Matching, choice, and entropy

Alfred Galichon

Based on joint works with:

K. Chiong, A. Dupuy, S.D. Kominers,

B. Salanié, M. Shum, and S. Weber

New York University, Courant Institute

November 12, 2014

This talk is based on the following papers:

– A. Dupuy, and AG "Taxation and Sorting: the Second Tax Incidence Effect"

– Chiong, K., AG, and M. Shum, "Duality in dynamic discrete choice models"

– AG, "DARUM: Deferred Acceptance for Random Utility Models"

– AG, S. D. Kominers and S. Weber, "Costly Concessions: Estimating the Unintended Consequences of Policy Intervention in Matching Markets"

– AG and B. Salanié, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models"

Introduction

What is structural econometrics? what is it for?

- Estimation of preferences

- Counterfactuals

Today's setting: employees and firms.

Exogenous quantities: employee's preferences and productivity, scarcity of employee's skills and positions.

Endogenous quantities (equilibrium): wages, matching patterns (who works for whom)

Estimation: Employee's preferences and productivity are unobserved. On would like to infer them based on observed equilibirum quantities: wage and matching patterns.

Counterfactuals: what happens if one increases taxes?

This talk

1. The assignment probem: optimality and equilibrium

2. Linear taxes: comparative statics

3. Estimation: entropy of choice

4. Nonlinear taxes: equilibrium vs optimality

1 The assignment problem: optimality and equilibrium

Assume employee's characteristics is indexed by x, and firm's characteristics by y.

There are n_x employees of type x and m_y firms of type y.

Employee x working with firm y enjoys job amenity α_{xy} and has productivity γ_{xy}. Their joint output is

$$\Phi_{xy} = \alpha_{xy} + \gamma_{xy}.$$

One looks for the equilibrium *matching* μ_{xy} (number of employees of type x working with firms of type y), and the equilibrium *wage* w_{xy}.

For now, let's assume away taxes: if x works for y, then the payoff of x and y are respectively

$$\alpha_{xy} + w_{xy} \text{ and } \gamma_{xy} - w_{xy}$$

and each of them gets 0 if they remain unassigned.

Optimal assignment. One can compute the optimal assignment (central planner's solution) as

$$S(\Phi) = \max_{\mu \in \mathcal{M}} \sum_{xy} \mu_{xy} \Phi_{xy}$$

where the set of feasible matchings $\mathcal{M}$ is the set

$$\mathcal{M} = \{\mu \geq 0 : \sum_y \mu_{xy} \leq n_x \text{ and } \sum_x \mu_{xy} \leq m_y\}.$$

The expression of S coincides with its dual

$$S(\Phi) = \min_{u,v \geq 0} \{\sum_x n_x u_x + \sum_y m_y v_y : u_x + v_y \geq \Phi_{xy}\}$$

and by complementary slackness, for μ and (u, v) solution of these programs, $\mu_{xy} > 0$ implies $u_x + v_y = \Phi_{xy}$.

This is a manifestation of **Monge-Kantorovich duality** in optimal transportation. Conjugacy between u and v:

$$\begin{aligned} u_x &= \max_y \{\Phi_{xy} - v_y\} \\ v_y &= \max_x \{\Phi_{xy} - u_x\}. \end{aligned}$$

Equilibrium assignment. Any w_{xy} satisfying

$$\gamma_{xy} - v_y \leq w_{xy} \leq u_x - \alpha_{xy}$$

is an equilibrium wage; indeed

$$u_x = \max_y \{\alpha_{xy} + w_{xy}, 0\} \text{ and } v_y = \max_x \left\{\gamma_{xy} - w_{xy}, 0\right\}.$$

Hence, in this setting, **optimality coincides with equilibrium** (Shapley and Shubik 1972, Becker 1973).

This talk

1. The assignment probem: optimality and equilibrium

2. Linear taxes: comparative statics

3. Estimation: entropy of choice

4. Nonlinear taxes: equilibrium vs optimality

2 Linear taxes: comparative statics

Our purpose: simulate policy intervention ("comparative statics"). Here, we want to understand the effect of a raise in income tax.

Let w_{xy} be the gross wage; there is a (proportional) tax tw paid by the employee.

