Matching, Sorting and Wages

Preliminary

Thibaut Lamadon* Jeremy Lise† Costas Meghir‡
Jean-Marc Robin§

Barcelona, June 2014

*University College London and IFS
†University College London and IFS
‡IFS and Yale
§Science-Po and University College London

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 ▸ pic
  - 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 ▸ pic
  - 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 ▸ AKM
  - 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 **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

$$\text{output} := \int f(x,y)h(x|y,\epsilon)\,\mathrm{d}y$$

  - and owns masses $v(y)$ of open vacancies
  - the measure $\int h(x|y)\,\mathrm{d}x + v(y)$ 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

$$\text{output} := \int f(x, y) h(x|y, \epsilon) \, \mathrm{d}y$$

  - and owns masses $v(y)$ of open vacancies
  - the measure $\int h(x|y) \, \mathrm{d}x + v(y)$ 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, ϵ) 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. $\kappa\mu g(z)h(x|y')$ 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. δ job is destroyed, firm **does not** retain the vacancy
3. with pr. $\lambda \kappa g(z') v(y')$ worker meets another firm (y', z')

Notations: Values, Surplus

- The firm and the worker sequentially agree on a wage w

$$\begin{aligned} \mathcal{U}(x) &: \text{life time utility when unemployed} \\ \mathcal{V}(y) &: \text{present value of a vacancy} \\ \mathcal{W}(x,y,w) &: \text{worker lifetime utility when employed at } (y,w) \\ \mathcal{P}(x,y) &: \text{present value of a match} \end{aligned}$$

- and the surplus of a match

$$\mathcal{S}(x,y) := \mathcal{P}(x,y) - \mathcal{U}(x) - \mathcal{V}(y)$$

- $\mathcal{S}$ is not a function of w because utility is transferable

Matching outcomes

when unemployed meets an offer:

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

$$\mathcal{W}(x,y,w) = \beta\left(\mathcal{S}(x,y) - z\right) + \mathcal{U}(x)$$

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 $\mathcal{S}(x,y') - z' \geq \mathcal{S}(x,y)$, worker gets full (x,y) surplus

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

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

$$\mathcal{W}(x,y,\omega) = \mathcal{S}(x,y') - z'$$

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 $\mathcal{U}(x), \mathcal{V}(y)$ and $\mathcal{S}(x,y)$ such that:

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

Equilibrium properties

1. If $f_x > 0$ then $\mathcal{U}(x) \nearrow$ in x
2. If $f_y > 0$ then $\mathcal{V}(y) \nearrow$ in y
3. Bertrand competition gives:

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

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 $\mathbb{E}_t$ and $\mathbb{P}$ for any observables
    - for ex: $\mathbb{E}[R_t|J]$, $\mathbb{E}_t[R_t|E_t > E_{t-1}]$ or $\mathbb{P}\{E_t < E_{t-1}\}$

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 ▸ details
2. get $\mathcal{U}(x)$ ▸ details β, κ, δ ▸ details and $G(z)$ ▸ details
3. get a measure of y for each firm and $v(y), \mu$ ▸ details
4. identify $\mathcal{S}(x, y)$ ▸ details
5. construct $\mathcal{V}(y)$ and identify $f(x, y)$ ▸ details
6. identify $c(n, y, \epsilon)$ ▸ details

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 $\mathcal{V}(y)$ (value of poaching)
- use $\mathcal{S}$, $\mathcal{U}$ and $\mathcal{Q}(l|y)$ as a moments

Simulation for small sample performance

- $\sim 40,000$ workers, 10 years quarterly, 50 worker and firm types
- $f(x,y) = \left(.5\Phi^{-1}(x)^{\sigma} + .5\Phi^{-1}(y)^{\sigma}\right)^{1/\sigma}$ with 2 parametrization $\sigma = \{\text{pam:} -1\,,\ \text{nam: } 2\}$ pic
- $\hat{x}$ and $\hat{y}$ pic (SNR: x:0.96, y:0.95)
- $\hat{\mathcal{U}}(x)$ and $\hat{\mathcal{S}}(x,y)$ pic pam nam
- estimating complementarity : pic

