Introduction

What we want to learn about:
- are better workers employed at more productive firms?
- what is the production function?
- is the allocation efficient, what prevents efficiency?
- can policies improve total output?

What we observe:
- matched employer-employee data
- a panel data with employment status, wage, firm identifier \( \{e_{it}, w_{it}, j_{it}\}_{it} \)

The difficulties:
- firm and worker productivities are not directly observed
- allocation is endogenous, sorting on unobservables?
- wage might depend on employment history
Literature on assignment models

- Labor market as an assignment model
  - mass of workers \((x)\) and mass of jobs \((y)\)
  - production function \(f(x, y)\)

- Becker (1974): friction-less
  - assignment is one-to-one
  - do not observe mismatch, can’t differentiate firm/worker effect
  - Choo and Siow (2006); Galichon and Salanié (2011) add preference heterogeneity

- Shimer and Smith (2003): derives condition for sorting in the presence of search frictions
  - agents settle for sub-optimal matches
  - still complementarities in production lead to PAM
Literature on identification in the presence of frictions

• Eeckhout and Kircher (2011)
  • wages are not monotonic in $y$, linear decomposition cannot identify the sorting pattern
  • without discounting, only strength of sorting can be estimated

• Hagedorn, Law, and Manovskii (2014)
  • uses property that wages rank workers within firms
  • provides a non-parametric estimation technique
  • demonstrates that full production function and sign of sorting can be recovered in practice even with small discounting

• Bagger and Lentz (2014)
  • model with endogenous search effort, no capacity constraint
  • shows identification, estimates the model on Danish data

• This paper:
  • introduces OTJ search in a model with capacity constraint
  • wages do not directly rank workers within firms, we need to work with present values
This paper

1. present an equilibrium search model that includes:
   - two sided heterogeneity
   - on the job search with Bertrand Competition
   - job creation and job filling
   - sorting due to capacity constraint and complementarity in production

2. develop constructive identification

3. simulation and preview of data
Model
Environment

- measure 1 of \textbf{workers} indexed by fixed ability \(x \in [0, 1]\)
  - risk neutral, discount at rate \(r\)
  - \(u(x)\) workers are unemployed
  - \(1 - u(x)\) workers are employed in a firm

- measure 1 of \textbf{firms} indexed by fixed technology and job creation cost \((y, \epsilon) \in [0, 1]^2\)
  - each firm employs measures \(h(x|y, \epsilon)\) of workers
  - output := \(\int f(x, y)h(x|y, \epsilon) \, dy\)
    - and owns masses \(v(y)\) of open vacancies
    - the measure \(\int h(x|y) \, dx + v(y)\) is endogenous
Environment

- measure 1 of workers indexed by fixed ability \( x \in [0, 1] \)
  - risk neutral, discount at rate \( r \)
  - \( u(x) \) workers are unemployed
  - \( 1 - u(x) \) workers are employed in a firm

- measure 1 of firms indexed by fixed technology and job creation cost \( (y, \epsilon) \in [0, 1]^2 \)
  - each firm employs measures \( h(x|y, \epsilon) \) of workers
    
    \[
    \text{output} := \int f(x, y) h(x|y, \epsilon) \, dy
    \]
  - and owns masses \( v(y) \) of open vacancies
  - the measure \( \int h(x|y) \, dx + v(y) \) is endogenous
Job/Vacancy creation

- firms can create a per period flow $n$ of vacant jobs at convex cost $c(n, y, \epsilon)$
- define $\mathcal{V}(y)$ as the present value of a vacancy
- firm $(y, \epsilon)$ optimally sets $n$:
  $$\sup_n n \cdot \mathcal{V}(y) - c(n, y, \epsilon)$$
- cost is independent of current size
- once created vacancies are added to the firm vacancy stock
Timing and meeting probabilities for un-matched agents

**timing for unemployed worker** $x$

1. receives flow value of unemployment $b(x)$
2. with pr. $\lambda g(z)v(y)$ finds an offer from firm $y$ with training cost $z$

**timing for vacancy** $y$

1. with pr. $\mu g(z)u(x)$ meets an unemployed worker $x$ with training cost $z$
2. with pr. $\kappa \mu g(z)h(x|y')$ meets a worker $x$ employed at $y'$ with training cost $z$
Timing and meeting probabilities for un-matched agents