Employee x's and firm y's indirect surpluses are now respectively

$$\begin{aligned} u_x &= \max_y \{\alpha_{xy} + (1-t)\, w_{xy}, 0\}, \\ v_y &= \max_x \left\{\gamma_{xy} - w_{xy}, 0\right\}. \end{aligned}$$

Rewrites this as

$$\begin{aligned} \frac{u_x}{1-t} &= \max_y \left\{\frac{\alpha_{xy}}{1-t} + w_{xy}, 0\right\}, \\ v_y &= \max_x \left\{\gamma_{xy} - w_{xy}, 0\right\}. \end{aligned}$$

Denoting

$$\lambda = \frac{1}{1-t} > 1,$$

it becomes apparent that:

Theorem 2.1. *The equilibrium matching μ_{xy} is optimal for fictitious surplus*

$$\tilde{\Phi}_{xy} = \lambda \alpha_{xy} + \gamma_{xy}.$$

That is, μ_{xy} is solution to

$$\begin{aligned} & \max_{\mu \geq 0} \sum_{xy} \mu_{xy} \tilde{\Phi}_{xy} \\ s.t. \qquad & \sum_{y} \mu_{xy} \leq n_x \\ & \sum_{y} \mu_{xy} \leq n_x. \end{aligned}$$

Note that this is not the only instance where equilibrium is the solution of an optimization problem; Wardorp equilibria in congested network flow problems are another one. Later on we shall see instances where equilibrium has no formulation as an optimization problem.

Two consequences are in order:

1. When the tax rate increases, the relative weight of the amenity term (versus the productivity term) increases in importance in the determination the sorting. We will quantify this intuition next.

2. We get the following result:

Proposition 2.1. *When $\alpha = 0$ (no amenity term), then the tax has no effect on the optimal sorting.*

Comparative statics. There are two terms of interest here:

- The *amenity term*, which comes from α. The aggregate amenity is $\sum_{xy} \mu_{xy} \alpha_{xy}$, which we denote in vector notations by

$$A = \alpha \cdot \mu = \sum_{xy} \mu_{xy} \alpha_{xy}.$$

- The *productivity term*, which comes from γ. The aggregate productivity is likewise given by

$$\Gamma = \gamma \cdot \mu = \sum_{xy} \mu_{xy} \gamma_{xy}.$$

Our measure of social welfare will be

$$W = A + \Gamma$$

or just Γ, depending on whether we are taking account employee's happiness or not.

We would like to understand how these quantities evolve when one raises the tax rate (i.e. when one raises λ).

Note that the former analysis implies that, formally,

$$\mu = \nabla S\left(\tilde{\Phi}\right) = \nabla S\left(\lambda\alpha + \gamma\right) \tag{2.1}$$

Because S is positive homogenous of degree one, we get by Euler's identity that $S = \tilde{\Phi} \cdot \mu = \tilde{\Phi} \cdot \nabla S\left(\tilde{\Phi}\right)$, where the notation $\mu \cdot \tilde{\Phi}$ denotes $\sum_{xy} \mu_{xy} \tilde{\Phi}_{xy}$ and differentiating twice yields

$$\left(D^2 S\left(\tilde{\Phi}\right)\right)\tilde{\Phi} = 0 \tag{2.2}$$

that is, $\tilde{\Phi}$ is in the nullspace of $D^2 S\left(\tilde{\Phi}\right)$. Note that by convexity of S, $D^2 S\left(\tilde{\Phi}\right)$ is a semidefinite positive.

Differentiation obtains $d\mu/d\lambda = \left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\alpha$. Therefore

$$\frac{dA}{d\lambda} = \alpha\cdot\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\alpha \geq 0$$

By (2.2), we get $\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\left(\lambda\alpha+\gamma\right) = 0$, hence

$$\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\gamma = -\lambda\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\alpha$$

hence

$$\begin{aligned}\frac{d\Gamma}{d\lambda} &= \gamma\cdot\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\alpha \\ &= -\lambda\alpha\cdot\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\alpha \leq 0\end{aligned}$$

and

$$\frac{dA}{d\lambda}+\frac{d\Gamma}{d\lambda} = \left(1-\lambda\right)\alpha\cdot\left(D^2 S\left(\lambda\alpha+\gamma\right)\right)\alpha \leq 0.$$

This allows us to recover a qualitative result of Jaffe and Kominers (2013):

Proposition 2.2. *When the tax rate increases (i.e. when λ increases), then:*

(i) aggregate amenity A increases

(ii) aggregate productivity Γ decreases

(iii) total welfare $W = A + \Gamma$ decreases.

