Identifying Equilibrium Models of Labor Market Sorting

Marcus Hagedorn
University of Oslo

Tzuo Hann Law
Boston College

Iourii Manovskii
University of Pennsylvania

September 24, 2016

Basic matching literature

On the labor market, a **matching function** can be thought of as a function that tells the odds of a match between a worker and a firm.

Most recent search models contain explicit microfoundations about matching behavior.

Two landmark contributions that provided crucial insights on matching are the following:

- Becker (1973): complementarity and substitutability between inputs determine the strength and direction of sorting in a competitive matching model.
- Shimer and Smith (2000, SS): search frictions can explain why we observe some **mismatch** between workers and firms with respect to Becker's frictionless framework.

Taking the model by Shimer and Smith to the data is not straightforward...

Empirical challenges: labor market case

- what we observe in the data: workers' and firms' identity and matches, unemployment and employment stocks, flows in and out of unemployment, wages (*transfers*);
- what we do **not** observe in the data: the workers' and firms' productivity (*types*), the (flow) output produced by a match, the value of unemployment for a worker and of a vacancy for a firm and the surplus generated by a match.

Challenges: *how to identify workers' and firms' productivity? how to recover the match surplus? how to estimate the parameters of the production function? how to assess the welfare loss due to frictional unemployment and mismatch?*

Empirical challenges: marriage market case

- what we observe in the data: spouses' identity and matches, individual characteristics, stocks of married and singles, flows in and out of marriage, household allocation (hours worked, domestic work, savings, consumption);
- what we do **not** observe in the data: non-monetary gains from marriage and total marriage (flow) output, the number of meetings among singles, the value of singlehood and the surplus generated by a match.

Challenges: *how to explain marriage between spouses with very different observables? how to explain high (or low) marriage rates, is it because of high (low) meeting rates or because singles are (not) selective? how to recover the match surplus? how to disentangle monetary and non-monetary gains from marriage?*

Labor market: useful literature

Unobservables seem to play a crucial role in sorting. **Abowd et al. (1999)** suggest estimating workers' and firms' fixed effects:

$$\log w_{ijt} = x_{ijt}\beta + \delta_i + \psi_j + \varepsilon_{ijt}$$

where ε_{ijt} is orthogonal to the fixed effects.

- **Unobservables and non-linearity**: with this specification, it is not possible to test for complementarities of (unobservable) inputs and skills.
- **Bargaining power**: bargaining theory suggests wages depend on the counterparts' reservation utilities, which are not separately identified in closed form.

Identification: general idea

The model is identified if there exists a unique **function**:

Joint distribution of observables $\longrightarrow$ Primitives of the model

HLM are especially interested in the identification:

$$\begin{matrix} \text{Matches } (i,j) \\ \text{Wages } (i,j) \end{matrix} \longrightarrow \begin{matrix} \text{Type of } i \\ \text{Type of } j \\ \text{Output } (i,j) \end{matrix}$$

Environment

- At any moment in time, a worker is either *employed* (e) or *unemployed* (u); a worker is characterized by his productivity $x \in [0, 1]$.
- A firm is either *producing* (p) or *vacant* (v); a firm is characterized by its productivity $y \in [0, 1]$.
- The flow output of a match (x, y) is given by the twice differentiable *production function* $f : [0, 1]^2 \to \mathbb{R}_+$.
- Assume $f_x > 0$ and $f_y > 0$. What matters for sorting is a worker's and firm's rank (uniform marginal distributions).

Important: no assumption on Increasing (or Decreasing) Differences between (x, y) for f $\to$ the goal is to understand the sign and strength of the cross-derivative f_{xy} from the data.

Distributions

- The number of (x, y) matches is $m(x, y)$.
- The number of unemployed x is $n_u(x)$.
- The number of vacancies y is $n_v(x)$
- Aggregate measures are: unemployment $N_u = \int n_u(x)dx$, number of vacant firms $N_v = \int n_v(y)dy$.

