Bonn Econon Discussion Papers

Discussion Paper 04/2011

Exploring the Causes of Frictional Wage Dispersion

by

Volker Tjaden and Felix Wellschmied

May 2011

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Kaiserstrasse 1
D-53113 Bonn

Financial support by the
Deutsche Forschungsgemeinschaft (DFG)
through the
Bonn Graduate School of Economics (BGSE)
is gratefully acknowledged.

Deutsche Post World Net is a sponsor of the BGSE.

Exploring the Causes of Frictional Wage Dispersion*

Volker Tjaden
Bonn Graduate School
of Economics

Felix Wellschmied
Bonn Graduate School
of Economics

May 26, 2011

Abstract

Standard search models are inconsistent with the amount of frictional wage dispersion found in U.S. data. We resolve this apparent puzzle by modeling skill development (learning by doing on the job, skill loss during unemployment) and duration dependence in unemployment benefits in a random on the job search model featuring two-sided heterogeneity. The model's key parameters are calibrated using micro data on employment mobility and wages from the *Survey of Income and Program Participation* (SIPP). Our model is consistent with the amount of frictional wage dispersion found in the data. Skill development on the job is the most important driver behind this result. Meanwhile, firm heterogeneity never accounts for more than 20% of overall wage inequality within an age cohort.

Key Words: Frictional wage dispersion, Search model, Heterogeneity
JEL Classification: J24, J31, J64

*We are grateful for comments from Christian Bayer, Jörg Breitung, Thomas Hintermaier, Alexander Kriwoluzky, Moritz Kuhn, Monika Merz, and Konstantinos Tatsiramos

1 Introduction

Large residual wage variation among observationally identical workers is a pervasive phenomenon in empirical studies of wage determination. In standard search models of the labor market, we think of workers sampling job offers from heterogeneous firms and a search friction prohibiting instantaneous matching. Refusing an offer entails opportunity costs in the form of foregone wage earnings and the risk to be receiving no offer next period. Hence, we would attribute wage dispersion not explained by worker characteristics to these search frictions.

Meanwhile, as was recently pointed out by Hornstein et al. (2007) (HKV, henceforth), empirically observed frictional wage dispersion is far too large to be consistent with standard specifications of these models. Their estimate on U.S data would imply optimal mean unemployment durations of twenty to thirty times the three months we see in the data. In our view, this failing hints towards important other aspects of the worker's decision problem which have been neglected so far.

A number of well documented empirical facts stand out as candidates for resolving this apparent puzzle. First, employment carries additional benefits such as experience gains and unemployment additional costs such as skill losses which are not captured by a standard search model. Second, if agents can efficiently search for better job prospects while already in employment, accepting a job carries much less finality and this should make them more willing to enter into relatively poor matches. Finally, there might be no puzzle after all and frictional wage dispersion would just be an artefact of a misspecified reduced form estimation. For instance, if worker characteristics have a strong stochastic and time varying component which is unobservable to the econometrician this will cause a bias. Imperfect sorting of worker types across firms can have similar effects.

The main contribution of our paper is to quantify the relative importance of the above mentioned channels by calibrating a structural model to individual level data on employment and wages. This approach allows us to account for endogenous worker responses to obtain structural parameters. The size of frictional wage dispersion in our model is compatible with the corresponding estimate from HKV. Finally, we can put frictional wage dispersion into the bigger picture of overall wage dispersion. This allows us to address the long standing question of firm versus worker effects in accounting for wage inequality. In our appendix, we also demonstrate that the HKV estimator is successful in identifying the magnitude of frictional wage dispersion even though our model specification would imply it to be biased. This rules out the possibility that frictional wage dispersion is just a misidentification issue.

Summarizing our main findings, it turns out that the agents' forward looking behavior regarding their own productivity development on the job is the single most important factor in explaining frictional wage dispersion. We also argue for the importance of realistically modeling the efficiency of on the job search when trying to asses its contribution to wage dispersion. As we demonstrate, an empirically pervasive phenomenon are job to job movements resulting in nominal wage losses, which has large effects on the implied search efficiencies. The contribution of skill losses in unemployment and limited duration of unemployment benefits is much smaller. Nonetheless, it is only the combination of all factors that can explain empirically observed residual wage dispersion.

Lastly, we find the contribution of frictional wage dispersion explained by firm heterogeneity to be modest. While it causes substantial wage differences between similar workers, it nonetheless never accounts for more than 20% of wage dispersion within an age cohort and this share decreases as workers age. Instead, most of wage inequality is attributable to initial worker characteristics. This finding also holds important implications for policy makers interested in reducing overall wage inequality. A large contribution of firm productivity dispersion would have suggested that increasing search efficiencies might go a long way in compressing the wage distribution. Instead, our results hint towards improvements in general education and skill upgrades for older workers already in the workforce as more promising routes to pursue.

Most closely related to our paper are two recent contributions by Burdett et al. (2009) and Carrillo-Tudela (2010). The former generalize the Burdett and Mortensen (1998) model[1] by including labor market experience. Their goal is to show that the

[1] Mortensen (2003) shows that the basic Burdett Mortensen (1998) model with endogenous search

model creates reasonable wage dispersion in the sense of HKV and an equilibrium wage distribution with a fat right tail. Carrillo-Tudela (2010) extends this model by introducing firm heterogeneity and heterogeneity in job offer arrival rates for the employed and the unemployed. Also related is a strand of literature which tries to rationalize overall wage heterogeneity by on the job search models. Bontemps et al. (1999) and Bontemps et al. (2000) set up on the job search models and structurally estimate them on French panel data. Postel-Vinay and Robin (2002) introduce worker heterogeneity and use French linked employer-employee data for estimation. All of the papers mentioned in this paragraph have in common that they attribute search on the job a prominent role in explaining wage heterogeneity. In the light of our empirical findings on the efficiency of on the job search, we think they might overstate the importance of that channel.

The remainder of the paper is structured as follows: We present our model in Section 2. Section 3 discusses our own empirical work and parametrization. Section 4 presents and analyzes our results. Section 5 concludes. Additional information on the empirical part and the numerical algorithm is relegated to an appendix.

2 The Model

2.1 The Labor Market

A firm is a match producing with the worker's idiosyncratic log productivity A_t and firm specific log productivity Γ_t[2]. We assume that search is random and the labor market is guided by matching function $m = \xi u^{\iota} v^{1-\iota}$ where v are vacancies and u are the unemployed. As usually, an unemployed worker contact rate $q(\theta)$ and a job offer probability $p(\theta)$ can be derived from that matching function. Let $\chi(A_t, \Gamma_t, \phi_t)$ and $\psi(A_t, \phi_t, \varpi)$ be measures of employed and unemployed agents over idiosyncratic productivity, firm specific productivity, the life-cycle state (ϕ) and an

effort implies implausible large monopsony power in wage posting when estimated on Danish data. He advocates a bargaining wage approach instead, which is also our choice.

[2]Our model does not distinguish between firm specific productivity and match specific productivity. We use the term firm productivity, but actually mean the sum of the two.

indicator for unemployment benefit entitlement ϖ in period t. Firm productivity is drawn from the distribution $F \sim N(0, \sigma_F^2)$. Once a match is formed it produces output y_t according to

$$y_t = exp(A_t + \Gamma_t)$$

2.2 The Household Problem

Household period income is given by:

$$I_t(A_t, \Gamma_t, \phi_t) = \begin{cases} w_t(A_t, \Gamma_t, \phi_t) & \text{if } \textit{employed} \\ b + Z & \text{if } \varpi = u_1 \\ Z & \text{if } \varpi = u_2 \end{cases}$$

b is the UI payment and Z is the leisure value of unemployment. If the agent is in state u_1 he receives UI, but with probability λ_l he looses the benefit entitlement and moves to state u_2. After match destruction, an agent is always entitled to benefits[3].

In modeling productivity development we are guided by the finding of Dustmann and Meghir (2005), who show that the first two years of labor market experience raise wages substantially (6-10% per year), while the return to experience is close to zero afterwards (0-1.2%)[4]. We therefore introduce a very stylized life cycle dimension where agents transit through two life-cycle states (ϕ) with stochastic transition probabilities $p = (p_1, p_2)$. When the second shock hits, the agent dies and is reborn as an unemployed labor market entrant in state u_2 and with idiosyncratic productivity drawn from the distribution $N \sim N(\mu_N, \sigma_N^2)$.

The evolution of worker productivity depends on the agent's employment status

[3]Low et al. (2010) and Ljungqvist and Sargent (2008) assume that entitlement is conditional on the separation being a forced one. Our interpretation of productivity shocks is a different one from theirs making this distinction not feasible.

[4]Dustmann and Meghir (2005) use German data, but have the advantage of identifying effects by using displaced workers. For US data, Altonji and Williams (1998) come to similar results.

and in case the agent is employed it also depends on his life-cycle state:

$$A_{t+1} = \begin{cases} max(A_t + \nu(\phi) + \epsilon_t, pmin) & \text{if } employed \\ max((1-\delta)A_t + \epsilon_t, pmin) & \text{if } unemployed \end{cases}$$

δ represents skill depreciation while being unemployed, $pmin$ is a subsistence level of productivity and $\nu(\phi)$ is a drift term that depends on the life-cycle state. ϵ is a productivity shock with $\epsilon \sim N(0, \sigma_\epsilon^2)$. We think of wage shocks as anything altering productivity such as demand shocks for specific skills or health shocks. The fact that net productivity growth can be negative means that our model also features wage cuts on the job[5].

