Separable Matching

Alfred Galichon[1] Bernard Salanié[2]

[1]New York University and Sciences Po

[2]Columbia University

October 2023
Berkeley

Early 2000s consensus:

Estimation is hard!

Estimating matching models is hard. . .

the choice is between brute-force methods (simulating and fitting)

and ad hoc, unjustified regressions.

Now:

Estimation is easy!

At least for separable, TU matching markets.

How Easy?

Take the most popular "marriage model": Choo and Siow 2006

A match between a man m with observed characteristics $x = 1, \ldots, X$
and a woman w with observed characteristics $y = 1, \ldots, Y$

generates joint utility

$$\boldsymbol{\phi}_{xy} \cdot \boldsymbol{\beta}_0 + \varepsilon_{my} + \eta_{xw}$$

and the ε, η are iid standard Gumbel (type I EV).

Suppose we only observe numbers of matches ($\hat{\mu}_{xy}$) — not singles.

The Brute-force Method

1. pick values $\boldsymbol{\beta}$
2. draw ε and η vectors for each man and each women
3. solve for the optimal assignment
4. aggregate to get $(\mu_{xy}(\boldsymbol{\beta}))$
5. compare with the observed $(\hat{\mu}_{xy})$
6. iterate until happy.

It does give valid estimates,
but it is laborious, and not very illuminating.

Ad hoc Regressions

1. regress $\hat{\mu}_{xy}$ on x and y dummies, maybe other covariates
2. do wet-finger interpretation.

What do the coefficients of the regression mean?

Estimation by Generalized Least-squares

1. estimate $\text{Var}\,\hat{\boldsymbol{\mu}}$
2. then $S^* = \left(\text{Var}\,(2\log\hat{\boldsymbol{\mu}})\right)^{-1}$
3. solve
$$(\phi' S^* \phi)\,\hat{\boldsymbol{\beta}} = 2\phi' S^* \log\hat{\boldsymbol{\mu}}.$$
4. if the model is well-specified,
$$\|\phi\hat{\boldsymbol{\beta}} - 2\log\hat{\boldsymbol{\mu}}\|^2_{S^*}$$
is (asymptotically) a χ^2 with $X \times Y - \dim \boldsymbol{\beta}_0$ degrees of freedom.

$\hat{\boldsymbol{\beta}}$ is a consistent and asymptotically normal estimator of $\boldsymbol{\beta}_0$

and we also get a specification test (basically the sum of square residuals).

Matching: TU, one-to-one, bipartite

one-to-one and bipartite: each match is a couple with one partner in each of two given subpopulations

Call it "(heterosexual) marriage" with "men" and "women".

A match of man m with woman w must be voluntary
$\rightarrow$ it must make them both better off than any other match, or singlehood ("partnered with 0").

If m ends up with utility u_m and w with v_w, we must have

$$u_m + v_w = \Phi_{mw}$$

where Φ_{mw} is the sum of the (transferable) utilities they get when together.

Moreover,

- $u_m \geq \Phi_{mw} - v_w$ for any other woman w, and for $w = 0$
- $v_w \geq \Phi_{mw} - u_m$ for any other man m, and for $m = 0$.

Stability Equations

$$u_m + v_w \geq \Phi_{mw} \text{ for all } m, w$$

with equality if m, w are matched "in equilibrium".

"Equilibrium" solves the dual $\min \sum_m u_m + \sum_w v_w$ under stability.

The primal is $\mu_{mw} \in [0, 1]$ that maximizes $\sum_{m,w} \mu_{mw} \Phi_{mw}$ under the margin constraints

$$\sum_w \mu_{mw} + \mu_{m0} = 1 \text{ for all } m$$

$$\sum_m \mu_{mw} + \mu_{0w} = 1 \text{ for all } w.$$

Econometrics means unobserved heterogeneity

Now we want to write $\Phi_{mw} = Q(x_m, y_w, \zeta_{mw})$ where the econometrician observes

- all x_m and y_w
- whether any m and w end up being matched
- but not the ζ_{mw}.

Problem: we know that estimating even one-sided choice models require strong assumptions and/or a lot of data

here we have two-sided choice.

we need to simplify (restrict) the ζ_{mw}.