Production and search

- Matched workers earn flow payoff $w(x, y)$ and firms $\pi(x, y)$, so that $w(x, y) + \pi(x, y) = f(x, y)$ (*transferable utility*).
- Free entry of vacancies at fixed cost c. The productivity of a vacancy is randomly drawn *after* its creation (*no selection*).
- Search is *time-consuming* but no *explicit cost*.
- Search is *random*: meetings occur at a flow probability rate

$$\frac{M(n_u, n_v)}{N_u N_v} \quad \text{with} \quad M : [0, 1]^2 \rightarrow [0, min(N_u, N_v)]$$

- The probability of a vacant firm to meet a worker of any type is $\lambda_v \equiv \frac{M(n_u, n_v)}{N_u}$ and the probability of an unemployed to meet a firm of any type is $\lambda_u \equiv \frac{M(n_u, n_v)}{N_v}$.

Match formation and dissolution

- Not all meetings lead to a match! Strategies can be described by *acceptance sets*: $A^w(x)$ is the set of firms an agent x is ready to match with upon a meeting ($A^f(y)$ for firms).
- Denote $V_e(x,y)$ and $V_u(x)$ a worker's value of being employed in a firm y and unemployed respectively. Denote $V_p(x,y)$ and $V_v(y)$ a firm's value of hiring a worker x and holding a vacancy respectively. The total surplus of a match is

$$S(x,y) = \underbrace{(V_e(x,y) - V_u(x))}_{\text{surplus of worker } x} + \underbrace{(V_p(x,y) - V_v(y))}_{\text{surplus of firm } y}.$$

Match formation and dissolution

- Workers and firms bargain over wages in order to split the surplus generated by a match. Call β the worker's bargaining weight: under transferable utility, *generalized Nash-bargaining* leads to

$$V_e(x,y)-V_u(x) = \beta S(x,y) \qquad V_p(x,y)-V_v(y) = (1-\beta)S(x,y)$$

- The *matching set* $B^w(x)$ of a worker is the set of firms that accept a worker's candidature (same for the firm). Notice that $A^w(x) = B^w(x)$ as there is no disagreement between the two counterparts. We define the *matching rule*:

$$\alpha(x,y) \equiv \mathbb{1}\{x \in B^f(y) \wedge y \in B^w(x)\} = \mathbb{1}\{S(x,y) \geq 0\}$$

- There is a constant flow probability $\delta > 0$ that any match is destroyed.

Steady State Search Equilibrium (SE)

At the steady state, for every pair (x, y), the number of destroyed matches must be equal to the number of newly formed matches at any moment in time (*flow balance conditions*).

A SE can be represented as a triple $((V_u, V_v), (B^w, B^f), (n_u, n_v))$*, where (i)* (V_u, V_v) *solves the implicit system given by the Bellman equations and the match surplus equation for given* (B^w, B^f)*,* (n_u, n_v)*; (ii)* (B^w, B^f) *is optimal given the match surplus; (iii)* (n_u, n_v) *solves the implicit system given by the steady-state flow balance conditions.*

Identifying workers' type

Matches and **wages** are observed.

1. The **value of unemployment** is *increasing* in the worker's type: a more productive worker can imitate the acceptance strategy of a less productive one, but can demand a higher wage for the same job.
2. Hence, the **reservation wage** is also *increasing* in the worker's type.
3. Within a firm (i.e. for given y), the **worker's wage** is *increasing* in the worker's type.

However, an unemployed worker might accept (1) a high wage in a low-productivity firm as well as (2) a low wage in a high-productivity firm.

NB: look how this changes with on-the-job search... (e.g. Postel-Vinay and Robin, 2002)

Identifying workers' type

Result

i) *$V_u(x)$, $V_e(x, y)$ and $w(x, y)$ are increasing in x (for given y).*

ii) *The lowest accepted wage (reservation wage) $w(x, y^{min}(x))$ is increasing in x.*

iii) *The highest accepted wage $w(x, y^{max}(x))$ is increasing in x.*

iv) *The adjusted average wage $\bar{w}^{adj}(x)$ is increasing in x.*

v) *...but the average wage $\bar{w}(x)$ is not increasing in x.*

Results 1 (part i) suggest that the *ranking* of workers can be inferred if labor **mobility across firms** is sufficiently high: "movers" across different firms allow to create a global ranking of workers.

