in Table 3. Corroborating Figure 5, the data show that the model places too much density at the lower end of the wage distribution. Looking first at the 50-10 and 90-50 percentile ratios for our model, we note that the estimates when tenure is not included are closer to the data. However, we also note that when tenure is included, as expected, there is more weight placed at the upper end of the income distribution. Unlike at the lower end of the wage distribution, the 90-10 and 99-10 percentile ratios when tenure is included are closer to the data than regular on-the-job search model. This model is useful for examining fat tails at the upper end of the income distribution, and therefore has the potential to be useful when studying income inequality.

The estimated parameters for the job finding rate, λ, and the job destruction rates, δ_o and δ_i, are roughly in line with the literature. Christensen et al. find a job finding rate of 0.5833 and job destruction rate of 0.2873 for a standard on-the-job search model, while Lise (2012) finds a job finding rate of 0.615, a job destruction rate of 0.221 for those with a high school education and a job destruction rate of 0.0989 for those with a college education, when savings are introduced into the model. In general, the job finding rates are higher than in other models, while the job destruction rates are generally lower.

This is partially explained by the relationship between the separation rate and the share of workers with insider employment. The higher the share of insiders in the labor market, the lower the overall separation rate, as insiders have both a lower job destruction rate and will not change employment from on-the-job search. Therefore, in order to offset the lower separation rate, it must follow that outsiders have a comparatively higher separation rate, because the overall separation rate is the weighted average of insiders' and outsiders' respective separation rates. This can come in the form of either a higher job finding rate λ or a higher job destruction rate δ_o, or both. In the estimates in Table 2, we note that only λ is elevated. This is because of the relationship between the job destruction differential and the fatness of the right tail. As the differential between the two job destruction rates increases, the density in the upper end of the distribution increases. To intuitively understand this relationship, let us take the extreme case, where insiders receive complete employment security, and $\delta_i = 0$. Outsider workers will continue to churn between outside employment and unemployment until they receive insider employment. At this point, they are permanently insiders, so in steady state, all workers are insiders. This means the that steady state distribution is static, with no flow between employment and unemployment, and the density is focused solely in the upper tail. The extreme case is illustrative, showing that a large difference between job destruction rates increases upper tail, which further depresses the job separation rate. Therefore, the estimate for δ_o is relatively lower, as a balance between the insider share of workers and the separation rate.

Given the large wage dispersion that this model predicts, it is interesting to investigate how well it predicts income inequality, specifically if it predicts an increase in the postwar period. We can check this by varying the parameters to which we are matching the model, and noting what changes in dispersion, if any, such changes predict. Because matching the moments of the realized wage distribution over time, say from the 1980s, would tautologically predict an increase in wage dispersion over that period, that leaves the other two moments, the separation rate and share of insiders in the economy. For similar reasons, varying the share of insiders in the labor market is not very interesting, because a higher share of insiders will of course increase the share of workers at the upper end of the wage distribution. We therefore focus on the separation rate, which is interesting because of the interplay it has with $\lambda, \delta_o, \delta_i$ and i, the share of insiders, as outlined above.

Shimer (2012) shows that the separation falls dramatically over time, from 0.24 for workers when averaged over 1948-2010, falling to 0.17 when narrowed to 1967-2010, and 0.10 for 1987-2010. The fall is even more dramatic if the 2008 financial crisis is excluded, falling to 0.05 for 1987-2007. We vary the separation rate to which the model is optimized, and note changes in the distribution. In order to emphasize the effect that the separation rate has on the distribution, we alter the weighting matrix A to place additional weight (100x) on the separation rate. The results are shown in Tables 4 and 5 and Figure 6, in the Appendix.

Interestingly, it does not appear that there is a monotonic relationship between separation rate and mean of the wage offer distribution υ_F, job finding rate λ or outsider job destruction rate δ_o. The pessimistic interpretation is that this is evidence of multiple local minima in the search function, and that the calculated values are not the most optimal values. The optimistic interpretation is that there is indeed a non-monotonic relationship between separation rate and the listed parameters. However, it is still possible to glean some interesting relationships from this study. The first, regarding δ_i, shows that insiders become more attached to the labor market over this period, which is indicative of greater benefits to reaching insider employments. Furthermore, the cost of training to become an insider, E, also decreased with separation rate. The cost can be interpreted as any number of costs, from the monetary cost to disutility of training. Overall, a decline in E indicates that training, or some sort of further education, became more accessible over this period. This general trend is consistent with data from the U.S. Bureau of Labor Statistics, which show that the share of the labor force over 25 years old with a college education increased from 0.20 in 2000 to 0.24 in 2016.

