Matching, Sorting and Wages

Jeremy Lise∗ Costas Meghir† Jean-Marc Robin‡

February 3, 2008
Preliminary and Incomplete

Abstract

This paper develops an empirical model of employer-employee matching with search frictions and heterogeneous agents to address the empirical question: How much sorting is there in the labour market with respect to unobserved worker and firm characteristics? The only empirical evidence we currently have on this issue comes from estimates of Mincer type wage regressions which include a fixed worker effect and a fixed firm effect. The correlation between the estimated worker and firm fixed effects can be calculated and, across a surprising number of countries and data sets, has been found to be either zero or negative. This result has been widely interpreted as indicating a lack of sorting in the labour market, which is consistent with a lack of complementarity between workers and firms in production. In this paper we show that the sign of the correlation between estimated fixed effects is not necessarily informative on the degree of sorting, and depends crucially on the production function. Indeed, we provide examples in which the production function is supermodular, inducing strong positive sorting between workers and firms, and yet the correlation between estimated fixed effects is negative.

Our empirical model is a generalization of the matching model with two-sided ex-ante heterogeneity of Shimer and Smith (2000) to include on-the-job search, counter-offers from firms whose workers are being poached and endogenous job destruction due to firm productivity shocks. Given matched employer-employee data the primitives to be estimated are the production function, the search frictions, the shock process on firm productivity, and the underlying distributions of worker and firm types. Given these primitives, we can calculate the degree of sorting in the market, as well as the welfare loss due to workers and firms consummating “sub optimal” matches.

∗University College London and IFS
†University College London and IFS
‡Paris School of Economics and University College London
1 Introduction

This paper develops an empirical model of employer-employee matching with search frictions and heterogeneous agents. Matching models of the labour market have become standard in the macroeconomic literature since the seminal works of Diamond (1982), Mortensen (1982) and Pissarides (1990). However, matching models with heterogeneous workers and firms are a relatively new subject of interest. Marriage models with heterogeneous agents in a frictional environment are studied in Sattinger (1995), Lu and McAfee (1996), Shimer and Smith (2000), and Atakan (2006).\textsuperscript{1} To the best of our knowledge, there has not yet been any empirical applications of assignment models with transferable utility in a frictional environment with heterogeneous agents.

There is a large body of empirical evidence showing that wages differ across industries, thus indicating that a matching process is at work in the economy (see for example Krueger and Summers, 1988). Static, competitive equilibrium models of sorting (Roy models) have been estimated by Heckman and Sedlacek (1985) and Heckman and Honore (1990), and Moscarini (2001) and Sattinger (2003) explore theoretical extensions of Roy models with search frictions. However, few characteristics of workers and jobs are recorded in available data sets. In general, workers differ by the numbers of years of education and experience, and jobs differ by the type of industry. There is thus an enormous amount of differences between workers and between jobs that are not accounted for by observables in the data.

How much sorting is there with respect to these unobserved characteristics? We are aware of only one piece of empirical evidence on that question. Abowd, Kramarz and Margolis (2000) and Abowd, Kramarz, Lengermann and Roux (2003) use French and U.S. matched employer-employee data to estimate a static, linear log wage equation with employer and worker fixed effects (by OLS). They find a small, and if any negative, cross-sectional correlation between firm and worker fixed effects. Abowd, Kramarz, Lengermann and Perez-Duarte (2004) docu-

\textsuperscript{1}Sattinger develops a framework but does not prove the existence of an equilibrium. Lu and McAfee prove the existence for a particular production function ($f(x, y) = xy$). Shimer and Smith prove the existence of an equilibrium in a more general setup and derive sufficient conditions for assortative matching. Atakan shows that Becker’s (1973) complementarity condition for positive sorting is sufficient if there exist explicit search costs.
ment the distribution of these correlations calculated within industries. In the U.S. 90% of these
 correlations range between -15% and 5%, and in France between -27% and -5%. The slight
 shift toward negative numbers is troubling. The authors acknowledge the possible existence of
 a negative bias if job-to-job mobility is limited. Nevertheless, Abowd et al. are certainly right
to conclude that there is no evidence of positive correlation between person and firm effects.
Whether this indicates a lack of positive sorting based on unobservables is a different story. The
person and firm effects which are estimated from the linear log wage equation are complicated
transformations of the underlying individual-specific, unobserved characteristics. A structural
model is thus required to recover the true underlying distribution.\(^2\)

The aim of this paper is to develop a structural matching model of heterogeneous workers
with frictions and readdress the empirical question raised by Abowd et al. (2004). The model
is similar to Shimer and Smith’s but differs in important aspects. Workers and Firms differ in
only one, continuous dimension of heterogeneity. Firms’ production function have constant
returns to scale in labour, so matches can be described by pairs of workers and employers.
Contrary to Shimer and Smith, we make workers and employers essentially different. Workers’
characteristics do not change over time whereas firms’ contributions to match productivity
fluctuate. Workers search on the job and may leave their employers to form better matches,
whereas firms cannot replace employees. Finally, we try to model a process of job entry and
exit.

2 An Overview of the Model

In what follows we provide an intuitive explanation of how the model is setup, and the motiva-
tion for each component. We then follow this up with a formal description of the model.