Let ω be the exogenous match separation rate. Match shocks leave worker productivity unaffected but cause match dissolution. Examples can be demand shocks or financing shocks to the firm. Matches may also be dissolved endogenously as result of a negative productivity innovation.

Our model allows employed workers to search for better job prospects, while being employed. The ability to search on the job is one of the important potential channels for generating frictional wage dispersion. The more efficient on the job search, the less final is an accepted position and the higher the incentive to accept employment at low productivity firms. We follow Jolivet et al. (2006) in modeling some job to job transitions as forced movements. One can think of such transitions occurring due to family reasons, or being the result of mismeasurement in the data due to time aggregation. An employed worker receives a job offer with probability λ and can in general decide to stay with his old match, or form a new match. However, when receiving an outside offer, with probability λ_d the offer is a forced movement and the outside option, instead staying with the old match, is unemployment. In our empirical section we show how we can infer the structural parameters λ and λ_d from micro data on job transitions and wages.

At this point we need to make an assumption on how wage bargaining takes place.

[5] Postel-Vinay and Robin (2002) argue show that this is an important empirical feature of wage development.

We assume firms cannot commit to a wage path and each period wages are renegotiated by Nash-Bargaining. We assume a worker always quits into unemployment before making a job to job transition, hence his outside option being unemployment with benefit entitlement when bargaining with the new firm[6]. Postel-Vinay and Robin (2002) show that when firms can commit to a wage path, transiting to a higher productive match implies an option value, which can lead workers to accept an initial wage cut. We rule out such behavior, because it appears not to be borne out by the data. Recently, Flinn and Mabli (2008) show that standard Nash-Bargaining fits key data moments better in an estimated DSGE model. In Appendix B.2.2 we deliver some reduced form evidence that supports this point. We show that future wage growth is uncorrelated to the initial wage cut accepted by workers, a statistic clearly at odds with the idea of initial wage cuts being accepted because of an option value[7].

The timing within one model period is as follows:

- At the beginning of the period, the employed workers negotiate a wage with their firm and production takes place.
- End of period transitions occur. First, some unemployed transit from u_1 to u_2.
- The employed and unemployed experience productivity transitions according to their laws of motion.
- Life cycle transitions take place. Agents die and are reborn.
- Exogenous job destruction occurs. Agents becoming unemployed cannot search for employment within this period.
- On the job offers realize.
- Employed agents decide whether to quit and the unemployed with job offers decide whether to accept the job.

[6]The same outside option would result when assuming the bargaining game from Moscarini (2005) where the firms enter into an auction for the worker.

[7]Moreover, Mortensen (2003) also criticizes the mechanism outlined by Postel-Vinay and Robin (2002), arguing that it seems infeasible in many circumstances.

We can thus define the value of employment for each life-cycle state (V_ϕ^E) and the value of unemployment ($V_{\phi,\varpi}^U$) depending on worker's idiosyncratic productivity, firm productivity, and benefit entitlement. We state the Bellman equations describing the problems of agents in the first life-cycle state as an example. The value of employment is the fixed point to:

$$\begin{aligned}V_1^E(A_t,\Gamma_t) = w_t(A_t,\Gamma_t,1) + \beta E_t\{(1-\omega)&\\ [(1-p_1)[(1-\lambda)H(1)+\lambda[(1-\lambda_d)\Omega_E(1)+\lambda_d\Lambda(1)]]&\\ +p_1[(1-\lambda)H(2)+\lambda[(1-\lambda_d)\Omega_E(2)+\lambda_d\Lambda(2)]]]&\\ +\omega[(1-p_1)V_{1,u_1}^U(A_{t+1})+p_1V_{2,u_1}^U(A_{t+1})]\}&\end{aligned}$$

Furthermore, there are two value functions for the unemployed with and without benefit entitlement. Conditional on receiving benefits, the value of unemployment solves:

$$\begin{aligned}V_{1,u_1}^U(A_t) = b + Z + \beta E_t\{(1-\lambda_l)&\\ [(1-p_1)[p(\theta)\Omega_U(1,u_1)+(1-p(\theta))V_{1,u_1}^U(A_{t+1})]&\\ +p_1[p(\theta)\Omega_U(2,u_1)+(1-p(\theta))V_{2,u_1}^U(A_{t+1})]]&\\ +\lambda_l[(1-p_1)[p(\theta)\Omega_U(1,u_2)+(1-p(\theta))V_{1,u_2}^U(A_{t+1})]&\\ +p_1[p(\theta)\Omega_U(2,u_2)+(1-p(\theta))V_{2,u_2}^U(A_{t+1})]]\}&\end{aligned}$$

Once benefits expire, the agents flow value is reduced to the utility of leisure / home production:

$$\begin{aligned}V_{1,u_2}^U(A_t) = Z + \beta E_t\{(1-p_1)[p(\theta)\Omega_U(1,u_2)+(1-p(\theta))V_{1,u_2}^U(A_{t+1})]&\\ +p_1[p(\theta)\Omega_U(2,u_2)+(1-p(\theta))V_{2,u_2}^U(A_{t+1})]\}&\end{aligned}$$

E_t is the expectation operator given all information in period t. For clarity of presentation, we have defined the following auxiliary variables: $\Omega_E(x)$ and $\Omega_U(x,\varpi)$ are the expected values of receiving a job offer for the employed and unemployed conditional on life-cycle state and benefit entitlement. $H(x)$ represents the decision whether to quit the job voluntarily and $\Lambda(x)$ represents optimal behavior after a forced job movement.

$$\begin{aligned}
\Omega_E(x) &= \max \int \{V_x^E(A_{t+1}, \Gamma_{t+1}), V_{x,u_1}^U(A_{t+1}), V_x^E(A_{t+1}, \Gamma_{t+1})\} dF \\
\Omega_U(x, \varpi) &= \max \int \{V_x^E(A_{t+1}, \Gamma_{t+1}), V_{x,\varpi}^U(A_{t+1})\} dF \\
H(x) &= \max\{V_x^E(A_{t+1}, \Gamma_{t+1}), V_{x,u_1}^U(A_{t+1})\} \\
\Lambda(x) &= \max \int \{V_x^E(A_{t+1}, \Gamma_{t+1}), V_{x,u_1}^U(A_{t+1})\} dF
\end{aligned}$$

Note, that frictional wage dispersion is created by the dispersion of the firm specific productivity distribution F. Moreover, observe that future values of employment and unemployment depend on the future idiosyncratic states. This forward looking behavior of the rational agent makes unemployment a less desirable state.

2.3 The Firm Problem

An entering firm's problem is described by its value to post a vacancy (V^I). An open vacancy entails flow costs of φ each period. We assume vacancies are homogeneous ex ante and the realization of the idiosyncratic productivity reveals only upon meeting a worker and entering into wage negotiations. If a worker is contacted, Γ is drawn from F[8]. There are three ways to fill a vacancy. First, an unemployed agent might be contacted, occurring with probability $q(\theta)$. Second, the firm might headhunt a worker that is employed and make him a job offer, which happens at rate $\frac{\lambda(1-\lambda_d)}{v}$. Or third, a worker might be offered the vacancy by a forced job movement, occurring at rate $\frac{\lambda\lambda_d}{v}$. Note that in any case the ex ante acceptance probability depends on the productivity of the vacancy. Given that firm and worker productivities are complements, higher productivity vacancies attract also lower productivity workers and are less likely to lose parts of their workforce to other firms. We relegate the further description of

[8] This can be rationalized by assuming that there is a match specific component in productivity. This is also the assumption made by Yamaguchi (2010).

(V^I) to Appendix A as it provides little additional intuition.

The value of a filled vacancy V_x^J depends on the life-cycle state of the matched employee and a firm employing someone in life-cycle state 1 has value

$$
\begin{aligned}
V_1^J(A_t, \Gamma_t) = y_t - w(A_t, \Gamma_t, 1) + \beta(1-\omega)E_t\{ \\
&(1-\lambda)[(1-p_1)\Phi(1) + p_1\Phi(2)] \\
&\quad + \lambda(1-\lambda_d)\eta(\Gamma_{t+1})[(1-p_1)\Phi(1) + p_1\Phi(2)]\}
\end{aligned}
$$

where $\eta(\Gamma)$ is the probability that the worker stays with the firm when contacted from an outside firm, which is increasing in Γ. Moreover, we have defined the auxiliary variable $\Phi(x)$ indicating the match continuation choice conditional on the life-cycle state and productivities:

$$\Phi(x) = \max\{0, V_x^J(A_{t+1}, \Gamma_{t+1})\}$$

2.4 Equilibrium

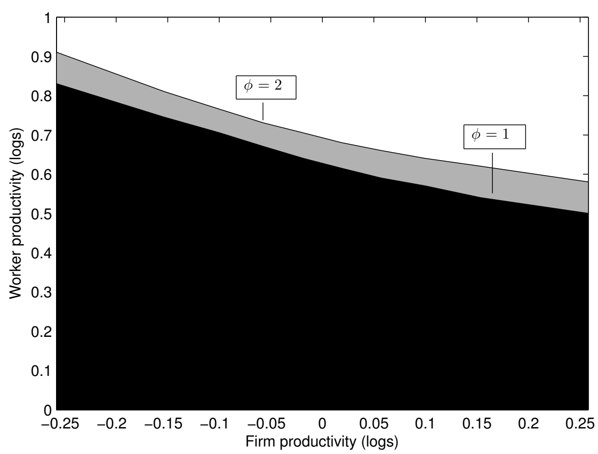
A stationary equilibrium consists of

- Value functions for the employed, unemployed and the firm value.
- Free entry drives profits for newly posted vacancies to zero: $V^I = 0$.
- Wages solve
  $$max_w : \{\alpha log(V_x^E - V_{x,u_1}^U) + (1-\alpha)log(V_x^J)\}$$
  where α is the bargaining power of workers and we made use of the fact that the value of a vacancy is zero.
- A policy function that is consistent with the value functions and that maps worker productivity, firm productivity, benefit entitlement, and the life-cycle state into a decision, whether a match is formed or not.