6 Concluding Remarks and Future Work

In this paper, we construct a search theoretic model that allows workers to opt for increased attachment to the labor market in order to explain wage dispersion. We solve for the invariant wage distribution analytically. We also use a Markov chain to numerically solve a discretized version of the model by solving for the transition probability between unemployment and the different wage states. Finally, we apply this model to U.S. wage data, using generalized method of moments to estimate the parameters. Compared to the data, our model overestimates wage dispersion, placing excess density at the lower end of the distribution. However, the addition of density in the upper tail relative to the traditional on-the-job search model matches the fat tails found in the data as shown by the ratio of wages at the upper end and lower end of the wage distribution, particularly 90-10 and 99-10.

In constructing this model, we make some strong assumptions about the homogeneity of workers and the reasons that an insider-outsider division exists. Specifically, we assume that all workers face the same wage offer distribution, and the choice to between insider and outsider is made after the wage offer. In this way, we explain wage dispersion as essentially based on the distribution of wage offers, as insider status only occurs when workers arbitrarily receive a sufficiently high wage. When insiders lose their job, they then lose the skills they previously gained, and enter the general unemployment pool. This of course ignores the role of ability on wage dispersion. Workers with higher ability are more likely to go to college, so higher education is a signal of ability, rather than an arbitrary occurrence.

There are a couple of adjustments to this model that could partially alleviate some of the problems found above. One is to adjust the wage offer distribution for the insiders. As is, when insiders lose employment, they face the same wage offer distribution as unemployed outsiders. An adjustment could be made so that insiders face a higher wage offer distribution, and are therefore more likely to become an insider again. This could partially account for the idea that workers maintain their skill set over time. If the wage offer distribution increased after each bout of insider employment, this could account for further wage dispersion, and potentially explain growth in wages over time. Another potential adjustment to the model is to have heterogeneous agents, skilled and unskilled, from the beginning, where skilled workers face some combination of a lower job destruction rate, higher job finding rate and higher wage offer distribution. This model has the potential to more realistically explain wage dispersion because it takes natural ability into account. In practice, however, this model is less interesting because workers are unable to move between the skilled and unskilled states. The model would in essence be an aggregation of two on-the-job search models.

There are also some changes to the data calibration that could make the findings of this paper more economically interesting. As mentioned above, education is an imperfect calibration target for the share of insider workers, because it ignores ability. However, the European labor market, with its extensive job protection measures, is much more appropriate for the insider-outsider model, as shown in Lindbeck and Snower (2002) and Blanchard and Summers (1986). The labor market could then be segregated by industry, where workers in industries with high job protection are classified as insiders and the others are outsiders. While this method relies on extensive data collection, it allows for the comparison of workers who are more roughly equal in ability. This of course assumes that the industry of employment is independent of ability.

Another aspect of the model that does not match the data is the discontinuity in the distribution of accepted wages $G(w)$, which implies a point-mass at $G(w_T)$. This point mass is due to the lower job destruction rate as an insider. The probability of receiving a job offer at $w_T - \epsilon$ and w_T are the same, for $\epsilon > 0$, but the job destruction rate differs, resulting in a point mass at w_T. One possible solution for the discontinuity in the wage distribution is the introduction of heterogeneous classes of workers to introduce a spectrum of insiders. These different classes of workers would receive job offers according to a different distribution, have different job finding or job destruction rates. The wage at which the discontinuity occurs would differ by class of workers, so the aggregation as the number of classes $N \to \infty$ would result in no point-mass in the distribution of accepted offers.