In the economy there is a fixed number of individuals and a fixed number of occupations or
production lines. Individuals have linear utility in income and the employers managing occupa-
tions care about profits. Individuals may be matched with an occupation and thus working, or

\(^2\)Abowd et al. refer to Shimer (2005) as an explanation of this empirical fact. Nevertheless, Shimer’s model
predict positive sorting, however imperfect. Abowd et al.’s explanation of the puzzle is not convincing.
they may be unemployed job seekers. Occupations on the other hand may be in three different states. First they may be matched with a worker, in which case they are producing. Second, they may be vacant and waiting for a suitable worker. Finally production lines may be inactive, and thus potential entrants in the labour market.

Both workers and occupations are heterogeneous with respect to a productivity relevant characteristic; output depends on the characteristic of both sides. Crucially though, productivity follows a first order Markov process, which leads to the value of the match changing, with consequences for worker mobility, job creation, and job destruction that are at the centre of our model.

When an occupation and a worker meet and the total match surplus is positive, the sharing of this surplus is agreed upon based on Nash asymmetric bargaining, where workers and firms have some exogenous bargaining power. Finally we close the model by a free entry condition: all production lines, whether active or not have a productivity relevant parameter, which they know. This determines whether they will want to enter the market and post a vacancy. The marginal occupation has zero surplus from entering the market and posting a vacancy.

To derive the implications of the model, solve it and ultimately apply it to data we proceed by defining the value functions for the workers and the firms. We use the Nash bargain to define the pay policy of the firm and hence derive how wages depend on heterogeneity and on productivity shocks. We use the free entry condition to determine the distribution of active matches and firms. In what follows we describe formally the model.

3 The Formal Description of the Model

3.1 Setup

Each individual worker is characterised by a permanent productivity relevant characteristic which we denote by $x$. We assume that $x$ has bounded support defined by $x \in [\underline{x}, \overline{x}]$. The distribution of worker heterogeneity in the population is $L(x)$ and we assume it possesses density $\ell(x)$. There are $\mathcal{L}$ individuals of which $\mathcal{U}$ are unemployed. We denote the unemployment
rate by \( \bar{\pi} \equiv \bar{U}/\bar{L} \). We also denote by \( U(x) \) the (endogenous) distribution of \( x \) among the unemployed.

Occupations are characterised by a productivity parameter \( y \) with bounded support \([\underline{y}, \bar{y}]\). The (stationary) distribution of occupation productivity in the population of firms, whether vacant, matched, or inactive is denoted by \( N(y) \) and possesses a density \( n(y) \). There are \( \bar{N} \) posts in the economy and the (endogenous) distribution of vacant posts is \( V(y) \) with density \( v(y) \). The number of vacancies is denoted by \( \bar{V} \). The number of inactive posts, i.e. potential posts for which firms have not advertised a vacancy, is \( \bar{I} \). The endogenous distribution of \( y \) among these posts is \( I(y) \) and similarly the density is denoted by \( i(y) \).

In a given match, \( y \) fluctuates according to a jump process. \( \delta \) is the instantaneous arrival rate of jumps and \( Q(y'|y) \) is the (Markov) transition probability for \( y \).

A match between a worker \( x \) and a firm \( y \) produces a flow of output \( f(x, y) \).

We denote the distribution of existing matches by \( H(x, y) \): (with density: \( h(x, y) \)). We can relate the density of individual productivities to the density of active matches as well as the density of productivities for the unemployed by

\[
(L - U) \int h(x, y) \, dy + Uu(x) = L\ell(x). \tag{1}
\]

Similarly we can write an equivalent relationship between the distribution of firm productivities, active matches, unfilled vacancies and inactive firms

\[
(N - V - I) \int h(x, y) \, dx + Vv(y) + Ii(y) = Nn(y). \tag{2}
\]

In both cases the relationship is essentially an accounting identity. Finally, matches can end both endogenously, as we characterise later and exogenously. We denote by \( \xi \) the instantaneous rate of exogenous job destruction.

We now discuss the process by which workers get to know about vacant occupations.

We assume that the unemployed workers search for work at a fixed intensity \( s_0 \). The search intensity for an employed worker is \( s_1 \). The process of search leads to a total number of meet-
ings, that as usual depends on the number of posted vacancies as well as on the number of total searchers in the economy, weighted by their search intensities. This matching function is denoted by \( M \left( s_0 \bar{U} + s_1 \left( \bar{L} - \bar{U} \right), V \right) \). Given this we can define by \( \lambda \) the instantaneous probability that a searching worker applies to a vacant job, i.e.

\[
\lambda = \frac{M \left( s_0 \bar{U} + s_1 \left( \bar{L} - \bar{U} \right), V \right)}{s_0 \bar{U} + s_1 \left( \bar{L} - \bar{U} \right)}.
\]

Thus \( \lambda_0 = s_0 \lambda \) is the contact rate for the unemployed and \( \lambda_1 = s_1 \lambda \) is the contact rate for employees. Finally it is useful to define the labour market tightness parameter as the number of vacancies relative to the effective number of workers searching, i.e.

\[
\theta = \frac{V}{s_0 \bar{U} + s_1 \left( \bar{L} - \bar{U} \right)}.
\]

### 3.2 Match formation and rent sharing

The value of an \((x, y)\) match to a worker of ability \(x\), matched with an occupation of productivity \(y\), at wage \(w\) is denoted as \( W_1 (w, x, y) \). The value of an \((x, y)\) match to a firm of productivity \(y\) matched to a worker of ability \(x\), paying a wage \(w\) is denoted as \( \Pi_1 (w, x, y) \). Let \( W_0 (x) \) denote the present value of unemployment for a worker with characteristic \(x\), and let \( \Pi_0 (y) \) denote the present value of a vacancy. Define the “surplus” of a match \((x, y)\) as

\[
S(x, y) = \Pi_1 (w, x, y) + W_1 (w, x, y) - \Pi_0 (y)^+ - W_0 (x),
\]

where \( \Pi_0 (y)^+ = \max \{ \Pi_0 (y), 0 \} \). Notice that it may well be that a match \((x, y)\) yields a positive output \(f(x, y)\) but the cost of a vacancy exceeds the expected profit. In this case, \( \Pi_0 (y) < 0 \). However, from the point of view of the employer, any negative profit implies the same decision: exiting the market, that is pay no vacancy cost and get zero chance of hiring a worker.

