On the firms’ component of wage dispersion

Adrian Masters
amasters@albany.edu
Department of Economics, SUNY Albany
1400 Washington Ave, room BA109
Albany, NY 12222, USA

June 10, 2014

Abstract

This paper presents an environment in which firms’ productive heterogeneity passes through to wage dispersion via sequential search and endogenous worker effort levels. Despite small gains from trade, the model is able account for more than two thirds of the measured firm component of wage dispersion. The implied narrow range of worker utility effectively pins down the lowest wage in the distribution and higher wages simply compensate workers for their extra effort.

JEL codes: E24, J64

Keywords: Wage dispersion, sequential search

1 Introduction

It has been well established that substantial wage dispersion exists among otherwise similar workers (Acemoglu [2002], Lemieux [2006], Autor et al
Of that wage dispersion, Abowd and Kramarz [1999] attribute about one third to firm effects. This paper investigates the importance of endogenous worker effort for understanding how variations in firm productivity translate into variations in worker pay.

In a competitive market for homogeneous labor the law of one price will hold so that, despite firm heterogeneity, workers all get paid alike. Hornstein et al [2011] (henceforth HKV) further show that search frictions alone do not help much in this regard. Unless the value to unemployment is unrealistically low, a worker will never take a low wage offer when a much higher one could arrive at any moment. Here, endogenous work effort is introduced to the Diamond-Mortensen-Pissarides (DMP) framework with ex ante productive heterogeneity among firms. High productivity firms offer high wages but require greater effort from their workers. The workers face a simple labor supply-type problem in which they trade off disutility of work for higher pay. While, consistent with HKV, the utility dispersion among workers remains small, the model can generate any degree of wage dispersion. A model simulation suggests that at least two thirds of the firm effect identified by Abowd and Kramarz [1999] can be attributed to endogenous work effort.

In the model, entrepreneurial firms sample ideas according to a Poisson process. The ideas represent the total factor productivity (TFP) of the job when created. With a viable idea in hand, the firm acquires physical capital and looks for a worker. Because the ideas differ, the capital investments differ and so do the contracts that the firms offer workers. Effort is observable to the firm and it enters the production function so firms post wage/effort contracts and workers direct their search accordingly.

In search environments, disutility of work and utility of leisure (or income from non-market activities) are typically interchangeable. The former, however, can be match specific and, therefore, correlated with the current wage. This is what happens when work effort is endogenized. The point is that what
matters to workers is utility rather than the wage per se. In the model presented here utility dispersion is small but, because higher wage jobs demand more effort, any degree of wage dispersion can be supported. When effort is exogenous, however, almost none of the dispersion in TFP passes through to wages. With endogenous effort, high wage jobs remain more sought after than low wage jobs but, because the former require greater work effort, the differential in terms of the utility they provide is small.

To demonstrate the quantitative relevance of the model I simulate it using realistic parameter values. Most of the parameters are taken from the literature on search and matching or real business cycles. Others are chosen to meet certain market outcomes. One such outcome is the extent of productivity dispersion across firms. Syverson [2004] provides information on the dispersion of TFP across establishments within narrowly defined industries. In an attempt to account for factor utilization he includes materials and energy as well as capital and labor as inputs. He finds an inter-quartile ratio (the ratio of 75th to 25th percentile) for TFP, depending on the industry, of between 1.34 and 1.56. I use a target for the inter-quartile range of observable productivity dispersion (i.e. inclusive of unobserved effort), of 1.45.

An important factor in the determination of the extent to which productive heterogeneity shows up as wage dispersion is the sensitivity of work effort to the wage. Unfortunately, effort is hard to observe. Leamer and Thornberg [2000] and Lundberg [1985] show that hours of work and wages are positively correlated and are jointly determined so that hours might serve as a proxy for effort. Many of the studies that demonstrate large degrees of wage dispersion, however, already control for hours. Bils and Cho [1994] provide analysis of factor utilization over the business cycle. They use data from time-and-motion studies in the UK to obtain a Frisch elasticity of effort of around one third. That is to say, abstracting from any wealth effects, a
3% increase in the hourly wage is required to bring about a 1% increase in labor services per hour from the worker. This will be used to pin down the elasticity of the cost of effort function.

A parameter of the sequential search model that strongly affects its ability to generate realistic wage dispersion is the utility of leisure (combined here with the disutility of work). The higher it is, the less unpleasant is unemployment and the more likely a worker is to turn down low wage offers. Only if unemployment is highly unpleasant, can the standard sequential search model be reconciled with the data. A discussion as to what the appropriate level of the utility from leisure should be is available from the business cycle literature. There, a high value is desirable to reduce the fluctuations of the wage relative to that of unemployment. Hagedorn and Manovskii [2008] argue that what matters is the utility of leisure for the marginal worker which they put at 95% of the worker’s productivity. Mortensen and Nagypal [2007], however, point out that in the DMP framework it is the average worker’s utility of leisure that controls vacancy creation. Hall [2009] generalizes the DMP model so that he can parameterize the labor supply and vacancy creation elasticities. He estimates the implied Frisch system and argues that the appropriate figure for the ratio of consumption when unemployed to employed (the replacement ratio) is around 0.85.

After adopting parameters from the literature and taking account of the normalizations that show up in the model there are two free parameters. These are the scale parameter of the effort cost function and the utility from leisure. They are pinned down by inter-quartile range of observable productivity dispersion (1.45) and the replacement ratio (0.85). The mean to minimum ($M_m$) ratio, as introduced by HKV, that emerges for wages is 1.25. If in the same calibration, endogenous effort is shut down, the $M_m$ ratio for wages drops to 1.004. From the discussion in HKV, the $M_m$ ratio from all sources of wage dispersion for observably similar workers should be
around 1.8. Given that, from Abowd and Kramarz [1999], firm heterogeneity should explain about one third of this, the implication is that endogenous effort explains close to all of the firm effect.