Further,

$$\frac{d\Gamma}{d\lambda} = -\lambda \frac{dA}{d\lambda}.$$

As we have seen, this analysis crucially depends on the estimation of α and γ, which we focus on next.

This talk

1. The assignment probem: optimality and equilibrium

2. Linear taxes: comparative statics

3. **Estimation: entropy of choice**

4. Nonlinear taxes: equilibrium vs optimality

3 Estimation: entropy of choice

So far, we have assumed that α and γ were known and given and we have deduced the equilibrium matching μ and wage w. The econometrician's problem is actually the converse problem: given μ and w, how to infer α and γ.

To answer this problem, we need to introduce unobserved heterogeneity.

Employee x considers between working for firm $y \in \mathcal{Y}_0 = \mathcal{Y} \cup \{0\}$, where 0 is not working.

Assume a utility shock which is a vector $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_{|\mathcal{Y}|}) \sim \mathbf{P}_x$ such that the utility of option y is $U_{xy} + \varepsilon_y$, while the outside option yields utility ε_0.

The preferred option is the one which attains the maximum in

$$\max_{y \in \mathcal{Y}} (U_{xy} + \varepsilon_y, \varepsilon_0)$$

where $U_{xy} = \alpha_{xy} + w_{xy}$.

The expected utility (conditional to belonging to this group) of an employee in group x is

$$G_x(U_{x.}) = \mathbb{E}\left[\max_{y \in \mathcal{Y}} (U_{xy} + \varepsilon_y, \varepsilon_0)\right].$$

G_x is a convex function of U. Note that

$$\frac{\partial G_x}{\partial U_{xy}}(U_{x.}) = \mu_{y|x} \tag{3.1}$$

where $\mu_{y|x}$ is the probability that an employee of type x chooses firm y.

Consider the Legendre-Fenchel transform of G_x

$$G_x^*(\mu_{.|x}) = \max_U \left(\sum_{y \in \mathcal{Y}} \mu_{y|x} U_{xy} - G_x(U) \right). \tag{3.2}$$

One has

$$\frac{\partial G_x^*}{\partial \mu_{y|x}}(\mu_{.|x}) = U_{xy} \tag{3.3}$$

Interpretation. Conjugacy implies that if $\mu_{.|x}$ and $U_{x.}$ are related by (3.1), then

$$G_x(U_{x.}) = \sum_{y \in \mathcal{Y}} \mu_{y|x} U_{xy} - G_x^*(\mu_{.|x}). \tag{3.4}$$

Thus letting Y_i^* be the optimal choice of y given $\varepsilon_{i.}$, one has

$$-G_x^*(\mu_{.|x}) = G_x(U_{x.}) - \sum_{y \in \mathcal{Y}} \mu_{y|x} U_{xy} = \mathbb{E}\left[\varepsilon_{iY_i^*}\right]. \tag{3.5}$$

Hence, $G_x^*\left(\mu_{.|x}\right)$ is interpreted as (minus) the expected amount of heterogeneity needed to rationalize the choice probabilities of an agent of type x, and we call it the *generalized entropy* of the corresponding discrete choice problem.

Examples. We take a few examples of distributions where G_x and/or G_x^* can be computed in closed form.

Our first example is the classical logit model, which is obtained as a particular case of the results in the previous section when the $\mathbf{P}_x$ distributions are iid standard type I extreme value:

Example 3.1 (logit). *Assume further that $\mathbf{P}_x$ are the distributions of i.i.d. standard type I extreme value random variables. Then*

$$
\begin{aligned}
G_x(U_{x.}) &= \log\left(1+\sum_{y\in\mathcal{Y}}\exp(U_{xy})\right) \\
\textit{and } G_x^*\left(\mu_{.|x}\right) &= \mu_{0|x}\log(\mu_{0|x})+\sum_{y\in\mathcal{Y}}\mu_{y|x}\log\mu_{y|x}.
\end{aligned}
$$

where $\mu_{0|x}=1-\sum_{y\in Y}\mu_{y|x}$. The fact that G_x^* is an actual entropy explains our choice of terminology "generalized entropy".