- Stationary distributions of the employed and unemployed over worker productivities, employment states, life cycle states, benefit entitlement states and firm productivities.

For a better understanding of the model, Figure I presents the equilibrium policy functions for workers with benefit entitlement. In the black area, match formation would yield negative surplus. In the white area all matches are formed and in the gray area match formation depends on the life-cycle of workers. Firms with low productivity only match up with workers of high idiosyncratic productivity. Therefore, the outflow rates for workers with high productivity are larger than for those with low productivities. The profile for workers in the first life-cycle is strictly below the profile of workers in the second life-cycle, representing their additional gains from taking up employment.

Figure I: Worker policy functions



Notes: The graph displays the policy functions. In the black area, no matches are formed. In the gray area, match formation depends on the life-cycle of the worker and in the white area all matches are formed.

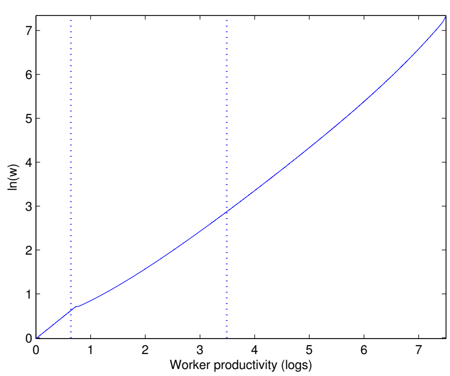
2.5 Approximating the Wage Schedule

To facilitate our subsequent analysis and to make our approach more comparable to standard microeconometric specifications we approximate the equilibrium wage schedule by a linear function. From the Nash-bargaining solution it is obvious that log wages are not a linear function in worker and firm productivity. Figures II and III plot $ln(w)$ over worker and firm productivity for agents in life-cycle state 2, holding the productivity of the other fixed at its mean value. The plots indicate that these functions can still be reasonably well approximated by a linear function. We asses this more formally by fitting a linear OLS regression to an economy generated by the true non-linear dynamics of our model. To be more specific, we simulate 50000 workers for 2 years from the stationary distribution, using our non-linear model. We then project the resulting data into a linear space employing the following regression:

$$ln(w_{i,t}) = \beta_0 + \beta_1 A_{i,t} + \beta_2 \phi_{i,t} + \beta_3 \Gamma_{i,t} + a_{i,t} \tag{1}$$

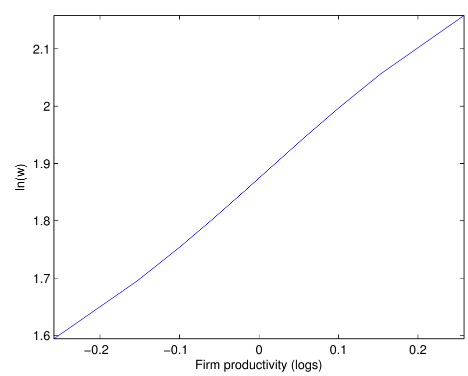
Note, assuming the law of large number holds, the error term $a_{i,t}$ measures the approximation error that results from the linear projection. Success is interpreted as getting predicted wages that are close to the actual wages. R^2 is above 0.996, suggesting that the fit of the linear regression model is quite well. Hence, we continue to work from now on with the linear approximation (1) to our true non-linear model.

Figure II:
Log wages over individual productivity



Notes: The graph displays the equilibrium log wage schedule, holding firm productivity and its median level and the life-cycle state fixed. The first 95% of all workers employed at such matches are within the dashed bounds.

Figure III:
Log wages over firm productivity



Notes: The graph displays the equilibrium log wage schedule, holding worker productivity and the life-cycle state fixed.

3 Parametrization

We take a dual strategy in assigning parameter values to our model. For a number of parameter values we take numbers from other studies. This allows us to make our results easily comparable. Also, for many of those parameters (discount factor and bargaining share, for example) our results are robust to variations. We will come back to this point below. The particular focus of our paper requires us to take great care in calibrating worker and firm productivity uncertainty and flow rates in and out of employment and between firms. Wherever possible, we therefore estimate our calibration targets for the related parameters using a single data set in order to insure consistency. The data set best suited for our analysis is the Survey of Income and Program Participation (SIPP) which is a representative panel of the non-institutionalized US civilian population. Although the SIPP provides very detailed and extensive coverage, we cannot estimate all of the productivity parameters on the basis of our data set. We therefore take additional information from other micro studies carefully discussing each of our choices. This section proceeds as follows: We first introduce the SIPP and explain sample selection. We then discuss our calibration regarding non-distributional parameters (preferences, institutions, flow rates). Finally, we discuss productivity distributions and how we estimate idiosyncratic and firm productivity uncertainty. Our calibration is summarized in Table 1. Some additional information regarding our data work is given in Appendix B.

3.1 Data Source and Sample Creation

Our empirical analysis aims to accurately identify job-to-job transitions and accompanying wage changes as well as wage dynamics on the job. We therefore require longitudinal monthly wage information which identify employer and occupation changes. The data set which best meets these requirements is the Survey of Income and Program Participation (SIPP). The SIPP is a representative sample of the non-institutionalized civilian US population maintained by the US Census Bureau. Its main goal is to track income dynamics and welfare program participation of households and individuals. The level of detail it provides in individual records allows us to accurately identify an individual's main job and hourly wages on that job.

In our analysis, we use the 1993 cohort from the SIPP which covers the years 1993-1995 (which also includes some observations from 1992)[9]. During that time, an individual completes at most 9 interviews. We use observations from individuals aged 16-70 for which we require complete information for the period of the interview on the individual's employment status, age and employer id. On top of that, we only consider an individual's primary job[10]. These restrictions leave us with 1,084,679 person/month observations.

The SIPP is a collection of panels of which a new one starts every year. In constructing the panels, the Census Bureau randomly assigns people to rotation groups which are then interviewed subsequently on a four-month basis. One completed rotation is called a wave. During the interviews, the respondents give information on their labor market status for each week in the past four months separately which is then used to assign one of eight possible activity statuses. While this form of reporting allows for a very precise labor market classification it also constitutes one of the sample's few drawbacks. Not only is it very hard to compare unemployment measures based on this classification to those based on other more widely used ones like for instance the ones in the Current Population Survey (CPS). It has also been shown to downward bias estimates of transition flows between employment and unemployment[11]. Because of these well known biases, we use estimates from corresponding CPS cohorts. Both panels are representative samples from the same population and so this should be unproblematic.

3.2 Non-Distributional Parameters

Model period is one month. The length of a period is of importance, because it puts an upper bound on the job offer probability $p(\theta)$ and the minimum duration of an

[9] We use the CEPR SIPP extracts available for download at *http://www.ceprdata.org/sipp/sipp_data.php.*

[10] As primary job we consider the position where the largest share of hours worked is spent.

[11] See Mazumder (2007) for a discussion.

unemployment spell. The first point is well supported by the data[12], but the second constraint is likely to be binding[13].

We calculate the EU and UE rate of the US non-institutionalized population from CPS data for the years 1994-1995 following Fallick and Fleischmann (2004) for reasons discussed above. The exogenous job destruction rate is set such that the total destruction rate d, the sum of endogenous and exogenous movements from employment to unemployment, is 1.43% per month. We attach to ξ a value that implies a monthly job finding rate of 0.271.

We can use SIPP data to calibrate the parameters guiding on the job search. Information on EE movements and wage changes identify λ and λ_d. We adjust λ to imply that 2.51% of workers switch employers every period. As discussed previously, in order to correctly model the efficiency of on the job search, it is important to know how many of these movements result in wage improvements. Our identifying assumption for telling voluntary and involuntary movements apart is that voluntary movements always result in wage increases. In our data set, 33% of all EE movements result in a nominal wage loss. We set the percentage of forced movements (λ_d) to 0.41 to match this statistic. In Appendix B, we provide further details on our identification of EE movements. We also supply additional evidence that wage cuts after job to job movements are a pervasive phenomenon in all subgroups of the population.

There is a large debate on the appropriate values of α, ι and θ, because of their importance for business-cycle fluctuations. Fortunately, in our stationary distribution analysis these parameters do not affect our results, because they only affect the job finding rate. Therefore, changing the parameters leads only to a recalibration of ξ. Hence, we normalize $\alpha = \iota = 0.5$ and use φ to match a labor market tightness of

[12]Holzer (1988) reports based on NLSY data that in the previous month 34% of the unemployed received at least one job offer and 12% received more than one offer. We are therefore confident that on average the unemployed worker does not receive more than one job offer per month.