A small literature has emerged following HKV that attempts to identify additional sources of wage dispersion from otherwise identical workers. In Ortego-Marti [2012] the basic source of wage dispersion is match-specific productivity as in Pissarides [2000] Chapter 6. He also incorporates skill decay while unemployed. The skill decay makes workers more eager to take jobs which lowers the reservation match productivity at which hiring occurs. This increases wage dispersion even among those with the same realized level of skill. Burdett et al [2011] use learning-by-doing to obtain a similar effect in the Burdett and Mortensen [1998] on-the-job search framework. Unemployment has an opportunity cost to workers in terms of lost productivity gains which lowers their reservation wages. Models in this vein, then, either make unemployment worse or employment better than in the contexts considered by HKV. The model presented here takes a very different tack. Wage dispersion comes from ex ante productivity dispersion. Endogenous effort means that utility dispersion can remain low while wage dispersion reflects the dispersion in effort levels.

The remainder of the paper proceeds as follows. Section 2 lays out the theoretical model. A numerical simulation of the model appears in Section 3. Section 4 provides some discussion of the results and Section 5 concludes.

## 2 Model

### 2.1 Environment

Time is continuous with an infinite horizon. There is mass one of ex ante identical workers who face death according to a Poisson arrival rate $\delta$. Those
who die are replaced by newborns so that $\delta$ represents both the birth and death rates. Workers are risk neutral. They get utility from the consumption good which they receive in the form of wages. Workers get disutility, $c(e)$ from working with effort level $e$. The cost function $c(.)$ is twice differentiable, strictly increasing and strictly convex with $c(0) = c'(0) = 0$ and $\lim_{e \to \infty} c'(e) = \infty$. While unemployed they receive a flow utility from leisure $b$.

There are also a large number of ex ante identical entrepreneur/employers who I will call firms. The firms can create vacancies by the following process. At flow cost $a$, they sample ideas, $z$, at an arrival rate $\gamma$ from a continuous distribution $G$. The support of $G$ is $(0, \bar{z}]$.\footnote{It will become clear that this mechanism for vacancy creation is isomorphic to one in which firms pay an up-front cost of $a/\gamma$ in order to sample $z$ from $G(.)$. As both $a$ and $\gamma$ will be shown to be subject to a normalization the distinction is moot.} Once a firm realizes an idea, $z$, it decides whether or not to pursue it. Pursuing an idea means that the firm acquires physical capital $k$. The cost of capital is normalized to 1.\footnote{As with the advertising cost in the standard DMP model, the costs $a$ and $k$ here can be viewed as utility costs to the firm or as physical costs measured in the consumption good. In the latter case it is assumed that firms each have a large number of on-going and potential projects and they pay for $a$ and $k$ out of the profits from existing employment relationships.} A firm with idea, $z$, specific capital stock, $k$, and matched with a worker who provides effort level $e$ produces $zf(k,e)$ units of the consumption good. Ideas are therefore synonymous with match total factor productivities (TFPs). The function $f(.,.)$ is a standard neoclassical production function that has constant returns to scale, is twice differentiable, increasing in both arguments, concave and satisfies the Inada conditions.\footnote{That is $f(0,0) = f(0,e) = f(k,0) = 0$ for all $k,e > 0$, $\lim_{k \to 0} f_k(k,e) = \infty$, $\lim_{e \to 0} f_e(k,e) = \infty$, $\lim_{k \to \infty} f_k(k,e) = 0$, $\lim_{e \to \infty} f_e(k,e) = 0$.}

While it is standard in the search and matching literature for firms to occupy a single worker, in the current context the assumption is not without
consequence. If each entrepreneur had an infinite capacity for workers only
the most productive firm would survive. Here the assumption represents an
extreme version of the Lucas [1978] span-of-control restriction in which the
second worker is not productive. Recall that the focus of the paper is on
how differences in firm productivities pass through to wage dispersion. In
the absence of any span-of-control restriction there would be no productivity
dispersion which would be contrary to the data.

Jobs are subject to destruction at the rate $\lambda$. For the purpose of simula-
tions there will be a common discount rate $r \geq 0$. To simplify the exposition
I set $r = 0$ for the analytical part of the paper.

Consistent with competitive search, there is a continuum of potential
trading locations. The rate at which workers and firms encounter each other
in any location is governed by a matching function, $M(u, v)$ where $u$ is the
unemployment rate among workers associated with that location and $v$ is
the number of vacant jobs per worker associated with that location. The
matching function has constant returns to scale (c.r.s), is twice differentiable,
increasing in both arguments and concave. The Poisson arrival rate at which
a worker meets firms is then $M(u, v)/u = M(1, \theta)$ where $\theta \equiv v/u$. It is con-
venient to define the function $m(\theta) \equiv M(1, \theta)$. Then $m(.)$ inherits concavity
from $M(.,.)$ and c.r.s means $m(\theta) \geq \theta m'(\theta)$. In order to prevent the trivial
equilibrium (in which no vacancies are created) from being stable and rule
out other corner-type solutions, I will further assume that $\lim_{\theta \to 0} m'(\theta) = \infty$
and that $\lim_{\theta \to \infty} m'(\theta) = 0$.

In a decentralized economy, trading locations are occupied by markets.
The set of active markets will depend on which aspects of jobs become com-
mon knowledge. Here, firms post a wage/effort pair, $(w, e)$, which workers
observe and use to direct their search. Markets are therefore indexed by
$\omega = (w, e, \theta) \in \Omega = \mathbb{R}_+^3$. The set of active markets is $\Omega_A \subset \Omega$. The equilib-
rium concept will be competitive search as developed by Moen [1997]. I seek
steady-state symmetric equilibria. Symmetry, here, means that all firms of the same type enter the same market and all workers enter any active market with the same probability as each other.