→ a sufficiently long panel allows to infer reservation wages for each type x from their observed lowest accepted wages (see how to estimate the marginal distributions).

Identifying firms' type

Similarly, it is necessary to find some statistic that is monotonically growing in firm's type y.

Result

$V_v(y)$, $V_p(x, y)$ are increasing in y (for given x).

Result

The share of surplus a firm expects from filling a vacancy

$$(1-\delta)(1-\beta)\lambda_v \int \frac{n_u(x')}{N_u} S^+(x', y) dx'$$

is increasing in y.

...while these values are not observed, we can use this properties to prove another result in the following slide.

Identifying firms' type

Results 2 and 3 lead to

Result

The expected wage premium over the reservation wage payed by a firm y to newly hired workers

$$\Omega(y) \equiv (1-\delta)\lambda_v \int \frac{n_u(x')}{N_u}\alpha(x,y)\left(w(x',y) - w(x',y^{min}(x))\right) dx'$$

is increasing in y.

The statistic $\Omega(y)$ can be more intuitively rewritten as

$$\Omega(y) = (\textit{prob. of filling a vacancy } y) \times (\textit{av. wage premium paid by firm } y)$$

which, after ranking workers, can be inferred from (1) unemployment stocks and new hires by worker's type, (2) the total number of vacancies V, (3) workers' reservation wages and (4) a firm j's wages paid to newly hired workers (see Section 3.4.3).

Identifying the value functions

- The unemployment value is found using $V_e(x, y^{min}(x)) = V_u(x)$ and is proportional to the reservation wage.
- Once computed $V_u(x)$, observing wages and employment spells for workers of type x in firms of type y, it is possible to estimate $V_e(x, y)$ for each pair (x, y), which is the average present discounted value of the income flows during the employment spells of different length observed in the data.
- Knowing $V_e(x, y)$ and $V_u(x)$, one obtains $S(x, y)$ for a given β.
- Knowing $S(x, y)$ and $\Omega(y)$, one obtains $V_p(x, y)$ and $V_v(y)$ for a given β.

Production function

The *wage equation* in the SS model is

$$w(x,y) = \beta f(x,y) + \beta r V_v(y) + (1-\beta) r V_u(x)$$

from where, for given $V_v(y)$ and $V_u(x)$, it is possible to obtain the production function $f(x,y)$.

The *non-parametric estimation* of the production function $f(x,y)$ is possible over the subset of $[0,1]^2$ where wages $w(x,y)$ are observed. Studying the cross-derivative f_{xy} is informative about the nature of input complementarities.

On-the-job search

The model is extended to include *on-the-job search*. A crucial adjustment is that results 1 must be changed as follows:

Result

Within a firm, wages of workers hired from unemployment are increasing in the worker's type x.

The on-the-job search version of the model is taken to the data with the 1993-2007 LIAB data set on German employer-employee matching.

Ranking inference

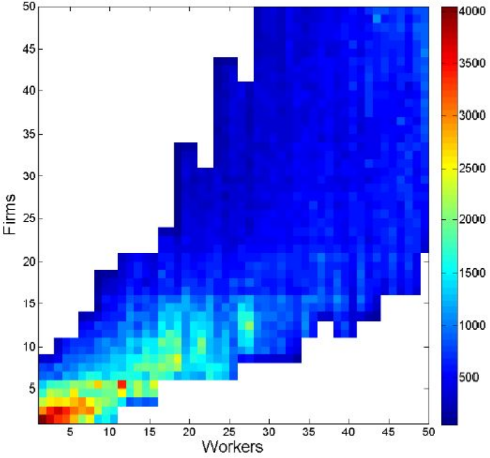


Figure 6: Estimated Match Density.