[13]Clark and Summers (1979) report that based on the CPS 60% of all unemployed spells end within one month, while at any point in time, 69% of all unemployed have been out of a job for two months or more. These two figures can only coincide when a considerable fraction of unemployment spells end within less than one month. Therefore, our model cannot by construction match the high outflow rates within the first month. However, time disaggregation below one month is rather costly, because our numerical algorithm uses value function iteration which converges at a rate of $1 - \beta$.

0.6, which is the estimate by Hagedorn and Manovskii (2008).

Consistent with findings from Siegel (2002) for average bond and stock returns, we set β to imply a yearly interest rate of 4%. Next, we consider the flow value of unemployment. We choose unemployment benefits (b) to target a replacement rate of 25% of the mean wage[14]. As argued in Hall and Milgrom (2008) this provides an parsimonious describtion of the system. The same source suggests a value of leisure (Z) of 46% of the median wage which is inferred from micro consumption data. HKV claim that the random matching model does not permit frictional wage dispersion close to those observed in the data for any positive replacement rates. In total, our calibration implies a replacement rate of 71% of the median wage, which is substantial. Last, we fix the probability for an unemployed worker to loose his benefit entitlement such that average benefit entitlement is six months, which is the standard length in the US system outside of economic crisis.

In the presence of tenure and selection effects, it would be very hard (and potentially produce unreliable results) to estimate mean experience gains from our data set. We therefore use life-cycle transition rates and drift terms in productivity during employment to match statistics found by Dustmann and Meghir (2005). Productivity is assumed to grow at an annual rate of 8% when employed during the first life-cycle state and at a rate of 1% during the second. The transition probability between life cycle states (p) is set such that agents spend on average 24 months in the first state and 480 in the second. Following Olivetti (2006), an unemployed worker experiences 2% skill depreciation per year[15]. The subsistence level of log-productivity ($pmin$) is normalized to zero, as it does not carry any additional information that is not contained in the distribution of initial productivities.

Getting the life-cycle properties of productivity development right is of particular importance given the focus of our paper. We therefore compare additional non-

[14] This way of modeling UI implies that the effective replacement rate for low wage earners is much higher than for high wage earners. This is also true in the data, because the UI system has both an upper and lower bound. Yet, our model is likely to overstate this effect. Still, we opt not to correct for it as we believe it to be reasonably small and in order to keep the dimension of the state space tractable.

[15] This is also in line with Ljungqvist and Sargent (2008), who impose that skill depreciation is twice the rate of skill accumulation.

targeted reduced form moments that are functions of the above parameters in our model and the data. Average workers' wage gains over time are informative about the validity of our experience profile and the on the job search behavior of workers. We have to rely to outside estimations for this statistic, because we do not observe workers sufficiently long in our data. Topel and Ward (1992) estimate that after 10 years of potential experience, US male workers have experienced wage gains of about 50%. Because their paper uses a different sample from another tim period and because we neglect other possibly important drivers of wages, one would not expect the model to match this statistic perfectly. Nevertheless, the model yields quite a good fit to this statistic (56% average wage growth).

3.3 Distributional Parameters

We now describe the way we estimate the variance of idiosyncratic productivity shocks σ_ϵ^2 and firm productivity dispersion σ_F^2. Neither statistic is directly observable in the data because of measurement error. Additionally, agents endogenously select themselves into and out of employment and into employment with firms of specific productivity levels in response to idiosyncratic productivity developments. Instead, we identify them as follows: We derive a model statistic that is an (unknown) function of σ_ϵ^2 and σ_F^2 respectively. Taking our model to be the data generating process, we estimate the same statistic in our data controlling for measurement error. We can then adjust our model parameter until the identifying moment matches the empirical estimate.

3.3.1 Measuring Idiosyncratic Productivity Uncertainty

Remember equation (1), our linearized approximation to the wage function:

$$ln(w_{i,t}) = \beta_0 + \beta_1 A_{i,t} + \beta_2 \phi_{i,t} + \beta_3 \Gamma_{i,t} + a_{i,t}$$

Next, consider $\Delta ln(w_{i,t}^w)$, the change in log wages of workers being employed with the same employer in two consecutive months. Endogenous responses to productivity shocks (workers quitting after bad productivity shocks) implies that we will only observe a self selected subgroup of $\Delta ln(w_{i,t}^{w})$ every period. However, we do not have to control for this effect explicitly, because it is present both in our model and in the data. We therefore introduce the concept of the observed wage $w_{i,t}^{obs}$:

$$\Delta ln(w_{i,t}^{obs}) = \beta_1(\nu + \epsilon_{i,t}^{obs}) \tag{2}$$

ϵ^{obs} follows a distribution of unknown functional form. It is, however, an object which we observe in the data and whose moments we can use to identify σ_ϵ^2 in our model. Regressing out the constant from (2), we can calculate the resulting prediction error variance whose data counterpart we will later use to identify idiosyncratic productivity uncertainty:

$$E(\Delta ln(\hat{w}_{i,t}^{obs})\Delta ln(\hat{w}_{i,t}^{obs})) = \beta_1^2\sigma_{\epsilon^{obs}}^2 \tag{3}$$

Turning to our SIPP data, we assume that wages are generated by:

$$ln(w_{i,t}) = \alpha_0 + \alpha_1 d_t + \alpha_2 Z_{i,t} + \beta_2\Gamma_i + e_{i,t}$$

where d_t captures aggregate states, such as TFP and $Z_{i,t}$ is a vector of idiosyncratic components. We split the unobservable $e_{i,t}$ into two parts:

$$e_{i,t} = r_{i,t} + \beta_1 A_{i,t}$$

As in the model, $A_{i,t}$ is assumed to follow a random walk with drift while $r_{i,t}$ captures measurement error. Following Meghir and Pistaferri (2004), we assume measurement error to follow an $MA(q)$ process (i.e. $r_{i,t} = \Theta(q)\iota_{i,t} = \iota_{i,t} - \sum_{j=1}^{q}\theta_j\iota_{i,t-j}$). Analogous to before, we work with the observed wage process $w_{i,t}^{obs}$.

We first regress log-differences in observed within-firm wages on a constant, a period dummy to control for business cycle effects, an industry dummy[16], a month dummy to control for seasonality and an interaction between the industry and month

[16] We use the 23 major industry classification system.

dummy to control for month specific industry effects such as collective bargaining. Again, we call the corresponding residuals of this regression $\Delta ln(\hat{w}_{i,t}^{obs})$. In order to derive a moment condition analogous to equation (3), we need one more identifying assumption:

$$E(\epsilon_{i,t}^{obs}\epsilon_{i,t-j}^{obs}) = 0 \ \forall j \neq 0$$

Our model economy indicates almost no endogenous quits and so we think this assumption not overly restrictive. Following Meghir and Pistaferri (2004), we can now derive a moment estimator for on the job wage variance:

$$E(\Delta ln(\hat{w}_{i,t}^{obs})(\sum_{j=q-1}^{q+1} \Delta ln(\hat{w}_{i,t+j}^{obs}))) = \beta_1^2 \sigma_{\epsilon^{obs}}^2 \tag{4}$$

Given that studies on annual wage growth typically assume iid measurement error, we fix q at 12. Finally, we adjust σ_ϵ^2 in our model economy, until (3) and (4) coincide. All endogenous sorting that causes the observed productivity distribution in the data to differ from the true one is also present in our model.

3.3.2 Measuring Firm Productivity Dispersion

Using wage changes from workers experiencing a job to job transition, we use a similar identification strategy as before to identify firm productivity dispersion. Define the change in log wages of individual i after a job to job transition as

$$b_{i,t} = ln(w_{i,t}) - ln(w_{i,t-1})$$

After regressing out a constant, we can again define residual observed wage changes as:

$$\hat{b}_{i,t}^{obs} = \beta_1 \epsilon_{i,t}^{obs} + \beta_2(\Gamma_i^{obs} - \Gamma_{i,-1}^{obs})$$

here Γ_i^{obs} is the productivity of the current employer and $\Gamma_{i,-1}^{obs}$ is the productivity of the previous one. We can now identify firm productivity dispersion via the excess variance of job movers relative to job stayers:

$$E(\hat{b}_{i,t}^{obs}\hat{b}_{i,t}^{obs}) - E(\beta_1^2 \epsilon_{i,t}^{obs^2}) = \beta_2^2(\Gamma_i^{obs} - \Gamma_{i,-1}^{obs})^2 + 2\beta_1\beta_2\epsilon_{i,t}^{obs}(\Gamma_i^{obs} - \Gamma_{i,-1}^{obs}) \tag{5}$$

Under the assumption that measurement error for job movers is not more severe than for job stayers[17] and after again controlling for industry and time effects, we can construct the same statistic in our data and use it to match firm productivity dispersion. Again, we are confident that our estimation provides a good fit to the data. A worker who switches employer on average experiences a wage gain of 2.8% which looks good compared to the model estimate of 3.1%.

3.3.3 Initial Worker Productivities

Finally, we have to calibrate the distribution of initial productivities for which we assume normality. Without matched employer-employee data it is not possible to separately identify the variance of initial individual productivities σ_N^2. We take this value from Woodcock (2008), who estimates it to be 0.2 based on an US linked employer-employee data set[18]. The mean of the distribution of initial productivities carries information about the unemployment duration distribution, once $pmin$ is fixed. However, the effect turns out to be extremely small for our calibration. Changing the mean productivity from 1.9 to 3.5 changes the fraction of unemployed with a spell of one month or less by only 0.005. It leaves the fraction of unemployed who endogenously quit after a bad productivity shock virtually unchanged[19]. Hence, we fix μ_N at 2.