**timing for unemployed worker** \( x \)

1. receives flow value of unemployment \( b(x) \)
2. with pr. \( \lambda g(z) v(y) \) finds an offer from firm \( y \) with training cost \( z \)

**timing for vacancy** \( y \)

1. with pr. \( \mu g(z) u(x) \) meets an unemployed worker \( x \) with training cost \( z \)
2. with pr. \( \kappa \mu g(z) h(x \mid y') \) meets a worker \( x \) employed at \( y' \) with training cost \( z \)
Timing and meetings within match

timing for match \((x, y)\) at wage \(w\):

1. collects output \(f(x, y)\) pays wage \(w\) to the worker
2. with pr. \(\delta\) job is destroyed, firm does not retain the vacancy
3. with pr. \(\lambda \kappa g(z')v(y')\) worker meets another firm \((y', z')\)
The firm and the worker sequentially agree on a wage $w$

- $U(x)$: lifetime utility when unemployed
- $V(y)$: present value of a vacancy
- $W(x, y, w)$: worker lifetime utility when employed at $(y, w)$
- $P(x, y)$: present value of a match

and the surplus of a match

$$S(x, y) := P(x, y) - U(x) - V(y)$$

- $S$ is not a function of $w$ because utility is transferable.
Matching outcomes

When unemployed meets an offer:
- worker $x$ meets firm $y$ and draws training cost $z$
- the match is created if $S(x, y) - z \geq 0$
- wage $w$ is set by generalized Nash bargaining:

$$W(x, y, w) = \beta (S(x, y) - z) + U(x)$$

When employed worker receives outside offer:
- worker $x$ employed by $y$ at $w$ meets firm $(y', z')$
- $y$ and $(y', z')$ enter Bertrand competition
- **poaching** if $S(x, y') - z' \geq S(x, y)$, worker gets full $(x, y)$ surplus

$$W(x, y', \omega) = S(x, y) + U(x)$$

- **wage raise** if $S(x, y') - z' \geq W(x, y', \omega) - U(x)$

$$W(x, y, \omega) = S(x, y') - z'$$
Equilibrium

Given primitives $f(x, y)$, $G(z)$, $c(n, y, \epsilon)$, $r, \beta, \mu, \lambda, \kappa, b, \delta$, a **Stationary Search Equilibrium** is characterized by distributions $h(x|y, \epsilon), u(x), v(y, \epsilon)$, firm job creating $n(y, \epsilon)$ and values $U(x), V(y)$ and $S(x, y)$ such that:

- $V(y), U(x), S(x, y)$ are the present values of a vacancies, unemployed worker and match surplus
- $n(y, \epsilon)$ solves optimal vacancy creation given $V(y)$
- $v(y), u(y)$ and $h(x|y)$ are implied by meeting rates, transition probabilities, $S(x, y)$ and $n(y, \epsilon)$
Equilibrium properties

1. If $f_x > 0$ then $U(x) \nearrow$ in $x$

2. If $f_y > 0$ then $V(y) \nearrow$ in $y$

3. Bertrand competition gives:

$$(r + \delta)S(x, y) = f(x, y) - rU(x) - (r + \delta)V(y)$$
Identification
Identification

- Consider random process $\Gamma_t = (X, E_t, R_t, J_t, Y_t), t \geq 1$ generated by the model
  - $(X, Y_t)$ are unobserved worker and firm types
  - $E_t$ is the employment status
  - $(R_t, J_t)$ are wages and firm ID whenever $E_t = 1$

- The econometrician is given $E_t$ and $P$ for any observables
  - for ex: $E[R_t|J]$, $E_t[R_t|E_t > E_{t-1}]$ or $P\{E_t < E_{t-1}\}$
Identification

• Assume that
  • $f(x, y)$ is differentiable and $f_x > 0, f_y > 0$
  • $c(n, y, \epsilon)$ is differentiable, convex in $n$ and $c(0, y, \epsilon) = 0$
  • $G(z)$ has full support on $[0, \infty)$ and is parametrized
  • $r$ is given
  • $\mathbb{E}_t$ and $\mathbb{P}$ are known for observables $(E_\tau, R_\tau, J_\tau)_{\tau > t}$
  • the total number of vacancies is known