We assume that incumbent employers match outside offers. A negotiation game is then
played between the worker and both firms as in Cahuc, Postel-Vinay, Robin (2006). If a worker \( x \) currently paired to a firm \( y \) finds an alternative occupation \( y' \) such that \( S(x, y') > S(x, y) \), the worker moves to the alternative occupation and the new employer signs with the worker a contract that is worth the value of the total surplus of the previous \((x, y)\) match, plus a fraction \( \beta \) of the quasi rent:

\[
W_1^*(x, y'|y) - W_0(x) = S(x, y) + \beta [S(x, y') - S(x, y)]
\]

or, equivalently,

\[
W_1^*(x, y'|y) - W_0(x) = \beta S(x, y') + (1 - \beta) S(x, y).
\]  

Notice that the present value of the new wage contract \( W_1^*(x, y'|y) \) does not depend on the last wage contracted with the incumbent employer. It depends on the total surplus of the previous match and the total surplus of the current match, both of which are independent of the previous wage.

Next, consider the case where \( W_1^* - W_0(x) < S(x, y') \leq S(x, y) \), where \( W_1^* > W_0(x) \) is the current value of the wage contract that \( x \) and \( y \) have agreed upon. The worker uses the external offer to obtain a wage rise up to \( W_1^*(x, y'|y) - W_0(x) = S(x, y') \).

If \( S(x, y') \leq W_1^* - W_0(x) \), nothing happens. The worker gains nothing from the competition between \( y \) and \( y' \).

When an unemployed worker \( x \) finds a vacant occupation \( y \), then the match is formed if and only if \( S(x, y) > 0 \), and the surplus is split according to Nash bargaining:

\[
W_1^*(x, y|0) - W_0(x) = \beta S(x, y).
\]

This equivalence result between Nash bargaining and rent splitting is not subject to Shimer’s (2006) critique because the continuation value for workers when the match is destroyed—including when the worker changes occupation through on-the-job search—is not a function of
the last negotiated contract.

We now denote by $\mathcal{M}_0(x)$ the set of occupations $y$ that are acceptable for an unemployed worker:

$$\mathcal{M}_0(x) = \{ y | S(x, y) > 0 \}.$$

Let also $\mathcal{M}_1(x, y)$ be the set of occupations $y'$ such that match $\{x, y'\}$ can be formed and is preferred to an acceptable current match $\{x, y\}$:

$$\mathcal{M}_1(x, y) = \{ y' | S(x, y') > S(x, y) \}.$$

The complement of $\mathcal{M}_0(x)$ is denoted by $\overline{\mathcal{M}}_0(x) = [\underline{y}, \overline{y}] \setminus \mathcal{M}_0(x)$ and of $\mathcal{M}_1(x, y)$ by $\overline{\mathcal{M}}_1(x, y) = [\underline{y}, \overline{y}] \setminus \mathcal{M}_1(x, y)$.

### 3.3 Value functions

The next step in solving the model is to characterise the value functions of workers and firms in turn. These define the decision rule for each agent.

Let $r$ denote the discount rate. A match is destroyed when it is hit by a productivity shock that changes $y$ into some unacceptable $y'$, or some other adverse shock (at rate $\xi$). Because it may be welfare improving to reduce equilibrium unemployment, we allow for a tax $\tau$ (experience rating) on endogenous separations.

**Unemployed workers.** The present value of unemployment to a worker of type $x$ is $W_0(x)$ that satisfies the option value equation:

$$rw_0(x) = b + \lambda_0 \int_{\{y | S(x, y) \geq 0\}} [W_1^*(x, y|0) - W_0(x)] dV(y),$$

where we denote $V(A) = \int_A dV(y')$.  

7
Employed workers. Expected flow value of an \((x, y)\) match paying wage \(w\) to an employed worker

\[
r W_1 (w, x, y) = w + \left[ \delta Q \left( \mathcal{M}_0 (x) | y \right) + \xi \right] \left[ W_0 (x) - W_1 (w, x, y) \right] \\
+ \delta \int \{ y' | W_1 (w, x, y) - W_0 (x) > S(x, y') \geq 0 \} \left[ W^*_1 (x, y'|0) - W_1 (w, x, y) \right] dQ (y'|y) \\
+ \lambda_1 \int \{ y'|S(x,y)>S(x,y')>W_1 (w,x,y)-W_0 (x) \} \left[ W^*_1 (x, y'|y') - W_1 (w, x, y) \right] dV (y') \\
+ \lambda_1 \int \{ y'|S(x,y')>S(x,y) \} \left[ W^*_1 (x, y'|y) - W (w, x, y) \right] dV (y')
\]

(7)

The first term on the right hand side is the wage, followed by the cost of negative employment shocks, namely the possibility of an adverse shock to productivity making the match not viable and the possibility of exogenous job destruction. The second line represents the impact of a productivity shock on the value of employment, given that the match is not destroyed. Note that when \(S(x, y') > W^*_1 - W_0 (x)\) then the current contract remains acceptable by both parties. Either the firm or the worker will thus refuse to renegotiate. When \(W^*_1 - W_0 (x) > S(x, y') > 0\) then the firm is better off firing the worker, and both accept to renegotiate as \(S(x, y') > 0\). The final two terms represent the possibility of raising on offer from a \(y'\)-firm.