The logit model is the simplest example which fits into McFadden's Generalized Extreme Value (GEV) framework, where G_x obtains in closed form. As a more complex example of a GEV distribution, we turn to a nested logit model.

Example 3.2 (Nested Logit). *Suppose for instance that men of a given group x are concerned about the social group of their partner and her education, so that the type $y = (s, e)$. We can allow for correlated preferences by modeling this as a nested logit in which educations are nested within social groups. Let $\mathbf{P}_x$ have cdf*

$$F(w) = \exp\left(-\exp(-w_0) - \sum_s \left(\sum_e \exp(-w_{se}/\sigma_s)\right)^{\sigma_s}\right)$$

This is a particular case of the Generalized Extreme Value (GEV) framework. The numbers $1/\sigma_s$ describe the correlation in the surplus generated with partners of different education levels within social group s. Then (dropping the x indices for notational simplicity, so that for instance μ_s denotes the number of matches with women in social *group s)*

$$\begin{aligned} G_x(U_{x.}) &= \log\left(1+\sum_s\left(\sum_e \exp(U_{se}/\sigma_s)\right)^{\sigma_s}\right), \text{ and} \\ G_x^*\left(\mu_{.|x}\right) &= \mu_0 \log \mu_0 + \sum_s (1-\sigma_s)\,\mu_s \log \mu_s \\ &\quad + \sum_s \sigma_s \sum_e \mu_{se} \log \mu_{se}. \end{aligned}$$

The corresponding choice probabilities are given by

$$\mu_{se} = \frac{\left(\sum_{e'} \exp(U_{se'}/\sigma_s)\right)^{\sigma_s}}{1+\sum_{s'}\left(\sum_{e'} \exp(U_{s'e'}/\sigma_{s'})\right)^{\sigma_{s'}}} \frac{\exp(U_{se}/\sigma_s)}{\sum_{e'} \exp(U_{se'}/\sigma_s)}.$$

We investigate a model that does not belong in the GEV class: the Random Scalar Coefficient (RSC) model, where the dimension of $\zeta_x(y)$ and $\boldsymbol{\varepsilon}_i$ is one. This model yields simple closed-form expressions, even though it does not belong to the Generalized Extreme Value (GEV) class.

Example 3.3 (Random Uniform Scalar Coefficient (RUSC) models). *Assume that for each man i in group x,*

$$\varepsilon_{iy} = \varepsilon_i \zeta_x(y),$$

where $\zeta_x(y)$ is a scalar index of the observable characteristics of women which is the same for all men in the same group x, and the $\varepsilon_i \sim U([0,1])$'s are iid random variables. We call this model the Random Uniform Scalar Coefficient (RUSC) model. For any $x \in \mathcal{X}$, let S^x be the square matrix with elements $S^x_{yy'} = \max(\zeta_x(y), \zeta_x(y'))$ for $y, y' \in \mathcal{Y}_0$. Define T^x by $T^x_{yy'} = S^x_{y0} + S^x_{0y'} - S^x_{yy'} - S^x_{00}$, and let $\sigma^x_y = S^x_{00} - S^x_{y0}$. Then G^*_x is quadratic with respect to $\boldsymbol{\mu}_{\cdot|\mathbf{x}}$:

$$G^*_x(\mu_{\cdot|x}) = \frac{1}{2}(\boldsymbol{\mu}_{\cdot|\mathbf{x}}{}' T^x \boldsymbol{\mu}_{\cdot|\mathbf{x}} + 2\sigma^x . \boldsymbol{\mu}_{\cdot|\mathbf{x}} - S^x_{00}).$$

Aggregate demand. We now consider the previous problem when there are a number of men of type $x \in \mathcal{X}$. Call n_x the number of men of type x.