• Then $f(x, y), G(z), c(n, \epsilon), \beta, \mu, \lambda, \kappa, b, \delta$ are identified
Overview

Constructive Identification:

1. get a measure of $x$ for each worker
2. get $U(x)$, $\beta, \kappa, \delta$, and $G(z)$
3. get a measure of $y$ for each firm and $v(y), \mu$
4. identify $S(x, y)$
5. construct $\mathcal{V}(y)$ and identify $f(x, y)$
6. identify $c(n, y, \epsilon)$
Estimation strategy in practice

Two important limitations:

- in practice the time dimension is short (10 to 20 years)
- using $S(x, J)$ requires a lot of $x$ workers in each firm $J$

We use a simplified algorithm for the estimation as an auxiliary model

- parametrize production function
- drop the second term in $V(y)$ (value of poaching)
- use $S$, $U$ and $Q(l|y)$ as a moments
Simulation for small sample performance

- ~ 40,000 workers, 10 years quarterly, 50 worker and firm types
- \( f(x, y) = (0.5\Phi^{-1}(x)^\sigma + 0.5\Phi^{-1}(y)^\sigma)^{1/\sigma} \) with 2 parametrization \( \sigma = \{ \text{pam: } -1, \text{ nam: } 2 \} \)
- \( \hat{x} \) and \( \hat{y} \) (SNR: x:0.96, y:0.95)
- \( \hat{U}(x) \) and \( \hat{S}(x, y) \)
- estimating complementarity:
Auxiliary model on the data

- Matched employer-employee data from Sweden
  - **today**: only male, college graduates under 50
  - 10 years, 424k individuals, 19k firms,
  - 265k j2j transitions, 158k u2e transitions

- Applying simplified procedure:
  - $U(x)$
  - $S(x, y)$
  - $h(x, y)$
Conclusion

- developed a model with 2 sided heterogeneity,
  - rich wage dynamics with OTJ
  - both job creation and job filling
- provided a constructive identification proof and preliminary simulation results
- direct non-parametric estimation seems difficult with 10 years of data
  - use rank aggregation (Hagedorn, Law, and Manovskii, 2014) to get more precise measurement
  - use NP as auxiliary OR use simulated method of moments
Parametrization Surplus

![Diagram of parametrization surplus](image)
Estimated $x$ versus true
Estimated $x$ versus true
Estimated $\mathcal{U}(x)$
Estimated $S(x, y) + U(x)$ using $\bar{w}(x, y)$
Estimated $\mathcal{S}(x, y) + \mathcal{U}(x)$ using $\bar{w}(x, y)$
Lessons from linear wage equation

- Abowd, Kramarz, and Margolis (1999); De Melo (2009)

\[
\log w_{it} = \beta X_{it} + \theta_i + \psi_{J(i,t)} + \epsilon_{i,t}
\]

- within 10 years panel, explains $\sim 85\%$ of earnings dispersion

- Firm share: \[ \frac{\text{var}(\psi_j)}{\text{var}(\psi_j) + \text{var}(\theta_i)} \approx 20\% \]

- Allocation to firms appears to be random \[ \text{Cov}(\theta_i, \psi) \approx 0 \]

- Workers cluster together \[ \text{Cov}(\theta_i, \bar{\psi}_{J(i,.)}) > 0 \]
Estimation for different countries