This equation allows us to compute the optimal wage contract \(w\) given \((x, y, y')\).

Vacant firms. The expected flow value of an unmatched post is:

\[
r \Pi_0 (y) = -c + \frac{\lambda_0 U}{V} \int \{ x | S(x, y) > 0 \} \left[ \Pi^*_1 (x, y|0) - \Pi_0 (y) \right] dU (x) \\
+ \lambda_1 \frac{(L - \U)}{V} \int \{ (x, y') | S(x, y) > S(x, y') \} \left[ \Pi^*_1 (x, y'|y') - \Pi_0 (y) \right] h (x, y') dxdy' \\
+ \delta \int \left[ \Pi_0 (y')^+ - \Pi_0 (y) \right] dQ (y'|y),
\]

(8)

where \(c\) is a fixed cost of posting the vacancy; the second term is the flow of benefits from matching with a previously unemployed worker; the third term is the flow of benefits from poaching a worker who is already matched with another firm; the final term reflects the impact
of a change in productivity from $y$ to $y'$.

We assume that only occupations with $\Pi_0(y) > 0$ are active (vacant). If the option value of posting a vacancy is not greater than the cost $c$, then a vacancy is not posted and the occupation has no chance of being filled by a worker. Assuming that $\Pi_0(y)$ is increasing in $y$, we define the threshold level of firm productivity $y^*$ by

$$\Pi_0(y^*) = 0,$$

the marginal firm active in the market is just indifferent between posting a vacancy and remaining inactive.

**Producing firms** the present flow value of an $(x, y)$ match to a firm paying wage $w$

$$r\Pi_1(w, x, y) = f(x, y) - w + \xi [\Pi_0(y)^+ - \Pi_1(w, x, y)]$$

$$+ \delta \int_{\{y' | S(y' | x) < 0\}} [\Pi_0(y')^+ - \Pi_1(w, x, y) - \tau] \, dQ(y'|y)$$

$$+ \delta \int_{\{y' | W_1(w, x, y) - W_0(x) > S(x, y') \geq 0\}} [\Pi_1^*(x, y'|0) - \Pi_1(w, x, y)] \, dQ(y'|y)$$

$$+ \delta \int_{\{y' | S(y, y') \geq W_1(w, x, y) - W_0(x) \geq 0\}} [\Pi_1(w, x, y') - \Pi_1(w, x, y)] \, dQ(y'|y)$$

$$+ \lambda_1 \int_{\{y' | S(x, y') > S(x, y) \geq W_1(w, x, y) - W_0(x)\}} [\Pi_1^*(x, y'|y') - \Pi_1(w, x, y)] \, dV(y')$$

$$+ \lambda_1 \int_{\{y' | S(x, y') > S(x, y)\}} [\Pi_0(y)^+ - \Pi_1(w, x, y)] \, dV(y')$$

(10)

### 3.4 The match surplus

Making use of the definition of the match surplus (3), the surplus sharing rules (4) and (5), and the value functions (6), (7), (8), and (10), we can write the surplus of an $(x, y)$ match, for
For low levels of firm productivity, \( y < y^* \), the surplus functions is defined by

\[
[r + \delta + \xi + \beta \lambda_1 V (M_1 (x, y))] S (x, y) = f (x, y) + \delta \int_{\{y' | S(x, y') \geq 0\}} S (x, y') dQ (y'|y) \\
- \delta Q (M_0 (x) |y) \tau + \lambda_1 \beta \int_{\{y' | S(x, y') > S(x, y)\}} S(x, y')dV(y') \\
- b - \lambda_0 \beta \int_{\{y | S(x, y) \geq 0\}} S(x, y')dV(y') \\
+ c - \frac{\lambda_0 U}{V} (1 - \beta) \int_{\{x' | S(x', y) \geq 0\}} S(x', y)dU(x') \\
- \frac{\lambda_1 (L - U)}{V} (1 - \beta) \int_{\{x', y' | S(x', y') > S(x, y)\}} [S(x', y) - S(x', y')] dH(x', y').
\]

(11)

For low levels of firm productivity, \( y < y^* \), the surplus functions is defined by

\[
[r + \delta + \xi + \beta \lambda_1 V (M_1 (x, y))] S (x, y) = f (x, y) + \delta \int_{\{y' | S(x, y') \geq 0\}} S (x, y') dQ (y'|y) \\
- \delta Q (M_0 (x) |y) \tau + \lambda_1 \beta \int_{\{y' | S(x, y') > S(x, y)\}} S(x, y')dV(y') \\
- b - \lambda_0 \beta \int_{\{y | S(x, y) \geq 0\}} S(x, y')dV(y') \\
- \delta \int [\Pi_0 (y')^+ - \Pi_0 (y)^+] dQ(y'|y).
\]

3.5 Steady-state flow equations.

We now proceed to define the steady state flow equations. The first relates to matches. The total number of matches in the economy will be

\[ L - U = N - V - I. \]

(12)

Existing matches, characterised by the pair \( \{x, y\} \), can be destroyed for a number of reasons. First, there is exogenous job destruction, at rate \( \xi \); second, with probability \( \delta \), the job component of match productivity changes to some value \( y' \) different from \( y \), and the worker may move to unemployment or may keep the job; third, the worker may change job, with probability \( \lambda_1 V (M_1 (x, y)) \)–i.e., a job offer has to be made (at rate \( \lambda_1 \)) and has to be acceptable \((y' \in M_1 (x, y))\). On the inflow side, new \( \{x, y\} \) matches are formed when some unemployed
or employed workers of type \(x\) match with vacant occupations \(y\), or when \(\{x, y\}'\) matches are hit with a productivity shock and exogenously change from \(\{x, y\}\) to \(\{x, y\}\).