Consider the aggregate social welfare function

$$G(U) = \sum_{x \in \mathcal{X}} n_x G_x(U_{x\cdot})$$

and call $\nabla G(U) = (\partial G(U) / \partial U_{xy})$ the vector of partial derivatives of G. One has

$$\partial G(U) / \partial U_{xy} = \partial G_x(U_{x\cdot}) / \partial U_{xy} = n_x \mu_{y|x} = \mu_{xy}$$

where μ_{xy} is the number of men of type x wishing to marry a woman of type y. Hence the total "demand" for type y is

$$\sum_{x \in \mathcal{X}} \mu_{xy} = \sum_{x \in \mathcal{X}} \partial G(U) / \partial U_{xy}.$$

Note that automatically, if $\mu = \nabla G(U)$, then

$$\sum_{y \in \mathcal{Y}} \mu_{xy} \leq n_x$$

for all x; however, there is no restriction as to $\sum_{x \in \mathcal{X}} \mu_{xy}$. The problem is that, in a matching problem, there is a fixed number of women of type $y \in \mathcal{Y}$ too. Hence we need to describe the market constraints that adjust demand to supply.

Letting

$$G^*(\mu) = \max_U \sum_{xy} \mu_{xy} U_{xy} - G(U)$$

be the Legendre-Fenchel transform of G, one has that

$$\mu = \nabla G(U)$$

is inverted into

$$U = \nabla G^*(\mu).$$

But recall that $U_{xy} = \alpha_{xy} + w_{xy}$. Thus α_{xy} is identified by

$$\alpha_{xy} = \frac{\partial G^*(\mu)}{\partial \mu_{xy}} - w_{xy}.$$

Firm's problem. A similar formula holds on the firms' side. Assuming that the firm's problem is to choose employee's type y that solves

$$\max_x \left\{ V_{xy} + \eta_y, 0 \right\}$$

where $V_{xy} = \gamma_{xy} - w_{xy}$, and $\eta \sim \mathbf{Q}_y$, and letting

$$\begin{aligned} H_y(V) &= \mathbb{E}\left[\max_x \left\{ V_{xy} + \eta_y, 0 \right\} \right] \\ H(V) &= \sum_{y \in \mathcal{Y}} H_y(V) \end{aligned}$$

one has

$$\mu_{xy} = \frac{\partial H(V)}{\partial V_{xy}} \text{ and } V_{xy} = \frac{\partial H^*(\mu)}{\partial \mu_{xy}}$$

thus γ is identified by

$$\gamma_{xy} = \frac{\partial H^*(\mu)}{\partial \mu_{xy}} + w_{xy}.$$

Hence we have solved the problem of estimating α and γ.

Equilibrium computation. One can use the previous results for computing an equilibrium. Indeed, recall

$$\begin{aligned}
\alpha_{xy} &= \frac{\partial G^*(\mu)}{\partial \mu_{xy}} - w_{xy} \\
\gamma_{xy} &= \frac{\partial H^*(\mu)}{\partial \mu_{xy}} + w_{xy}
\end{aligned}$$

thus

$$\Phi_{xy} = \frac{\partial G^*(\mu)}{\partial \mu_{xy}} + \frac{\partial H^*(\mu)}{\partial \mu_{xy}}$$

thus μ solves

$$\max_{\mu} \sum_{xy} \mu_{xy} \Phi_{xy} - \mathcal{E}(\mu)$$

where

$$\mathcal{E}(\mu) = G^*(\mu) + H^*(\mu)$$

is the generalized entropy of matching.

Some remarks are in order:

- This problem is a simulated annealing version of the optimal assignment problem. In some cases (when $\mathcal{E}$ is an actual entropy) it is extremely fast to compute via alternated projections.

- Indeed, a variant of this problem (with no unassigned agent) yields the system

$$\begin{aligned} \sum_y \mathsf{exp}\left(\Phi_{xy} - u_x - v_y\right) &= n_x \\ \sum_x \mathsf{exp}\left(\Phi_{xy} - u_x - v_y\right) &= m_y \end{aligned}$$

  which is called the *Schrödinger system*, investigated by Berstein, Fortet, Beurling, Föllmer, Rüschendorf inter al. See survey in Léonard (2012). Link with stochastic processes ("reciprocal diffusions").