Focusing on the estimation techniques, there are a couple of methods that could potentially improve the accuracy of the parameters. The first is regarding the weighting matrix A in Equation (10). In this paper, we use an identity weighting matrix. This is potentially problematic because of the different scales on which the different moments are measured, varying from 10^{-1} to 10^4. This of course means that larger factors, such as mean and variance, are implicitly weighted more heavily. Two-step (or n-step) estimation of the weighting matrix would allow a more efficient estimator. This is done by taking a positive definite matrix A_0, such as the identity matrix, and solving for the optimal parameters $\hat{\theta}_1$. It is then possible to solve for A_1 in terms of $\hat{\theta}_1$. Then, using this estimated value of A_1, it is possible to solve for $\hat{\theta}_1$. This can then continue until convergence, for the n-step estimator, which allows for a consistent and efficient estimator of the parameters, $\hat{\theta}_n$.

Finally, it is possible that the search algorithm used finds local minima rather than the global minimum. This is potentially problematic, because it means that the estimated parameters are dependent on the initial parameters used θ_0. One way to get around this is to essentially repeat the Nelder-Mead simplex algorithm k times, varying the initial values across the ranges of the parameters. This is a somewhat brute force method of ensuring that the global minimum is found. In this paper, we estimated six parameters, so repeating the search algorithm k times requires more time or computing power than feasible for this project.

7 Appendix

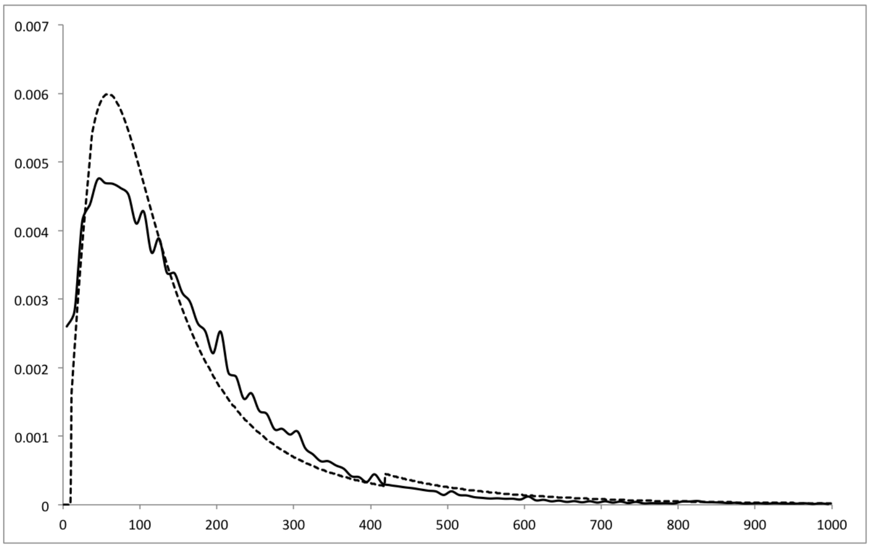


Figure 5: The plot above compares the model distribution from the estimated parameters with the histogram of CPS wage data.

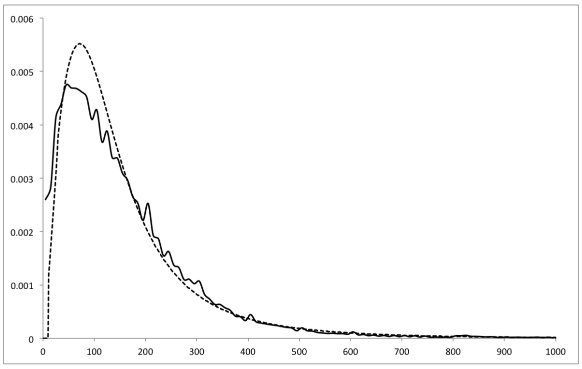
Separation rate:	0.24	0.17	0.10
υ_F	4.180	4.175	4.189
σ	0.948	0.946	0.944
λ	0.714	0.739	0.692
δ_o	0.128	0.199	0.139
δ_i	0.091	0.087	0.086
E	80.776	78.495	75.728

Table 4: The estimated parameters that most closely match the data, with varying targets for separation rate. The parameters υ_F and σ, mean and standard deviation, respectively, define the distribution of the log-normal job offer distribution. λ is the job finding rate, δ_o and δ_i are the respective job destruction rates for outsiders and insiders, and E is the cost of becoming an insider.

