Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models

Alfred Galichon & Bernard Salanié, Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models, *The Review of Economic Studies*, Volume 89, Issue 5, October 2022, Pages 2600–2629

B4 村山拓未

Abstract

- **Research focus:** One-to-one matching model with transferable utility and general unobserved heterogeneity
- **Method:** Extended the separability assumption from Choo and Siow (2006)
- **Results:**
  - Shows that equilibrium matching maximize a social gain function, balancing complementarities in observable characteristics and matching on unobserved traits
  - Derives simple closed-form formulas to identify joint matching surplus and equilibrium utilities for all participants, given any known distribution of unobserved heterogeneity
- **Contributions:**
  - Provides efficient algorithms for computing stable matching and estimating parametric models
  - Revisits Choo and Siow's empirical application, demonstrating the potential of a more general approach

Novelty, Utility, Reliability

Novelty:

- Extends Choo and Siow's (2006) separability assumption to a more general framework

Utility:

- Provides practical solutions for identifying matching surplus and utilities.
- Offers efficient algorithms for stable matching and parametric model estimation.

Reliability:

- Conducts empirical approach

1. Introduction

Models of matching with transferable utilities

- Model the marriage problem as a matching problem (Becker, 1973)
- ”Assignment game” (Shapley and Shubik, 1972)
  - Models of matching with transferable utilities
- Applications of the model
  - Competitive equilibrium in good markets with hedonic pricing
  - Trade
  - The labor market
  - Industrial organization

Becker’s theory and its problem

Becker’s theory:

- The type of the partners are one-dimensional and complementary in producing surplus (Special case)
- Social optimum shows *positive assortative matching*:
  - higher types pair up with higher types

↕

The data:

- Matches are observed between partners with quite different characteristics

Choo and Siow's model

- Solution for Becker's problem:
  - Allow the matching surplus to incorporate latent characteristics – **heterogeneity**

Choo and Siow's model

- Conditions:
  - The unobserved heterogeneities enter the marital surplus quasi-additively
  - These heterogeneities are independent and identically distributed as standard type I extreme value terms
- Examples:
  - Evaluate the effect of the legalization of abortion on gains to marriage
  - Use Canadian data to measure the impact of demographic changes
- The idea has been used in various later studies

Choo and Siow‘s model

3 assumptions of their model:

1. The unobserved heterogeneities on the two side of a match do not interact in producing matching surplus (Separability assumption)
2. They are distributed as iid type I extreme values (Distributional assumption)
3. Populations are large

Contributions of the paper

1. **Extended idea of Choo and Siow's model**
   - Choo and Siow's distributional assumption is very special
     - Generate a MNL model
     - Specific restrictions on cross-elasticities
   - The authors show:
     - Choo and Siow's distributional assumption can be completely dispensed with
     - Choo-Siow framework can be extended to encompass much less restrictive assumptions on the unobserved heterogeneity
2. **Complete empirical approach**
   - Identification
   - Parametric estimation
   - computation
3. **Revisit the original Choo and Siow (2006) dataset on marriage patterns by age**

Other approaches

- **Market with transferable utilities**
  - Fox (2010, 2018)
  - Bajari and Fox (2013) – spectrum auctions
  - Fox *et al.* (2018) – identify the complementarity between unobserved characteristics
  - Gualdani and Sinha (2019) – partial identification issues in nonparametric matching models
- **Market with non-transferable utilities**
  - Menzel (2015) – investigation of large non-transferable utilities markets
  - School assignment, where preferences on one side of the market are highly constrained by regulation
  - Agarwal (2015) – matching in the US medical resident program

2. Framework and Roadmap

A bipartite matching market with transferable utility

- A bipartite, one-to-one matching market with transferable utility
- Maintains some of the basic assumptions of Choo and Siow (2006)
  - Utility transfers between partners are unconstrained
  - Matching is frictionless
  - No asymmetric information among potential partners
- An application to the heterosexual marriage market

Setting

$\mathcal{I}$	Set of men	$i \in \mathcal{I}$
$\mathcal{J}$	Set of women	$j \in \mathcal{J}$
$\mathcal{X}$	Set of groups of men	$x \in \mathcal{X}$
$\mathcal{Y}$	Set of groups of women	$y \in \mathcal{Y}$
n_x	Mass of men in group x	$\sum_x n_x + \sum_y m_y = 1$,
m_y	Mass of women in group y	$\boldsymbol{r} = (\boldsymbol{n}, \boldsymbol{m})$

- The analyst can observe groups are defined by the intersection of the characteristics
- The analyst cannot observe men and women of a given group differ along some dimensions