This talk

1. The assignment probem: optimality and equilibrium

2. Linear taxes: comparative statics

3. Estimation: entropy of choice

4. **Nonlinear taxes: equilibrium vs optimality**

4 Nonlinear taxes: equilibrium vs optimality

Let us summarize the previous findings.

1. In the absence of taxes, we have

$$\begin{aligned} U_{xy} &= \alpha_{xy} + w_{xy} = \frac{\partial G^*(\mu)}{\partial \mu_{xy}} \\ V_{xy} &= \gamma_{xy} - w_{xy} = \frac{\partial H^*(\mu)}{\partial \mu_{xy}} \end{aligned}$$

thus μ is a solution of

$$\max_{\mu} \sum_{xy} \mu_{xy} \left(\alpha_{xy} + \gamma_{xy} \right) - (G^* + H^*)(\mu).$$

2. When there is a linear tax $(1-t)\,w_{xy}$ on the employee's income,

$$
\begin{aligned}
U_{xy} &= \alpha_{xy} + (1-t)\,w_{xy} = \frac{\partial G^*(\mu)}{\partial \mu_{xy}} \\
V_{xy} &= \gamma_{xy} - w_{xy} = \frac{\partial H^*(\mu)}{\partial \mu_{xy}}
\end{aligned}
$$

thus μ is a solution of

$$
\max_{\mu} \sum_{xy} \mu_{xy} \left(\frac{\alpha_{xy}}{1-t} + \gamma_{xy} \right) - \left(\frac{G^*}{1-t} + H^* \right)(\mu)
$$

which is still an optimization problem, even though it does not coincide with the social welfare.

We now investigate the effect of a nonlinear tax. We will assume that the gross wage $\gamma_{xy} - V_{xy}$ will be related to the net wage $U_{xy} - \alpha_{xy}$ by

$$\mathbf{\Psi}_{xy}\left(V_{xy} - \gamma_{xy}, U_{xy} - \alpha_{xy}\right) = 0$$

where $\mathbf{\Psi}_{xy}$ is continuous and increasing in both its arguments. Then equation becomes

$$\mathbf{\Psi}_{xy}\left(\frac{\partial G^*(\mu)}{\partial \mu_{xy}} - \gamma_{xy}, \frac{\partial H^*(\mu)}{\partial \mu_{xy}} - \alpha_{xy}\right) = 0. \quad (4.1)$$

When $\mathbf{P}_x$ and $\mathbf{Q}_y$ are iid Gumbel distributions, G^* and H^* are proper entropy functions, and

$$\frac{\partial G^*(\mu)}{\partial \mu_{xy}} = \log \frac{\mu_{xy}}{\mu_{x0}} \text{ and } \frac{\partial H^*(\mu)}{\partial \mu_{xy}} = \log \frac{\mu_{xy}}{\mu_{0y}}$$

and thus (4.1) rewrites

$$\mathbf{\Psi}_{xy}\left(\log \frac{\mu_{xy}}{\mu_{x0}} - \gamma_{xy}, \log \frac{\mu_{xy}}{\mu_{0y}} - \alpha_{xy}\right) = 0,$$

which defines implicitely

$$\mu_{xy} = M_{xy}\left(\mu_{x0}, \mu_{0y}\right).$$

Thus μ_{x0} and μ_{0y} are solution to

$$\begin{aligned}
\mu_{x0} + \sum_y M_{xy}\left(\mu_{x0}, \mu_{0y}\right) &= n_x \\
\mu_{0y} + \sum_x M_{xy}\left(\mu_{x0}, \mu_{0y}\right) &= m_y
\end{aligned}$$

which can be computed via alternative fitting.

Note that:

- Ther is no variational formulation for this problem.

- The variant of this problem with no unassigned agent yields the system

$$\begin{aligned}
\sum_y M_{xy}\left(u_x, v_y\right) &= n_x \\
\sum_x M_{xy}\left(u_x, v_y\right) &= m_y
\end{aligned}$$

  which is a nonlinear version of the Schrödinger system described above. Solution is a manifold of dimension 1 such that there is a complete ordering on $(u, -v)$.

5 Conclusion

- Structural estimation

- Counterfactual policy experiment

- Equilibrium vs Optimality