DISCUSSION PAPER SERIES

IZA

IZA DP No. 4031

Wage Dispersion and Wage Dynamics Within and Across Firms

Carlos Carrillo-Tudela
Eric Smith

February 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Wage Dispersion and Wage Dynamics Within and Across Firms

Carlos Carrillo-Tudela
University of Leicester
and IZA

Eric Smith
Federal Reserve Bank of Atlanta
and University of Essex

Discussion Paper No. 4031
February 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4031
February 2009

ABSTRACT

Wage Dispersion and Wage Dynamics Within and Across Firms*

This paper examines wage dispersion and wage dynamics in a stock-flow matching economy with on-the-job search. Under stock-flow matching, job seekers immediately become fully informed about the stock of viable vacancies. If only one option is available, monopsony wages result. With more than one firm bidding, Bertrand wages arise. The initial and expected threat of competition determines the evolution of wages and thereby introduces a novel way of understanding wage differences among similar workers. The resulting wage distribution has an interior mode and prominent, well-behaved tails. The model also generates job-to-job transitions with both wage cuts and jumps.

JEL Classification: J31, J63, J64

Keywords: wage dispersion, wage dynamics, job search, stock-flow matching

Corresponding author:

Carlos Carrillo-Tudela
Department of Economics
University of Leicester
Astely Clarke Building
University Road
Leicester, LE1 7RH
United Kingdom
E-mail: cct9@leicester.ac.uk

* We would like to thank Randall Wright, Ken Burdett, Jean Marc Robin for their many helpful insights. This paper has also benefited from comments and suggestions of participants in seminars at the University of Essex, University of Leicester, University of Exeter, Humboldt University (Berlin), University of Basel, University of Konstanz, Rutgers University, University of Notre Dame, Michigan State University, FBR Philadelphia, FBR Atlanta and at the SED (2007), EEA (2007) and the Midwest Macro (2008) conferences. The views expressed are those of the authors and not necessarily of the Federal Reserve Bank of Atlanta or the Federal Reserve System. All errors are our responsibility.

1 Introduction

Competition determines wages and prices. In models of frictionless markets where competition is most fierce, the outcome is straightforward. The law of one price holds. As Mortensen (2003) and others point out, this result is completely at odds with overwhelming empirical evidence. Similar goods and similar workers are simply not paid the same price and the same wage.

In models of markets with search, however, matching frictions limit the extent of competition. Several studies have shown that this restriction can generate equilibrium price dispersion for identical goods and wage dispersion among similar workers. The Burdett and Mortensen (1998) model, the benchmark of this approach in the labour market, directly links wage dispersion to the determinants of labour turnover and wage mobility. Job and worker flows along with wage dispersion thus become key phenomena in making the search framework relevant for labour market analysis.

Although the search literature claims numerous insights and successes, some empirical findings remain elusive and difficult to reconcile. In particular, the benchmark model predicts that the density distribution function for wages is upward sloping, that there are no job-to-job transitions with wage cuts, wages do not increase without outside offers. These predictions are in stark contrast with observed behaviour.

The objective of this paper is to demonstrate that these and other inconsistencies may stem from the underlying specification of search frictions rather than from the general search approach. The conventional approach posits a 'black-box' random matching function to approximate search frictions. This paper adopts an alternative approach, the stock-flow methodology, which offers not only a more rigorous and plausible microfoundation for search frictions but also a more empirically valid picture of matching dynamics.[1]

In a labour market with stock-flow matching, when a seller, i.e. a worker, goes on the market in search of a partner, he or she immediately becomes fully informed about the number of suitable buyers in the stock, i.e. the stock of job vacancies. If lucky, the worker finds several viable options. If the worker is unlucky, the market turns up few or possibly no viable opportunities. In the event that no acceptable vacancies exist in the marketplace, the worker must wait to match from the flow of new jobs.

Consider wage determination in this set-up with on-the-job search.[2] After job search reveals the number of currently available jobs, all suitable firms bid for the worker's services. If only one option is currently available, the firm offers a monopsony payoff that claims all of the gains to trade for the firm. On the other hand, with more than one firm involved, competitive Bertrand bidding occurs. This time, the worker extracts the gains to trade.

[1]The matching framework used here is most closely related to the matching models of Taylor (1995), Coles (1999) and Lagos (2000). Emerging empirical evidence indicates this framework has more validity than random matching. See Coles and Smith (1998), Petrongolo and Pissarides (2001), Andrews, Bradley and Upward, (2001), Gregg and Petrongolo (2005), Coles and Petrongolo, (2008), Kuo and Smith (2009).

[2]Taylor (1995) and Coles and Muthoo (1998) examine wages in this set-up without on-the-job search.

At the outset of the employment relationship, wage dispersion obtains and depends on the number of competitive bidders found at that time.

Now suppose that at any time after a firm and worker pair up, the firm can update its offer. In other words, a new wage is offered in each instant. The worker can either accept the latest offer or go again to the market to elicit bids. The firm updates its wage offer knowing that as time proceeds, firms come and go and the number of prospective bidders in the market evolves randomly. The birth and death of job opportunities generate turnover but the worker and the employer do not directly observe this turnover unless the worker actively engages in on-the-job search. The worker must physically visit the market to learn the actual number of bidders.

This process provides a new source of wage progression with tenure at a firm. Employers who want to avoid bidding with the (anticipated) firms in the market can keep the worker away from the market with a sufficiently high wage offer. Such an offer outbids the evolving threat of on-the-job search, not the actual firms.

No-search wages face two countervailing forces from turnover in the market. Previous bidders gradually leave the market and new options enter the market. Outside options therefore can rise or fall depending on this birth and death process. Wages not only differ at the outset, they also evolve in different patterns. For monopsony wages, the unfortunate history (from the worker's perspective) fades and the outside option improves. Low initial wages rise over time. For competitively bid wages, the more favorable history that led to high initial wages fades and eventually a less attractive expectation of the number of new firms matters more. Although wages start at different points and evolve in different patterns, they ultimately converge with long tenures.

Job availability and turnover jointly determine wage dispersion and wage dynamics. Coupled with the mechanics of job search presented here, these factors also lead to job-to-job transitions with wage cuts and wage rises. Worker job search follows the forced dissolution of a match. The probability of finding a suitable bidder depends on the most recent number of bidders and the duration of the previous match. For many separated workers, especially with those who had short tenures and many previous bidders, there is a high probability of finding at least one match. When a vacancy is available, a job-to-job transition takes place since all information is immediately revealed and acted upon. The wage at this new job will again depend on the realized number of bidders. For some, the realization will exceed the expectations that were pinning down the wage at dismissal. Wages in the new job will rise above the old wage. For others, the realization will be below the anticipated value and the new wages in the job-to-job transition will involve a wage cut. This feature of the model fits recent evidence on wages in job-to-job transitions and the importance of reallocation shocks in explaining labour turnover. See, for example, Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005), Jolivet, Postel-Vinay and Robin (2006) and Nagypal (2008).

Note as well that some workers who lose their jobs are unlucky and do not find a new employer immediately. These workers become unemployed and must wait for new opportunities to arise. Since jobs arrive slowly, monopsony wages follow unemployment. Finding no or only one employment opportunity is most likely for those workers who had short spells of employment with employers who were paying monopsony wages. In other words, workers with the lowest pay are prone to unemployment and wage cuts when a new job is found. Unemployment can thus exhibit state dependence with workers seemingly trapped in low pay - no pay doldrums not due to their abilities but due to their history of finding it difficult to locate suitable employment. This feature of the model is consistent with recent evidence on the joint wage and employment dynamics of low pay workers. See Stewart (2008).

Initial wages and their subsequent progression within a firm combine to create a distribution of wages at a point in time. Although it is difficult to formulate and evaluate an explicit expression for the distribution, numerical methods reveal sensible shapes for a range of parameters. In a homogeneous environment, the cross section of wages is dispersed around an interior mode with prominent tails on both sides. Skewness exists but varies left or right with the underlying fundamentals of the model. The model can also generate reasonable mean-min ratios and thus overcoming the lack of frictional wage dispersion found in standard search models by Hornstein, Krusell and Violante (2007).

The next section describes the general framework and the process governing vacancy turnover. Sections 3 and 4 analyze the worker's and firm's problem and solve for optimal wages. Section 5 and 6 derive wage and employment dynamics. Sections 7 and 8 describe the steady state distribution of wages for homogeneous and heterogeneous workers respectively. Section 9 compares our results with standard sequential search models. The last section concludes.

2 Stock-Flow Dynamics

Consider a continuous time economy with a mass of segregated markets containing homogeneous workers and firms.[3] Both agents are risk neutral, discount the future at rate $r > 0$ and maximize expected lifetime payoffs.

From time to time, firms enter (via a Poisson process) a given market with a project that produces output using a constant returns to scale technology. Let λdt denote the number of firms that enter a market over a time interval of length dt. In all projects, any worker generates a constant flow revenue x per unit of time.

A worker's life is tied to the existence of their particular market. A new market opens with all workers present and exists for exponentially distributed duration with parameter μ. To maintain a constant number of markets, a proportion μdt are created each period. While the given market exists, a worker is either employed or unemployed at any point in time.

[3]Unlike standard matching models (e.g. Pissarides, 2001) we do not endogenize the number of vacancy/firms in the economy. In our framework the number of firms is fixed ex-ante as in Burdett and Mortensen (1998).

Unemployed workers receive payoff $b < x$ per unit of time. Wages of employed workers are described below.

Given the production technology, a firm is willing to hire as many workers as it finds productive for any given project. Workers do not interfere with each other within the firm or in the market.[4] The worker simply cares about the arrival rate of employment opportunities. As a result, it is sufficient to analyze the decisions and outcomes of a representative worker in each market.

Workers find out about the existence of new projects in their market through search. Following the stock-flow matching approach (see Coles and Smith, 1998), information about the availability of buyers (firms) and sellers (workers) in a given market is centralized in a job centre. A firm posts project requirements as vacancies at the job centre as soon as they occur and maintain this listing until the project terminates.