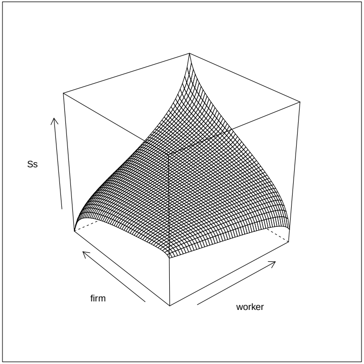
Auxiliary model on the data

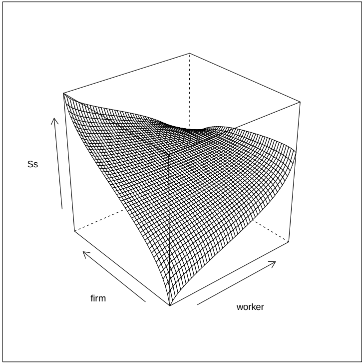
- Matched employer-employee data from Sweden
  - **today:** only male, college graduates under 50
  - 10 years, 424k individuals, 19k firms,
  - 265k j2j transitions, 158k u2e transitions
- Applying simplified procedure:
  - $\mathcal{U}(x)$ pic
  - $\mathcal{S}(x, y)$ pic
  - $h(x, y)$ heat 3D

Conclusion

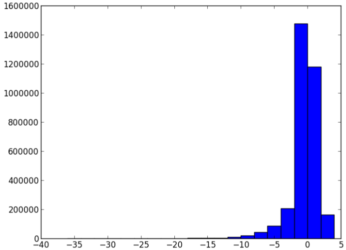
- developed a model with 2 sided heterogeneity,
  - rich wage dynamics with OTJ
  - both job creation and job filling
- provided a constructive identification proof and preliminary simulation results
- direct non-parametric estimation seems difficult with 10 years of data
  - use rank aggregation (Hagedorn,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