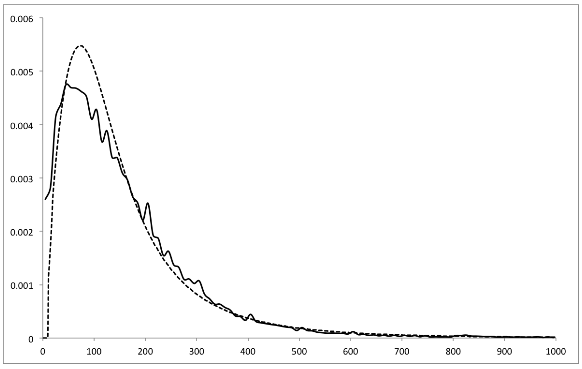
	Data			
Target separation rate:		0.24	0.17	0.10
$E(W)$	157.7	157.6	157.7	157.7
$E(W^2)$	17579.1	17579.1	17579.1	17579.1
$E(W^3)$	2.01	2.22	2.21	2.21
$E(W^4)$	9.19	9.72	9.68	9.68
Separation rate	–	0.142	0.137	0.134
Labor force w/ B.A.	0.24	5.22E-06	3.22E-06	4.56E-06

Table 5: The moments used to estimate the parameters and the corresponding values that minimize the loss function in the model. The first four moments in the data are calculated from the distribution of wages, W, in 2015. The separation rate is taken from Shimer (2012). The share of the labor force with a college education is available via the U.S. Bureau of Labor Statistics.

(a) Calibrated to 0.24 separation rate



(b) Calibrated to 0.17 separation rate



(c) Calibrated to 0.10 separation rate

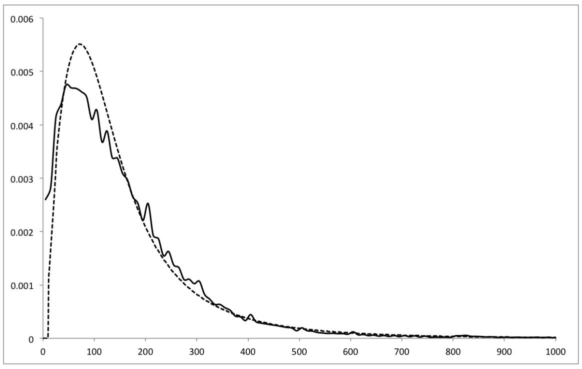


Figure 6: The plots above show the estimated parameters when varying the target separation rate, compared with the histogram of CPS wage data. We change the weighting in the weighting matrix A to emphasize the effects of the change, though, as is evident above, the differences between the different targets are minimal. Tables 4 and 5 show the differences in estimated parameters and moments.

Figure 7: Matrix P shows the transition probability between states of unemployment and employment at wages $[w_R, w_{max}]$. U denotes the state of unemployment and w_* denotes the state at a given wage. The matrix is $n \times n$, so the rows have labels corresponding to the columns. The row is the current location and the column is the probability of moving to the corresponding state.

$$P =$$

$$\begin{array}{ccccccc} U & w_R & \dots & w_{T-1} & w_T & \dots & w_{max} \end{array}$$

$$\begin{bmatrix}
1-\lambda s(w_R)\left[1-F(w_R)\right] & \lambda s(w_R) f(w_{R+1}) & \dots & \dots & \dots & \dots & \lambda s(w_R) f(w_{max}) \\
\delta_o & 1-\delta_o-\lambda s(w_{R+1})\left[1-F(w_{R+1})\right] & \dots & \dots & \dots & \dots & \lambda s(w_{R+1}) f(w_{max}) \\
\vdots & \vdots & \ddots & \dots & \dots & \dots & \vdots \\
\delta_o & 0 & \dots & 1-\delta_o-\lambda s(w_{T-1})\left[1-F(w_{T-1})\right] & \dots & \dots & \lambda s(w_{T-1}) f(w_{max}) \\
\delta_i & 0 & \dots & 0 & 1-\delta_i & 0 & 0 \\
\vdots & \vdots & \dots & \dots & 0 & \ddots & 0 \\
\delta_i & 0 & 0 & \dots & 0 & \dots & 1-\delta_i
\end{bmatrix}$$

References

[1] Abowd, John, Francis Kramarz and David Margolis. 2003. *High Wage Workers and High Wage Firms.* Econometrica.