Workers check the posted list from time to time. When the worker checks the list of posted vacancies, there are no frictions or delays in processing the information. All information regarding the viability of a position is immediately made clear and common knowledge. As such, there are no impediments to exchange after the worker finds out about the existence of gains to trade at the job centre.

Given the worker pays a search cost $c > 0$, he or she enters the job centre and immediately observes the number of vacancies in the market. A complete information auction for the worker's services follows given the number of viable job opportunities. Consider an unemployed worker. Three relevant cases can arise in this setting:

- No viable vacancies are posted in which case the worker remains unemployed and chooses another date to re-visit the job centre.
- Only one viable vacancy is available in which case the firm is a monopsonist for this particular worker and does not face competition in the auction. The firm then makes a take-it-or-leave-it offer to the worker such that it extracts the entire gains to trade thereby making the worker indifferent between employment and unemployment.
- Two or more vacancies are viable in which case the firms competitively bid against each other for the worker. Bertrand bidding implies that the worker captures the entire gains to trade.

During employment, the worker can decide to re-visit the job centre in search of a better offer. If on-the-job search occurs, the outcome is common knowledge - both the worker and the firm become informed about the number of available employment opportunities for the worker in the job centre.[5] Given the number of viable opportunities found in the job centre,

[4] See Coles (1999) for an analysis of congestion externalities in a model of stock-flow matching.

[5] Common knowledge rules out the possibility that a worker visits the job centre and calls for an auction only if conditions are favorable. As demonstrated below the firm can infer worker behavior from its wage offer.

a new auction occurs. The new or re-negotiated wage depends on the number of vacancies in the job centre at that moment.

The worker may also visit the job centre following a job separation. A match may end from (endogenously induced) job search or due to an exogenous job destruction shock, which follows a Poisson process with rate δ. In what follows matched pairs will break up only due to exogenous separation shocks. Moreover, for analytical tractability we assume that job destruction occurs when a firm withdraws its vacancies from the job centre.

2.1 Vacancy Turnover

Let $t \geq 0$ denote the duration since the worker last visited the job centre. Suppose at this last visit, the worker left behind $k \geq 0$ job opportunities. More specifically, if the worker came away from this visit to the job centre either unemployed or hired monopsonistically, then $k = 0$. If a competitive auction occurred (an auction with more than one bidder), there were $k + 1$ bidders and $k \geq 1$. One bidder won the auction and the other k remain behind. Note that when a market opens, the worker has no history and no firm has arrived yet so that $t = k = 0$.

Given k, the probability of finding $n \geq 0$ projects or vacancies after a duration t since the last visit to the job centre follows from queuing theory. In particular, this probability can be obtained from the combination of two random variables:[6]

- $RV1$: The number of vacancies that survive until time t, given the number of vacancies that were left in the job centre at the time of last visit.
- $RV2$: The number of new and un-sampled vacancies in the job centre at time t, given an entry and exit process of job opportunities.

Since δ describes the rate at which vacancies in the job centre are destroyed, the properties of the Poisson process imply that $RV1$ follows a Binomial distribution with parameters $\left(k, e^{-\delta t}\right)$, where $e^{-\delta t}$ is the probability that a vacancy has not left the job centre by time t. From the forward Kolmogorov differential equations describing this birth-death process,[7] $RV2$ has a Poisson distribution with parameter

$$\phi(t) = \frac{\lambda}{\delta}(1 - e^{-\delta t}).$$

As $RV1$ and $RV2$ are independently distributed, the probability density of their sum is given by the convolution of their respective density functions. Hence, the probability that there are $n \geq 0$ vacancies at time $t \geq 0$, given $k \geq 0$ vacancies were in the job centre at the

[6] For details see Prabhu (1965).

[7] For a derivation see Lemma 1 in Smith (2007).

time of last visit is described by

$$P_n(t,k) = \sum_{i=0}^{\min\{k,n\}} \binom{k}{i} (e^{-\delta t})^i (1-e^{-\delta t})^{k-i} \left[\frac{\phi(t)^{n-i} e^{-\phi(t)}}{(n-i)!} \right], \tag{1}$$

where the $\min\{.\}$ operator applies since the convolution is not defined for negative factorials. Moreover, as $t \to \infty$ each of these transition probabilities converge to a unique value

$$P_n = \left(\frac{\lambda}{\delta}\right)^n \frac{e^{-\frac{\lambda}{\delta}}}{n!}, \tag{2}$$

that is independent of k. The ergodic distribution of n has mean and variance equal to λ/δ.

3 Worker Payoffs

Worker payoffs critically depend on the timing and expectations of events in the job centre. Let $E(t,k)$ represent the expected utility of an employed worker given spell length t and job opportunities k as described above. Suppose this worker decided to visit the job centre. Once there, the worker calls an auction and compares the resulting payoff to that of unemployment. In the latter case the worker may revisit the job centre at a later stage knowing that at the time of the last visit there were $k+1$ vacancies available. Let $U(t,k)$ denote the expected value of unemployment having observed k vacancies in the job centre t periods ago. Since visiting the job centre at any point in time resets t to zero, the expected payoff to on-the-job search for a worker with pair t,k is thus given by

$$W(t,k) = -c + \sum_{i=0}^{\infty} P_i(t,k) \max\{E(0,i), U(0,i+1)\}.$$

On the other hand, if a worker with pair t,k experienced a separation shock and decided to visit the job centre immediately, the expected payoff of search is instead given by

$$V(t,k) = -c + P_0(t;k)U(0,0) + \sum_{i=1}^{\infty} P_i(t,k) \max\{E(0,i-1), U(0,i)\}.$$

Note that the difference between W and V arises since the current job of an employed worker is always available as a fall back option.

To focus on the employment relationship the following analysis adopts the limiting case in which $c \to 0$. This assumption is useful as it implies that workers visit the job centre immediately after an exogenous job separation. Moreover, in this limiting case any unemployed worker revisits the job centre every period and the expected value of unemployment given that the worker can call an auction with i bidders is then given by

$$(r+\mu)U(0,i) = b + \lambda[\max\{E(0,i), U(0,i+1)\} - U(0,i)] - i\delta[U(0,i) - U(0,i-1)], \quad (3)$$

for all $i \geq 0$.

Now consider an employed worker that visited the job centre t periods ago and observed $k+1$ vacancies such that he/she is currently offered a wage $w(t,k)$. Suppose this worker decides not to search this instant. As $c \to 0$, it follows that

$$E(t,k) = w(t,k)dt + \frac{1-\mu dt}{1+rdt}\left[\delta dt V(t+dt,k) + (1-\delta dt)\max\{E(t+dt,k), W(t+dt,k)\}\right] + O(dt^2).$$

Moreover, if the worker also decides to forgo search in the next instant, $E(t+dt,k) \geq W(t+dt,k)$, we get

$$(r+\mu)E(t,k) = w(t,k) + \dot{E}(t,k) - \delta[E(t,k) - V(t,k)]. \quad (4)$$

The firm can offer a sufficiently high wage such that $E(t,k) \geq W(t,k)$ and the worker chooses or is effectively bribed not to visit the job centre at time t. If it chooses to go this way, the firm would optimally offer the lowest possible wage that satisfies this criteria. Hence, given t,k a worker's no-search wage implies

$$E(t,k) = W(t,k).$$

We now characterize the wage the worker demands in order not to search and not visit the job centre, denoted as $w^w(t,k)$, and then subsequently consider whether the firm is willing to offer this wage.

3.1 Initial Payoffs

As search resets t, the first step is to characterize the value functions at zero durations. If an employer is met during a visit to the job centre, any ensuing employment will not have zero spell length. The worker does not expect discrete changes in market options - prospects do not change quickly enough - so that from (4)

$$(r+\mu)E(0,k) = w^w(0,k) + \dot{E}(0,k) - \delta[E(0,k) - V(0,k)].$$

If the firm finds itself in an auction for the worker, the winning bid will make the worker indifferent between accepting and the next best alternative. When the firm is the only bidder for the worker, the firm only needs to entice the worker out of unemployment. When more than one firm bids for the worker, the winning bid entices the worker away from the competition. Noting that once employed the worker does not loose the option of calling an auction once again and that $x > b$, it follows that $E(0,k) \geq U(0,k+1)$ for all k. Using

$E(t,k) = W(t,k)$ and noting that the above payoffs then imply $V(0,k) = E(0,k-1)$ when $k \geq 1$, $E(0,k)$ can be described by

$$(r+\mu)E(0,k) = w^w(0,k) + \sum_{i=0}^{\infty} \dot{P}_i(0,k)E(0,i) - \delta[E(0,k) - E(0,k-1)], \tag{5}$$

for all $k \geq 1$.

After a visit to the job centre, all information is revealed and the distribution of available firms is degenerate at k. The birth and death process of potential jobs at that instant involves a one step instantaneous transition such that $\dot{P}_i(0,k) = 0$ for all $i \notin \{k-1, k, k+1\}$. By differentiating (1) with respect to t, we obtain that $\dot{P}_{k-1}(0,k) = \delta k$, $\dot{P}_k(0,k) = -\delta k - \lambda$ and $\dot{P}_{k+1}(0,k) = \lambda$, which greatly simplifies (5). Conditional on starting wages, $w^w(0,k)$, initial payoffs are described by the following recursive system of second order difference equations:

$$(r+\mu)U(0,0) = b + \lambda[E(0,0) - U(0,0)], \tag{6}$$

$$(r+\mu)E(0,0) = w^w(0,0) + \lambda[E(0,1) - E(0,0)] - \delta\left[E(0,0) - U(0,0)\right], \tag{7}$$

$$(r+\mu)E(0,k) = w^w(0,k) + \lambda[E(0,k+1) - E(0,k)] - \delta(k+1)[E(0,k) - E(0,k-1)] \tag{8}$$

for $k \geq 1$, where the equation for $U(0,0)$ follows from (3). The boundary condition is that a worker with many firms gets productivity payments indefinitely:

$$\lim_{k\to\infty}(r+\mu)E(0,k) = w^w(0,\infty) = x. \tag{9}$$

3.2 Worker's No-Search Wages

Substituting the expressions for $W(t,k)$ and $V(t,k)$ into $E(t,k)$ and solving for $w^w(t,k)$ gives the following no-search wages for all (t,k),

$$w^w(t,k) = \sum_{i=0}^{\infty}[(r+\mu+\delta)P_i(t,k) - \dot{P}_i(t,k) - \delta P_{i+1}(t,k)]E(0,i) - \delta P_0(t,k)U(0,0), \tag{10}$$

where equation (1) describes $P_i(t,k)$ and equations (6)-(9) describe $E(0,k)$ and $U(0,0)$. Hence, given starting wages $w^w(0,k)$, $w^w(t,k)$ is completely characterised by (10) for all t, k.

