The Econometrics of Matching

Pierre-André Chiappori   Bernard Salanié
Columbia University

Chicago, June 2012

Bringing Matching Theory to Data

Basic, one-dimensional theory yields stark predictions (assortative matching) that only approximately hold in the data
In the data, almost every "matching cell" is non-empty.
To reconcile theory and data :

- can add frictions (see Robin and Shimer)
- or assume a reduced-form relationship (the rank-order property of Fox)
- or model directly matching on both observables and unobservables (our main focus here.)

Main Themes

- More emphasis on TU than on NTU
- Focus on one-to-one rather than many-to-(one, many) (unlike Fox)
- These models are hard to identify, depending on what we observe/what we assume.

Matching assumed frictionless here : no search cost, perfect information among participants, transfers (if any) are costless.

Common Structures

- Populations : X et Y, measures F et G ("men" and "women")
- Matching : measure μ on $XY = (X \times Y) \cup (X \times 0) \cup (0 \times Y)$, given marginals F et G
- normalize utilities of singles to 0
- NTU : match (x, y) brings (algebraic, ordinal) gain $U(x, y)$ to x and $V(x, y)$ to y
- (perfectly) TU : it creates (algebraic) joint surplus $s(x, y)$, endogenously shared between the partners.

Common equilibrium concept : stability

- no one is better off divorcing their current match and remaining single
- NTU : no (x, y) pair is better off divorcing their current matches and rematching together
- TU : no (x, y) pair can find a way of sharing $s(x, y)$ that improves over both of their current outcomes.

Consequences for NTU : coupled equations

NTU : stable matchings solve

$$u(x) = \max_z \{U(x,z) | V(x,z) \geq v(z)\}$$

and

$$v(y) = \max_z \{V(z,y) | U(z,y) \geq u(z)\}$$

for some pair of functions u and v.

And for TU

TU : stable matchings solve

$$u(x) = \max_z \{s(x, z) - v(z)\}$$

and

$$v(y) = \max_z \{s(z, y) - u(z)\}$$

for some pair of functions u and v.

The Data

We observe a realized matching, and perhaps a noisy measure of outcomes (e.g. children, divorce, labor market dynamics. . .)
We partition payoff-relevant types into observed (by the econometrician) and unobserved characteristics :

$$x = (I, \varepsilon), \;\; y = (J, \eta)$$

We only observe $\bar{\mu}(I, J) = E\left(\mu(x, y)|I, J\right)$ and perhaps (say in TU)

$$\bar{s}(I, J) = E\left(\tilde{s}(x, y)|I, J\right)$$

with $\tilde{s}$ a noisy measure of s (and $E \equiv E_\mu$.)

The Empirical Content of Matching Models

The theory implies the coupled equations, and only that : given s, F, G, and a solution (u, v), any μ

- whose support is contained in the $\arg\max_z$
- which integrates to the margins F and G

is a stable matching.

Four approaches

- Choo-Siow : impose separability (unobserved characteristics do not interact)
- Chiappori-Oreffice-Quintana-Domeque : assume agreement on “attractiveness indices” (lower-dimensional matching)
- Fox : assume a rank-order property
- Hitsch-Hortacsu-Ariely : build on the coupled equations.

Conditional just-identification in TU models

Choo-Siow (2006, multinomial logit) + Galichon-Salanié (2011, general case) : even if we assume
separability (S) : $s(x,y) = \bar{s}(I,J) + \varepsilon_I(J) + \eta_J(I)$,
and we (assume that we) know the distributions $\mathcal{P}_I$ of $\varepsilon_I(.)$ and $\mathcal{Q}_J$ of $\eta_J(.)$,
the function $\bar{s}$ is just-identified from data on the observed matching $\bar{\mu}$ on one cross-section.

Separability

=absence of complementarities on unobserved characteristics : if

$$x = (I, \varepsilon)\ ;\ x' = (I, \varepsilon')$$

and

$$y = (J, \eta)\ ;\ y' = (J, \eta')$$

then $s(x, y) + s(x', y') = s(x, y') + s(x', y)$.
How strong ? depends on data, and on question asked from it.

Separability buys us discrete choice

Then

$$u(x) = \max_z \left(s(x,z) - v(z) \right)$$

gives

$$u(x) = \max_J \left(\bar{s}(I,J) + \varepsilon_I(J) - \min_\eta \left(v(J,\eta) - \eta_J(I) \right) \right)$$

Denote $V(I,J)$ the $\min_\eta$ and $U(I,J) = \bar{s}(I,J) - V(I,J)$;

$$u(x) = \max_J \left(U(I,J) + \varepsilon_I(J) \right).$$

Just-identification

See Salanié, tomorrow : in a large market, if we "know" the $\mathcal{P}_I$ and $\mathcal{Q}_J$, we can compute easily

$$\bar{s}(I,J) = \dots \;;\; u,v,U,V = \dots \text{ etc.}$$

Large market : large numbers of individuals for any observed type
Ensures proportions ≡ probabilities, and uniqueness of market equilibrium payoffs u and v.

Beyond just-identification : testing

Just identification rules out testing, unless

- we restrict the specification of $\bar{s}(I, J)$; e.g. Galichon-Salanié : expand

$$\bar{s}(I, J) = \sum_{k=1}^{K} \lambda^k \bar{s}_k(I, J)$$

  for some unknown λ's and known basis functions $\bar{s}_k$.
- we pool data on several "markets" and restrict $\bar{s}(I, J)$ and/or the variation of the $\mathcal{P}_I$ and $\mathcal{Q}_J$ across markets. E.g., Chiappori-Salanié-Weiss :
  stability of complementarities and heteroskedasticity across cohorts.