Setting 2

μ_{xy}	Mass of the couples where the man belongs to group x, and where the woman belongs to group y	$\boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}(\boldsymbol{r})$
$\mathcal{M}$	Set of μ	$\mathcal{M}(\boldsymbol{r})=\{\boldsymbol{\mu} \geq 0: \forall x \in \mathcal{X}, \sum_{y \in \mathcal{Y}} \mu_{xy} \leq n_x ; \forall y \in \mathcal{Y}, \sum_{x \in \mathcal{X}} \mu_{xy} \leq m_y\}$
μ_{x0}	Mass of single men of group x	$\mu_{x0} = n_x - \sum_{y \in \mathcal{Y}} \mu_{xy}$
μ_{0y}	Mass of single women of group y	$\mu_{0y} = m_y - \sum_{x \in \mathcal{X}} \mu_{xy}$
$\mathcal{X}_0$	Set of marital choices available to male agents	$\mathcal{X}_0 = \mathcal{X} \cup \{0\}$
$\mathcal{Y}_0$	Set of marital choices available to female agents	$\mathcal{Y}_0 = \mathcal{Y} \cup \{0\}$
$\mathcal{A}$	Set of marital arrangement	$\mathcal{A} = (\mathcal{X} \times \mathcal{Y}) \cup (\mathcal{X} \times \{0\}) \cup (\{0\} \times \mathcal{Y})$

Separability

Assumption 1 (Separability)

- **Joint utility of a match:** $\widetilde{\Phi}_{ij} = \Phi_{xy} + \varepsilon_{iy} + \eta_{xj}$
  - $\widetilde{\Phi}_{ij}$: Joint utility for man i (group x) and woman j (group y)
  - Φ_{xy}: Base utility between group x and y
  - ε_{iy}: Random term specific to man i
  - η_{xj}: Random term specific to woman j
- **Utility of single individuals:**
  - Single man i: $\widetilde{\Phi}_{i0} = \varepsilon_{i0}$
  - Single woman j: $\widetilde{\Phi}_{0j} = \eta_{0j}$
- **Distribution and finite expectations:**
  - Conditional on $x_i = x$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{iy})_{y\in\mathcal{Y}_0}$ has distribution $\boldsymbol{P}_x$
  - Conditional on $y_i = y, \boldsymbol{\eta}_j = (\eta_{xj})_{x\in\mathcal{X}_0}$ has distribution $\boldsymbol{Q}_y$
  - $max_{y\in\mathcal{Y}_0}|\varepsilon_{iy}|$ and $max_{x\in\mathcal{X}_0}|\eta_{xj}|$ have finite expectations under $\boldsymbol{P}_x$ and $\boldsymbol{Q}_y$, respectively

- Allow for "matching on unobservables"
- Rule out sorting on unobserved characteristics on both sides of the market
  - E.g. some unobserved preference of man i for some unobserved characteristics of woman j

$\boldsymbol{P}_x$ and $\boldsymbol{Q}_y$ are not only limited to the extreme value class

2.3 Objectives and a roadmap

Final goal

Develop inference tools for matching problems with transferable utility and separable unobserved heterogeneity

Steps

1. Two-sided matching problem resolves into a collection of one-sided problems of lower complexity (given separability)
2. Provide new results on discrete choice (one-sided) models
3. Stable matching solves a convex optimization problem
4. Use convex duality to identify the matching surplus
5. New computational methods to solve for the stable matching and to estimate underlying parameters

3. Social Surplus and Identification in the One-Side Case: Discrete Choice Models

Splitting the surplus

Proposition 1: Splitting the surplus

- Under Assumption 1, there exist $\boldsymbol{U} = \left(U_{xy}\right)$ and $\boldsymbol{V} = \left(V_{xy}\right)$ for $(x, y) \in \mathcal{A}$, with $U_{x0} = V_{0y} = 0$, such that at any stable matching (μ_{xy})

1. **Men's matching decision:**
   - A man i of group x marries a woman of group $y^* \in \mathcal{Y}$ if y^* maximizes $U_{xy} + \varepsilon_{iy}$ over $y \in \mathcal{Y}_0$
   - If the maximum is achieved at $y = 0$, the man remains single
   - Man $i's$ utility $\tilde{u}_i$ is the value of the maximum
2. **Women's matching decision:**
   - A woman j of group y marries a woman of group $x^* \in \mathcal{X}$ if x^* maximizes $V_{xy} + \eta_{xj}$ over $x \in \mathcal{X}_0$
   - If the maximum is achieved at $x = 0$, the woman remains single
   - Woman $j's$ utility $\tilde{v}_j$ is the value of the maximum
3. **Surplus splitting condition:**
   - $U_{xy} + V_{xy} \geq \Phi_{xy}$ for all $(x, y) \in \mathcal{A}$, with equality if $\mu_{xy} > 0$