Table 3: Sign and Strength of Sorting

	HLM	AKM
Corr(W-rank,F-rank)	0.7547	0.055

Nonparametric estimation of the production function

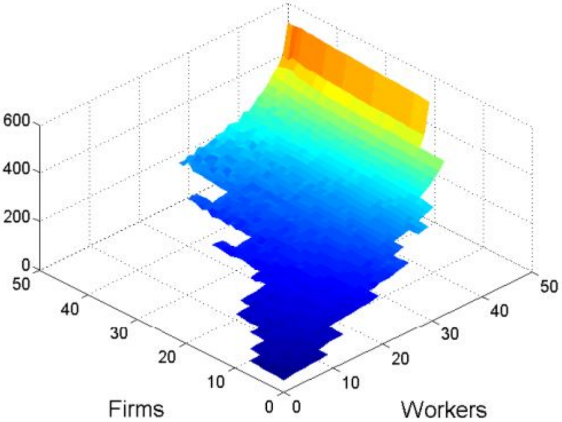
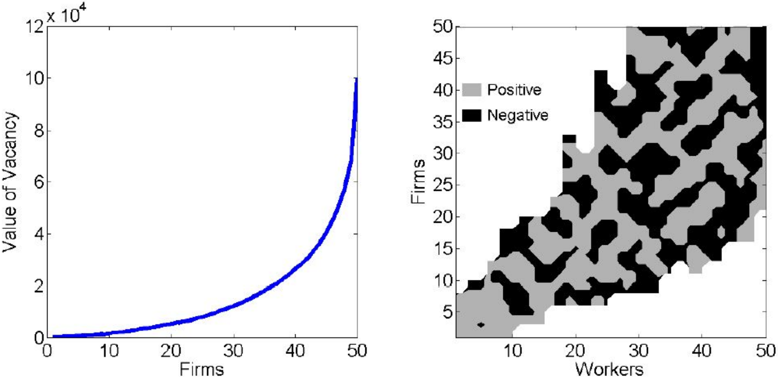


Figure 7: Estimated Production Function.

Value of vacancies and supermodularity



(a) Estimated Value of Vacancy.

(b) Cross-Partial Derivative of Estimated Production Function.

Figure 8: Features of the Estimated Production Function.

Frictions and mismatch

- Comparison with a *counterfactual frictionless equilibrium*: output loss due to *misallocation* equal to 1.83%.
- An additional 8.47% loss is due to frictional unemployment: frictional unemployment mostly concerns low-productivity workers and vacancies.
- Forcing the matching to be on the main diagonal (rank correlation equal to one) results in a slight output loss (0.23 %), since the production function is **not** supermodular on its whole support.

Marriage market: what's different

- Much less mobility: very hard to estimate agents' "fixed effects".
- How do counterparts transfer utility? Domestic work, public good expenditure, within-couple transfers.

Motivation

- A large empirical literature on within-household allocation (starting from Chiappori, 1988).
- Singlehood, divorce and remarriage are becoming more common → we want to characterize agents' reservation utilities as endogenous.
- The *sharing rule* depends on agents' attractiveness on the marriage market.

Environment

Similar model to HLM (Goussé et al., 2015, see), but written in continuous time (little difference).

- At any moment in time, agents are either *married* (1) or *single* (0).
- Men are characterized by a vector of traits $x \in \mathcal{X}$; women by $y \in \mathcal{Y}$.
- The flow output of a match (x, y) is now given by $f(x, y) + z$, with $z \sim G$.
- $n_m(x)$ is the number of single men of type x; $n_w(y)$ of single women y; $m(x, y)$ of couples (x, y).

Marriages

As in SS, assume $M(n_x, n_y) = \lambda N_m N_w$: meetings among singles follow a Poisson process with arrival rate λ.

As in HLM, assume output is divided such that

$$f(x, y) = u_m(x, y) + u_w(x, y).$$

Generalized Nash bargaining implies that, upon a meeting (before drawing z), a man and a woman marry with probability

$$\alpha(x, y) \equiv Pr\{S(x, y, z) > 0\}$$

and their expected surplus is

$$\bar{S}(x, y) \equiv \mathbb{E}_z \left[\max\{S(x, y, z), 0\}\right].$$

Marriage and endogenous divorce

Out of $\lambda n_m(x)n_w(y)$ meetings between types (x, y), only

$$MF(x, y) \equiv \lambda\alpha(x, y)n_m(x)n_w(y)$$

decide to marry.

A married couple (x, y) must draw a new z' when a shock hits, following a Poisson process with arrival rate δ. This means that

$$DF(x, y) \equiv \delta(1 - \alpha(x, y))m(x, y)$$

will end up divorcing.