2.2 Value functions

Let $V_j$ be the value to a firm of type $(z, k)$ of employing a worker in market $\omega = (w, e, \theta)$. Then

$$\lambda V_j = z f(k, e) - w + \delta [V_f - V_j]$$

where $V_f$ is the value to the firm of holding a vacancy in market $\omega$. Thus

$$\lambda V_f = \frac{m(\theta)}{\theta} [V_j - V_f].$$

Solving for $V_f$ implies

$$V_f = V_f(\omega; z, k) \equiv \frac{m(\theta) [zf(k, e) - w]}{\lambda[(\delta + \lambda)\theta + m(\theta)]},$$

(2)

For a worker who enters the same market, let $V_e$ be the value of employment. Then

$$\delta V_e = w - c(e) + \lambda(V_u - V_e)$$

where $V_u$ is the value to looking for work in market $\omega$. So,

$$\delta V_u = b + m(\theta)(V_e - V_u).$$

Solving for $V_u$ implies

$$V_u = V_u(\omega) \equiv \frac{m(\theta) [w - c(e)] + (\delta + \lambda)b}{\delta(\delta + \lambda + m(\theta))}.$$
2.3 Equilibrium

Because workers are identical, for multiple markets to coexist, it must be the case that there exists some $V_u^* > 0$ such that $V_u(\omega) = V_u^*$ for all $\omega$ in $\Omega_A$. Given $z$, and their capital stock, when firms decide which market to enter they take this indifference of workers as a constraint to their problem. Let

$$\omega^*(z, k) \in \arg\max_{\omega} V_f(\omega; z, k) \text{ subject to: } V_u(\omega) = V_u^*.$$  \tag{5}

Then, let $V_v(z)$ be the value of pursuing an idea type $z$ to a firm. Thus,

$$V_v(z) = \max_k V_f(\omega^*(z, k); z, k) - k.$$  \tag{6}

Firms will pursue any idea, $z$, for which $V_v(z) \geq 0$. They will drop any idea, $z$, for which $V_v(z) < 0$.

Definition 1 An equilibrium is a list, $Z_A, k^*(Z_A), \Omega_A, V_u^*$ such that, given $V_u^*$,

(i) $Z_A$ is the set of ideas for which $V_v(z) \geq 0$.

(ii) $k^*(Z_A)$ is the set of values of $k$ associated with each $z \in Z_A$, such that $k$ solves Problem (6).

(iii) $\Omega_A$ is the set of markets $\omega = \omega^*(z, k^*(z))$ such that $z \in Z_A$. Then $V_u^* = V_u(\omega^*(z, k^*(z)))$ for every $z \in Z_A$.

As the infimum to the support of $G$ is 0, for any $V_u^* > 0$ there exists some $\hat{z}$ for which $V_v(\hat{z}) = 0$. It is immediate from the envelope theorem and equation (2) that (as long as $\omega$ remains in the interior of the positive orthant) $V_v$ is strictly increasing in $z$ so that $\hat{z}$ is unique and $Z_A = [\hat{z}, \bar{z}]$. Free-entry to the market for ideas therefore means

$$a = \gamma \int_\hat{z}^{\bar{z}} V_v(z) dG(z).$$  \tag{7}

For a given value of $z$, Problems (5) and (6) can be combined into
This demonstrates that the above environment leads to identical outcomes to one in which the firms also advertise their capital stock. Under a wage contract, the firm is residual claimant and investment is efficient regardless of the extent to which it becomes common knowledge.\footnote{See Masters [2011].} By eliminating the wage from equations (2) and (4), problem (8) reduces to

\begin{equation}
V_v(z) = \max_{k, \omega} V_f(\omega; z, k) - k \text{ subject to: } V_u(\omega) = V_u^*.
\end{equation}

The first order conditions in $k$ and $e$ are respectively,

\begin{align}
m(\theta)z f_k(k, e) - \lambda [m(\theta) + (\delta + \lambda)\theta] &= 0 \tag{10} \\
z f_e(k, e) - c'(e) &= 0 \tag{11}
\end{align}

Clearly, problem (9), is strictly concave in $k$ and $e$ implying that for any given value of $\theta$ there is unique pair $(k(\theta), e(\theta))$ that solve equations (10) and (11). Substituting these functions back into the maximand in problem (9) generates a continuous function of $\theta$ which will have a maximum on the positive real line. Using equations (10) and (11), it is straightforward to show that $k'(\theta) < 0$ so that by inspection of problem (9), any optimizing value of $\theta$ must be finite.

\textbf{Claim 2} Suppose the functional forms, $m(\theta) = \bar{m}\theta^\eta$, $\bar{m} > 0$, $0 < \eta < 1$; $c(e) = \bar{c}e^\sigma$, $\bar{c} > 0$, $\sigma > 1$; $f(k, e) = k^\alpha e^{1-\alpha}$, $0 < \alpha < 1$. Then for each $z$ such that $V_v(z) \geq 0$, problem (8) has a unique solution.

\textbf{Proof.} See appendix. \hfill \blacksquare

The value of $\hat{z}$ is obtained from equation (7). The value of $V_u^*$ is obtained by setting $V(\hat{z}) = 0$. It should be clear that if $b$ is large enough there can be no
gains from trade. That is, there is some $b_{\text{max}}$ such that for $b \geq b_{\text{max}}$, $\hat{z} \geq \bar{z}$. In such cases only the trivial equilibrium, which corresponds to autarky, exists. Under the functional form restrictions to be imposed for the numerical analysis and with $b < b_{\text{max}}$, then, there is a unique non-trivial equilibrium. An earlier version of this paper, Masters [2013], shows that this equilibrium also coincides with the Planner’s allocation.\footnote{In that paper $b = 0$ but the extension to strictly positive values of $b$ is straightforward.}

Under these functional form restrictions, it is straightforward to show that (given $z$ and $V_u^*$) $e$, $k$ and $w$ increase with $z$ and $\theta$ decreases with $z$. So, despite the increased effort required of high wage workers, their jobs are the most sought after. Workers are indifferent across markets because they anticipate a longer search period in the higher wage markets. Workers trade the utility derived from jobs against the time it takes to get one – the unemployment rate is higher in the high wage markets. High wage jobs therefore fill more quickly. The model is consistent with the data in this regard. Holzer et al [1991] find that for jobs requiring similar skill sets, applicant queue lengths increase with the wage.