Social surplus in discrete choice models

One-sided discrete choice problems

- An individual chooses from a set of alternatives $y \in \mathcal{Y}_0$
  - Utilities are $U_y + \varepsilon_y$
  - Assume the vector $\boldsymbol{\varepsilon} = (\varepsilon_y)_{y \in \mathcal{Y}_0}$ has a distribution $\mathbb{P}$; without loss of generality
  - $U_0 = 0,\ U = (U_1, \dots, U_{|Y|})$

The *ex ante* indirect surplus
= weighted sum of the mean utilities + generalized entropy of choice

Two characterizations of generalized entropy function

1. The convex conjugate of the *ex ante* indirect utility
2. The solution to an optimal transport problem (Galichon, 2016)

Generalized entropy of choice

The average utility of the agent

$$G(\boldsymbol{U}) = \mathbb{E}_P max_{y \in \mathcal{Y}_0} \left(U_y + \varepsilon_y\right) \qquad (3.1)$$

$$= \mathbb{E}_P \left(U_{Y_i^*} + \varepsilon_{i,Y_i^*}\right)$$

$$= \sum_{y \in \mathcal{Y}} \mu_y U_y + \mathbb{E}_P\left(\varepsilon_{i,Y_i^*}\right) \qquad (3.2)$$

- The expectation is taken over the random vector $\boldsymbol{\varepsilon} = \left(\varepsilon_{0,\ldots,} \varepsilon_{|\mathcal{Y}|}\right) \sim \boldsymbol{P}$
- The function G is known as the *Emax operator* in the discrete choice literature

$Y_i^* \in \mathcal{Y}_0$ is the optimal choice of individual i

μ_y is the proportion of individuals who choose alternative y

Legendre-Fenchel transform of G

$$\boldsymbol{\mu} = \left(\mu_1, \ldots, \mu_{|\mathcal{Y}|}\right)$$

$$G^*(\boldsymbol{\mu}) = \begin{cases} sup_{\widetilde{U}=(\widetilde{U}_1,\ldots,\widetilde{U}_{|\mathcal{Y}|})}\left(\sum_{y \in \mathcal{Y}} \mu_y \widetilde{U}_y - G(\widetilde{\boldsymbol{U}})\right), whenever \ \sum_{y \in \mathcal{Y}} \mu_y \leq 1 \\ +\infty \text{ , } otherwise \end{cases} \qquad (3.3)$$

- The domain of G^* is the set of $\boldsymbol{\mu}$ that can be interpreted as vectors of choice probabilities of alternatives in $\mathcal{Y}$
- G^* is a convex function

Generalized entropy of choice

Definition 1.

The function $-G^*$ is the generalized entropy of choice

$$G(\boldsymbol{U}) = sup_{\tilde{\boldsymbol{\mu}}=(\tilde{\mu}_1,\ldots,\tilde{\mu}_{|\mathcal{Y}|})}\left(\sum_{y\in\mathcal{Y}}\tilde{\mu}_y U_y - G^*(\tilde{\boldsymbol{\mu}})\right) \quad (3.4)$$

$$G(\boldsymbol{U}) + G^*(\boldsymbol{\mu}) = \sum_{y\in\mathcal{Y}}\mu_y U_y$$

$$G^*(\boldsymbol{\mu}) = -\mathbb{E}_P\left(\varepsilon_{iY_i^*}\right)$$

The theory of convex duality implies that since G is convex, it is reciprocally the Legendre-Fenchel transform of G^*

Assume that $\boldsymbol{\mu}$ attains the supremum in (3.4)

- $-G^*$ is just the average heterogeneity that is required to rationalize the conditional choice probability vector $\boldsymbol{\mu}$

Characterization of the generalized entropy of choice

Theorem 1 (Characterization of the generalized entropy of choice)

Statement:

Let $\boldsymbol{\mu} = (\mu_1, \dots, \mu_{|\mathcal{Y}|})$ with $\sum_{y\in\mathcal{Y}} \mu_y \leq 1$, and define $\mu_0 = 1 - \sum_{y\in\mathcal{Y}} \mu_y$. Let $\mathcal{M}(\boldsymbol{\mu}, \boldsymbol{P})$ denote the set of probability distributions π of the random joint vector $(\boldsymbol{Y}, \boldsymbol{\varepsilon})$, where $\boldsymbol{Y} \sim (\mu_0, \boldsymbol{\mu})$ is a random element of $\mathcal{Y}_0$, and $\boldsymbol{\varepsilon} \sim \boldsymbol{P}$ is a random vector of $\mathbb{R}^{|\mathcal{Y}_0|}$.