4 Firm Payoffs

A wage offer below the no-search threshold $w^w(t,k)$ triggers a visit to the job centre where all information is revealed. If the worker searches and finds another potential employer, the firm finds itself competing in an auction where it can revise its offer. In this case the worker captures the entire expected payoff to matching. If the worker does not find another firm, the worker accepts the existing employer's offer provided it is preferred to unemployment.

For this reason, should the current firm offer a search inducing wage, the firm needs to only offer a wage that makes the worker as well off as unemployed, that is $w^w(0,0)$. If this wage is accepted, the firm captures the match benefits. To assess whether the risks of re-negotiation following the worker's visit to the job centre are worth the reward, we now consider the wage that the firm is willing to pay to avoid uncertainty and keep the worker from engaging in on-the-job search.

Let $\Pi(t,k)$ denote the firm's continuation payoff given spell length t and job opportunities k from employing a worker who does not search on the job. The payoff to inducing the worker to visit the job centre is then given by the expectation of job opportunities, knowing that the firm is always available as a fall back option:

$$S(t,k) = \sum_{i=0}^{\infty} P_i(t,k)\Pi(0,i).$$

Since a competitive auction implies the worker obtains all the gains to trade, $\Pi(0,k) = 0$ for $k \geq 1$. The expected value of inducing the worker to search therefore reduces to

$$S(t,k) = P_0(t,k)\Pi(0,0).$$

That is, the expected payoff of inducing the worker to search is given by the firm's chances of not encountering competition and offering the monopsony wage $w^w(0,0)$.

To derive a firm's continuation payoff recall that when a δ-shock arrives the job is destroyed and the firm obtains zero profits. Further, when a worker receives a μ-shock and leaves the labour market the firm obtains zero profits for this worker's job.[8] Standard dynamic programming arguments then imply if the worker accepts the wage $w(t,k)$ then

$$\Pi(t,k) = [x - w(t,k)]dt + \frac{1-\mu dt}{1+rdt}\left[(1-\delta dt)\max\{\Pi(t+dt,k), S(t+dt,k)\}\right] + O(dt^2),$$

Provided $\Pi(t+dt,k) \geq S(t+dt,k)$ and letting $dt \to 0$, $\Pi(t,k)$ is described by

$$(r+\mu+\delta)\Pi(t,k) = x - w(t,k) + \dot{\Pi}(t,k)$$

for all $k, t \geq 0$.

4.1 Firm's No-Search Wages

As discussed above, firms offer one of only two wages at each point in time. Either the firm pays the lowest wage such that the worker prefers not to go on the market and search for other competing firms or the current firm risks job search and offers the lowest possible wage

[8] This occurs in the sequential search literature. Moreover, this outcome is consistent with the free entry condition for vacancy creation found in matching models in which the expected value of a vacancy is equal to zero.

that the worker is willing to accept. Letting $\Pi(t,k) = S(t,k)$ gives the wage that makes a firm indifferent from keeping the worker or inducing him to search:

$$w^f(t,k) = x - \Big[(r+\mu+\delta)P_0(t,k) - \dot{P}_0(t,k)\Big]\,\Pi(0,0) \tag{11}$$

for all $k, t \geq 0$. For the firm to actually make this (or some lower wage offer), it must be the case that the worker decides not to search at the offered wage. What the firm is willing to pay in equation (11) must be sufficient to keep the worker away from search as described in (10), that is $w^w(t,k) \leq w^f(t,k)$. Otherwise, the firm prefers to induce the worker to search.

Let $Z(0,i)$ denote the sum of worker and firm expected payoffs in a match when search occurs and the worker leaves $i = n-1$ unmatched viable opportunities in the job centre at the moment of the pairing. Define expected combined payoffs in the next worker-firm pairing at t by

$$\bar{Z}(t,k) = \mathbb{E}_i Z(0,i) = \sum_{i=0}^{\infty}[E(0,i) + \Pi(0,i)]P_i(t,k).$$

If a visit to the job centre takes place, any firm existing in the market, including the current firm, can wind up with the worker in the auction occurring after search, that is in the next match. As the current firm can always re-hire the worker for any given k (recall firms are homogeneous), the expected gains to trade with the current firm immediately after search must equal the expected gains to trade with any other firm in the market.

Given there is a positive cost of visiting the job centre, it is efficient for the worker and the firm to save the search cost and split $\bar{Z}(t,k)$ within the current employment match at any given t. As $c \to 0$, these match specific rents dissipate and $\bar{Z}(t,k)$ also comes to equal the expected payoffs in a match before search. Moreover, at $w^f(t,k)$ the worker is by construction receiving

$$\bar{Z}(t,k) - S(t,k) = \bar{Z}(t,k) - P_0(t,k)\Pi(0,0)$$

in the current match. Since $\bar{Z}(t,k) - P_0(t,k)\Pi(0,0)$ describes the worker's expected payoff from search, he is indifferent between search and accepting $w^f(t,k)$. Hence, we have established the following result.

Proposition 1: $w^w(t,k) = w^f(t,k)$ *for all* t *and* k.

In this economy, on-the-job search is a wasteful, rent seeking activity. Once a match is formed, search does not generate any further gains to trade or match specific rents. The participation constraints of both agents bind at the same wage. The market does not fundamentally change when the worker visits the job centre. There are no new opportunities generated by a visit - existing opportunities are merely realized. Search does not change the expected gains to trade at any given point in time, it just reallocates the division of these benefits. Since workers and firms share the same risk neutral, intratemporal preferences, and since all firms are identical, there is no potential role for meaningful on-the-job search.

5 Within Firm Wage Dynamics

Given Proposition 1, we can circumvent directly solving the system of equations describing $E(0,k)$ and $U(0,0)$ in (10) and instead simply look at wage dynamics and wage distributions using (11). We start by analyzing how wages evolve within a firm for any given k.

A firm that does not face any competition in the job centre offers the worker an expected payoff $E(0,0) = U(0,1)$ such that it makes the worker indifferent from accepting employment or staying unemployed. The firm extracts the entire gains from the match. Using (11), (7) and (3) it follows that with such an offer the firm's expected payoff

$$\Pi(0,0) = \frac{x-b}{r+\mu+\delta+\lambda}.$$

The firm obtains the present value of the gains from forming a match, appropriately discounted by the interest rate, the rate at which jobs are destroyed and the rate at which the firm stops being a monopsonist in the job centre. From equation (1), we also obtain

$$P_0(t,k) = (1-e^{-\delta t})^k e^{-\phi(t)} = \left(\frac{\delta}{\lambda}\right)^k \phi(t)^k e^{-\phi(t)}$$

$$\dot{P}_0(t,k) = P_0(t,k)\frac{\phi'(t)}{\phi(t)}[k-\phi(t)]. \tag{12}$$

Substituting out these expressions in (11) gives the set of wages that makes the worker and firm indifferent from visiting the job centre for all t, k:

$$w(t,k) = x - \left(\frac{\delta}{\lambda}\right)^k \phi(t)^k e^{-\phi(t)} \left[(r+\mu+\delta) - \frac{\phi'(t)}{\phi(t)}[k-\phi(t)]\right] \frac{x-b}{r+\mu+\delta+\lambda}. \tag{13}$$

In the absence of voluntary quits or layoffs, t also denotes the elapsed time since the match was formed whereas equation (13) describes the wages firms pay their workers over this duration given that there were k other firms at the hiring stage competing for the worker.

At any time t, wages are inversely related to the probability the current employer is the only one that can use the worker's services and positively related to the change of this probability during the next instant. The worker is prepared to accept a lower wage and avoid re-negotiating the terms of employment when there is a higher probability that the employer can become a monopsonist. However, if this probability is higher in the next instant, ceteris paribus, the firm prefers a visit to the job centre to take place during the next instant rather than during the current one. In this case, the worker requires a higher wage today in compensation for any potential loss. Although Proposition 1 guarantees no voluntary visits to the job centre take place, optimal no-search wages at any t for any given k are determined by the trade-off between these two opposing effects.

5.1 Initial wages

When a separation shock occurs, the current job ends, the worker re-visits the job centre and observes n available job opportunities with probability $P_n(t,k)$. If the worker does not find at least one viable job opportunity in the job centre ($n = 0$), unemployment occurs, and the worker gets payoff $U(0,0)$. If an employment opportunity exists ($n \geq 1$), the worker immediately becomes employed with wages given by $w(0, n-1)$.

If the worker finds exactly one job opportunity, he or she immediately takes it and gets payoff $E(0,0)$. The firm offers the minimum amount necessary to entice the worker into employment. Equating (3) and (7) the starting wage offered by the firm is

$$w(0,0) = b.$$

Since the worker does not loose the option of calling an auction once employed and job destruction occurs at the same rate as vacancies leave the job centre, $E(0,0) = U(0,1)$ implies the firm needs to offer a starting wage of b to employ the worker.

Some displaced workers will be fortunate and find that more than one job have accumulated in the market. If the worker finds exactly two job opportunities, $n = 2$, the starting wage is

$$w(0,1) = x + \delta \frac{x-b}{r+\mu+\delta+\lambda},$$

while for $n \geq 3$, the starting wages are

$$w(0, n-1) = x.$$

Interestingly, $w(0,1) > w(0,k)$ for $k \geq 2$. Immediately after the worker is employed and k firms are left behind, only two events can happen in the job centre. One of the sampled vacancies is destroyed or a new vacancy arrives. Suppose $n = 2$ so that $k = 1$. The current employer becomes a would-be monopsonist in the next instant of duration dt with probability δdt. In this case, inducing the worker to visit the job centre, the employer obtains $\Pi(0,0)$ whereas the worker gets $U(0,1)$. The worker extracts this potential loss at the start of the employment relationship in the form of an initial wage above productivity.