3 Simulation

3.1 Functional forms

The numerical analysis uses the following functional forms: $m(\theta) = \bar{m}\theta^\eta$, $\bar{m} > 0$, $0 < \eta < 1$; $c(e) = \bar{c}e^\sigma$, $\bar{c} > 0$, $\sigma > 1$; $f(k, e) = k^\alpha e^{1-\alpha}$, $0 < \alpha < 1$. The distribution $G(.)$ is assumed to be uniform on $(0, \bar{z}]$. Recall that the focus of the paper is on the extent to which the modeled mechanism allows dispersion in firm productivity to pass through to wages. Of some additional interest is how the mechanism adapts the shape of the distribution. The uniform distribution has two helpful properties. First, whatever is the threshold,
\( \hat{\bar{z}} \), of productivity above which ideas become viable, the distribution of viable ideas remains uniform. Second, as we know the density of productivities is flat, any shape that emerges in the density of wages is a consequence of the mechanism itself rather than it being exogenously imposed.

### 3.2 External parameters

The basic unit of time is 1 year. So, the common discount rate, \( r \), is set to 0.04. The job destruction rate, \( \lambda \), is set to 0.2 to reflect the expected life of a job as about 5 years as identified by Cole and Rogerson [1999]. The death rate, \( \delta \), (or rate of labor-force exit) for workers was set to 0.05. This corresponds to an expected duration of 20 years which would be the average expected remaining participation of workers with 40 year working lives. Blanchard and Diamond [1989] and Pissarides and Petrongolo [2001] discuss matching function parameters. From there I set \( \bar{m} = 12 \) and \( \eta = 0.5 \).

The output elasticity of capital, \( \alpha \), is potentially an issue here as much of the focus of the paper is on the intensity of factor (specifically labor) utilization. Micro-data studies such as Syverson [2004] include energy and materials consumption to proxy for variation in factor utilization. But, whatever the true utilization rate is, it should be reflected in payment to factors. Because frictions are small, factor payments will be close to their marginal products. Restricting attention to payments to labor and capital in aggregate data should, therefore, internalize the variations in utilization. Following the real business cycle literature, then, \( \alpha = 0.33 \).

Bils and Cho [1994] provide analysis of factor utilization over the business cycle. They use data from time-and-motion studies in the UK to obtain a Frisch elasticity of effort with respect to the wage of around one third. That is to say, abstracting from any wealth effects, a 3% increase in the hourly

\[^{6}\text{As the model is cast in continuous time the choice of time unit is completely arbitrary.}\]
wage is required to bring about a 1% increase in labor services per hour from the worker. This implies a value for \( \sigma \) of 4.\(^7\) All of the externally obtained parameters are summarized in Table 1.

<table>
<thead>
<tr>
<th>( \bar{m} )</th>
<th>( \eta )</th>
<th>( \sigma )</th>
<th>( \alpha )</th>
<th>( r )</th>
<th>( \lambda )</th>
<th>( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.5</td>
<td>4</td>
<td>0.33</td>
<td>0.04</td>
<td>0.2</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 1: External parameter values for the Leading Example.

### 3.3 Normalizations

The parameterized model exhibits certain normalizations that restrict the set of independent adjustments that can be made. The chosen distribution of \( z \) has one parameter, \( \bar{z} \). Any change to \( \bar{z} \), however, can be fully corrected for by an adjustment to \( \bar{c} \). To see why, consider problem (9). As long as keeping \( zf(k, e) - c(e) \) constant, while equation (11) continues to hold, implies that \( zf_{k}(k, e) \) remains unchanged, there will be no impact of a change in \( z \) on any of the variables of interest. After substituting parametric forms we, therefore, impose that, for any \( z \),

\[
zk^\alpha e^{1-\alpha} - \bar{c}e^\sigma = \chi
\]

(12)

\[
(1 - \alpha)zk^\alpha e^{-\alpha} - \sigma \bar{c}e^{\sigma-1} = 0
\]

(13)

\(^7\)If the instantaneous utility from a wage \( w \) and effort level \( e \) is \( U(w, e) \), then it is straightforward to show that the Frisch elasticity of effort is

\[
\varepsilon_F = \frac{U_e}{e U_{ee} - \frac{U_{ee}^2}{U_{ww}}}. \]

Here, \( U(w, e) = w - ce^\sigma \). Linearity with respect to the wage reduces the above definition to

\[
\varepsilon_F = \frac{U_e}{e U_{ee}} = \frac{1}{\sigma - 1}.
\]
where $\chi$ is an arbitrary constant. Equations (12) and (13) can be viewed as a system in $\bar{c}$ and $e$. Solving for $\bar{c}$ from (13) yields

$$\bar{c} = \frac{(1 - \alpha)zk^\alpha e^{1-\alpha} - \sigma}{\sigma}.$$ 

Into (12) implies

$$zk^\alpha e^{1-\alpha} = \left(\frac{\sigma}{\sigma - 1 + \alpha}\right)\chi$$

which means match output does not change as long as $e$ and $\bar{c}$ are adjusted according to equations (12) and (13). As $z_f(k,e) = \alpha zk^\alpha e^{1-\alpha}$ this will not change either. For any given adjustment, $z_0$ to $z_1$ in TFP, the implied adjustments to $\bar{c}$ and $e$ are,

$$\frac{\bar{c}_0}{\bar{c}_1} = \left(\frac{z_0}{z_1}\right)^{\frac{\sigma}{1-\alpha}} \quad \text{and} \quad \frac{e_0}{e_1} = \left(\frac{z_1}{z_0}\right)^{\frac{1}{1-\alpha}}.$$ 

Nothing else changes. For computational purposes I will normalize $\bar{c} = 1$ and use equation (7) to determine $\bar{z}$.