[2] Autor, David, Lawrence Katz and Melissa Kearney. 2008. *Trends in U.S. Wage Inequality: Revising the Revisionists.* Review of Economics and Statistics.

[3] Ball, Laurence. 2009. *Hysteresis in Unemployment: Old and New Evidence.* NBER Working Paper Series.

[4] Blanchard, Olivier and Lawrence Summers. 1986. *Hysteresis and the European Unemployment Problem.* NBER Macroeconomics.

[5] Burdett, Kenneth. 1978. *A Theory of Employee Job Search and Quit Rates.* The American Economic Review.

[6] Christensen, Bent Jesper, Rasmus Lentz, Dale Mortensen, George Neumann and Axel Werwatz. 2005. *On-the-Job Search and the Wage Distribution.* Journal of Labor Economics.

[7] Davis, Steven, R. Jason Faberman and John Haltiwanger. 2006. *The Flow Approach to Labor Markets: New Data Sources and Micro–Macro Links.* Journal of Economic Perspectives.

[8] Flood, Sarah, Miriam King, Steven Ruggles and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

[9] Hall, Alistair. 2005. *Generalized Method of Moments.* Oxford: Oxford University Press.

[10] Hall, Robert. 2006. *Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years.* NBER Macroeconomics.

[11] Heathcote, Jonathan, Fabrizio Perri and Giovanni Violante. 2010. *Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006.* Review of Economic Dynamics.

[12] Hobijn, Bart and Aysegul Sahin. 2009. *Job-finding and separation rates in the OECD.* Economics Letters.

[13] Krueger, Alan, Judd Cramer and David Cho. 2014. *Are the Long-Term Unemployed on the Margins of the Labor Market?.* Brookings Papers on Economic Activity.

[14] Krueger, Alan and Lawrence Summers. 1988. *Efficiency Wages and the Inter-Industry Wage Structure.* Econometrica.

[15] Krueger, Dirk and Fabrizio Perri. 2005. *Does Income Inequality Lead to Consumption Equality? Evidence and Theory.* Federal Reserve Bank of Minneapolis.

[16] Lagarias, Jeffrey, James Reeds, Margaret Wright and Paul Wright. 1998. *Convergence Propoerties of the Nelder-Mead Simplex Method in Low Dimensions.* Society for Industrial and Applied Mathematics.

[17] Lee, David. 1999. *Wage Inequality in the United States during the 1980s: Rising Dispersion or Falling Minimum Wage?.* Quarterly Journal of Economics.

[18] Lindbeck, Assar and Dennis Snower. 2002. *The Insider-Outsider Theory: A Survey.* IZA Discussion Paper No. 534.

[19] Lise, Jeremy. 2013. *On-the-Job Search and Precautionary Savings.* Review of Economic Studies.

[20] Ljungqvist, Lars and Thomas Sargent. 2004. *Recursive Macroeconomic Theory.* Cambridge, Massachusetts: MIT Press.

[21] Mortensen, Dale. 2003. *Wage Dispersion: Why Are Similar Workers Paid Differently, Zeuthen lecture book series.* Cambridge, Massachusetts: MIT Press.

[22] Piketty, Thomas and Emmanuel Saez. 2001. *Income Inequality in the United States, 1913-1998.* Quarterly Journal of Economics.

[23] Rogerson, Richard, Robert Shimer and Randall Wright. *Search-Theoretic Models of the Labor Market: A Survey.* Journal of Economic Literature.

[24] Shimer, Robert. 2012. *Reassessing the ins and outs of unemployment* . Review of Economic Dynamics.

[25] Stokey, Nancy and Robert Lucas. 1989. *Recursive Methods in Economic Dynamics.* Cambridge, Massachusetts: Harvard University Press.

[26] Taschereau-Dumouchel, Mathieu. 2015. *The Union Threat.* Social Science Research Network (accessed February 15, 2016).

This paper represents my own work in accordance with University regulations.