Solving the model

What is exogenous (primitives):

- household technology (f), match-quality characteristics (δ and G) and bargaining weight (β);
- marginal distributions of x and y;
- probability of meeting (λ).

What is endogenous (output):

- marriage surplus (S) and value of staying single (V_m^0 and V_w^0) (see equations);
- probability of matching (α);
- number of singles (n_m, n_w), married (m), newlyweds and divorces by type.

Identification (Goussé et al., 2015)

What is observed:

- marginal distributions for (x, s) and y;
- number of couples (x, s, y), number of single men (x, s) and of women y;
- number of weddings (x, s, y) and divorces (x, s, y) per period;

What is estimated:

- the arrival rate of meetings (λ) and match-quality shocks (δ);
- the expected surplus $(\bar{S})$ and reservation utilities $(U^0,\ V^0)$;
- the production function (f).

What is not estimated:

- the bargaining weight (β), the distribution of match-quality (G), the discount rate (r).

Identification: first step

Nice empirical moments, we can "easily" compute from the data:

- **marriage rates**, $MF(x,y)/(n_m(x)n_w(y)) = \lambda\alpha(x,y)$;
- **divorce rates**, $DF(x,y)/m(x,y) = \delta(1-\alpha(x,y))$.

Identification problem: do people marry frequently because they meet frequently or because they are not very selective? Do people divorce rarely because shocks are rare or because they enjoy high surplus?

Information on the matching rule are present in both marriage and divorce rates. The following relationship identifies (λ, δ):

$$\frac{1}{\lambda}\frac{\hat{MF}(x,y)}{\hat{n}_m(x)\hat{n}_w(y)} + \frac{1}{\delta}\frac{\hat{DF}(x,y)}{\hat{m}(x,y)} = 1 \quad \forall(x,y).$$

With $(\hat{\lambda}, \hat{\delta})$, one can back out $\hat{\alpha}$ for each (x,y) (see a possible way of proceeding).

Identification: second step

There exists a *bijection*

$$\alpha \rightarrow f,$$

the household production function is thus identified over the whole support $\mathcal{X} \times \mathcal{Y}$ (see details).

Identification relies on the idea that $\alpha \uparrow \bar{S}$. Once we have disentangled α from marriage and divorce rates, we can find $\bar{S}$ and f.

Identification: third step

The most interesting and complicated part (Goussé et al., 2015, see).

Assume spouses bargain over hours worked. Then,

$$h_m(x,y) = H_m(x,y,V_m^0,V_w^0,\beta,G)$$
$$h_w(x,y) = H_w(x,y,V_m^0,V_w^0,1-\beta,G)$$

which helps us identifying β and G, because we observe the outcome of the negotiation (h_m, h_w).

However, we need to make clever assumptions on f to obtain a nice (possibly linear) specification for H_m and H_w.

Bellman Equations BACK

Define $1/(1+r) \in (0,1)$ as the *discount factor* and $x^+ \equiv \max(x,0)$ the *positive part* of x:

$$V_u(x) = \frac{1}{1+r}\left[V_u(x) + \beta(1-\delta)\lambda_u \int \frac{n_v(y')}{N_v} S^+(x,y')dy'\right]$$

$$V_v(y) = \frac{1}{1+r}\left[V_v(y) + (1-\beta)(1-\delta)\lambda_v \int \frac{n_u(x')}{N_u} S^+(x',y)dx'\right]$$

$$V_e(x) = w(x,y) + \frac{1}{1+r}\left[V_u(x) + \beta(1-\delta)S(x,y)\right]$$

$$V_p(y) = \pi(x,y) + \frac{1}{1+r}\left[V_v(y) + (1-\beta)(1-\delta)S(x,y)\right]$$

Estimation BACK

1. There might be *inconsistency* in workers' ranking: it is assumed it is due to *measurement error* in wages. Drawing from *social choice literature*, the estimated aggregate ranking is chosen to minimize the number of disagreements. In the process, HLM assign a weight to each workers' pair: the ranking is more likely to be inconsistent for close wages, because of measurement error.
2. Once workers are ranked, they are divided into bins of equal size: hereafter, each stands for a type x.
   The longer the panel, the fewer the bins. Reservation wages are estimated for each type.
3. Result 4 implies that firms can be ranked using the statistic $\Omega(y)$. Firms are also divided into bins of equal size: hereafter, each stands for a type y.