A second normalization comes in the choice of $a$ and $\gamma$. The only place that either enters the equilibrium conditions is in equation (7) where both do and then only their ratio matters. With a lower cost, $a$, of searching for ideas more entrepreneurs enter so that to generate the same number of jobs their arrival rate of ideas, $\gamma$, has to be proportionately lower. The value of $\gamma$ is set to 1.

A less obvious additional normalization is that the ratio $a/\gamma$ itself can also be set arbitrarily. This is because it determines the opportunity cost of a vacancy. So, if $a$ doubles, it should be clear that the variables of interest will be restored to their former values if output doubles, the cost of effort doubles and the utility of leisure doubles. The last requirement is easily met by doubling $b$. Constant returns to scale in production mean output will double if the production parameters, $z$ and $\alpha$, remain unchanged but both $k$ and $e$ double. For effort to double requires that $\bar{c}$ shrinks by a factor of
8. Given $\sigma = 4$ the total cost of effort, $\bar{\sigma} e^\sigma$, then doubles as required. Thus some variables (e.g. the aggregate capital stock and the aggregate level of effort) change but the variables of interest (e.g. welfare, the unemployment rate and the $Mm$ wage ratio) are fully restored to their initial values. While a normalization like this eliminates a parameter that could be used to match empirical targets, it avoids the thorny issue of what value either $a$ or $\gamma$ ought to be. The value of $a$ will be set to 1.

### 3.4 Empirical targets

The remaining parameters $\bar{\sigma}$, and $b$ are determined by consideration of empirical targets. These pertain to the dispersion in TFP, the effective replacement ratio for workers and the unemployment rate. Syverson [2004] provides information on the dispersion of TFP across establishments within narrowly defined industries. In an attempt to account for factor utilization he includes materials and energy as well as capital and labor as inputs. Of course both energy and raw materials can be substitutes for labor. Syverson’s use of these as proxies requires that they be perfect complements to capital and labor services. To the extent that they are not, his number overestimates the extent of true TFP dispersion. He finds an inter-quartile ($IQ$) ratio for TFP of between 1.34 and 1.56. I use a the midpoint of his range, 1.45, as a target for the $IQ$ ratio of the observable distribution of TFP. As effort is not measured, the observable distribution includes effort variation so that the implied true TFP variation will be smaller. If I applied the target value for to true TFP the model would, therefore, generate a larger degree of wage dispersion.

As identified by Sattinger [1985] and highlighted by HKV, the relative value of employment to unemployment is an important determinant of wage dispersion in sequential search models. Only if unemployment is extremely
unpleasant, will workers accept low wage offers rather than wait for a better one. Attempts to measure the difference in wellbeing between employment and unemployment, however, do not find the kinds of numbers typically required to generate observed wage dispersions. At one extreme, Hagedorn and Manovskii [2008] argue that what matters is the marginal entrant worker whose flow utility from leisure (or non-market activity) should be 95.5% of his marginal product. Shimer [2005], on the other hand, sets that figure at 40%. Attempts to reconcile these numbers have come from both theoretical and empirical analyses. Mortensen and Nagypal [2007] for instance point out that what matters in the DMP framework is the average worker whose value of leisure will be much lower than that of the marginal worker. Hall [2009] estimates a Frisch system embedded in a generalized DMP framework. What emerges is a "replacement ratio" – the ratio of consumption (i.e. utility) while unemployed to that while employed of around 85%. I adopt this target figure.

In the current model, the relevant replacement ratio, $R_r$, is the utility of leisure plus the mean cost of effort divided by the mean wage. So,

$$R_r = \frac{b + \int_{\hat{z}}^{z} c(e^*(z)) \frac{dG(z)}{1-G(z)}}{\int_{\hat{z}}^{z} w^*(z) \frac{dG(z)}{1-G(z)}}$$

where $e^*(z)$ and $w^*(z)$ are the equilibrium values of the effort level and wage at the firm with TFP $z$.\(^8\)

The target unemployment rate is 5.5%. This is the typical figure used to characterize the average state of the US labor market for the 20 years or so prior to the financial crisis of 2007.

\(^8\)All of the important outcome measures are identical if $b$ is eliminated, the functional form of the cost function is $C + \bar{c}e^\sigma$, and the value of $C$ is set to that of $b$ in the current formulation. In which case

$$R_r = \frac{\int_{\hat{z}}^{z} c(e^*(z))dG(z)}{\int_{\hat{z}}^{z} w^*(z)dG(z)}.$$
With two free parameters it is not possible to hit all three of the empirical targets. The values of $\bar{c} = 0.217$ and $b = 0.426$ are chosen to obtain an IQ range for observable TFP of 1.45 and a replacement ratio of 0.85. These imply an unemployment rate of 5.93%. That without trying I get so close to the target of 5.5% is a testament to the general framework and the careful work that precedes mine in identifying the external parameters.\footnote{An earlier version of the paper, Masters [2013], includes and extension of the model which allows for ex post bargaining with firms posting only their capital stock. This introduces an additional parameter, the bargaining power of the worker. Changes in the bargaining power have almost no effect on the other variables but strongly impact the level of unemployment through the effect on vacancy creation. (The current model corresponds to a bargaining power of 0.5.) Given the proximity we already have, the target value of unemployment would be achieved with a small decrease in the worker’s bargaining power.}

The equilibrium $Mm$ ratio for wages is 1.253. In summarizing the empirical literature on wage dispersion, HKV obtain a range of values for the $Mm$ ratio of wages among similar workers of 1.7 to 1.9. Thus, at least according to the $Mm$ ratio measure, the model almost fully accounts for the one third of wage dispersion attributable to firm effects as identified by Abowd and Kramarz [1999]. The IQ ratio for true TFP is 1.33.