In contrast, when $k = n - 1 \geq 2$, there is no threat that the employer could be a monopsonist immediately after employment begins. If one of the vacancies left behind is destroyed, the current employer does not improve its prospects of being a monopsonist. In the immediate instant after hiring, with probability one there is at least one other vacancy willing to compete for the worker. In this case, Bertrand competition implies the worker receives a starting wage equal to productivity.

5.2 Wages over time

Now consider the evolution of wages over the duration of a match. When $k = 0$, equation (12) implies that the probability that the current employer remains a monopsonist decreases over time. The outside option of the worker therefore improves and the firm must increase wages to prevent the worker from visiting the job centre. Differentiation of (13) for $k = 0$ further establishes that

$$\dot{w}(t,0) = e^{-\phi(t)} \left[\phi'(t)(r+\mu+\delta+\phi'(t)) - \phi''(t)\right] \frac{x-b}{r+\mu+\delta+\lambda} > 0, \tag{14}$$

for all t since $\phi''(t) < 0$.

On the other hand, for any $k \geq 1$, equation (12) implies that the probability the current employer becomes a monopsonist increases over time. When $k = 1$, the eventual destruction of the vacancy left behind in the job centre outweighs the effect of new vacancy turnover on expected competition. The worker's outside option and hence $w(t,1)$ decreases with t, as the threat of any potential loss brought about by an induced visit to the job centre increases over time. Differentiating (13) for $k = 1$ shows that

$$\dot{w}(t,1) = -\delta e^{-(\delta t+\phi(t))} \left[(r+\mu+2\delta+\phi'(t))(1-\phi(t)) + \phi'(t)\right] \frac{x-b}{r+\mu+\delta+\lambda},$$

which is negative for all t if $\lambda = \delta$.

Now suppose $k \geq 2$. Differentiating (13) establishes that

$$\dot{w}(t,k) = \left(\frac{\delta}{\lambda}\right)^k \phi(t)^{k-2} e^{-\phi(t)} \phi'(t) \left[\phi'(t)[[k-\phi(t)]^2 - k] - \phi(t)[k-\phi(t)](r+\mu+2\delta)\right] \frac{x-b}{r+\mu+\delta+\lambda}. \tag{15}$$

In this case, wages are not monotonic but exhibit a "humped" shape progression, increasing and then decreasing as the employment relation evolves over time. This feature can be further characterized for the special case where the arrival of projects equals the job destruction rate in the job centre.

Proposition 2: *For $\lambda = \delta$, $w(t,k)$ is strictly quasi-concave with respect to t for all $k \geq 2$.*

Proof: See Appendix

Quasi-concavity in this special case follows from the relative importance that $P_0(t,k)$ and $\dot{P}_0(t,k)$ have on (13) as t increases. As mentioned earlier, the former strictly increases with t. However, the rate of change of this probability has a greater impact at early stages of the employment relation. The firm must offer higher wages to keep the worker from visiting the job centre early on so that it can later reap the benefits of a higher $P_0(t,k)$. Eventually $P_0(t,k)$ settles towards its long run value. The impact of $\dot{P}_0(t,k)$ diminishes and the worker accepts lower wages to avoid a firm-induced visit to the job centre.

Although wages start and evolve very differently as the employment spell becomes long, all wages limit to the same value:

$$\lim_{t\to\infty} w(t,k) = x(1 - e^{-\lambda/\delta}) < x. \tag{16}$$

For any initial k, the birth and death process governing vacancy turnover implies that the distribution of n converges to a unique distribution with a mean and variance equal to λ/δ as $t \to \infty$. History fades over time so that eventually all workers face the same prospects in the job centre. Since the effect of k is only transitory, wages converge to a unique wage, $w(\infty)$ which increases with the expected number of vacancies in the job centre and converges to x as $\lambda \to \infty$.

5.3 Numerical examples

To further gauge the properties of the wage profile, we calibrate the model to match salient features of the US economy. The time period is one month. For expositional ease we focus on $k \in \{0, 1, 2, 5, 10, 20\}$.[9] Following standard practice, $r = 0.0041$. We normalize the value of productivity to $x = 1$ and let $b = 0.4$. Following Shimer (2005) $\lambda = 0.45$, which yields an average unemployment duration of 2.2 months. We set $\delta = 0.02$ to generate an average job tenure of 50 months consistent with Hornstein, Krusell and Violante (2007).

This parameterization generates a job-to-job transition rate consistent with the one reported by Nagypal (2008).[10] On the other hand, these parameters imply that the probability after 50 months that a worker becomes unemployed following a separation is essentially zero. For values of μ based on average working lives, the model is unable to generate sufficient employment to unemployment transitions to be consistent with the data.[11] The unemployment rate is too low. This limitation results from the constant returns to scale production technology along with the matching mechanics. Workers do not crowd out each other, that is turn over jobs, when they become employed. Job turnover in the job centre is the same as the destruction rate of occupied jobs. This assumption seems natural in the context of the model, but unemployment is unlikely given the above parameters. After a sufficient length of time, jobs are likely to exist in a given market.

In contrast, the standard matching model adopts complete crowding out by assuming one vacancy corresponds to one job and hence the turnover of vacancies depends directly on the hiring of workers, not the job destruction process. Allowing for different vacancy and job turnover rates in this paper considerably complicates the analysis. The life of a market, however, may not correspond to a worker's entire period in the labor force. Individuals can from time to time find themselves displaced into unemployment and switching careers. As

[9] Numerical simulations with more values of k yield the same pattern as shown here.

[10] Nagypal (2008) estimates that the probability an employed worker experiences a job-to-job transition in any given month is approximately 0.022.

[11] Consider a worker with $k = 0$ and $t = 1/\delta = 50$. Since $P_0(t,k)$ is decreasing in k for all t, $P_0(1/\delta, 0)$ gives an upper bound for the probability a worker becomes unemployed at 50 months following a separation shock. Using the parameter values described above, $P_0(1/\delta, 0) = e^{-[\lambda/\delta(1-e^{-1})]} = 0.0000006655$.

such, we set $\mu = 0.02$ to match the average separation rate into unemployment reported in Hornstein, Krusell and Violante (2007).[12] Using the unemployment rate from (18) derived below, we obtain an average unemployment rate of 4.3%.[13]

Figure 1.a illustrates the wage dynamics for this parameterization. As implied by (14), workers hired from unemployment ($k = 0$) are offered a wage profile that increases with tenure. In this case, returns to tenure occur mostly within the first year of employment. For workers that were able to call a competitive auction ($k \geq 1$), wages are essentially flat and equal to their productivity. In this case returns to tenure are nearly zero. This follows from the degree of expected competition implied by the calibration. In particular, the ergodic distribution of $P_n(t, k)$ described by (2) implies an average number of vacancies in the job centre of 22. Effectively, the probability that the current employer becomes a monopsonist in the future is zero and $w(\infty) = 1$. In turn, (13) implies wages converge to this long run value relatively quickly.

Figures 1.b and 1.c depict the wage dynamics for lower degrees of expected competition. In particular, Figure 1.b considers a $\delta = 0.2$. In this case (2) implies an average of 2 vacancies in the job centre. Figure 1.c then considers the case in which $\delta = \lambda = 0.45$. These figures clearly show the impact of vacancy turnover and expected competition on the shape of the wage profiles. As established and depicted in Figure 1.c, $\delta = \lambda$ yields wages that are strictly increasing for $k = 0$, strictly decreasing for $k = 1$ and follow a humped, quasi-concave shape at early stages of the employment relation for $k \geq 2$. When $k = 2$, equation (15) implies $\dot{w}(0, 2) > 0$ and wages increase immediately after the match is formed. When $k > 2$, (15) implies $\dot{w}(0, k) = 0$ and wages stay closer to x for a longer period of time.[14] In Figure 1.b, on the other hand, wages resemble a combination of the patterns depicted in Figure 1.a and 1.c. For $k = 1$ wages fall below $w(\infty)$ for some period of time before rising back again to achieve their stationary value. When $k \geq 2$, (13), (15) and (16) together imply that wages become flatter and converge faster to x.

6 Individual Wages and Employment Dynamics

Although on-the-job search is practically free, Proposition 1 establishes that employed workers will not visit the job centre until their current match is destroyed. As a result, job-to-job and employment-to-unemployment transitions occur only through involuntary reallocations arising from separation shocks. A worker who visits the job centre and encounters zero bidders becomes unemployed. A realization of $n \geq 1$, leads to immediate re-employment and an apparent job-to-job transition.

[12] Generating employment to unemployment flows in this way is consistent with many sequential search models. Moreover, the shapes of the wage profiles shown in Figure 1 do not change if instead we set $\mu = 0.00185$ to match an average working lifetime of 45 years.

[13] The MATLAB codes are available from the following web site: http://www.le.ac.uk/ec/cct9

[14] This property follows from $\phi(t)^{k-2}$ in (15). The rate at which wages grow at initial stages of the employment relation is inversely related to k.

Recent empirical evidence supports the emphasis here on exogenous separation shocks as a source of employment transitions. Fallick and Fleischman (2004), Nagypal (2005, 2008) report that only a small fraction (less than 5 percent) of employed workers are actively searching. Jolivet, Postel-Vinay and Robin (2006) find that "relative to involuntary mobility (reallocation shocks and lay-offs), voluntary mobility is a rather rare event" in many European countries and in the US. Shimer (2007) and Ebrahimy and Shimer (2006) demonstrate that a stock-flow matching model with reallocation shocks is able to replicate cyclical patterns of unemployment and vacancy time-series in the US economy.