[17] As discussed previously, we are excluding those individuals who are holding multiple jobs after a transition to rule out this source of additional reporting error. We have also constructed three-month-averages of wages after a movement to mitigate other sources of reporting error in the following the transition. This, however, did not affect our estimates.

[18] As we do, he treats initial ability as a random variable and estimates the variance by a random effect model. Storesletten et al. (2004) come to practically identical results.

[19] The reason is that increasing the mean productivity level increases equilibrium wages almost proportionally. This leads to an increase in unemployment benefits to keep the replacement rate constant and therefore the threshold levels for hiring move almost proportionally out.

Table 1: Calibration

Variable	Target
$\beta = 0.9967$	4% yearly interest rate
$\varphi = 62.5$	$\theta = 0.6$
$\alpha = \iota = 0.5$	Normalization
$b = 2.195$	$\frac{b}{w_{mean}} = 25\%$
$Z = 4.04$	$\frac{Z}{w_{mean}} = 46\%$
$\lambda_l = 0.16$	6 month benefit duration
$\omega = 0.0142$	$d = 0.0143$
$\xi = 0.37$	UE flow of 0.271
$\lambda = 0.0413$	EE flow of 0.0251
$\lambda_d = 0.405$	33.3% of EE are forced
$\nu(1) = 0.0067$	8% yearly productivity growth
$\nu(2) = 0.00083$	1% yearly productivity growth
$p_1 = 0.04$	2 years in 1st life-cycle
$p_2 = 0.002$	40 years in 2nd life-cycle
$\delta = 0.00167$	2% yearly skill depreciation
$pmin = 0$	Normalization
$\sigma_\epsilon = 0.0547$	Equation (4)=0.0022
$\sigma_F = 0.147$	Equation (5)=0.051
$\sigma_N = 0.445$	-
$\mu_N = 2$	Normalization

Notes: The first column states the calibrated variable and the value, the second states the target, and the third states the source. SIPP refers to the 1993 Survey of Income Program Participants and CPS refers to the 1994-1995 Current Population Survey.

4 Results

We now present the main results of our paper. First, we demonstrate that our model is successful in generating frictional wage dispersion of the size suggested by HKV. We then show that the forward looking behavior of agents with respect to skill development and benefit duration and the ability of on the job search are both of key importance to understand why identical workers accept very different wages. We demonstrate that skill development while being employed is the single most important factor that drives a wedge between the value of employment and unemployment.

An important related result is that ignoring the possibility of forced movements and inferring on the job search efficiency only from observed job to job movements leads to an implausibly high contacting rate and misleading interpretation of the results. We then turn to the second main contribution of our paper. Having identified firm dispersion, productivity development and the distribution of workers over firms we can reassess the importance of each of these factors for overall wage inequality. Our results indicate that about 90% of wage inequality is explained by dispersion in worker productivities.

4.1 Frictional Wage Dispersion

Table 2: Frictional Wage Dispersion

Percentile	Model	Hornstein et al. (2007)
Min wage	2.06	3.11
1^{st}	1.42	1.9
5^{th}	1.39	1.45
10^{th}	1.27	1.32

Notes: The table displays the simulated mean-min ratios from our baseline model and the measure reported by Hornstein et al. (2007) based on PSID data. The minimum wage is measured as the absolute minimum, at the 1st percentile, the 5th percentile and the 10th percentile respectively of the frictional wage distribution respectively.

HKV suggest to measure frictional wage dispersion by looking at the predicted ratio of the mean to the minimum wage. As minimum wage they try out the 1^{st}, 5^{th} and 10^{th} percentile of the frictional wage distribution respectively[20]. They estimate this ratio from different sources of US data and find a plausible range between 1.32 and 3.11 with their favorite estimate being around 1.7. This estimate controls for worker productivity by a fixed effect regression, which includes observable time varying worker characteristics as further controls. We evaluate success of our model by its ability to generate a sizable mean to min ration in wages net of worker effects.

[20] The authors reason that the reported absolute minimum wage is likely to be too small, due to reporting errors and they try to correct for this by looking at percentiles.

We argue in Appendix C that their estimator is likely to be upward biased, due to stochastic productivity. We use our structural model to asses the size of the bias and show it to be small.

Our model has no algebraic solution to this statistic, we therefore use (1) to compute

$$ln(\hat{w}_{i,t}) = \beta_3 \Gamma_{i,t}$$

After transforming log predicted wages $ln(\hat{w}_{i,t})$ back to their level, we compute the ratio of mean to minimum frictional wage. The results are displayed in Table 2. The results vary with the percentile considered, but they clearly indicate that the model creates sizable frictional wage dispersion in the range suggested by HKV[21]. More specific, the model predicts that the observed median wage is 2.06 times the minimum wage of exactly identical workers, which is well supported by their estimates given their argument that the data contains additional noise.

4.1.1 Explaining the Drivers of Frictional Wage Dispersion

We now turn to evaluate the sources of frictional wage dispersion in our model and their relative importance. We evaluate the role of forward looking behavior by the agents and the role of on the job search consecutively. In each experiment we recalibrate the model such that the flow rates and the flow value of unemployment remain as in our baseline model.

In Table 3 we solve different versions of our model specification subsequently excluding skill depreciation in unemployment, learning on the job and finite duration of unemployment benefits. It turns out that skill accumulation on the job is the single most important factor that drives a wedge between the value of employment and unemployment. The large average on the job wage growth that we observe in the data greatly amplifies the value of employment compared to non-employment. As potential experience gains are equal in all firms, being employed at all becomes the crucial. The same argument applies to skill depreciation in unemployment. Even

[21] As a robustness check whether our estimator is able to filter out individual effects, we resolve the model and simulate it for a lower value of the variance of initial productivities. The computed mean min ratio is indeed the same.

though its effect is smaller than the effect of skill accumulation on the job, it turns out to be an important feature in understanding why agents accept relatively low reservation wages at job entry. The introduction of finite unemployment benefits has the smallest effect.

Table 3: Contributions to Frictional Wage Dispersion

Specification	FWD
No skill depreciation ($\delta = 0$)	1.83
No learning on the job ($\nu(\phi) = 0$)	1.15
Infinite UI ($\lambda_l = 0$)	1.96

Notes: The table displays the frictional wage dispersion (FWD) for four different model specifications that differ from our baseline model by one parameter restriction.

Table 4 presents job offer probabilities for the employed and the resulting frictional wage dispersion for different assumptions about on the job search. Consider first a model which replicates empirical EE flow sizes but neglects that some transitions are followed by wage cuts. In such a setting, the imputed job offer arrival rate on the job increases by a factor of 17 compared to our baseline specification. Going back to our baseline specification, this means that workers are a less selective when receiving on the job offers in the presence of force movements. There are two reasons for this. First, if a job offer is a forced one, moving is almost always preferred to quitting into unemployment. Second, forced job movements decrease the rate at which agents climb up the productivity latter of firms, making future job offers more likely to be better than today's offer. Therefore, search on the job is less efficient in a model featuring forced job movements. Hence, the value of employment decreases relative to the value of unemployment, which again decreases frictional wage dispersion. In total, the model without forced job movements predicts a mean to min ratio of 4.85, which would largely overstate frictional wage dispersion given realistic values for other parameters.

This is not to say that on the job search has no important role. The bottom panel of Table 4 illustrates this point. Once we disallow all on the job search, agents do not accept low productive jobs anymore and our measure of frictional wage dispersion falls well below its empirical counterpart. This is despite the presence of our skill accumulation processes and finite UI payments. In the end, only the combination of all effects allows us to replicate empirically observed frictional wage dispersion. So it is a combination of effects that makes both unemployment less attractive (in an absolute sense) and employment more attractive compared to a standard search model.

Table 4: Frictional Wage Dispersion and On the Job Search

Specification	λ	FWD
No forced movements ($\lambda_d = 0$)	0.71	4.85
No on the job search ($\lambda = 0$)	0	1.1

Notes: The table displays the job offer probability on the job for two different model specifications and the resulting frictional wage dispersion (FWD) measured at the minimum of the frictional wage dispersion.

4.2 Wage Dispersion

In this final section, having established that our model reproduces empirically reasonable amounts of frictional wage dispersion, we want to assess its contribution to overall wage dispersion. The answer to this question has important implications for policymakers interested in reducing overall wage dispersion. If search frictions, manifested in productivity dispersion of firms and sorting of workers over firm types, were the main driver of wage inequality, policies should be aimed at increasing matching efficiency. If, however, worker heterogeneity either in the form of differing initial abilities or of heterogenous employment histories are the main drivers of wage dispersion, little efficieny and equity gains can be expected from such policies. Measures aimed at improving general education or aimed at skill updating for workers already in the labor force can then be expected to be far more beneficial.

In order to assure that our model can be used to make such statements, we first have to assure that it reproduces a wage distribution comparable to what we see in the data. Whereas our results presented so far were mainly just dependent on a good identification of σ_F, a good identification of σ_ϵ and σ_N is now of crucial importance as well. Our results regarding the overall wage distribution are reassuring and allow us to proceed. We then construct a simple variance decomposition to asses the contribution of search frictions to overall wage dispersion over the life cycle. We find that frictions can only explain a very moderate amount of wage dispersion. Their contribution never exceeds 20% and is decreasing in worker's age. These results favor education and training as policies to overcome wage hetergeneity.