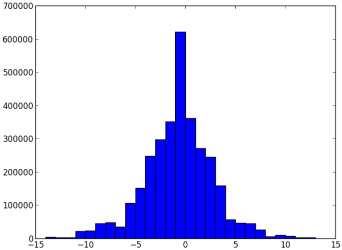
Parametrization Surplus



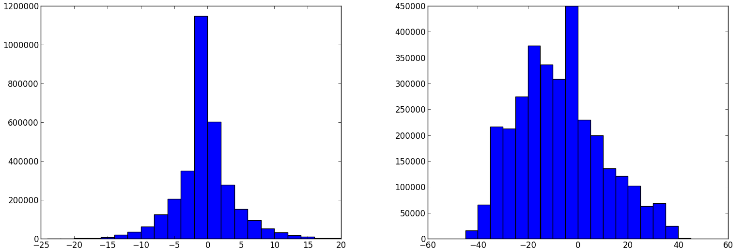


Estimated x versus true

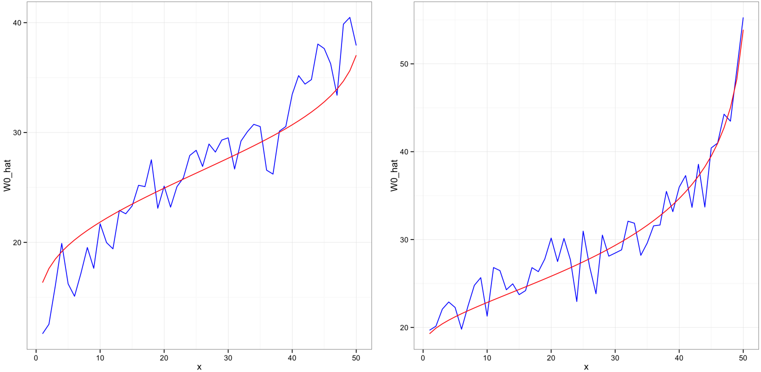




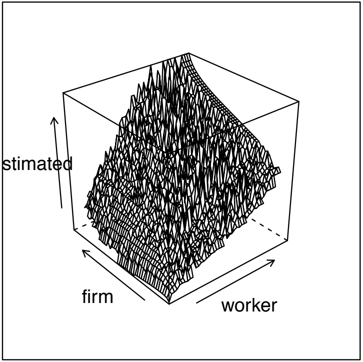
Estimated x versus true

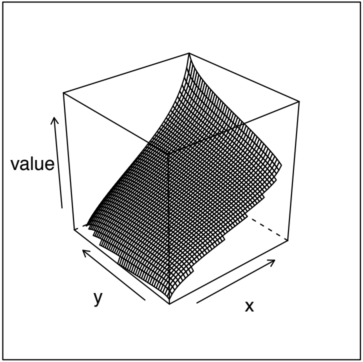


Estimated $\mathcal{U}(x)$

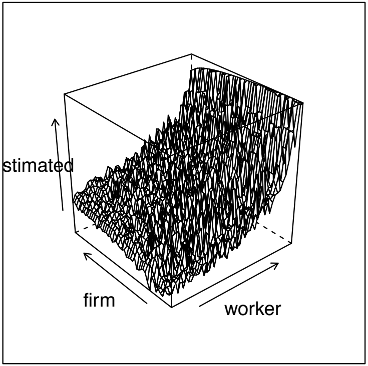


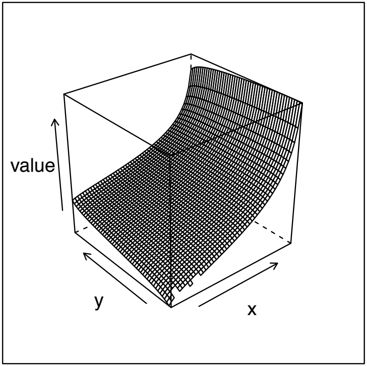
Estimated $\mathcal{S}(x, y) + \mathcal{U}(x)$ using $\bar{w}(x, y)$





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)

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

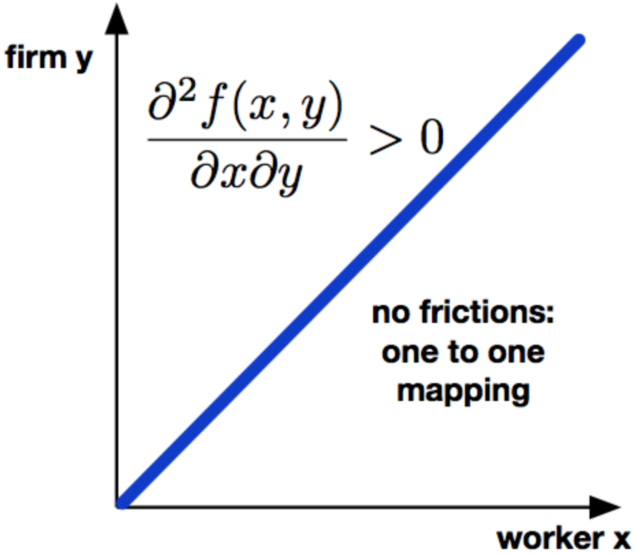
- within 10 years panel, explains $\sim 85\%$ of earnings dispersion details
- Firm share: $\frac{var(\psi_j)}{var(\psi_j)+var(\theta_i)} \simeq 20\%$
- Allocation to firms appears to be random $Cov(\theta_i, \psi) \simeq 0$
- Workers cluster together $Cov(\theta_i, \bar{\theta}_{J(i,\cdot)}) > 0$ back

Estimation for different countries

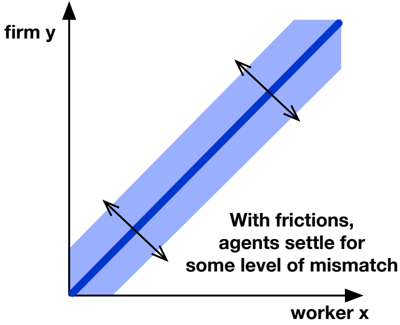
Country	US 1[a]	US 2	FR	GE	IT	DE[b]	BR
$Var(x\beta)$	0.03	0.14	0.02	—	0.01	—	0.02
$Var(\theta)$	0.29	0.23	0.21	0.05	0.05	0.08	0.40
$Var(\psi)$	0.08	0.053	0.08	0.013	0.01	0.00	0.18
$\frac{Var(\psi)}{Var(\theta+\psi)}$	0.22	0.19	0.32	0.22	0.23	0.03	0.31
$Corr(\theta,\psi)$	-0.01	-0.03	-0.28	-0.19	0.04	0.00	0.04[f]
$Corr\left(\theta,\widetilde{\theta}\right)$	—	—	—	—	0.17[c]	0.40[d]	0.52
R^2	0.89	0.9	0.84	—	—	0.85	0.93
Sample Statistics							
Years	90-99	84-93	76-87	93-97	81-97	94-03	95-05
Nobs	$37.7M$	$4.3M$	$5.3M$	$4.8M$	—	$6.9M$	$16.0M$
Nworkers	$5.2M$	$293K$	$1.2M$	$1.8M$	$1.7M$	$563K$	$2.0M$
Nfirms	$476K$	$80K$	$500K$	1821	$421K$	$53.6K$	$137K$
% 1st Group[e]	—	99.1%	88.3%	94.9%	99.5%	—	98.6%