Bellman equations BACK

Husband's surplus share:

$$(r+\delta)\left(V_m^1(x,y,z)-V_m^0(x)\right)=$$
$$=u_m(x,y)+\frac{z}{2}+\delta\beta\bar{S}(x,y)-rV_m^0(x)$$

Wife's surplus share:

$$(r+\delta)\left(V_w^1(x,y,z)-V_w^0(y)\right)=$$
$$=u_w(x,y)+\frac{z}{2}+\delta(1-\beta)\bar{S}(x,y)-rV^0(y)$$

Man's value of singlehood:

$$rU^0(x)=u^0(x)+\lambda\beta\int_{\mathcal{Y}}\bar{S}(x,y)n(y)dy$$

Woman's value of singlehood:

$$rV^0(y)=v^0(y)+\lambda(1-\beta)\int_{\mathcal{X}}\bar{S}(x,y)n(x)dx$$

Expected surplus BACK

Since marriage surplus is

$$S(x,y,z) = f(x,y) + z + \delta \bar{S}(x,y) - rV_m^0(x) - rV_w^0(y),$$

$\bar{S}(x,y)$ is found as the fixed point of the contraction mapping

$$(r+\delta)\bar{S}(x,y) = \\ = \mathbb{E}_z \left[0, \max\left\{ f(x,y) + z + \delta \bar{S}(x,y) - rV_m^0(x) - rV_w^0(y) \right\} \right].$$

Inflow-outflow balance BACK

The *steady-state assumption* implies the following relationships, necessary to maintain the number $m(x, y)$ of married couples (x, y) constant:

$$\underbrace{\lambda\alpha(x, y)n_m(x)n_w(y)}_{MF(x,y)} = \underbrace{\delta(1 - \alpha(x, y))m(x, y)}_{DF(x,y)}$$

- for given strategies α, it allows to compute the equilibrium number of couples $m_{x,y}(n)$ as a function of the vector of single agents' densities (n_m, n_w);
- for given values of (m, n_m, n_w) - observed in the data - and $(\hat{\lambda}, \hat{\delta})$ - estimated at the first step - it allows to compute the matrix $\hat{\alpha}$.

Bijection $\alpha \to f$ BACK

Additive separability of z implies $S(x, y, z) = -T(x, y) + z$. The matching rule implies

$$\alpha(x, s, y) = Pr\{-T(x, y) + z > 0\} = 1 - G(T(x, y)).$$

In addition, the expected surplus upon a meeting is

$$\bar{S}(x, y) = \int \max(-T(x, y) + z, 0) dG(z) = \mathcal{G}(T(x, y))$$

where $\mathcal{G}$ is continuous and strictly decreasing, and thus invertible.

After computing $\hat{\alpha}(x, y)$ at the first step,

$$\hat{\bar{S}}(x, y) = \mathcal{G}(1 - G^{-1}(\hat{\alpha}(x, y))).$$

Hence, it is straightforward to obtain $(\hat{V}_m^0(x), \hat{V}_w^0(y))$ and $\hat{f}(x, y)$.

John M Abowd, Francis Kramarz, and David N Margolis. High wage workers and high wage firms. Econometrica, 67(2): 251–333, 1999.

Gary S. Becker. A theory of marriage: Part i. Journal of Political Economy, 81(4):813–846, 1973.

Pierre-André Chiappori. Rational household labor supply. Econometrica, 56(1):63–90, 1988.

Marion Goussé, Nicolas Jacquemet, and Jean-Marc Robin. Assortative matching and search with labour supply and home production. 2015.

Fabien Postel-Vinay and Jean-Marc Robin. Equilibrium wage dispersion with worker and employer heterogeneity. Econometrica, 70(6):2295–2350, 2002.

Robert Shimer and Lones Smith. Assortative matching and search. Econometrica, 68(2):343–369, 2000.