Changes in wages accompany employment transitions. The immediate impact and any subsequent wage growth following an employment transition depends on workers' characteristics as embodied in x and b, worker histories as given by t and k, and realized turnover in the job centre. Transitions into unemployment, for example, always cause post displacement earnings losses. A worker unable to obtain at least one suitable vacancy that fits his or her skill set becomes unemployed. Unemployed workers waiting in the job centre find employment only with monopsonistic firms and hence begin employment at $w(0,0)$. Wages subsequently increase over time, evolving according to $w(t,0)$ until a new separation shock arrives.[15] Moreover, since (1) implies $P_0(t,k)$ is decreasing in k for all $t \in (0,\infty)$, a worker starting employment with $k=0$ will experience a higher probability of becoming unemployed when a new separation shock arrives.

Two implications about wage and employment dynamics of workers that experience unemployment spells arise. The first implication is that displacement into unemployment generates a transitory effect on wages. There is an initial re-employment earning loss that fades out with employment duration. As shown by Figure 1, the speed of this recovery or the persistence of a displacement shock on earnings is govern by the turnover rates of vacancies in the job centre.[16] Second, a current spell of unemployment increases the probability of experiencing unemployment in the future. Since $P_0(t,0)$ is decreasing in t and λ, the probability of becoming unemployed in the future is negatively related to the worker's employment duration and to the turnover of vacancies in the job centre. The overall effect of displacement on workers' earnings and employment status histories depend on the rate at which separations occur relative to the inflow of new vacancies. Since $w(t,0) \leq w(t,k)$, it follows that for sufficiently small values of λ/δ low pay and no pay outcomes can become tied closely together in workers' labour market histories. Low wage workers are more likely to become unemployed and do not have enough time to recover from the initial displacement

[15] If unemployed workers did not search continuously, some would find more than one job available at the job centre in which case their wages would be considerably higher.

[16] There is a sizable literature that analyses the long-term earning losses of displaced workers. Focusing primarily on the US economy, this literature argues that displacement has permanent effects on earnings. See Jacobson, LaLonde and Sullivan (1993) for an early example. Bender and Von Wachter (2006) argue that these studies overstate the earnings losses as these studies do not consider firm characteristics. Using matched employer-employee data from Germany, they find that initial post displacement earnings losses of young workers fade after 5 years, a finding in line with our prediction that re-employment earnings losses fade over time.

shock. For larger values of λ/δ, the effect of a displacement shock will tend to dissipate relatively fast. By the time a new displacement shock arrives, a larger λ/δ implies that these workers will be more likely to be immediately re-employed.[17]

Now consider a separation shock that leads to a reallocation into another job. In this case, the initial wage in the next match can be lower or higher than the wage earned before separation. Starting from low monopsony wages, $w(t, 0)$, job-to-job transitions will involve an immediate wage increase provided a realization of $n \geq 2$. An immediate wage decrease follows when $n = 1$.[18] Job-to-job wage transitions following on from employment with more competitively determined wages can also rise or fall. Given the wage progressions described earlier for $k \geq 1$, job-to-job transitions can involve lower wages for not only workers with any k who observe $n = 1$, but also for workers who observe $n \geq 2$. Similarly, wage increases following a separation shock can rise provided $n \geq 2$.

Empirical evidence supports the proposed wage effect of separation shocks. Involuntary separations that lead to job transitions with both wage increases and decreases are widespread in many labour markets. Jolivet, Postel-Vinay and Robin (2006) find that "a very substantial share (between 25% and 40%, with substantial variation across countries) of job-to-job transitions are associated with wage cuts" whereas Connolly and Gottshack (2008) show that taking a pay cut when switching jobs does not necessarily imply higher wage growth in the future. That is, accepting a wage cut does not always improve a worker's position in the earning distribution, but reflects an adverse shock.

7 Wage Distribution

Wages vary over time and across individuals. To calculate the steady state wage distribution, we first derive the distribution of workers for all t and k. In what follows we normalize the mass of workers to one. In a steady state, the total number of employed workers who observed k vacancies in the job centre during their last visit, $N(k)$, as well as the total number of unemployed and employed workers are all constant over time.[19]

Let u denote the steady state number of unemployed workers and let $n(t, k)$ denote the proportion of employed workers that left k vacancies in the job centre and currently have an employment spell t. Recall that a Poisson process with parameter μ governs the birth and

[17] Stewart (2008) finds evidence for low pay no pay cycles in the UK. Jolivet, Postel-Vinay Robin (2006) suggest that the turnover rate, λ/δ, in the UK is four times smaller than in the US, whereas Heckman and Borjas (1980) and Lynch (1989) do not find evidence of state dependence in unemployment occurrence for the US. The predictions of our model seem to find empirical support.

[18] Wage cuts associated with job-to-job transitions are a somewhat difficult feature to obtain from traditional equilibrium search models. Two notable exceptions are Postel-Vinay and Robin (2002) and Burdett and Coles (2007).

[19] Analyzing the economy in a steady state does not imply assuming that wages are at their long tenure value. These two are fundamentally different concepts.

death process of markets and hence workers. Since $n(t,k) = n(0,k)e^{-(\mu+\delta)t}$, it follows that

$$N(t,k) = (1-u)\int_0^t n(\tau,k)d\tau = (1-u)\frac{n(0,k)}{\mu+\delta}(1-e^{-(\mu+\delta)t}), \tag{17}$$

denotes the number of employed workers that left k vacancies and have an employment spell no greater than t. By construction, in steady state $N(\infty,k) = N(k)$.

Stationarity implies that the inflow of workers into $N(k)$ equals the outflow. Since outflow is composed of those workers of type k that experienced separation and observed $s \neq k+1$ suitable vacancies, the number of workers that leave $N(k)$ is

$$(1-u)n(0,k)\int_0^\infty (\mu + \delta[1 - P_{k+1}(t,k)])e^{-(\mu+\delta)t}dt.$$

On the other hand, the number of workers that enter $N(k)$ is given by all those workers of type $s \neq k$ that experienced a separation and found $k+1$ suitable vacancies:

$$(1-u)\sum_{s\neq k} n(0,s)\left[\int_0^\infty \delta P_{k+1}(t,s)e^{-(\mu+\delta)t}dt\right].$$

Equating the last two expressions establishes a system of k simultaneous linear equations,

$$n(0,k) = \left[1 - \int_0^\infty \delta P_{k+1}(t,s)e^{-(\mu+\delta)t}dt\right]^{-1}\sum_{s\neq k} n(0,s)\int_0^\infty \delta P_{k+1}(t,s)e^{-(\mu+\delta)t}dt$$

for all $k \geq 0$.

The number of unemployed workers in steady state is then given by

$$\lambda u = (1-u)\sum_{s=0}^{\infty} n(0,s)\left[\int_0^\infty [\mu + \delta P_0(t,s)]e^{-(\mu+\delta)t}dt\right].$$

The inflow into unemployment comes from those workers in $N(k)$ that experienced displacement and did not find a suitable job, whereas the outflow from unemployment consists of those workers that found a suitable vacancy with probability λ. Solving for u gives

$$u = \frac{\sum_{s=0}^{\infty} n(0,s)\left[\int_0^\infty [\mu + \delta P_0(t,s)]e^{-(\mu+\delta)t}dt\right]}{\lambda + \sum_{s=0}^{\infty} n(0,s)\left[\int_0^\infty [\mu + \delta P_0(t,s)]e^{-(\mu+\delta)t}dt\right]}. \tag{18}$$

Using the solutions for $n(0,k)$ for all k and (17) gives the solutions for $N(k)$ for all k.

Turning to the steady state wage distribution, let $G(w \mid k)$ describe the cumulative distribution of workers with characteristic k and any t earning a wage no greater than w. All employed (t,k) workers, $(1-u)n(t,k)$, earn wage $w(t,k)$ as described by (13). Defining $I(w(t,k) \leq w)$ as an indicator function[20] that takes the value of one if $w(t,k) \leq w$ and zero

[20]The indicator function is needed as wages are not necessarily monotone over t.

otherwise, it follows that

$$\begin{aligned} G(w \mid k) &= \frac{1-u}{N(k)} \int_0^{\infty} n(t,k) I(w(t,k) \leq w) dt \\ &= (\mu+\delta) \int_0^{\infty} e^{-(\mu+\delta)t} I(w(t,k) \leq w) dt, \end{aligned}$$

$G(. \mid k)$ is constructed by comparing w with $w(t,k)$ for each t and summing across workers for which the indicator function is one.

Aggregating across k gives the unconditional steady state distribution of wages of the economy,

$$\begin{aligned} G(w) &= \frac{1}{1-u} \sum_{k=0}^{\infty} G(w \mid k) N(k) \\ &= \frac{1}{\mu+\delta} \sum_{k=0}^{\infty} G(w \mid k) n(0,k), \end{aligned} \tag{19}$$

where $N(k)/(1-u)$ denotes the proportion of employed workers that left k vacancies in their last visit to the job centre.

Figure 2.a, 2.b and 2.c show kernel estimates of the earnings density, $dG(w)$, from simulated data of workers' employment histories with the parameter values described earlier.[21] The three densities shown here all exhibit an elusive property for search models with homogeneous agents - an interior mode with decreasing (for the most part) tails.[22] As established earlier, wages tend to converge to $w(\infty)$ relatively quickly, especially for those workers who observed a large number of vacancies in the job centre during their last visit. As a result, the model lies close to long term wages, $w(\infty)$. In addition, since workers face only three starting wages, we observe positive density estimates around those points.

The figures further illustrate that as δ increases, the density becomes more skewed to the right. Increasing the destruction rate of matches implies that workers visit the job centre more often during their working lives. Since the initial k has only a transitory effect and the process governing vacancy turnover reaches its steady state relatively fast, when a visit to the job centre takes place the worker is more likely to observe only one vacancy that suits his or her skills. In turn, this generates a higher probability of observing a worker with $k = 0$ and starting wage $w(0,0)$ and, hence, a greater mass in the left tail of the wage density. It also raises the relative probability of observing a worker with $k = 1$ relative to a worker with $k > 1$ making the spike in the right tail more prominent.

These wage densities can also be used to assess the amount of frictional wage disper-

[21] We simulated the data considering a sample of 10,000 workers. The vacancy turnover parameters are set to $\lambda = 0.45$ and $\delta = \{0.02, 0.2, 0.45\}$, and r, μ, x and b are as before.