back

Becker friction-less assignment



Becker friction-less assignment



- agents settle for lower than optimal match
- wages are not monotonic in y
- linear wage equation is mis-specified ▸ back

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

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

$$\overline{\mathrm{w}}(x, y) = f(x, y) - (r+\delta)\mathcal{V}(y) - 0 \quad \nearrow \text{ in } x$$

- define $\overline{R} := \max_t\{R_t : E_t = 1\}$ for each ω then

$$\forall \omega,\ X = Q_{\overline{R}}(\overline{R})$$

back

Identifying $\mathcal{U}(x)$

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

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

- and so

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

- where $W_t := \sum_{\tau=t}^{\infty} \frac{R_\tau}{(1+r)^\tau}$ and $R_{min}(x, J) := \min\limits_{\omega \in \Omega, t \in T} \{R_t : E_t > E_{t-1}, J_t = J, X = x\}$

back

identifying β

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

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

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

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

- combining gives:

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

Identifying δ and κ

- 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:

$$\kappa = \frac{\mathbb{P}\{R_t > R_{t-1} \cup J_t \neq J_{t-1} | X, R_{t-1} = R_{min}(J, X)\}}{\mathbb{P}\{E_t > E_{t-1} | X\}},$$

back

$G(z)$

- 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:

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

- but assuming that $G(z)$ is parametrized "globally" it is enough

back

Identifying firm type

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

$$\hat{\mathcal{V}}(J) = (1-\beta)\int \mathcal{G}\left[\mathcal{S}(x,J)\right] u(x)\,\mathrm{d}x + \kappa \iint \mathcal{G}\left[\mathcal{S}(x,J) - S\right] f_{Sx}(S,x)\,\mathrm{d}x\,\mathrm{d}S.$$

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

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

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

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

- total number of vacancies identifies μ. back

Identifying $\mathcal{S}(x, y)$ and

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

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

▸ back

Identifying $\mathcal{V}(y)$ and $f(x, y)$

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

$$
\begin{aligned}
r\mathcal{V}(y) = (1-\beta) \int \mathcal{G}\left[\mathcal{S}(x,y)\right] \mu u(x)\, \mathrm{d}x \\
+ s_1 \iint \mathcal{G}\left[\mathcal{S}(x,y) - \mathcal{S}(x,y')\right] \mu h(x,y')\, \mathrm{d}x\, \mathrm{d}y'.
\end{aligned} \tag{1}
$$

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

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

back

Identifying $c(n, y, \epsilon)$

- since δ 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

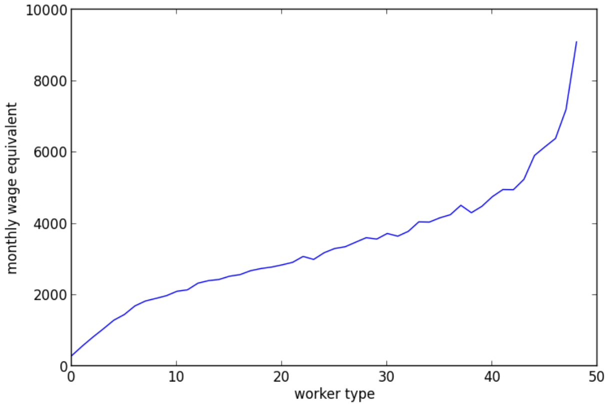
$$\frac{\partial c}{\partial n}(n, \epsilon) = \mathcal{V}(y)$$