In a steady state all these must balance leaving the match distribution unchanged. Thus formally we have for all \(\{x, y\}\) such that the match is acceptable, i.e. \(y \in \mathcal{M}_0 (x)\) or \(S(x, y) > 0\):

\[
[\delta + \xi + \lambda_1 V (\mathcal{M}_1 (x, y))] (L - U) h(x, y) = \delta (L - U) \int q(y|y') h(x, y') \, dy' \\
+ \left[ \lambda_0 U u (x) + \lambda_1 (L - U) \int_{\mathcal{M}_1(x,y)} h(x, y') \, dy' \right] v(y). \tag{13}
\]

We can obtain the total flow of workers between unemployment and employment by integrating out \(y\) and \(x\) in (13):

\[
\left[ \delta \int Q (\mathcal{M}_0 (x) | y) \, dH (x, y) + \xi \right] \cdot (L - U) = \lambda_0 \int V (\mathcal{M}_0 (x)) \, dU (x) \cdot U, \tag{14}
\]

which gives the steady state unemployment rate:

\[
\bar{u} = \frac{U}{L} = \frac{\delta \int Q (\mathcal{M}_0 (x) | y) \, dH (x, y) + \xi}{\delta \int Q (\mathcal{M}_0 (x) | y) \, dH (x, y) + \xi + \lambda_0 \int V (\mathcal{M}_0 (x)) \, dU (x)}
\]

Firms can choose to enter the labour market and post a vacancy, or they can choose to remain completely inactive and be potential entrants. The entry decision will depend on their current level of productivity \(y\), which in our model is stochastic. There is a productivity threshold level \(y^*\), determined endogenously and discussed in detail below, which separates firms active in the market (either matched or posting a vacancy) and those remaining inactive.

The following two equations, which follow directly from equations (2) and (9), define the number of vacancies and the number of inactive firms at each level of productivity \(y\).

\[
V_v(y) = \begin{cases} 
0, & y < y^*; \\
N n(y) - (L - U) \int h(x, y) \, dx, & y \geq y^*. 
\end{cases} \tag{15}
\]
\[ i(y) = \begin{cases} \overline{N}n(y) - (\overline{L} - \overline{U}) \int h(x,y)dx, & y < y^*, \\ 0, & y \geq y^*. \end{cases} \tag{16} \]

Note that there can be firms with productivity match \( y < y^* \) that are operating, i.e. matched with an employee because their productivity may have been reduced following a negative shock; at that point the value of remaining matched may continue to be positive, because the cost of posting a vacancy has been sunk and because firing costs are being avoided.

4 Equilibrium

Given values for the exogenous variables: \( \overline{L}, \delta, \xi, s_0, s_1, r, b, c, \) and \( \beta; \) the distributions of worker and firm types \( l(x) \) and \( n(y) \); the production function \( f(x,y) \); the transition function for productivity dynamics \( q(y'|y) \); and the matching function \( M(s_0 U + s_1 (L - U), V) \), the equilibrium involves the following endogenous objects

1. The distribution of active matches, \( h(x,y) \) (equation (13))

2. The latent value of posting a vacancy, \( \Pi(y) \) (equation (8))

3. The surplus function, \( S(x,y) \) (equation (11))

4. The number of unemployed workers, \( U \), (equation (14))

5. The number of vacant jobs, \( V \) (equation (12))

6. The threshold productivity for activity, \( y^*(equation (9)). \)

5 An Illustrative Numerical Example

In order to illustrate the properties and empirical implications of the model we solve and simulate under a particular set of parameters. For the numerical example we normalize \( \overline{L} = \overline{N} = 1 \) and assume that \( l(x) = n(y) \) and are \( U[0, 1] \); \( f(x,y) = xy \); the process governing the technological evolution is represented by a Gumbel copula with marginals equal to
Firm Productivity, $y$
Worker Productivity, $x$

(a) Matching set: $\{x, y | y \in M_0(x) \Leftrightarrow S(x, y) \geq 0\}$

(b) Stability of $x, y$ match. The contour lines indicate the probability the match will be destroyed given a technology shock: $Q(M_0(x) | y)$

Figure 1: Numerical Example: The Matching set

$n(y), q(y'|y) = n(y') \text{Gumble}(N(y'), N(y))$; $\delta = 0.01; \xi = 0.01; s_0 = 1.0; s_1 = 0.1$; $M(s_0 U + s_1 (L - U), V) = \alpha(s_0 U + s_1 (L - U))^{1-\gamma}$ with $\alpha = 0.3$ and $\gamma = 0.5; r = 0.03; b = c = 0; \text{and } \beta = 0.5$. The parameters are chosen so that the model nests the one used to produce the matching set in Figure 1 of Shimer and Smith (2000).