<table>
<thead>
<tr>
<th>Country</th>
<th>US 1(^{(a)})</th>
<th>US 2</th>
<th>FR</th>
<th>GE</th>
<th>IT</th>
<th>DE(^{(b)})</th>
<th>BR</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Var (x\beta))</td>
<td>0.03</td>
<td>0.14</td>
<td>0.02</td>
<td>—</td>
<td>0.01</td>
<td>—</td>
<td>0.02</td>
</tr>
<tr>
<td>(Var (\theta))</td>
<td>0.29</td>
<td>0.23</td>
<td>0.21</td>
<td>0.05</td>
<td>0.05</td>
<td>0.08</td>
<td>0.40</td>
</tr>
<tr>
<td>(Var (\psi))</td>
<td>0.08</td>
<td>0.053</td>
<td>0.08</td>
<td>0.013</td>
<td>0.01</td>
<td>0.00</td>
<td>0.18</td>
</tr>
<tr>
<td>(\frac{Var(\psi)}{Var(\theta+\psi)})</td>
<td>0.22</td>
<td>0.19</td>
<td>0.32</td>
<td>0.22</td>
<td>0.23</td>
<td>0.03</td>
<td>0.31</td>
</tr>
<tr>
<td>(Corr (\theta, \psi))</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.17(^{(c)})</td>
<td>0.40(^{(d)})</td>
</tr>
<tr>
<td>(Corr (\theta, \hat{\theta}))</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.85</td>
<td>0.93</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.89</td>
<td>0.9</td>
<td>0.84</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Sample Statistics

<table>
<thead>
<tr>
<th></th>
<th>90-99</th>
<th>84-93</th>
<th>76-87</th>
<th>93-97</th>
<th>81-97</th>
<th>94-03</th>
<th>95-05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nobs</td>
<td>37.7M</td>
<td>4.3M</td>
<td>5.3M</td>
<td>4.8M</td>
<td>—</td>
<td>6.9M</td>
<td>16.0M</td>
</tr>
<tr>
<td>Nworkers</td>
<td>5.2M</td>
<td>293K</td>
<td>1.2M</td>
<td>1.8M</td>
<td>1.7M</td>
<td>563K</td>
<td>2.0M</td>
</tr>
<tr>
<td>Nfirms</td>
<td>476K</td>
<td>80K</td>
<td>500K</td>
<td>1821</td>
<td>421K</td>
<td>53.6K</td>
<td>137K</td>
</tr>
<tr>
<td>% 1st Group(^{(e)})</td>
<td>—</td>
<td>99.1%</td>
<td>88.3%</td>
<td>94.9%</td>
<td>99.5%</td>
<td>—</td>
<td>98.6%</td>
</tr>
</tbody>
</table>
Becker friction-less assignment

\[ \frac{\partial^2 f(x, y)}{\partial x \partial y} > 0 \]

no frictions: one to one mapping
agents settle for lower than optimal match
wages are not monotonic in \( y \)
linear wage equation is mis-specified
Identifying worker type

- \( \frac{\partial f}{\partial x}(x, y) \geq 0 \) implies that \( \mathcal{U}(x) \) is increasing in \( x \)

- when the experienced worker extracts full surplus, the wage satisfies

\[
(r + \delta)\mathcal{S}(x, y) = \overline{w}(x, y) - (r + \rho)\mathcal{U}(x) + 0
\]

\[
\overline{w}(x, y) = f(x, y) - (r + \delta)\mathcal{V}(y) - 0 \quad \uparrow \quad \text{in } x
\]

- define \( \overline{R} := \max_t \{ R_t : E_t = 1 \} \) for each \( \omega \) then

\[
\forall \omega, \ X = Q_{\overline{R}}(\overline{R})
\]
Identifying $\mathcal{U}(x)$

- provided that $z \sim G(z)$ support is large enough, lowest accepted wage will happen for 0 surplus:

$$\mathcal{W}(x, y, \underline{w}_{u2e}(x, y)) - \mathcal{U}(x) = \beta (S(x, y) - z^*) = 0$$

- and so

$$\mathcal{U}(x) = \mathbb{E}_J \mathbb{E}_t [W_t \mid X = x, E_t > E_{t-1}, J_t = J, R_t = R_{\min}(x, J_t)]$$

- where $W_t := \sum_{\tau=t}^{\infty} \frac{R_{\tau}}{(1+r)^{\tau}}$

and $R_{\min}(x, J) := \min_{\omega \in \Omega, t \in T} \{R_t : E_t > E_{t-1}, J_t = J, X = x\}$
identifying $\beta$

- when worker $x$ leaves firm $J$ to another firm, he gets the full surplus:

$$\mathcal{W} = S(x, J) + U(x) = \mathbb{E} \left[ W_t | J_t \neq J_{t-1} = J, X \right]$$