Beyond just-identification : the distribution of errors

The $\mathcal{P}_I$ and $\mathcal{Q}_J$ have many degrees of freedom ; how can we identify some of them/test the model ?
again : restrict the specification of the mean surplus, and/or pool data across markets.

Estimation in the separable framework

First estimate the margins on observed characteristics : $\hat{F}(I), \hat{G}(J)$.
If the $\mathcal{P}_I$ and $\mathcal{Q}_J$ are fully specified and the $\bar{s}(I, J)$ is unrestricted : apply the closed-form formula
If they depend on a parameter vector λ : can use the formula as the basis of a minimum-distance estimator,
or do constrained maximum-likelihood estimation.

Constrained MLE

$$\max_{\lambda}\left(2\sum_{IJ}\hat{\mu}_{IJ}\ln\mu^{\lambda}_{IJ}+\sum_{I}\hat{\mu}_{I0}\ln\mu^{\lambda}_{I0}+\sum_{J}\hat{\mu}_{0J}\ln\mu^{\lambda}_{0J}\right)$$

where μ^{λ} is the optimal matching for parameter vector λ.
Problem : how do we compute μ^{λ}?

Iterated Projection Fitting Procedure

Galichon-Salanié : the closed form formula links $\bar{s}^\lambda(I, J)$, μ^λ_{IJ}, μ^λ_{I0}, μ^λ_{0J} and λ.

The difficulty : make sure that μ^λ fits the margins $F(I)$, $G(J)$.

The solution : start from a well-chosen $\mu^{\lambda,0}$, project on the margins for men (F), then on the margins for women (G), and iterate

using the proper quasi-distance (a Bregman divergence) dictated by the specification of the errors for λ.

A Two-index Model of Surplus

Chiappori-Oreffice-Quintana-Domeque assume that there exists two *scalar* functions m and f such that

$$s(I, J, \varepsilon, \eta) = S(m(I), \varepsilon, f(J), \eta)$$

and $\varepsilon \perp\!\!\!\perp J | f(J)$, $\eta \perp\!\!\!\perp I | m(I)$.
e.g. $m(I)$ is the socially agreed (by women) attractiveness index for men.
Then $\mu(I, J) \equiv \nu(m(I), f(J))$;
and on realized matches : $J \perp\!\!\!\perp I | m(I)$ and $I \perp\!\!\!\perp J | f(J)$.

Estimation and Testing

Regressing observed characteristics of wife J on observed characteristics of husband I identifies the indifference curves of $m(I)$;

cross-equation restrictions : same index m for *all* components of $J \longrightarrow$ testable.

Generalizes to multi-index models.

Comparing Similar Markets

Suppose we observe many markets that

- are separated (no matching across markets)
- have the same surplus function $s(x, y)$
- but differ in their sample margins.

Each such market n has a list of observed characteristics C_n (number of men of type $I = 1$, of women of type $J = 3\dots$) and an associated matching μ_n (how many such men marry such women...)
and we can estimate $\Pr(\mu_n = \mu | C_n = C.)$

Fox's Rank-order Property

States that for every C and all μ^1, μ^2 feasible for C

$$\Pr(\mu_n = \mu^1 | C_n = C) \geq \Pr(\mu_n = \mu^2 | C_n = C) \text{ iff } \sum_{\mu^1} \bar{s}(I, J) \geq \sum_{\mu^2} \bar{s}(I, J).$$

Intuition : in single-agent choice problems, under weak conditions
$\Pr(d = d_k | U_1, \ldots, U_K)$ is an increasing function of mean utilities U_k.

Pluses and Minuses

Maximum score estimation can rely on selected stability conditions

is easy to reformulate beyond one-to-one matching (Fox Bajari 2009, Fox 2011)

But not founded on a microeconomic model of matching ; e.g. from Galichon-Salanié, with separability μ_n maximizes

$$\sum_{\mu} \bar{s}(I, J) + \mathcal{E}(\mu; C_n)$$

with $\mathcal{E}$ a generalized entropy term that comes from non-random matching over unobservables.

NTU version

Hitsch-Hortacscu-Ariely (*AER* 2010) : man x is matched with woman y iff

$$U(x,y) \geq U(x,z) \;\; \forall z \;\; \text{s.t. } V(x,z) \geq v(y).$$

Specify $U(x,y) = \bar{U}(I,J;\lambda) + \xi_{xy}$, and $v(y)$ as a fixed effect ; if ξ_{xy} is iid type I EV, gives a conditional logit model to recover λ. Very restrictive assumption (more than separability ; no (x,J) interaction for instance.)

Solving the System

Brute-force approach : specify $U(x, y; \lambda)$ and $V(x, y; \lambda)$, solve for $u(x; \lambda)$ and $v(y; \lambda)$,
get $\mu(x, y; \lambda)$ and take it to the data.
(TU version : only specify $s(x, y; \lambda)$.)
In general, using the coupled equations is computationally intensive
And identification is mostly unexplored.

Concluding remarks

A lot remains to be done :
applications
testing of identifying assumptions in the various approaches
bridging theory and data for “matching in contracts”
frictions and dynamics (as in the job search literature.)