[22] For example, in the Burdett and Mortensen (1998) model, where employed and unemployed workers search randomly for jobs and offers arrive sequentially, the wage density is strictly increasing and convex over the entire support. It is well known that adding firm heterogeneity to the Burdett and Mortensen's model produces a unimodal wage density with decreasing right tail.

sion, as defined by the dispersion that cannot be accounted for by workers' observable and unobservable characteristics. Hornstein, Krusell and Violante (2007) use the ratio between the average wage to the minimum observed wage (or reservation wage) to measure this type of wage dispersion. They show that standard search models generate a mean-min (Mm) ratio of 1.036, which is at least twenty time smaller than the observed Mm ratio in the US. In our model, the respective Mm ratios generated by our model are 2.45, 2.24 and 1.89.[23] As wages converge to $w(\infty)$ relatively fast and $w(\infty)$ increases with the rate of vacancy turnover, λ/δ, in our model the Mm ratio becomes positively correlated with the amount of expected competition in the job centre.[24]

8 Worker Heterogeneity

This section extends the model to include worker heterogeneity using the following structure. Let α denote a worker's ability and assume that the worker has productivity αx per unit of time across all projects he or she can perform. When unemployed this worker receives a payoff αb per unit, where $b \geq 0$ can be interpreted as productivity at home. Note that α represents the amount of efficiency units of labour the worker supplies per unit of time both at home production and at a firm. Let $H(.)$ denote cumulative distribution of abilities across the population of workers with support $[\underline{\alpha}, \overline{\alpha}]$. H is continuous and has density $h > 0$.

The expected gains to trade between a worker and firm now depend on $\alpha(x - b)$ rather than on $x - b$. Since the worker is equally productive at all firms, a visit to the job centre does not fundamentally change the market. The logic leading up to Proposition 1 continues to hold. As in the homogeneous case, search is a wasteful activity. No-search wages for workers of type α are then given by

$$w(t, k; \alpha) = \alpha x - \left(\frac{\delta}{\lambda}\right)^k \phi(t)^k e^{-\phi(t)} \left[(r + \mu + \delta) - \frac{\phi'(t)}{\phi(t)}[k - \phi(t)]\right] \frac{\alpha(x - b)}{r + \mu + \delta + \lambda},$$

for all t, k and $\alpha \in [\underline{\alpha}, \overline{\alpha}]$. Wages described by $w(t, k; \alpha)$ have the same properties as $w(t, k)$. Since the same evolution of earnings applies, the difference is one of levels.

In the Appendix we show that the distribution of wages for each worker's type, $G(w; \alpha)$ is given by

$$G(w; \alpha) = \sum_{k=0}^{\infty} \left[\int_0^{\infty} e^{-(\mu+\delta)t} I(w(t, k) \leq w) dt\right] n(0, k; \alpha),$$

[23] Hornstein, Krusell and Violante (2007) find that the estimated average Mm ratios range from 1.46 to 1.98, depending on the data used. Our simulations produce slightly higher values. Increasing the value of b generates smaller Mm ratios. For example setting $b = 0.71$ as in Hall and Milgrom (2007), the benchmark calibration generates $Mm = 1.41$.

[24] When $\delta = 0.45$, long term wages converge to $w(\infty) = 0.632$; while the average wage paid is $E(w) = 0.756$. When $\delta = 0.2$, $w(\infty) = 0.895$ and $E(w) = 0.896$; and when $\delta = 0.02$, $w(\infty) = 1$ and $E(w) = 0.98$. Given that $\lambda = 0.45$ and $w(0, 0) = 0.4$, the Mm ratio decreases with δ. Simulations show, however, that the variance and the coefficient of variation increase with δ and under these measures, frictional wage dispersion decreases with expected competition in the job centre.

where $n(t, k; \alpha)$ denotes the proportion of employed workers of type α that observed k vacancies and have an employment spell of $t \geq 0$. The aggregate wage distribution is in turn given by

$$G(w) = \int_{\underline{\alpha}}^{\overline{\alpha}} G(w; \alpha) h(\alpha) d\alpha.$$

We again simulate workers' employment histories and use a kernel estimator to obtain the wage distribution of the economy. Assuming that the distribution of workers' abilities follows a Gamma distribution with shape and scale parameter of 6 and 3.3, respectively,[25] Figures 3.a, 3.b and 3.c replicate the estimates using the same parameter values as in the homogeneous worker case. The wage densities have the same general properties as before. Since wages paid are directly linked to a worker's ability, however, introducing worker heterogeneity implies that the wage density closely follows the shape of the ability density. In this sense, the model with worker heterogeneity can accommodate a wage density that is more in-line with the unconditional empirical wage density.

9 Discussion

To gain further insight on wage determination under stock-flow matching we compare our wage mechanism to those found in two prominent models with on-the-job search, Burdett and Mortensen (1998) and Postel-Vinay and Robin (2002).[26] In the former model firms have incomplete information about worker reservation wages. Firms post a single wage which they maintain for as long as the worker remains employed in the firm. With homogeneous agents the unique equilibrium yields a continuous wage offer distribution and inefficient job-to-job transitions. In Postel-Vinay and Robin (2002), firms have complete information about their applicant reservation wages and are prepared to enter into a Bertrand bidding game if an employee receives an outside offer. Firms optimally hire workers at their minimum acceptable wage and renegotiate the wage every time the worker receives a favorable outside offer. As in Burdett and Mortensen (1998), once the worker is paid a new wage, the firm is not allowed to renege on this wage even though the firm has an incentive to do so. With homogeneous agents, wages increase to marginal productivity as soon as the worker receive an outside offer and the worker is indifferent between employment opportunities.

Our model is closer to Postel-Vinay and Robin (2002). Our workers can also renegotiate their wages and are also indifferent between employment opportunities. The first important difference, however, is that our wage determination mechanism takes into account the incentives of both agents to renegotiate the wage at each point in time given their beliefs about the degree of competition in the job centre. Our workers do not have to wait until an offer

[25] Given the values of the shape and scale parameters the shape of h roughly resembles the shape of the distribution of ex-ante worker heterogeneity estimated by Postel-Vinay and Robin (2002).

[26] In these models by assumption workers always search on the job. However, they allow variable search intensity. See Mortensen (2003).

arrives to renegotiate. They can at any time re-visit the job centre and induce an auction between their current employer and the available vacancies found in job centre. At the same time, firms can also induce renegotiation by offering the worker $w(0,0)$.

As in Postel-Vinay and Robin (2002), firms extract the entire gains to trade from the unemployed by offering a wage that makes them indifferent between accepting employment and continued search. Unlike Postel-Vinay and Robin, wages here subsequently increase smoothly with tenure. The firm increases its wage offer to outbid the evolving threat of on-the-job search. Calibration suggests that when δ is sufficiently low, wages paid to workers hired from unemployment converge to marginal productivity and stay at that level in subsequent jobs.

More generally, our model predicts that there is downward as well as upward wage mobility within a firm. Models that allow a worker and a firm to credibly renegotiate every time their outside options change also lead to this type of wage mobility within a firm. For example, Postel-Vinay and Turon (2008) proposed a modified version of the Postel-Vinay and Robin model that exhibits upward and downward wage mobility. In their framework the existence of match-specific rents implies wages stay constant over a period of time before increasing or decreasing. In contrast, as we let $c \to 0$, any match-specific rents between the firm and the employed worker dissipate. Wages are renegotiated at every point in time because outside options are continuously changing.

Although empirical evidence suggests that such wage fluctuations are not present when analyzing workers' average returns to tenure, evidence on the earnings process of workers seem to suggest the presence of temporary shocks. In particular, McCurdy (1982), Abowd and Card (1989) and Meghir and Pistaferri (2004), among others, have documented the presence of permanent and transitory shocks by fitting ARMA or ARCH models to earnings data. It is usually assumed that these shocks arise from changes in worker productivity. Our model suggests that temporary wage variation at early stages of the employment relationship can be related to the interaction between reallocation shocks and the degree of competition in the labour market at the moment at which the worker is hired. Moreover, with worker heterogeneity, wages further reflect individual productivity differences as well as the state of the aggregate economy. As such, wage dispersion and wage progression can be linked to both the temporary and permanent shocks found in the error structure of longitudinal wage regressions.

10 Conclusion

In the Burdett-Mortensen model with random matching, on-the-job search by workers reduces a firm's monopsony power by introducing intertemporal competition. Firms optimally respond to this competition by differentiating their wage policies. In contrast, this paper adopts stock-flow matching and consequently direct competition for workers. Here, search reveals all potential employment opportunities in a centralized marketplace and auctions determine initial wages. If there is one bidder, low initial wages obtain. If there is more than one bidder, higher Bertrand wages result.

In this framework, although on-the-job search by workers continues to generate the prospect of intertemporal competition between firms, competition is not realized. The optimal response by firms is to avoid direct competition by offering wages that in effect outbid expectations in the marketplace. Wages evolve as the worker's threat of on-the-job search varies with anticipated turnover from the birth and death rates of outside employers. Potential employers and hence bidders for the worker come and go in background. Given this set-up, two key endogenous factors emerge - the bidding history from the last visit to the market, represented above by k, and the degree of expected turnover of rival bidders, captured above through t. History fades over time so initially low wages rise with time whereas higher wages gradually decay. Wages eventually converge to a common long term level.

Several results emerge that conform with empirical regularities.

- This framework provides a novel source of wage dispersion with the appealing feature of an interior mode and prominent, well-behaved tails on both sides.

- The initial and expected threat of competition in the market determine the evolution of wages within a firm and thereby introduce a novel way of understanding why firms pay different wages to similar workers.

- Unlike under random search, when a separation shock occurs, some workers find jobs immediately so that job-to-job transitions occur after a dismissal. Workers who find many bidders will often experience a wage rise from the ensuing bidding. The unlucky who only have one suitor experience wage cuts.

- Workers dismissed from low wage jobs have poor job prospects and often wind up with poor outcomes - unemployment or even lower wages. It can appear as if these workers with bad histories are stuck in low-wage, no-wage cycles.