The parameter choice implies that $V = U$, and $y^* = y$. The key aspects of the equilibrium surplus function, $S(x, y)$, are represented by the matching sets in Figure 1, while the distribution of unemployed workers and of vacancies are depicted in Figures 3(a) and 3(c).

In Figure 1(a) we plot the equilibrium matching set: $\{x, y | S(x, y) \geq 0\}$. The asymmetry of the matching set is the result of on-the-job search and endogenous match destruction resulting from technology shocks. By comparison, if we set $s_1 = 0$ and $\delta = 0$ we have the environment studied in Shimer and Smith (2000), and replicate their Figure 1 here as Figure 2. In Figure 1(b) we illustrate the stability or fragility of active matches. The contour lines depict the probability that an active match will be destroyed in the event of a productivity shock, which arrives at rate $\delta$. This is the probability that after the technology shock the match will no longer be in the matching set: $Q(M_0(x) | y)$. The inner most contour represents a probability of 0.01 that
Figure 2: Shimer and Smith (2000) Figure 1. Nested in the current model when we shut off on-the-job search and productivity shocks: $s_1 = 0$ and $\delta = 0$.

A match on this contour would be destroyed due to a productivity shock. The lines continue in increments of 0.01 toward the boundary of the matching set, at which point the outermost contour represents 0.5. Matches in the large white portion of the matching set have an instantaneous destruction probability less than 0.023 ($\delta Q (\overline{M}_0 (x) | y) + \xi$) compared to a destruction probability of 0.035 for those at the boundary of the matching set.

Note the fact that for low values of worker productivity (below the 40th percentile) matches are only formed when $y$ is sufficiently similar to $x$. Additionally, for these low productivity workers, matches are destroyed when firm productivity becomes too low and when it becomes too high. For higher productivity workers (above the 40th percentile) there is a reservation firm productivity above which all matches are acceptable. Additionally, matches are only endogenously destroyed when the firm productivity falls too low.

In Figure 3 we plot several implications of the model at the current parametrization. The figures are striking in that neither the density of unemployed worker types nor the expected duration of unemployment are monotonic in skill. In Figure 3(a) we see that in the cross-section, unemployed workers are more likely to be low skilled than high skilled (although the relationship is not strictly monotonic). Additionally, in Figure 3(b) we see that low productiv-
ity workers can expect much longer unemployment durations than high productivity workers, although the shortest expected durations are enjoyed by those around the third decile of the productivity distribution (the non-monotonicity is masked here by the fact that the lowest productivity workers expect to experience extremely long periods of continuous unemployment).

The lowest productivity firms attract very few workers, and those who do accept jobs are quick to leave for better firms. This results in the high fraction of low productivity firms in the distribution of vacancies, and the very long vacancy durations before low productivity firms find a suitable worker, displayed in Figures 3(c) and 3(d).

In the absence of frictions workers would be paid their marginal product and matching would be perfectly assortative (Becker; 1973). Given the uniform distributions of worker and firm productivity, the wage distribution would be uniform, truncated below by the workers’ reservation wage: \( w \sim U[b, 1] \). In the current environment the wage distribution is far from uniform, even though the distribution of worker and firm productivity are uniform, as seen here in Figure 3(e). Finally, we see in Figure 3(f) that, as expected, the average productivity of workers is increasing in the productivity of firms and vise-verse.

We can confirm some intuition about the distributional effects of some simple policy variables by considering Figures 4(a), and 4(b). Taking the version of the model displayed in Figure 1(a) as the base case, we consider the effect of increasing (decreasing) non-employment benefits, \( b \), and taxing (subsidizing) match destruction through \( \tau \). Both a higher non-employment benefit and a tax on match destruction have the same distributional effect in steady-state, they eliminate all matches for low productivity workers and firms. This is illustrated in Figure 4(a). Similarly, lowering non-employment benefits or subsidizing match destruction increases the set of feasible matches for low productivity workers and firms, illustrated in Figure 4(b). It should be noted that these “policy experiments” are done in partial equilibrium where we have not resolved for the appropriate level of firm activity, \( y^* \).
Figure 3: Numerical Example: Model Implications
5.1 Actual Verses Estimated Sorting

As discussed in the introduction, Abowd et al use a simple empirical measure of sorting can be obtained by estimating a log-wage equation in which wages are a linear function of a worker fixed effect, a firm fixed effect, and an orthogonal worker-firm effect

$$\log(w_{it}) = x_{it}\beta + \alpha_i + \sum_{j=1}^{J} d_{itj}\psi_j + u_{it},$$

where $x_{it}$ are observables of worker, $\alpha_i$ is the worker fixed effect, $\psi_j$ is the firm fixed effect, and $u_{it}$ is an orthogonal residual. The correlation between $\hat{\alpha}_i$ and $\hat{\psi}_{j(i)}$ in a given match is taken as an estimate of the degree of sorting. Note that in terms of asymptotics, OLS estimate of $\beta$ is consistent as $i \to \infty$ for fixed $T$ and OLS estimates of $\alpha$ and $\psi$ are consistent when $T \to \infty$ faster than $I$ and $J$. In practice, the data contains millions of workers, tens of thousands of firms, and fewer than ten years. In Table 1 we compare the actual degree of sorting implied by the model with the degree of sorting estimated by calculating the correlation between estimated worker and firm fixed effects. We repeat this exercise for three different production functions,
Figure 5: The implications of different production functions for the matching set.