Figure 1 plots the density of the implied wage distribution. As TFP, $z$, is uniformly distributed and firms match faster in high wage markets there are actually more workers earning high wages than low wages. However, convexity of the cost of effort function causes wages to be convex with respect to $z$ which spreads out the wage distribution towards the upper tail. The wage density is negatively sloped because the latter effect dominates the former. A long standing concern with on-the-job search is that the equilibrium wage density (see Burdett and Mortensen [1998]) is upward sloping. An implication here is that endogenizing work effort might be able to address this issue.
4 Discussion

4.1 Robustness

The target value for the $IQ$ ratio of observable TFP dispersion was not precisely obtained. Here I conduct an alternative parameterization of the model using the 5.5% unemployment rate along with the 0.85 replacement ratio as the empirical targets. The value of $\bar{c}$ becomes 0.049 and the value of $b$ becomes 0.71. Qualitatively, increases in both $\bar{c}$ and $b$ increase $Rr$ and increase unemployment. It is, therefore, necessary to increase one parameter and decrease the other leading to large adjustments in both to achieve the new empirical targets. Under these parameters, the $Mm$ wage ratio drops to 1.18. The $IQ$ ratio for observable TFP drops to 1.35 which is the lower bound of Syverson’s range. Under these parameters, then, the model accounts for
about two thirds of the firms’ component of wage dispersion.

4.2 Source of wage dispersion

Three features of the model contribute to the degree of wage dispersion: the distribution of productivities, variable effort levels, and frictions in the labor market. Here I consider what happens when each of these features is removed.

Clearly, if the distribution $G(.)$ were degenerate there could be no technology dispersion. Without technology dispersion, as long as the cost of effort function is strictly convex, there can be no wage dispersion.

Ideally, to obtain the level of wage dispersion in the version of the model with exogenous effort, it should be recalibrated to achieve the same outcomes with respect to the replacement ratio the $IQ$ ratio for observable TFP and unemployment rate. In principle we now have three parameters to achieve this goal. These are $\bar{c}$, $b$ and the exogenous level of effort, $\bar{e}$. Unfortunately, with exogenous effort all three parameters have identical effects. I, therefore, fixed $\bar{c}$ and $b$ at their former values, 0.217 and 0.426 respectively and adjust $\bar{e}$ to achieve a value for $Rr$ of 0.85. The value of $\bar{e}$ that emerges is 1.073 which causes unemployment to be 6.19%. Of course, with exogenous effort, the $IQ$ ratio for observable TFP equals that of true TFP which is unchanged at 1.33. Despite the higher rate of unemployment, the $Mm$ ratio for the wage drops to 1.004. Eliminating endogenous effort all but kills wage dispersion.

To shut down labor market frictions, distinct markets indexed by wage/effort pairs are maintained but workers get jobs instantaneously. Consequently workers are indifferent across the jobs themselves rather than simply the markets they occupy. Clearly, with zero unemployment, the model can longer come anywhere near that target. The fact that leisure is never realized does not prevent it from providing utility so the replacement ratio, $Rr$, is still well
defined. I recalibrate the model using $\bar{c}$ and $b$ to set the IQ ratio for observable TFP to 1.45 and $Rr$ to 0.85. The values are $\bar{c} = 0.23$ and $b = 0.418$. The $Mm$ ratio for wages is now 1.245. Which confirms the result from HKV that labor market frictions contribute little to wage dispersion. The fact that $1.253 - 1.245 = 0.008$, however, means that the contribution of frictions to the $Mm$ ratio of wages with endogenous effort is about twice its contribution with fixed effort.

The upshot is that ex ante productivity dispersion across firms is the primary source of wage dispersion. Endogenous effort and search frictions both provide a channel by which productive heterogeneity passes through to wages. Endogenous effort amplifies productivity dispersion through its impact on investment. Frictions in the labor market, cause dispersion in worker utility. The model simulation implies that the former channel is the most important. Simply amplifying the dispersion in TFP into the wage does not, however, lead to higher $Mm$ ratios. This is because the $Mm$ ratio is a relative measure of dispersion.

### 4.3 Comparison to HKV

In the above parameterization the Frisch elasticity of utility with respect to effort was set to one third. We saw that the model can generate significant wage dispersion compared to the model with fixed effort levels. With much larger Frisch elasticities of effort, however, any degree of wage dispersion can emerge. For instance setting $\sigma = 1.05$ which corresponds to Frisch elasticity of 20 generates (after recalibration to the other targets) a $Mm$ ratio for wages of 1.6 while the IQ ratio for true TFP dispersion drops to 1.014. Reading the technical appendix of HKV, it is reasonable to ask why my model can generate large $Mm$ ratios for the wage. After all, that wages have a larger absolute dispersion than utility or TFP does not mean it will show up in the
ratio of two points in the distribution. As HKV put it: in sequential search, the Mm ratio is independent of the distribution.

Consider a simplified version of my model.\(^{10}\) Search is random, there is no capital and no worker death but there is a common discount rate \(r\). Except where indicated, the remaining notation is consistent with the preceding analysis. Firms pay \(a\) to get an idea, \(z \sim G\), at Poisson arrival rate \(\gamma\) so that equation (7) holds. With an idea in place, firms look for workers. Output is simply \(p(z,e)\) and the cost of effort is \(c(e)\). Given the foregoing we know that optimal effort level in job type \(z\) is

\[
e(z) = \arg \max_e p(z,e) - c(e).
\]

Let \(w(z)\) be the wage paid to a worker in job type \(z\). Then, the reservation utility level (following HKV’s technical appendix) is

\[
\hat{w} - \hat{c} = \Phi \int_{\hat{z}}^z \left[w(z) - c(e(z)) - (\hat{w} - \hat{c})\right] dF(z)
\]

where \(\hat{w} = w(\hat{z})\), \(\hat{c} = c(e(\hat{z}))\), \(\Phi = m(\theta)/(r+\lambda)\) and \(F(z) = G(z)/(1-G(\hat{z}))\).