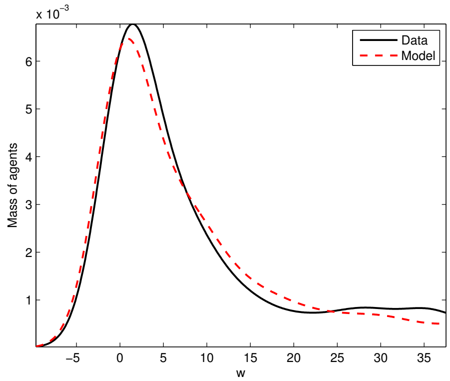
4.2.1 Overall Wage Dispersion in the Model and in the Data

Figure IV plots the kernel estimator of the aggregate density function of wages[22]. It features the characteristic right skew of the observed wage distribution in the data. We compute the coefficient of variation and the coeficient of excess kurtosis to be able to compare the theoretical distribution with the empirical distribution. The coefficient of variation and excess kurtosis in the model are 0.63 and 8.6, compared to 0.59 and 7.04 in the data, respectively.

Figure V displays the theoretical and empirical Lorenz curves of wages. Our model economy exhibits slightly more wage inequality, but the difference is negligible. Overall, the results make us confident that our model economy picks up the key moments of wage inequality present in the data.

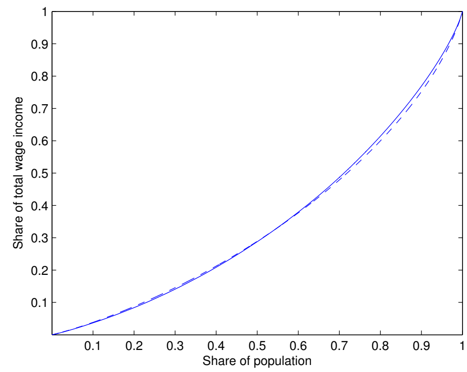
[22]We truncate our observed wage data at the bottom and top 1% wage observations to delete outliers. We do the same adjustment to our simulated data in this section.

Figure IV:
The Wage Distribution



Notes: The Figure displays the the theoretical PDF of workers over wages, smoothed with a kernel estimator.

Figure V:
Empirical and Theoretical Lorenz Curve



Notes: The straight line is the Lorenz curve of wages in SIPP data. The dashed line is the Lorenz curve from the theoretical model.

4.2.2 The Contribution of Frictional Wage Dispersion to overall Wage Dispersion

In this section we evaluate to what degree wage dispersion results from workers' productivity differences, firm differences, and worker selection into matches. For this purpose we simulate a panel of 10000 workers' histories for 16 years. Consider the following variance decomposition based on a slightly modified version of (1), which we estimate separately for each age cohort in our simulated data[23]

$$Var(ln(w_i)) = \beta_1^2 Var(A_i) + \beta_2^2 Var(\Gamma_i) + 2\beta_1\beta_2 Cov(A_i, \Gamma_i)$$

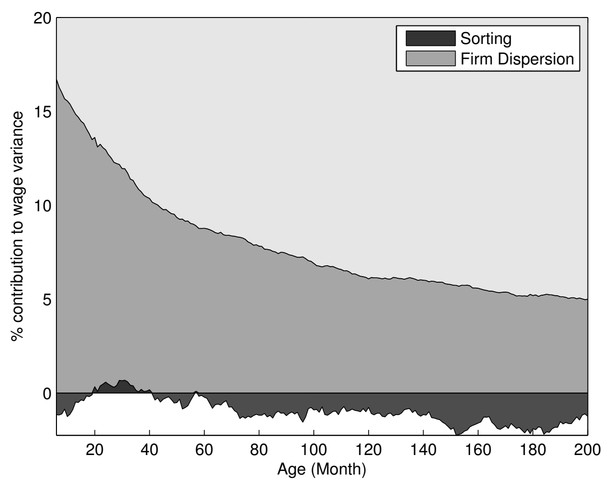
Figure VI displays the results. Sorting of worker to firm productivities has a mild negative effect. Moreover, already for labor market entrants firm heterogeneity explains only about 20 percent of overall log wage variance[24]. Our model identifies worker heterogeneity as the dominant factor in explaining variations in wages and this effect is increasing in age.

[23] The mean R^2 of all regressions is 0.99, giving us confidence that the approximation works fine.

[24] Abowd et al. (1999a) find for the state of Washington that firm effects explain around 24% of the variance in log wages. However, Abowd et al. (1999b) find much lower firm effects for France. Our estimates are somewhat in between these studies.

Note, that our finding that individual worker heterogeneity is the main driver of aggregate wage dispersion is not in contrast to the fact that a Mincer wage equation with worker fixed effects usually explains only little variation in wages. Individual productivity is only partially correlated to initial productivity[25] and all changes in productivity are time varying unobservables to the econometrician. The typical worker observables included in the Mincer wage equation can at best proxy for these variations. Also note, that our finding is perfectly in line with Hagedorn and Manovskii (2010), who use time series evidence to show that search frictions explain little of overall wage dispersion.

Figure VI: Contribution of Search Friction to Overall Wage Dispersion



Notes: The graph displays the contribution of sorting (dark gray area), firm effects (medium gray area) and worker effects (light gray area) on the variance of log wages, conditional on age.

5 Conclusions

Search theory emphasizes that identical workers can earn different wages in the market due to dispersion in firm payment schemes. However, Hornstein et al. (2007)

[25]The correlation between the initial ability and productivity after 16 years of labor market experience is 0.46 in our simulation.

show that the empirically observed size of these differences is too large to be consistent with observed worker flows within the usually applied model specifications. The high observed outflow rate from unemployment suggests that standard specifications overstate the benefits from additional search.

In our paper, we resolve this apparent puzzle by modeling skill development (learning by doing on the job, skill loss during unemployment) and duration dependence in unemployment benefits in a random on the job search model featuring two-sided heterogeneity. We demonstrate how to identify structural parameters of our model using data on job mobility and wages from the SIPP. Our expanded search model is successful in jointly generating frictional wage dispersion of the size suggested by the data and high outflow rates from unemployment.

The most important mechanism behind our results is experience gains during employment. It should be stressed, however, that only the joint presence of all the channels allows the model to successfully replicate the empirical amount of frictional wage dispersion. Another important result concerns the modeling of on the job search. The data suggest that around 1/3 of all observed job to job transitions result in nominal wage cuts. We argue, hence, that infering search efficiencies from a basic job ladder modell where all job movements are the result of optimal choices overstates the efficiency of on the job search.

Having identified the sources and size of frictional wage dispersion, we can asses its importance for overall wage inequality. Our model assigns less than 10 percent of overall wage dispersion to dispersion of firm productivities - a very modest contribution. Instead, large dispersion in worker skills at labor market entry drive large parts of the dispersion. This suggests that, although firm productivity dispersion can cause substantial wage differences between workers of identical capabilities, the effects of efficiency gains in search on overall wage inequality are likely to be small. Hence, we stress that to reduce wage inequlity the emphasis must be on factors influencing skill dispersion instead of increasing labor market search efficiencies.

References

Abowd, J. M., H. Finer, and F. Kramarz (1999a): "Individual and Firm Heterogeneity in Compensation: An Analysis of Matched Longitudinal Employer-Employee Data for the State of Washingtopn," mimeo, Cornell University.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999b): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–334.

Altonji, J. G. and N. Williams (1998): "The Effects of Labor Market Experience, Job Seniority, and job Mobility on Wages," *Research in Labor Economics*, 17, 233–276.

Bontemps, C., J.-M. Robin, and G. J. V. den Berg (1999): "An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms," *International Economic Review*, 40, 1039–74.

Bontemps, C., J.-M. Robin, and G. J. van den Berg (2000): "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, 41, 305–58.

Burdett, K., C. Carrillo-Tudela, and M. Coles (2009): "Human Capital Accumulation and Labour Market Equilibrium," Discussion Paper 4215, Institute for the Study of Labor (IZA).

Burdett, K. and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39, 257–73.

Carrillo-Tudela, C. (2010): "Job Search, Human Capital and Wage Inequality," mimeo, University of Leicester.

Center for Economic and Policy Research (2010): "SIPP Uniform Extracts, Version 2.1.5," .

Clark, K. B. and L. H. Summers (1979): "Labor Market Dynamics and Unemployemnt: A Reconsideration," *Brookings Papers on Economic Activity*, 10, 13–72.

Dustmann, C. and C. Meghir (2005): "Wages, Experience and Seniority," *Review of Economic Studies*, 72, 77–108.

Fallick, B. and C. A. Fleischman (2004): "Employer-to-employer flows in the U.S. labor market: the complete picture of gross worker flows," Finance and Economics Discussion Series 2004-34, Board of Governors of the Federal Reserve System (U.S.).

Flinn, C. and J. Mabli (2008): "On-the-Job Search, Minimum Wages, and Labor Market Outcomes in an Equilibrium Bargaining Framework," Carlo Alberto Notebooks 91, Collegio Carlo Alberto.

Hagedorn, M. and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98, 1692–1706.

——— (2010): "Search Frictions and Wage Dispersion," mimeo, University of Pensylvania.

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

Holzer, H. J. (1988): "Search Method Use by Unemployed Youth," *Journal of Labor Economics*, 6, 1–20.