Optimal transport interpretation:

$$-G^*(\boldsymbol{\mu}) = sup_{\pi\in\mathcal{M}(\boldsymbol{\mu},\boldsymbol{P})} \mathbb{E}_\pi(\varepsilon_{\boldsymbol{Y}}) \qquad (3.6)$$

- μ: Vector of choice probabilities for alternatives in $\boldsymbol{Y}$.
- π: Joint distribution of $(\boldsymbol{Y}, \boldsymbol{\varepsilon})$ with $\boldsymbol{Y} \sim (\mu_0, \boldsymbol{\mu})$ and $\boldsymbol{\varepsilon} \sim \boldsymbol{P}$.
- $\mathcal{M}(\boldsymbol{\mu}, \boldsymbol{P})$: Set of feasible joint distributions.
- $\varepsilon_{\boldsymbol{Y}}$: Surplus given by the chosen $\boldsymbol{Y}$

Explanation:

$-G^*(\boldsymbol{\mu})$ represents the value of the optimal transport problem between the distribution $(\mu_0, \boldsymbol{\mu})$ of $\boldsymbol{Y}$ and the distribution of $\boldsymbol{P}$ of $\boldsymbol{\varepsilon}$, where the objective is to maximize the expected surplus $\mathbb{E}_\pi(\varepsilon_{\boldsymbol{Y}})$.

Identification of discrete choice models

Theorem 2 (Identifying the mean utilities)

Given:

- $\boldsymbol{\mu} = (\mu_1, \dots, \mu_{|\mathcal{Y}|})$ with $\sum_{y \in \mathcal{Y}} \mu_y \leq 1$
- $U_0 = 0$ and $\boldsymbol{U} = (U_1, \dots, U_{|Y|})$
- Distribution ***P*** with full support, absolutely continuous w.r.t. the Lebesgue measure

Equivalent statements:

1. For every $y \in \mathcal{Y}$, $\mu_y = \frac{\partial G}{\partial U_u}(\boldsymbol{U})$ (3.7)
2. For every $y \in \mathcal{Y}$, $\mathrm{U}_y = \frac{\partial G^*}{\partial \mu_y}(\boldsymbol{\mu})$ (3.8)
3. There exists a scalar function $u(\boldsymbol{\varepsilon})$, integrable w.r.t. ***P***, such that $(u, \boldsymbol{U})$ are the unique minimizers of the dual problem to (3.6):

$$-G^*(\boldsymbol{\mu}) = min_{U,u} \int \bar{u}(\boldsymbol{\varepsilon}) d\boldsymbol{P}(\boldsymbol{\varepsilon}) - \sum_{y \in \mathcal{Y}} \mu_{y U_y}$$

$$\text{s.t.} \qquad \bar{u}(\boldsymbol{\varepsilon}) - \bar{U}_y \geq \varepsilon_y \ \forall y \in \mathcal{Y}, \forall \boldsymbol{\varepsilon} \in \mathbb{R}^{\mathcal{Y}_0}, \bar{U}_0 = 0.$$

- These conditions provide a way to uniquely identify mean utilities $\boldsymbol{U}$ from observed choice probabilities $\boldsymbol{\mu}$ under the given distribution ***P***

Daly-Zachary-Williams theorem

Fenchel duality theorem:
(3.7) and (3.8) are equivalent

- 1 is well-known in the discrete choice literature
- 2 and 3 provide a constructive method to identify U_y based on the conditional choice probabilities $\boldsymbol{\mu}$
  - As the solution to a convex optimization problem (2)
  - An optimal transport problem (3)

Examples

1. **Logit and nested Logit**
   - Two-layer nested logit model
     - Alternative 0 is alone in a nest
     - each other nest $n \in \mathcal{N}$ contains alternatives $y \in \mathcal{Y}(n)$
     - Correlation of alternatives within nest n is $1 - \lambda_n^2$ ($\lambda_0 = 1$ for the nest made of alternative 0)
   - Multinomial logit model (MNL)
     - When $\lambda_n = 1$ for every nest n
2. **Random coefficients multinomial logit and pure characteristics model**
   - Random coefficient logit model
     - Error term ε:

$$\boldsymbol{\varepsilon} = \boldsymbol{Z}\boldsymbol{e} + T\boldsymbol{\eta}$$

       - $\boldsymbol{e}$ is a random vector on $\mathbb{R}^d$ with distribution $\boldsymbol{P}_e$
       - $\boldsymbol{Z}$ is a $|\mathcal{Y}_0| \times d$ matrix
       - $T > 0$ is a scalar parameter
       - $|\mathcal{Y}|$ extreme value type-I (Gumbel) random variables, independent of $\boldsymbol{e}$
   - Pure characteristics model
     - When $T = 0$
     - Solution to the power diagram problem (Galichon, 2016)