Table 1: Actual and Estimated Sorting

<table>
<thead>
<tr>
<th>Production Function</th>
<th>(corr(x, y))</th>
<th>(corr(\mu, \eta))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f(x, y) = xy)</td>
<td>0.6749</td>
<td>0.5205</td>
</tr>
<tr>
<td>(f(x, y) = x + y)</td>
<td>-0.2203</td>
<td>-0.2795</td>
</tr>
<tr>
<td>(f(x, y) = 1 - (x - y)^2)</td>
<td>0.5502</td>
<td>-0.0693</td>
</tr>
</tbody>
</table>

\(f(x, y) = xy\), \(f(x, y) = x + y\), and \(f(x, y) = 1 - (x - y)^2\). The matching set for the first of these production functions is illustrated in Figure 1 while the second two are illustrated in Figure 5. The first and last of these produce positive assortative matching, while the second produces negative assortative matching. While the sign of the correlation between the estimated fixed effects correctly picks up the positive/negative sorting in the first two cases, it is of opposite sign in the last case. This estimated correlation is not necessarily informative on the degree of sorting in the model. Indeed, this suggests the need to estimate the production function in order to answer the question regarding the degree of sorting on unobservables.

In addition to this effect, Postel-Vinay and Robin (2006) demonstrate that in practice, empirical estimates of sorting which are based on worker and firm fixed effects introduce a negative bias, which will introduce a spurious negative correlation when calculating the correlation between worker and firm fixed effects. This is illustrated as follows; empirically, \(\beta\) and \(\psi_j\) are
estimated from the within transformation

\[ \log w_{it} - \log w_i = (x_{it} - \bar{x}_i)\beta + \sum_{j=1}^{J} (d_{jt} - \bar{d}_i')\psi_j + u_{it} - \bar{u}_i. \]

This makes it clear that we need to see workers change firm to identify the firm fixed effects \( \psi_j \). The worker fixed effects are estimated as

\[ \hat{\alpha}_i = \log w_i - \bar{x}_i\hat{\beta} - \sum_{j=1}^{J} \bar{d}_i'\hat{\psi}_j. \]

Notice, any statistical error affecting the estimate of the firm effect translates directly to the estimate of the worker effect, with a sign reversal. OLS estimates of firm and worker effects are likely to be imprecise and spuriously negatively correlated given short time dimension and limited worker mobility.

6 Conclusion

The model we have outlined in this paper is rich enough to generate positive, zero, and negative sorting on unobservables. Given matched employer and employee data on wages, value added, and worker employment transitions, we are in a position to readdress the empirical question of how much workers and firm sort based on unobservables.
A Extended Derivations

A.1 The Surplus function is not a function of the wage contract.

Making use of the definition of the match surplus (3), the surplus sharing rules (4) and (5), and the value functions (6), (7), (8), and (10), we can write

\[
\begin{align*}
rS(x, y) &= r \left[ \Pi_1(w, x, y) + W_1(w, x, y) - \Pi_0(y)^+ - W_0(x) \right] \\
&= f(x, y) - w + \xi \left[ \Pi_0(y)^+ - \Pi_1(w, x, y) \right] \\
&\quad + \delta \int_{\{y'|S(y', x) < 0\}} \left[ \Pi_0(y')^+ - \Pi_1(w, x, y) - \tau \right] dQ(y'|y) \\
&\quad + \delta \int_{\{y'|W_1(w, x, y) - W_0(x) > S(x, y') \geq 0\}} \left[ \Pi_1^*(x, y') - \Pi_1(w, x, y) \right] dQ(y'|y) \\
&\quad + \delta \int_{\{y'|S(x, y') \geq W_1(w, x, y) - W_0(x) \geq 0\}} \left[ \Pi_1^*(x, y') - \Pi_1(w, x, y) \right] dV(y') \\
&\quad + \lambda_1 \int_{\{y'|S(x, y') > S(x, y') \geq W_1(w, x, y) - W_0(x) \}} \left[ \Pi_1^*(x, y') - \Pi_1(w, x, y) \right] dV(y') \\
&\quad + \lambda_1 \int_{\{y'|S(x, y') > S(x, y') \geq W_1(w, x, y) - W_0(x) \}} \left[ W_1^*(x, y') - W_1(w, x, y) \right] dV(y') \\
&\quad + \lambda_1 \int_{\{y'|S(x, y') > S(x, y') \geq W_1(w, x, y) - W_0(x) \}} \left[ W_1^*(x, y') - W_1(w, x, y) \right] dV(y') \\
&\quad + \delta - \frac{\lambda_0 U}{V} \int_{\{y'|S(x', y') > S(x', y') \geq 0\}} \left[ \Pi_1^*(x', y') - \Pi_0(y) \right] dU(x') \\
&\quad - \frac{\lambda_1 (L - U)}{V} \int_{\{x', y'|S(x', y') > S(x', y') \}} \left[ \Pi_1^*(x', y') - \Pi_0(y) \right] h(x', y') dx'dy' \\
&\quad - \delta \int_{\{y'|S(y', x') \geq 0\}} \left[ \Pi_0(y')^+ - \Pi_0(y) \right] dQ(y'|y) \\
&\quad - b - \lambda_0 \int_{\{y'|S(x, y') \geq 0\}} \left[ W_1^*(x, y') - W_0(x) \right] dV(y').
\end{align*}
\]
Collecting terms we have

\[ rS(x, y) \]

\[ = f(x, y) - w + w + \xi \left[ \Pi_0(y)^+ - \Pi_1(w, x, y) + W_0(x) - W_1(w, x, y) \right] \]