The focus in this paper is on workers as sellers, and on firms as buyers of labour services. The framework and the results apply more broadly to other long term, repeated traders such sales reps and purchasing agents; health care providers and patients; men and women in the marriage market; and so on. It is sensible to investigate the patterns of trade in markets with long standing traders to see how they conform with the results established here.

References

[1] Abowd, J. and Card, D., (1989), "On the Covariance Structure of Earnings and Hours Changes", *Econometrica*, 57, 2, pp. 411-445.

[2] Andrews, M., Bradley, S. and Upward, R., (2001), “Estimating the Probability of a Match Using Microeconomic Data for the Youth Labour Market”, *Labour Economics*, 8, pp. 335-357.

[3] Bender, S. and Von Wachter, T., (2006), “In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers’ Careers”, *American Economic Review*, 96, 5, pp. 1679-705.

[4] Burdett, K. and Coles, M.G., (2007), “Equilibrium Wage-Tenure Contracts with Heterogeneous Firms”, mimeo, University of Essex.

[5] Burdett, K. and Mortensen, D., (1998), “Wage differentials, employer size, and unemployment”, *International Economic Review,* 39, pp. 257-273.

[6] Christensen, B.J., Lentz, R., Mortensen, D., Neumann, G., and Werwatz, A., (2005), “On the Job Search and the Wage Distribution”, *Journal of Labor Economics*, 23, 1, pp. 31-58.

[7] Connolly, H. and Gottschalk, P., (2008), “Wage Cuts as Investment in the Future Wage Growth: Some Evidence”, *LABOUR: Review of Labour Economics and Industrial Relations,* 22, pp. 1-22.

[8] Coles, M.G., (1999), “Turnover Externalities with Marketplace Trading”, *International Economics Review,* 40, pp. 851-867.

[9] Coles, M.G. and Muthoo, A., (1998), “Strategic Bargaining and Competitive Bidding in a Dynamic Market Equilibrium,” *Review of Economic Studies*, 65, pp. 235-260.

[10] Coles, M.G. and Petrongolo, B., (2008), “A Test Between Stock-Flow Matching and the Random Matching Function Approach”, *International Economic Review*, 45, 2, pp. 268-285.

[11] Coles, M.G. and Smith, E., (1998), “Marketplaces and Matching”, *International Economic Review,* 39, pp. 239-255.

[12] Ebrahimy, E. and Shimer, R., (2006), “Stock-Flow Matching”, mimeo, University of Chicago.

[13] Fallick, B. and Fleischman, C., (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.

[14] Gregg, P. and Petrongolo, B., (2005), “Non-random Matching and the Performance of the Labor Market”, *European Economic Review*, 49, pp. 1987-2011.

[15] Hall, R. and Milgrom, P., (2008), "The Limited influence of Unemployment on the Wage Bargain", *American Economic Review*, Vol. 98, No. 2, pp. 1653-1674.

[16] Heckman, J. and Borjas, G., (1980), "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence", *Economica*, 47, pp. 247-83.

[17] Hornstein, A., Krusell, P. and Violante, G., (2007), "Frictional Wage Dispersion in Search Models: A Quantitative Assessment", mimeo, New York University.

[18] Jacobson, L., Lalonde, R. and Sullivan, D., (1993), "Earning Losses of Displaced Workers", *American Economic Review*, 83, 4, pp. 685-709.

[19] Jolivet, G., Postel-Vinay, F. and Robin, J-M., (2006), "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US", *European Economic Review*, 50, pp. 877-907.

[20] Kuo, M. and Smith, E., (2009), "Marketplace Matching in Britain: Evidence from Individual Unemployment Spells", *Labour Economics*, 16, pp. 37-46.

[21] Lagos, R., (2000), "An Alternative Approach to Search Frictions", *Journal of Political Economy*, 108, pp. 851-73.

[22] Lynch, L., (1989), "The Youth Labor Market in the Eighties: Determinants of Reemployment Probabilities for Young Men and Women", *Review of Economics and Statistics*, 71, pp. 37-45.

[23] MaCurdy, T., (1982), "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis", *Journal of Econometrics*, 18, pp. 83 -114.

[24] Meghir, C and Pistaferri, L., (2004), "Income Variance Dynamics and Heterogeneity", *Econometrica*, 72, pp. 1-32.

[25] Mortensen, D., (2003), *Wage Dispersion: Why are Similar Workers Paid Differently?*, Zeuthen Lecture Book Series, Cambridge.

[26] Nagypal, E., (2005), "On The Extent of Job-to-Job Transitions", mimeo, Northwestern University.

[27] Nagypal, E., (2008), "Worker Reallocation over the Business Cycle: The Importance of Job-to-Job Transitions", mimeo, Northwestern University.

[28] Petrongolo, B. and Pissarides, C., (2000), "Looking Inside the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, 34, pp. 390-431.

[29] Pissarides, C., (2001), *Equilibrium Unemployment Theory*, 2nd ed. Cambridge, MA, MIT Press.

[30] Postel-Vinay, F. and Robin, J-M., (2002), "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70, 6, pp. 2295-2350.

[31] Postel-Vinay, F. and Turon, H., (2008), "On-the-job Search, Productivity Shocks, and the Individual Earnings Process", *International Economic Review*, forthcoming.

[32] Prabhu, N., (1965), "Queues and Inventories: A Study of their Basic Stochastic Processes", Wiley.

[33] Shimer, R., (2005), "The Cyclical Behavior of Equilibrium Unemployment and Vancancies", *American Economic Review*, Vol. 95, 1, pp. 25-49.

[34] Shimer, R., (2007), "Mismatch", *American Economic Review*, 97, pp. 1074-1101.

[35] Smith, E., (2007), "Limited Duration Employment", *Review of Economic Dynamics*, 10, pp. 444-47.

[36] Stewart, M., (2008), "The Inter-related Dynamics of Unemployment and Low-Wage Employment", *Journal of Applied Econometrics*, forthcoming.

[37] Taylor, C., (1995), "The Long Side of the Market and the Short End of the Stick: Bargaining Power and Price Formation in Buyers', Sellers' and Balanced Markets", *Quarterly Journal of Economics,* 110, pp. 837-855.

APPENDIX

Proof of Proposition 2:

Let $\lambda = \delta$ and define

$$\varphi_1(t,k) \equiv \phi'(t)[k(k-3) + 2ke^{-\delta t} + \phi(t)^2],$$

$$\varphi_2(t,k) \equiv \phi(t)[k - \phi(t)](r + \mu + 2\delta).$$

In this case, (15) implies

$$\dot{w}(t,k) = \phi(t)^{k-2} e^{-\phi(t)} \phi'(t) \left[\varphi_1(t,k) - \varphi_2(t,k)\right] \frac{x - b}{r + \mu + \delta + \lambda}$$

for all t and $k \geq 2$, where $\phi(t) \in [0,1]$.

Note that $\varphi_1(t,k) > 0$ and strictly decreasing for all t when $k \geq 3$ and that $\varphi_2(t,k) > 0$ and strictly increasing for all t when $k \geq 2$. When $k = 2$,

$$\varphi_1(t,2) = \lambda e^{-\delta t}[(1 + e^{-\delta t})^2 - 2],$$

and is strictly decreasing for all t, but $\varphi_1(t,2) > 0$ only for those $t < t' = -\frac{1}{\delta}\ln(\sqrt{2} - 1)$, $\varphi_1(t',2) = 0$ and $\varphi_1(t,2) < 0$, otherwise. Since $\varphi_1(0,k) = \lambda k(k-1) > 0$ and $\varphi_2(0,k) = 0$,continuity implies there exists a unique $t^*(k)$ such that $\varphi_1(t^*(k),k) = \varphi_2(t^*(k),k)$. Hence, $w(t,k)$ has a unique maximum at $w(t^*(k),k)$ where

$$\begin{aligned} \dot{w}(t,k) &> 0 \text{ for all } t \in (0, t^*(k)) \\ \dot{w}(t,k) &= 0 \text{ at } t = t^*(k) \\ \dot{w}(t,k) &< 0 \text{ for all } t > t^*(k), \end{aligned}$$

for all $k \geq 2$. Moreover, note that $\dot{w}(0,2) > 0$, $\dot{w}(0,k) = 0$ for all $k \geq 3$ and $\lim_{t\to\infty} \dot{w}(t,k) = 0$ for all $k \geq 2$. $\|$

Derivation of the wage distribution with heterogeneous workers:

To obtain the steady state distribution of wages, let $n(t,k;\alpha)$ denote the proportion of employed workers of type α that observed k vacancies and have an employment spell of $t \geq 0$. Using the same arguments as in (17), it follows that

$$N(t,k;\alpha) = [h(\alpha) - u(\alpha)]\frac{n(0,k;\alpha)}{\mu + \delta}(1 - e^{-(\mu+\delta)t}),$$

describes the number of employed workers of type α that left k vacancies and have an employment spell no greater than t. In this case, $u(\alpha)$ denotes the steady state number of unemployed workers of type α. In a steady state, $N(\infty,k;\alpha) = N(k;\alpha)$ is constant over time. Equating the flow of workers of type α that enter and leave this category we obtain a system of k simultaneous equations for each $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ such that

$$n(0,k;\alpha) = \left[1 - \int_0^\infty \delta P_{k+1}(t,k)e^{-(\mu+\delta)t}dt\right]^{-1} \sum_{s\neq k} n(0,s;\alpha) \int_0^\infty \delta P_{k+1}(t,s)e^{-(\mu+\delta)t}dt$$

for all $k \geq 0$, where $n(0,k;\alpha)$ describes its solutions. Integrating across α we obtain that $n(0,k) = \int_{\underline{\alpha}}^{\overline{\alpha}} n(0,k;\alpha)h(\alpha)d\alpha$ and hence $N(k) = \int_{\underline{\alpha}}^{\overline{\alpha}} N(k;\alpha)h(\alpha)d\alpha$. The steady state number of unemployed workers of type α is then given by