- and when hired from unemployment and $z = 0$

$$\mathcal{W} = \beta S(x, J) + U(x)$$

- combining gives:

$$\beta = \mathbb{E}_{Jx} \left[ \frac{\mathbb{E}_t \left[ W_t | E_t > E_{t-1}, X = x, R_t = R_{\text{max}}(X, J_t), J_t = J \right] - U(x)}{\mathbb{E}_t \left[ W_t | J_t \neq J_{t-1} = J \right] - U(x)} \right],$$
Identifying $\delta$ and $\kappa$

- separation rate is exogenous so
  \[
  \delta = \frac{\mathbb{P}\{E_t < E_{t+1}\}}{\mathbb{P}\{E_t = 1\}}
  \]

- and when collecting $\mathcal{U}(x)$ all meetings will a change:
  \[
  \kappa = \frac{\mathbb{P}\{R_t > R_{t-1} \cup J_t \neq J_{t-1} | X, R_{t-1} = R_{\text{min}}(J, X)\}}{\mathbb{P}\{E_t > E_{t-1} | X\}},
  \]
We use the variation in the value out of unemployment

\[ \mathcal{W} = \beta(S(x, J) - z) + \mathcal{U}(x), \quad z \sim G(z) \]

for \( z \in [0, \max_x, J S(x, J)] \) we get:

\[
G(z) = \mathbb{E}_{X, J} \mathbb{P} \left\{ \frac{\mathbb{E}_t [W_t|X, E_t > E_{t-1}, J_t = J, R_t = w] - \mathcal{U}(X)}{\beta} - S(X, J) > z | X, J \right\}.
\]

but assuming that \( G(z) \) is parametrized “globally” it is enough
Identifying firm type

- we know that $V(y)$ is increasing in $y$, we compute the following:

$$
\hat{V}(J) = (1 - \beta) \int G [S(x, J)] u(x) \, dx \\
+ \kappa \int \int G [S(x, J) - S] f_{Sx}(S, x) \, dx \, dS.
$$

- and $F_{Sx}$ is joint distribution of $(S, x)$ in the population

$$
F(S, x) = \mathbb{P}\{X \leq x \cup S(X, J) \leq S\}.
$$

- the rank of $V(J)$ gives the rank among active jobs, we finish by measuring the vacancy distribution

$$
v(y) \propto \mathbb{P}\{E_t > E_{t-1}|X = x, J_t = J\}/G[S(x, y)],
$$

- total number of vacancies identifies $\mu$. 

> back
Identifying $S(x, y)$ and

- we now know $x$ and $y$ we can average over $j2j$ transitions

\[
S(x, y) = \mathbb{E} [W_t | J_t \neq J_{t-1}, Y_{t-1} = y, X = x] - \mathcal{U}(x)
\]
Identifying $\mathcal{V}(y)$ and $f(x, y)$

- we can reconstruct $\mathcal{V}(y)$ fully from definition

$$r\mathcal{V}(y) = (1 - \beta) \int G[S(x, y)] \mu u(x) \, dx$$

$$+ s_1 \iint G[S(x, y) - S(x, y')] \mu h(x, y') \, dx \, dy'. \quad (1)$$

- and get $f(x, y)$ from the surplus definition

$$(r + \delta)S(x, y) = f(x, y) - r\mathcal{U}(x) - (r + \delta)\mathcal{V}(y)$$
Identifying $c(n, y, \epsilon)$

- since $\delta$ destroys the vacancy, when firm size is stable we have

$$\delta l(y, \epsilon) = n(y, \epsilon)$$

- where $l(y, \epsilon)$ is the stationary size, then the FOC gives

$$\frac{\partial c}{\partial n}(n, \epsilon) = \mathcal{V}(y)$$

- normalize $\epsilon \in [0, 1]$ and $c(n, y, \epsilon)$ decreasing in $\epsilon$

- given convexity of $c$, $\epsilon$ is the rank is the size distribution conditional on $y$

$$\frac{\partial c}{\partial n}(\delta Q_{l|y}(\epsilon), \epsilon) = \mathcal{V}(y)$$
$W_0$ in the data
$S$ in the data
$h$ in the data
$h$ in the data
recovering $\sigma$