\[ + \delta \int_{\{y' | S(y', x) < 0\}} \left[ \Pi_0(y')^+ - \Pi_1(w, x, y) + W_0(x) - W_1(w, x, y) - \tau + \Pi_0(y)^+ - \Pi_0(y)^+ \right] dQ(y'|y) \]

\[ + \delta \int_{\{y' | W_1(w, x, y) - W_0(x) > S(x, y') \geq 0\}} \left[ \Pi_1^+(x, y'|0) - \Pi_1(w, x, y) + W_1^+(x, y'|0) - W_1(w, x, y) \right. \]

\[ + \Pi_0(y)^+ - \Pi_0(y)^+ + \Pi_0(y')^+ - \Pi_0(y')^+ + W_0(x) - W_0(x) \] \[ \left. \right] dQ(y'|y) \]

\[ + \lambda_1 \int_{\{y' | S(x, y') \geq W_1(w, x, y) - W_0(x) \geq 0\}} \left[ \Pi_1^+(x, y'|y) - \Pi_1(w, x, y) + W_1^+(x, y'|0) - W_1(w, x, y) \right. \]

\[ + \Pi_0(y)^+ - \Pi_0(y)^+ + W_0(x) - W_0(x) \] \[ \left. \right] dV(y') \]

\[ + \lambda_1 \int_{\{y' | S(x, y') > S(x, y) \}} \left[ \Pi_0(y)^+ - \Pi_1(w, x, y) + W_1^*(x, y'|y) - W_1(w, x, y) + W_0(x) - W_0(x) \right] dV(y') \]

\[ + c - \frac{\lambda_0}{V} \int_{\{x' | S(x', y) > 0\}} \left[ \Pi_1^*(x', y'|0) - \Pi_0(y) \right] dU(x') \]

\[ - \frac{\lambda_1}{V} \int_{\{(x', y') | S(x', y) > S(x', y') \}} \left[ \Pi_1^*(x', y'|y') - \Pi_0(y) \right] h(x', y') \ dx' \ dy' \]

\[ - \delta \int \left[ \Pi_0(y)^+ - \Pi_0(y) \right] dQ(y'|y) \]

\[ - b - \lambda_0 \int_{\{y' | S(x, y') \geq 0\}} \left[ W_1^*(x, y'|0) - W_0(x) \right] dV(y'). \]
Making use of the definition of the surplus: \( S(x, y) = \Pi_1(w, x, y) + W_1(w, x, y) - \Pi_0(y)^+ - W_0(x) = \Pi_1^* + W_1^* - \Pi_0(y)^+ - W_0(x) \) we can write

\[
rS(x, y) = f(x, y) + \xi S(x, y) - \delta \int_{\{y'|S(y',x) < 0\}} [\Pi_0(y')^+ - \Pi_0(y)^+ - S(x, y) - \tau] dQ(y'|y) + \delta \int_{\{y'|W_1(w,x,y) > W_0(x) > S(x,y')\geq 0\}} [S(x, y') - S(x, y) + \Pi_0(y')^+ - \Pi_0(y)^+] dQ(y'|y) + \delta \int_{\{y'|S(x,y') \geq W_1(w,x,y) - W_0(x) > 0\}} [S(x, y') - S(x, y) + \Pi_0(y')^+ - \Pi_0(y)^+] dQ(y'|y) + \lambda_1 \int_{\{y'|S(x,y') > S(x,y)\}} [W_1^*(x, y|y') - W_0(x) - S(x, y)] dV(y') + c - \frac{\lambda_0 U}{V} \int_{\{x'|S(x',y') > 0\}} [\Pi_1^*(x', y|0) - \Pi_0(y)] dU(x') - \frac{\lambda_1 (\mathcal{L} - \mathcal{U})}{V} \int_{\{x',y'|S(x',y') > S(x',y')\}} [\Pi_1^*(x', y|y') - \Pi_0(y)] h(x', y') dx' dy' - \delta \int [\Pi_0(y')^+ - \Pi_0(y)] dQ(y'|y) - b - \lambda_0 \int_{\{y'|S(x,y') \geq 0\}} [W_1^*(x, y|0) - W_0(x)] dV(y').
\]

Finally, making use of the surplus splitting in equations (4) and (5) we have

\[
rS(x, y) = f(x, y) + \xi S(x, y) - \delta \int_{\{y'|S(y',x) < 0\}} [S(x, y) + \tau] dQ(y'|y) + \delta \int_{\{y'|W_1(w,x,y) > W_0(x) > S(x,y')\geq 0\}} [S(x, y') - S(x, y)] dQ(y'|y) + \delta \int_{\{y'|S(x,y') \geq W_1(w,x,y) - W_0(x) > 0\}} [S(x, y') - S(x, y)] dQ(y'|y) + \lambda_1 \int_{\{y'|S(x,y') > S(x,y)\}} [S(x, y) + \beta (S(x, y') - S(x, y)) - S(x, y)] dV(y') + c - \frac{\lambda_0 U}{V} (1 - \beta) \int_{\{x'|S(x',y') \geq 0\}} S(x', y) dU(x') - \frac{\lambda_1 (\mathcal{L} - \mathcal{U})}{V} (1 - \beta) \int_{\{x',y'|S(x',y') > S(x',y')\}} [S(x', y') - S(x', y')] h(x', y') dx' dy' - b - \lambda_0 \beta \int_{\{y'|S(x,y') \geq 0\}} S(x, y') dV(y').
\]

Equation (11) then follows directly.
References