Hornstein, A., P. Krusell, and G. L. Violante (2007): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," Working Paper 13674, National Bureau of Economic Research, Inc.

Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006): "The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US," *European Economic Review*, 50, 877–907.

Ljungqvist, L. and T. J. Sargent (2008): "Two Questions about European Unemployment," *Econometrica*, 76, 1–29.

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 100, 1432–1467.

Mazumder, B. (2007): "New Evidence on Labor Market Dynamics Over the Business Cycle," *Economic Perspectives*, 2007, 36–46.

Meghir, C. and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72, 1–32.

Mortensen, D. (2003): *Wage Dispersion: Why are Similar Workers Paid Differently*, MIT Press.

Moscarini, G. (2005): "Job Matching and the Wage Distribution," *Econometrica*, 73, 481–516.

Moscarini, G. and K. Thomsson (2007): "Occupational and Job Mobility in the US," *Scandinavian Journal of Economics*, 109, 807–836.

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70, 2295–2350.

Siegel, J. (2002): *Stocks for the Long Run*, McGraw-Hill.

Stinson, M. (2003): "Technical Describtion of SIPP Job Identification Number Editing in the 1990-1993 SIPP Panels," mimeo, U.S. Census Bureau.

Storesletten, K., C. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing ober the Life Cycle," *Journal of Monetary Economics*, 51, 609–633.

Topel, R. H. and M. P. Ward (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107, 439–479.

Woodcock, S. D. (2008): "Wage differentials in the presence of unobserved worker, firm, and match heterogeneity," *Labour Economics*, 15, 771–793.

Yamaguchi, S. (2010): "Job Search, Bargaining, and Wage Dynamics," *Journal of Labor Economics*, 28, 595–631.

A The Value of a Vacancy

Here we supply the calculation of vacancy value which for reasons of parsimony we excluded from the main text. To evaluate future profit prospects and acceptance probabilities, the entrepreneur needs to know the stationary distributions of the unemployed over productivity, benefit states and life-cycle states, which has density $f(\psi)$. Moreover, he needs to know the distribution of workers over their productivities, life-cycle states and other firms' productivities, which has density $f(\chi)$. Summarizing the workers' states in $s = (A, \Gamma, \phi)$, the value of posting a vacancy (V^I) is the expectation of firm value V_x^J over productivity and life-cycle states, minus the vacancy posting costs φ:

$$
\begin{aligned}
V^I = -\varphi + \beta E_t\{q(\theta)[\int\int\int V_x^J(s')f(\psi)d\psi q_1(s)dsdF] \\
+ \frac{\lambda(1-\lambda_d)}{v}[\int\int\int V_x^J(s')f(\chi)d\chi q_2(s)dsdF] \\
+ \frac{\lambda\lambda_d}{v}[\int\int\int V_x^J(s')f(\chi)d\chi q_3(s)dsdF]\}
\end{aligned}
$$

where q_1, q_2, q_3 are the probabilities that a worker will accept the job offer given that he is of type A and in life-cycle ϕ and the firm is of type Γ. These probabilities are strictly increasing in Γ, as a more productive firm finds it easier to attract workers. Note, we set the continuation value of a vacancy to zero, which is true in equilibrium, because of free entry into the market.

B More on the empirics of on the job search

B.1 Measuring job to job employment flows

In order to asses the efficiency of search on the job, it is crucial to accurately identify job to job transitions in the data. One of the biggest advantages in working with SIPP data is that workers are asked to report an employment status for each week of the reporting period separately. While a higher degree of time aggregation may mask intermittent unemployment spells, we can identify any unemployment spell lasting longer than one work week.

In a given month we count as employed someone who reports holding a job for the entire month. This definition includes paid as well as unpaid absences as result of vacations, illnesses or labor disputes. It does exclude, however, those who report having been on layoff for at least a week. There is no standard definition for job to job movements in empirical work. We therefore experiment with several different definitions. Our first measure is analogous to the definition in Fallick and Fleischman (2004) and equates job to job transitions with firm changes. We use a monthly employer identifier based on company names created by Stinson (2003). We refer to this definition by $EE1$. Given that a firm is a match in our model and given that employers may transit between jobs within a given firm, it might be useful to somewhat broaden the concept beyond employer id changes. For $EE2$ we therefore follow Moscarini and Thomsson (2007) in identifying job to job movements by changes in the three digit occupational code. Moreover, we define $EE3 = EE1 \cup EE2$ and $EE4 = EE1 \cap EE2$.

Table 5: EE flow rates based on different definitions

$EE1$	$EE2$	$EE3$	$EE4$	CPS
1.93	1.77	2.51	1.19	2.82

Notes: The Table shows percentage probabilities for worker job to job transitions from SIPP data from end of 1992 to 1995. For reference we also quote monthly averages from Fallick and Fleischmann (2004) for the years 1994-1995. The different flow definitions can be found in the text.

Table 5 lists EE flow rates based on the different definitions. For comparison, we also report averages from monthly estimates for the years 1994 and 1995 taken from Fallick and Fleischman (2004) who use CPS data. As can be seen, identifying EE movements by employer changes or changes in the occupational code alone yields roughly comparable flow sizes. However, only our broadest definition of job-to-job employment transitions comes close to the magnitude found using CPS. In order to ensure comparability of our results with studies based on CPS data and following the arguments made above, we calibrate our model baseline specification on the 2.51% based on definition $EE3$.

B.2 Wages and On the Job Search

We argue in the paper that the magnitude of job-to-job flows in itself is insufficient to evaluate the efficiency of on the job search. Instead, the question is how many of these job changes actually yield higher wages for the worker. In this section, we demonstrate that about a third of all job-to-job transitions result in lower nominal wages for the worker. In our model, we interpret these movements as forced ones which either mask the finding of a new job within notice period after having been layed-off or represent movements out of non-financial motives such as family reasons. As Postel-Vinay and Robin (2002) point out, these wage cuts might also be the result of optimizing behavior if the worker expects a steeper wage trajectory at his new employer. As we demonstrate below, this hypothesis is not borne out by our data. In addition, the phenomenon turns out robust to all sorts of data stratifications.

In the SIPP, respondents are asked, whether they are paid by the hour. If so, the reported hourly wage is recorded. Otherwise, we obtain hourly wages by dividing total monthly earnings by hours worked[26]. For the present purpose and all subsequent exercises, we drop any person/month observation for which we cannot determine an hourly wage. In addition, we drop observations without industry identifier, the self-employed and EE movements which result in the individual holding more than one job after transiting[27]. Finally, we exclude the .5 percent most extreme observations from both ends of the wage growth distribution to get rid of outliers[28].

[26]For further details see the CEPR SIPP User Notes.

[27]An individual working two jobs simultaneously may have trouble correctly attributing hours worked to the different jobs. This could potentially add noise to the data.

[28]For nominal log wage changes, this means excluding observations above a change in log wages of 0.818 and below -0.78.

B.2.1 Wage Gains from Employment Changes

First, we consider the mean change in log wages after a job to job transition. Our results depend somewhat on whether we consider nominal or real wage changes. Of course, the worker should only care about real wages in making his decision. Meanwhile, an argument can be made that in the presence of some wage rigidity, the worker expects a real wage loss on his current job as well and therefore compares nominal wages. Table 6 shows mean nominal and real wage gains for our different definitions of job to job movements.

Table 6: Aggregate Changes in Wages after EE

	Nominal			Real		
	Ave. change	Share loss	Ave. loss	Ave. change	Share loss	Ave. loss
$EE1$	0.0302	0.3420	-0.2522	0.0280	0.5303	-0.1651
$EE2$	0.0309	0.3618	-0.2503	0.0286	0.5032	-0.1826
$EE3$	0.0280	0.3337	-0.2353	0.0257	0.5306	-0.1505
$EE4$	0.0362	0.3908	-0.2815	0.0339	0.4879	´-0.2279

Notes: The Table shows statistics concerning wage changes after a job to job transition for real and nominal wages, respectively. The statistics under consideration are: The average change in log wages, the share of workers incurring a wage loss, and the average change in log wages, given that the observed change is a loss. We differentiate between four different measures of job to job transitions: $EE1$ identifies a job to job transition, if a worker is employed at a different firm between two consecutive months. $EE2$ identifies a job to job transition, if the worker's 3 digit occupation code changed between two consecutive months. $EE3 = EE1 \cup EE2$. $EE4 = EE1 \cap EE2$.

Wage gains after a job to job transition average only to about 3 percent. As shown in Table 6, this is because roughly 35 percent of these transitions actually yield nominal wage losses. The figure increases to about 51 percent when considering real wages. Wage losses are not just frequent, they are also sizable. Conditional upon taking a cut after an EE movement, losses average to 25 percent for nominal and 16 percent for real wages. Reassuringly, these figures are largely invariant to which definition we use. From now on, all statistics reported will therefore be based on $EE3$ only.

We also stratify our sample by different observable characteristics to show that the phenomenon we just described is not driven by a specific population sub group, but are a key characteristic of the entire labor market. The results are summarized in Table 7.