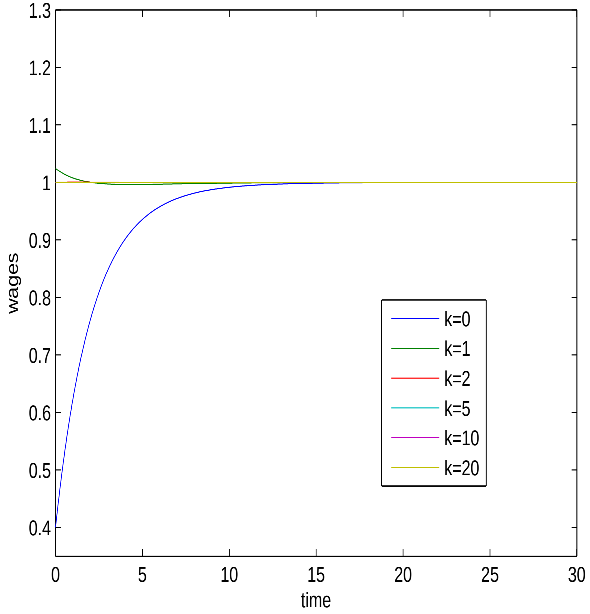
$$u(\alpha) = \frac{h(\alpha)\sum_{s=0}^\infty n(0,s;\alpha)\left[\int_0^\infty [\mu + \delta P_0(t,s)]e^{-(\mu+\delta)t}dt\right]}{\lambda + \sum_{s=0}^\infty n(0,s;\alpha)\left[\int_0^\infty [\mu + \delta P_0(t,s)]e^{-(\mu+\delta)t}dt\right]},$$

where $u = \int_{\underline{\alpha}}^{\overline{\alpha}} u(\alpha)d\alpha$.

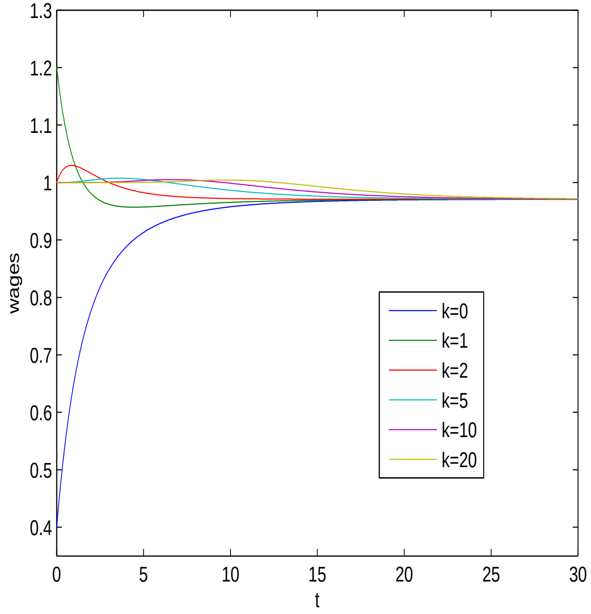
Using the same arguments as before we obtain that the distribution of wages for each worker's type, $G(w;\alpha)$ is then given by

$$G(w;\alpha) = \sum_{k=0}^\infty \left[\int_0^\infty e^{-(\mu+\delta)t} I(w(t,k) \leq w)dt\right] n(0,k;\alpha),$$

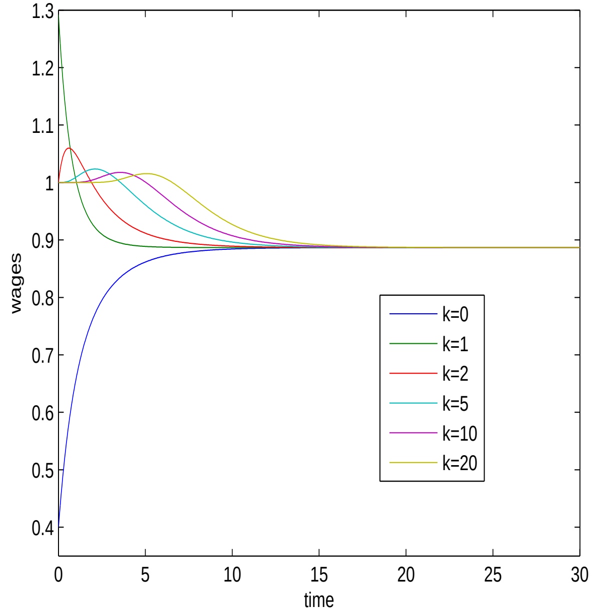
and $G(w) = \int_{\underline{\alpha}}^{\overline{\alpha}} G(w;\alpha)h(\alpha)d\alpha.\|$



(a) $\delta = 0.02$.

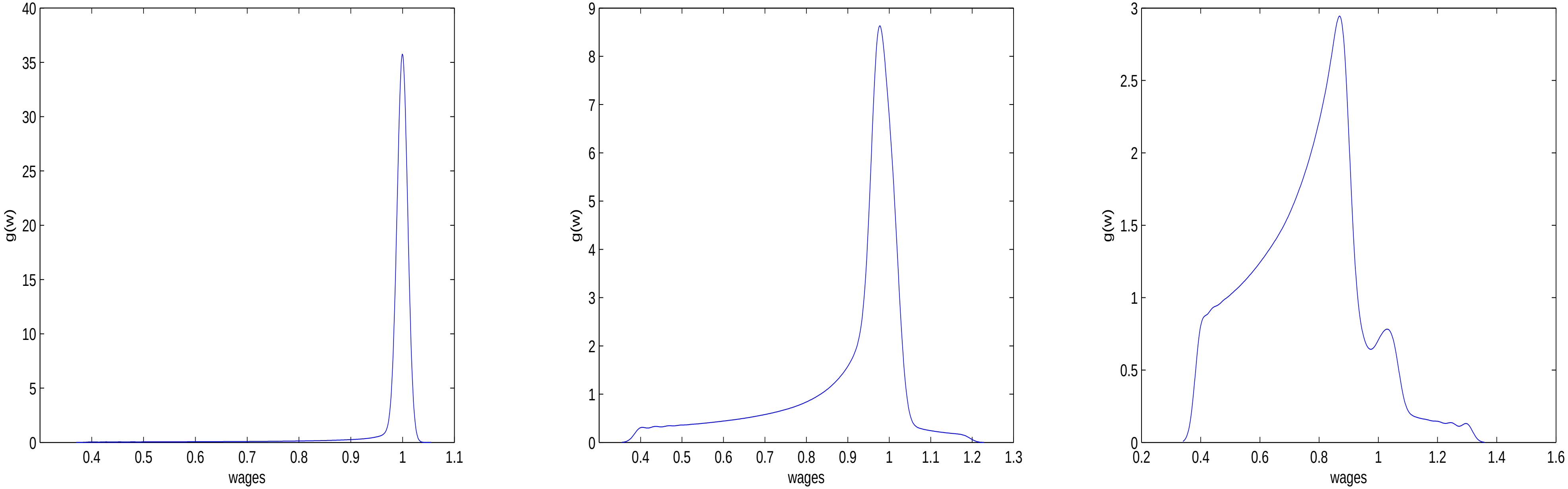


(b) $\delta = 0.2$.



(c) $\delta = 0.45$.

Figure 1: Within Firm Wage Dynamics

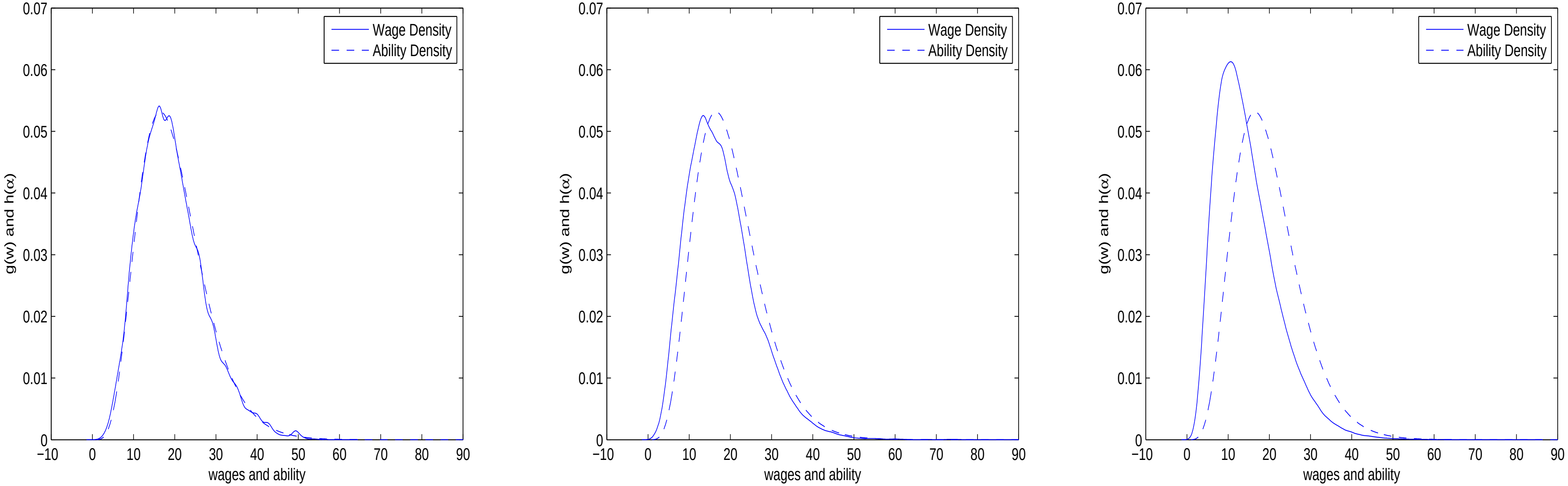


(a) $\delta = 0.02$.

(b) $\delta = 0.2$.

(c) $\delta = 0.45$.

Figure 2: Kernel Estimate of the Wage Density with Homogeneous Workers



(a) $\delta = 0.02$.

(b) $\delta = 0.2$.

(c) $\delta = 0.45$.

Figure 3: Kernel Estimate of the Wage Density with Heterogeneous Workers