- normalize $\epsilon \in [0, 1]$ and $c(n, y, \epsilon)$ decreasing in ϵ
- given convexity of c, ϵ is the rank is the size distribution conditional on y

$$\frac{\partial c}{\partial n}(\delta \mathcal{Q}_{l|y}(\epsilon), \epsilon) = \mathcal{V}(y)$$

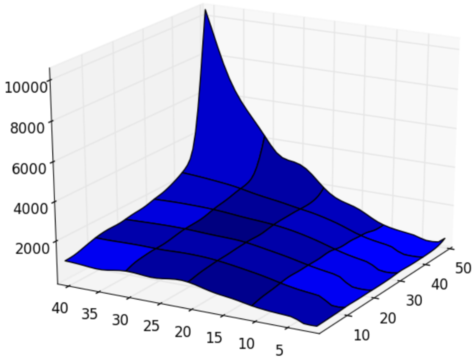
back

W_0 in the data



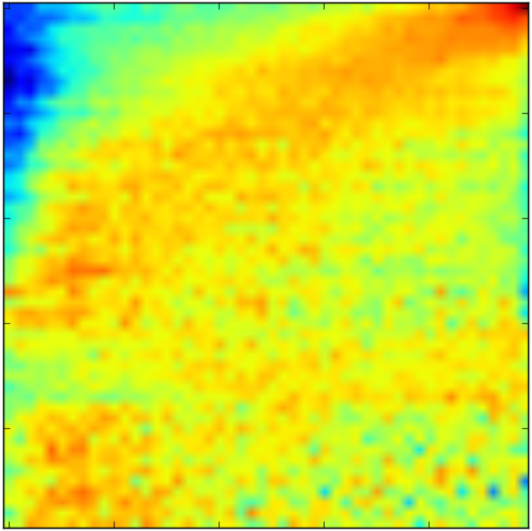
back

$\mathcal{S}$ in the data



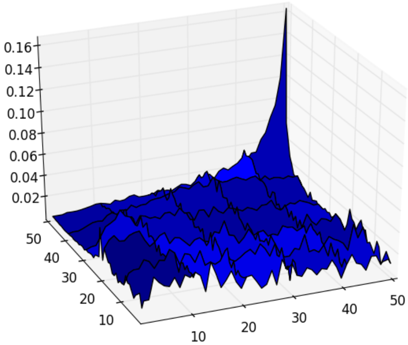
back

h in the data



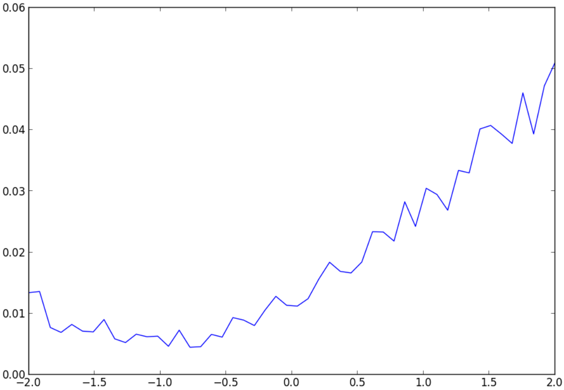
back

h in the data



back

recovering σ



▸ back

Abowd, J., F. Kramarz, and D. Margolis (1999): "High wage workers and high wage firms," *Econometrica*, 67(2), 251–333.

Bagger, J., and R. Lentz (2014): "An empirical model of wage dispersion with sorting," *University of Wisconsin mimeo*.

Becker, G. S. (1974): "A theory of marriage," *Journal of Political Economy*, pp. 299–351.

Choo, E., and A. Siow (2006): "Who marries whom and why," *Journal of Political Economy*, 114(1), 175–201.

De Melo, R. L. (2009): "Sorting in the labor market: Theory and measurement," *Working Paper*.

Eeckhout, J., and P. Kircher (2011): "Identifying sorting in theory," *The Review of Economic Studies*, 78(3), 872–906.

Galichon, A., and B. Salanié (2011): "Cupid's invisible hand: social surplus and identification in matching models," *Available at SSRN 1804623*.

Hagedorn, M., T. H. Law, and I. Manovskii (2014): "Identifying Equilibrium Models of Labor Market Sorting," *Working Paper*.

Shimer, R., and L. Smith (2003): "Assortative matching and search," *Econometrica*, 68(2), 343–369.