Table 7: Share of Wage Cuts after EE Movement in different subsamples

		Nominal		Real	
Stratify by:		Share loss	Nr. of Obs.	Share loss	Nr. of Obs.
Year					
	1993	0.3277	5971	0.5417	5973
	1994	0.3242	5039	0.5206	5038
	1995	0.3548	3915	0.5289	3916
Sex					
	Male	0.3355	8010	0.5271	8011
	Female	0.3316	6956	0.5349	6957
Age					
	16-25	0.3252	4353	0.5064	4354
	26-50	0.3394	8992	0.5335	8993
	51-70	0.3251	1621	0.5846	1621
Industry					
	Agriculture	0.3565	172	0.5622	172
	Manufacturing	0.3065	4781	0.5080	4782
	Trade	0.3512	4383	0.5391	4383
	Services	0.3501	1763	0.5727	1764
	Government	0.3389	3867	0.5276	3867
Income					
	Lowest 25%	0.2113	4002	0.4091	3979
	25-75%	0.3474	7343	0.5466	7358
	Top 25%	0.4479	3621	0.6380	3631

Notes: The Table shows the share of workers incurring a wage cut after a job to job movement, given different ways of splitting our sample. The column "Nr. of Obs." shows the number of measured job to job movements in the specific sub sample. Due to slightly different outlier identifications, this number does not need to match exactly between the cases of nominal and real wages.

We first split our sample into different years. The willingness of workers to accept a wage reduction upon transition might depend on the aggregate state of the economy. In the years 1993 to 1995, the time of our sample, the US economy was gradually moving out of the post-Gulf War I recession and unemployment was steadily falling throughout the sample period. Still, as indicated in the first panel of Table 7, there is now discernible time trend in the data. By 1995, unemployment had reached a historic low but workers still accepted a wage cut when making an EE movement about one third of the time.

Women are known to have less stable work relationships than men and might therefore be responsible for an overproportional share of loss making employment to employment transitions. Nonetheless, in the data both sexes have an equal probability of experiencing a wage cut after moving. The same holds for stratifications by age groups. Young workers have a loser attachment to the labor market and may initially experiment with different career paths or search for jobs with higher non-monetary benefits. But none of these phenomena cause the youngest age group to experience markedly more EE transitions with wage losses.

We try out two more relevant data subsets. The first concerns the industry the worker moves to. Some industries may offer substantial non-monetary benefits compared to others. Of course, this exercise is not only subject to selection issues, it is also well-known that wages show industry differentials. In consequence, we should be expecting to identify industry pairs where wages fall in expectations when moving from one industry to the other. In order to have sufficiently many observations for all subsamples, we group industries into four broad sectors using their three digit industry codes: Agriculture, Manufacturing, Trade, Private Services, Government. There are notable differences between sectors. Still, the share of workers incurring a wage cut after a job to job transition never falls below 30.65 percent.

Lastly, we stratify our sample by earnings. We split the main sample into its lowest and highest quartile and the observations in between. Again, there is a selection issue because high wage earners are most likely to incur a loss when they are forced to look for alternative employment. In a simple employment lottery, where all workers sample wages from the same random distribution, the probability of incurring a wage loss is an increasing function of the current wage. Nonetheless, low wage earners are far from insulated to wage losses when switching jobs and even in the lowest quartile, 21 percent of all EE transitions result in nominal wage losses.

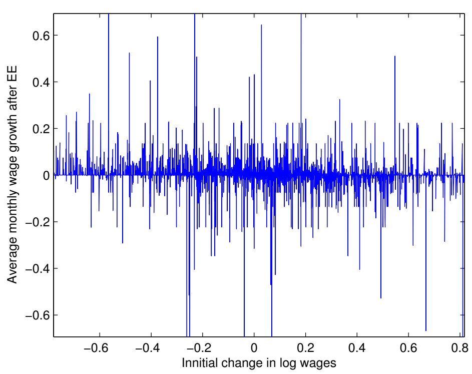
B.2.2 Are Wage Cuts the Result of Optimizing Behavior?

Postel-Vinay and Robin (2002) offer a competing explanation for the occurrence of wage cuts after EE movements. They lay out a model where wages can only be renegotiated by mutual agreement and the firm has all the bargaining power. Wage raises on the job occur as a result of counter-offers to bids by other firms. They demonstrate that in such a framework workers may accept wage cuts upon job to job transitions if the option value of working at the other firm is sufficiently high. Workers will only move to firms more productive than their current employer and very productive firms offer the potential of large future wage gains.

A testable implication of these types of models is that expected future wage growth with the new employer should be an increasing function of the wage cut accepted. As Figure VII demonstrates, it is not borne out by the data. Plotting initial wage changes against mean average wage growth in the ensuing employment spell, there appears to be no systematic relationship whatsoever.

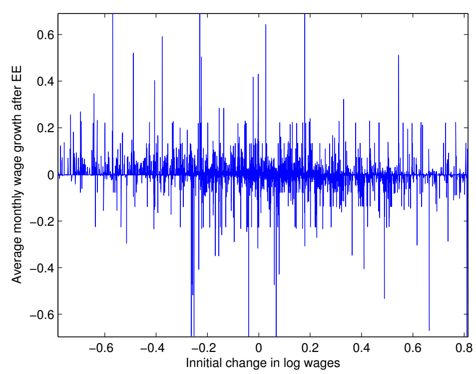
Figure VII: Expected Wage Growth as Function of Initial Wage Change

(A) Nominal Wages



Notes: The Figure plots on the X-axis the initial change in log nominal wages after a job to job movement and on the Y-axis the mean monthly wage growth of the corresponding employment spell. The correlation coefficient between the two statistics is -0.0839.

(B) Real Wages



Notes: The Figure plots on the X-axis the initial change in log real wages after a job to job movement and on the Y-axis the mean monthly wage growth of the corresponding employment spell. The correlation coefficient between the two statistics is -0.0827.

C Consistency of the HKV estimator

In Section 4 we assess the success of our structural model by comparing its results to reduced form estimates of HKV. However, our analysis implies that a part of the frictional wage dispersion estimated in HKV is the result of an upward bias. To understand this point, consider their econometric specification where they exploit the panel dimension of the PSID. They regress log wages on worker observables and recover the residuals of individual i in period t (γ_{it}). Frictional wage dispersion is then measured as $\tilde{w}_{it} = exp(\gamma_{it} - \bar{\gamma}_i)$, where $\bar{\gamma}_i$ is the average residual of worker i. This way the measure controls for both time varying observed and unobserved initial worker heterogeneity. This specification only yields an unbiased estimate, if workers' wages have no stochastic component that is unobserved by the econometrician. However, we estimate $\sigma_\epsilon = 0.0547$, which leads to an upward bias in frictional wage dispersion.

To asses the scale of this problem, we employ their econometric specification on our simulated data set. Table 8 displays the implied frictional wage dispersion, measured at different percentiles of the frictional wage distribution. In the two cases considered, the bias is mild. Additionally, in the case of the 1st percentile, the true degree of wage dispersion is even slightly underpredicted. We therefore conclude that for the variability of wages present in the data, the results from HKV yield a reliable estimate of frictional wage dispersion.

Table 8: Imputing Frictional Wage Dispersion

Percentile	Mm-ratio
Min wage	2.36
1^{st}	1.37

Notes: The table displays the simulated mean-min ratios from our baseline model, measured with the econometric model specified by Hornstein et al. (2007). The minimum wage is measured as the absolute minimum and at the 1st percentile respectively.

D Numerical Algorithm

The numerical algorithm consists of three nested loops and a simulation afterwards.

- We begin the algorithm by guessing a labor market tightness θ.

- Next, we guess the wage function over the states for the worker. Discretize the workers' log productivity by 1500 grid points. We find 7.5 to be a non binding upper bound. Discretize the distribution of log firm productivities by 10 equispaced grid points. The third dimension of the wage function are the two life-cycle states.

- Given the initial guesses, we can start the inner loop, which calculates the value functions using value function iteration. Expectations regarding next period's idiosyncratic productivity are calculated using Gaussian quadrature with 10 nodes for evaluating the productivity innovations and spline interpolation between productivity grid points.

- Taking the value functions of the workers we start the middle loop that updates the wage function. We compute the value of the firm by Nash-Bargaining: $V_x^J(s) = \frac{1-\alpha}{\alpha}(V_x^E(s) - V_x^U(s))$. Again using Gaussian quadrature and spline interpolation gives us the expected value of the firm next period. Using this and the value functions of the workers allows us to compute the policy functions.

- Solving the value of the firm function for wages yields the implied wage schedule for each grid point ($w_{computed}$). Wages are only determined by Nash-bargaining in equilibrium. However, worker heterogeneity implies that in equilibrium there will be certain potential matches whose surplus is negative. In order to be able to compute meaningful values of employment at these firms we set wages equal productivity or, put differently, we set the firm value to zero. Afterwards, we update wages by $w_{new} = \rho w_{initial} + (1-\rho) w_{computed}$ until convergence. ρ is the updating weight and we find 0.75 to work fine at the beginning and increase it to 0.9 towards convergence.

- The last loop computes the implied θ by setting the value of a vacancy to zero. We therefore need the stationary distributions of the employed and the unemployed. We compute these by distribution function iteration, using the policy functions. For the distribution function we a finer grid for worker productivities of 5000 grid points. Using the results update θ until convergence.

- The last step is the simulation, using the policy functions and equilibrium job offer rates. We use linear inter and extrapolation on the worker and firm productivity grid[29]

[29]We opted for linear interpolation at this step, as it considerably decreases the computational burden and does not appear to alter the results compared to spline interpolation. Also, spline extrapolation is known to be unreliable.