Because firms will not create vacancies that cannot match, this implies that

\[
\frac{\bar{w} - \bar{c}}{\hat{w} - \hat{c}} = \frac{1 + \Phi}{\Phi}
\]

where \(\bar{w}\) and \(\bar{c}\) are the mean wage and cost of effort under distribution \(F\). As in realistic environments, \(\Phi \gg 1\), this Mm ratio for utility will be close to 1 (1.07 for a 5.5% unemployment rate). For the Mm wage ratio to be large, we therefore require the Mm cost ratio, \(\hat{c}/\hat{c}\), also to be large.

Now define \(Y(z) = p(z,e(z)) - c(e(z)) - (r+\lambda)V_e(z)\). As the match surplus is \(Y(z) - rV_u\), \(\hat{Y} \equiv Y(\hat{z}) = rV_u = \hat{w} - \hat{c}\). With the bargaining power of the firms equal to \(\phi\) we obtain

\[
\phi (V_e - V_u) - (1 - \phi) (V_j - V_u) = 0
\]

\(^{10}\)This version of the model delivers, qualitatively similar results to the general framework but the TFP dispersion required to generate quantitatively meaningful values of the Mm wage ratio, while hitting the other calibration targets, is too large.
which, after some algebra, implies

\[
\frac{\bar{Y}}{\hat{Y}} = 1 + (1 - \phi)\Phi
\]

where \(\bar{Y}\) is \(E(Y(z)|z \geq \hat{z})\). This \(Mm\) ratio is also quite close to 1 (1.14 for \(\phi = 0.5\) and a 5.5% unemployment rate).

Now, assuming \(p(z, e) = ze\) and \(c(e) = \bar{c}e^\sigma\) it is straightforward to show that \(Y(z) = (\sigma - 1)c(e(z)) - (r + \lambda)V_v(z)\). So that

\[
\frac{\bar{Y}}{\hat{Y}} = \frac{(\sigma - 1)\hat{c} - (r + \lambda)\bar{V}_v}{(\sigma - 1)\hat{c}} = \frac{\bar{c} - (r + \lambda)\bar{V}_v}{(\sigma - 1)\hat{c}}
\]

where \(\bar{V}_v = E(V_v(z)|z \geq \hat{z})\).

Equation (14) highlights the distinction between this model and those considered by HKV. Allowing free-entry and ex post realization of \(z\), converts this model in to essentially that in their technical appendix. Setting \(V_v(z) = 0\) for all \(z\) reduces equation (14) to \(\bar{Y}/\hat{Y} = \hat{c}/\hat{c}\). Then as \(\sigma\) approaches 1 from above, it is simple to show that \(\bar{w}/\hat{w}\) approaches \(\bar{Y}/\hat{Y}\) from below. This implies an upper bound for \(\bar{w}/\hat{w}\) of 1.14. Simply incorporating cost of effort into their framework is not enough to generate the kind of wage dispersion that can emerge from the current model. Rather, here, high productivity firms are the sole proprietors of their technology and receive rents which are reflected in their continuation values, \(V_v(z)\), and are therefore excluded from the match surplus. The implied high value of \(\bar{V}_v\), means that with \(\bar{Y}/\hat{Y}\) fixed at 1.14, \(\hat{c}/\hat{c}\) will be large and therefore so will \(\bar{w}/\hat{w}\).

### 4.4 Competing sources of wage dispersion

As we have seen, the central mechanism by which productivity dispersion translates into wage dispersion in the model is through endogenous effort. An alternative theory of wage dispersion comes from on-the-job search. While
this can generate wage dispersion in the absence of firm heterogeneity, Burdett and Mortensen [1998] show that on-the-job search does interact with productivity dispersion. For clarity and modelling parsimony this paper has focussed on endogenous effort alone but from a quantitative perspective one might ask what is its relative contribution to the firms’ component of wage dispersion.

A circumstantial yet compelling argument for the role of endogenous effort comes from Postel-Vinay and Robin [2002]. They estimate a version of the Burdett and Mortensen [1998] model on a matched panel data set from France. They are able to decompose wage dispersion into firm, worker and search effects. They estimate their model separately for 7 classes of worker differentiated by the level of human capital associated with their jobs. The firm and search effects are significant for all skill levels but the individual worker effects disappear as the complexity of the tasks diminish. At every skill level they find that to fit wage dispersion, the discount rate has to be very large – in the order of 40% annually. With a more realistic, say 5%, discount rate the implied lifetime earnings differential from moving to a higher productivity firm would generate more movement in their model than appears in their data. Essentially, what is happening is that any remaining mis-specification of the model is showing up in the discount rate. As all of the jobs are in the Paris metropolitan area, it seems improbable that the missing component of the model is switching costs. As firms are permitted to match outside offers in the model, firm-specific human capital will show up in the measured wages. Disutility of work positively correlated with the firm effect is, therefore, a likely candidate for the missing component.
5 Conclusion

This paper explores the role of endogenous worker effort in addressing the relationship between firm level TFP dispersion and wage dispersion. By incorporating physical capital investment, it is able to provide a quantitatively important mechanism for understanding the size of firm effects in measured wage dispersion. Here, wage dispersion emerges due to differences in firm productivity and relies on a simple individual labor supply mechanism that permits high productivity firms to offer higher wages in return for greater effort from workers. In equilibrium, the $M_m$ ratio of worker utility and the gains from trade remain small. As the ratio is a relative measure of dispersion, however, amplification alone may not bring about a large $M_m$ ratio in wages. As sole proprietors of their superior technologies, high productivity firms receive rents which come partly from the high levels of effort they can require of workers in exchange for high wage jobs. This means that higher productivity firms receive a disproportionately larger share of the net revenue (revenue minus costs incurred). The narrow range of worker utility effectively pins down the lowest wage in the distribution and higher wages simply compensate workers for their extra efforts.