Separability

Much of the literature assumes separability:

if $x_m = x$ and $y_w = y$, then

$$\Phi_{mw} = \bar{\Phi}_{xy} + \varepsilon_{my} + \eta_{xw};$$

no interaction between the unobserved characteristics of m and w, conditional on $(x_m = x, y_w = y)$

allows for restricted matching on unobservables.

Separability as Dimension-Reduction

Choo-Siow 2006, Chiappori-Salanié-Weiss 2017, Galichon-Salanié 2022: in equilibrium, there exists U_{xy} and V_{xy} such that

- m with $x_m = x$ gets utility $u_m = \max_y (U_{xy} + \varepsilon_{my})$
- w with $y_w = y$ gets utility $v_w = \max_x (V_{xy} + \eta_{xw})$
- $U_{xy} + V_{xy} \geq \bar{\Phi}_{xy}$, with equality if some x and some y match.

Identification

Choose some distribution for $(\varepsilon_{m0}, \ldots, \varepsilon_{mY})$ for given x, etc

Then

$$\bar{\Phi}_{xy} = -\frac{\partial \mathcal{E}}{\partial \mu_{xy}}(\boldsymbol{\mu})$$

where the generalized entropy function $\mathcal{E}$ depends on the choice of distributions and on the group sizes.

it measures the total surplus generated by matching on unobservables.

Why?

Since m with $x_m = x$ maximizes $U_{xy} + \varepsilon_{my}$,
the expected utility of men of type x is

$$G_x(\boldsymbol{U}_{x\cdot}) = E_{\boldsymbol{\varepsilon}} \max(U_{xy} + \varepsilon_{my})$$

It is convex in $\boldsymbol{U}_{x\cdot}$, with gradient a.e.

$$\mu_{y|x} = \frac{\partial G_x}{\partial U_{xy}}(\boldsymbol{U}_{x\cdot})$$

and by convex duality

$$U_{xy} = \frac{\partial G_x^*}{\partial \mu_{y|x}}(\boldsymbol{\mu}_{\cdot|x})$$

where G_x^* is the Legendre-Fenchel transform of G_x.

Why, ctd

We do the same on women's side and we get,
if there are matches betwen x and y:

$$\bar{\Phi}_{xy} = U_{xy} + V_{xy} = \frac{\partial G_x^*}{\partial \mu_{y|x}}(\boldsymbol{\mu}_{\cdot|x}) + \frac{\partial H_y^*}{\partial \mu_{x|y}}(\boldsymbol{\mu}_{\cdot|y})$$

which defines (minus) the derivatives of the generalized entropy $\mathcal{E}$.

Minimum Distance Estimation

Let $\boldsymbol{\alpha}$ parameterize the distributions, and $\boldsymbol{\beta}$ for $\bar{\Phi}$

We have a mixed hypothesis:

$$\exists \boldsymbol{\lambda} \equiv (\boldsymbol{\alpha}, \boldsymbol{\beta}) \text{ s.t. for all } x, y, \ \bar{\Phi}^{\beta}_{xy} = -\frac{\partial \mathcal{E}_{\boldsymbol{\alpha}}}{\partial \mu_{xy}}(\mu).$$

1. we get $\hat{\boldsymbol{\mu}}$ from the data
2. we minimize a suitably weighted norm of the matrix

$$\bar{\boldsymbol{\Phi}}^{\boldsymbol{\beta}} + \frac{\partial \mathcal{E}_{\boldsymbol{\alpha}}}{\partial \boldsymbol{\mu}}(\hat{\boldsymbol{\mu}}).$$

Testing

If the weighted norm is chosen optimally,

then its value at the minimum over $\boldsymbol{\lambda}$ is a χ^2 if the model is well-specified.

$\rightarrow$ a "catch-all" specification test.

Simple Subcases

For many choices of the distributions (but not e.g. with random coefficients),

the derivatives of the generalized entropy $\mathcal{E}_{\boldsymbol{\alpha}}$ are linear in $\boldsymbol{\alpha}$

then one can minimize the weighted norm "profiled" on $\boldsymbol{\beta}$ only.

If moreover $\bar{\Phi}$ is linear in $\boldsymbol{\beta}$, we get quasi-generalized least squares (cf the opening example).

Extensions in the Paper

Many-to-one matching

Multipartite matching

Unipartite matching.