Analytical parsimony has been a motivating factor in restricting analysis to endogenous worker effort within a sequential search framework. The simulation is able to demonstrate the potential for this mechanism to explain a significant share of the firms’ component of measured wage dispersion. To enquire further into the quantitative importance of endogenous worker effort a possibility is to incorporate it into a structural model along the lines of Postel-Vinay and Robin [2002].
6 Appendix

6.1 Proof of Claim 2

To ensure applicability to the simulated model, the proof incorporates the discount rate, $r$. For expositional brevity I will solve the dual of (8):

$$V_u^* = \max_{k, \omega} V_u(\omega) \text{ subject to: } V_f(\omega; z, k) - k = V_v(z) \quad (15)$$

That reversing the problem is possible in this way is well known in the directed search literature (see Rogerson et al [2005]). To see that the result extends to the current environment consider the Lagrangians of each problem. They are,

$$L_f = V_f(\omega; z, k) - k + \mu_f [V_u(\omega) - V_u^*]$$

and

$$L_w = V_u(\omega) + \mu_w [V_f(\omega; z, k) - k = V_v(z)]$$

where $\mu_i, i = w, f$ are the Lagrange multipliers. After recognizing that $\mu_w = 1/\mu_f$ the first-order conditions are identical.

Substituting for the value functions problem (15) becomes,

$$V_u^* = \max_{k, \omega, \theta} \frac{m(\theta) [w - c(e)] + (r + \delta + \lambda)b}{(r + \delta)(r + \delta + \lambda + m(\theta))}$$

subject to

$$\frac{m(\theta) [zf(k, e) - w]}{(r + \lambda) [(r + \delta + \lambda)\theta + m(\theta)]} - k = V_v(z).$$

Eliminating the wage leads to

$$V_u^* = \max_{k, e, \theta} \frac{m(\theta) [zf(k, e) - c(e)] - (r + \lambda) [(r + \delta + \lambda)\theta + m(\theta)] (k + V_v(z))}{(r + \delta)(r + \delta + \lambda + m(\theta))}.$$  \quad (17)

The first order conditions for an interior solution with respect to $k, e$ and $\theta$ imply

$$m(\theta) [zf(k, e) - (r + \lambda)] - (r + \lambda)(r + \delta + \lambda)\theta = 0 \quad (18)$$
\[ zf_e(k, e) - c'(e) = 0 \quad (19) \]
\[ m'(\theta) [zf(k, e) - c(e) - b] - (r + \lambda) [r + \delta + \lambda + m(\theta) + m'(\theta)(1 - \theta)] (k + V_v(z)) = 0. \quad (20) \]

The objective function in Problem (17) is clearly concave in \((k, e)\) which means that for any given \(\theta\) there exists a unique solution, \((k(\theta), e(\theta))\) to the system of equations (18) and (19). Substituting these into LHS (20) and differentiating with respect to \(\theta\) yields

\[
m''(\theta) [zf(k, e) - c(e) - b] - (r + \lambda) m''(\theta)(1 - \theta)(k + V_v(z)) \\
+ m'(\theta) [zf_e(k, e) - c'(e)] \frac{de(\theta)}{d\theta} \\
+ \{m'(\theta)zf(k, e) - (r + \lambda) [r + \delta + \lambda + m(\theta) + m'(\theta)(1 - \theta)]\} \frac{dk(\theta)}{d\theta} \quad (21)
\]

Using (19) the second line of (21) is zero. Using (18) the contents of the curly brackets becomes

\[-(r + \lambda)(r + \delta + \lambda + m(\theta))/m(\theta).\]

Furthermore, if we impose (20) to be true then the first line of (21) becomes

\[
\frac{m''(\theta)(r + \lambda)(r + \delta + \lambda + m(\theta))}{m'(\theta)}(k + V_v(z)).
\]

Under this restriction (21) reduces to

\[
(r + \lambda)(r + \delta + \lambda + m(\theta)) \left[ \frac{m''(\theta)(k + V_v(z))}{m'(\theta)} - \frac{(m(\theta) - \theta m'(\theta))}{m(\theta)} \frac{dk(\theta)}{d\theta} \right]. \quad (22)
\]

Then if \(m(\theta) \equiv \bar{m}\theta^\eta\), with \(\eta \in (0, 1)\) (22) further reduces to

\[-(r + \lambda)(1 - \eta)(r + \delta + \lambda + \bar{m}\theta^\eta) \left[ \frac{(k + V_v(z))}{\theta} + \frac{dk(\theta)}{d\theta} \right]. \quad (23)\]

In general, \(\frac{dk(\theta)}{d\theta} < 0\) so (23) cannot be signed. However, if we further impose that \(c(e) \equiv \bar{c}e^\sigma\) with \(\sigma > 1\) and \(f(k, e) \equiv k^\alpha e^{1-\alpha}\) with \(\alpha \in (0, 1)\) then equations (18) and (19) can be solved for \(k\) and differentiated. This leads to

\[
\left[ \frac{k}{\theta} + \frac{dk(\theta)}{d\theta} \right] = \frac{\bar{m}\theta^\eta(\sigma - 1)(1 - \alpha) + (r + \delta + \lambda)[\eta(\sigma - 1 + \alpha) + \sigma\alpha]}{\theta[(r + \delta + \lambda)\theta + \bar{m}\theta^\eta](\sigma - 1)(1 - \alpha)} \frac{k}{\theta} > 0.
\]

26
So, as $V_v(z) \geq 0$, LHS (20) seen as a function of $\theta$ is downward sloping at any point for which equation (20) holds. As LHS (20) is continuous in $\theta$, it follows that after taking account of the effect of $\theta$ on $k$ and $e$, the maximand in (17) is quasiconcave in $\theta$. The unique solution to (20) represents a maximum.

7 References


Masters, A. [2013] “The value of leisure or disutility of work: Wage dispersion in a model of search with endogenous effort.” mimeo, University at Albany, SUNY.


29