4. Social Surplus and Identification in the Two-Side Case: Matching Models

Matching models

- Define G_x to be corresponding Emax function, based on the results of one-sided discrete choice

	Primary problem	Dual problem
Men's welfare	$G_x(\boldsymbol{U}_x) = E_{\boldsymbol{P}_x} max_{y \in \mathcal{Y}_0}(U_{xy} + \varepsilon_{iy})$	$G_x^*(\boldsymbol{v}) = max_{\boldsymbol{U} \in \mathbb{R}^{\mathcal{Y}}} \left(\sum_{y \in \mathcal{Y}} v_y U_y - G_x(\mathbf{U}) \right)$
Aggregate welfare (Given group numbers $n = (n_x)$)	$G(\boldsymbol{U}, \boldsymbol{n}) = \sum_{x \in \mathcal{X}} n_x G_x(\boldsymbol{U}_x)$	$G^*(\boldsymbol{\mu}, \boldsymbol{n})$ $= sup_{\boldsymbol{U} \in \mathbb{R}^{\mathcal{X} \times \mathcal{Y}}} (\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mu_{xy} U_{xy} - G(\boldsymbol{U}, \boldsymbol{n}))$

→ $G^*(\boldsymbol{\mu}, \boldsymbol{n}) = \sum_{x \in \mathcal{X}} n_x G_x^*(\frac{\boldsymbol{\mu_x}}{\boldsymbol{n_x}})$ = − generalized entropy of choice of all men

- Define $H_y(\boldsymbol{V}_y)$ as the Emax function on women's side
- Given group numbers $\boldsymbol{m} = (m_y)$, the aggregate welfare of women is $H(\boldsymbol{V}, \boldsymbol{m})$
- Dual problems of these are the generalized entropy of choice

Social surplus, equilibrium, and entropy of matching

- Social Surplus $\mathcal{W}$:

$$\mathcal{W} = G(\boldsymbol{U}, \boldsymbol{n}) + H(\boldsymbol{V}, \boldsymbol{m}) = \sum_{x \in \mathcal{X}} n_x G_x(\boldsymbol{U}_x) + \sum_{y \in \mathcal{Y}} m_y H_y(\boldsymbol{V}_y)$$

- Stable matching $\mu = (\mu_{xy})_{x \in \mathcal{X}, y \in \mathcal{Y}}$ $(\boldsymbol{U} + \boldsymbol{V} = \boldsymbol{\Phi})$
  - $G(\boldsymbol{U}, \boldsymbol{n}) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mu_{xy} U_{xy} - G^*(\boldsymbol{\mu}, \boldsymbol{n})$ $\because$ (3.4)
  - $H(\boldsymbol{V}, \boldsymbol{m}) = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mu_{xy} V_{xy} - H^*(\boldsymbol{\mu}, \boldsymbol{m})$ $\because$ (3.4)
- $\mathcal{W} = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mu_{xy} \Phi_{xy} + \varepsilon(\boldsymbol{\mu}, \boldsymbol{n}, \boldsymbol{m})$
  - $\varepsilon(\boldsymbol{\mu}, \boldsymbol{n}, \boldsymbol{m}) := -G^*(\boldsymbol{\mu}, \boldsymbol{n}) - H^*(\boldsymbol{\mu}, \boldsymbol{m})$
  - Generalized entropy of matching

Social surplus at equilibrium

Assumption 2
For all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, the distribution P_x and Q_y have full support and are absolutely continuous

Theorem 3 (Social surplus at equilibrium)

- Under assumptions 1 and 2, for any $\boldsymbol{\Phi}$ and $\boldsymbol{r} = (\boldsymbol{n}, \boldsymbol{m})$ the stable matching $\boldsymbol{\mu}$ maximizes the social surplus over all feasible matchings $\boldsymbol{\mu} \in \mathcal{M}(\boldsymbol{r})$

$$\mathcal{W}(\boldsymbol{\Phi}, \boldsymbol{r}) = max_{\boldsymbol{\mu} \in \mathbb{R}^{\mathcal{X} \times \mathcal{Y}}} (\textstyle\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mu_{xy} \Phi_{xy} + \varepsilon(\boldsymbol{\mu}, \boldsymbol{r})) \quad (4.5)$$

- Dual expression

$$\mathcal{W}(\boldsymbol{\Phi}, \boldsymbol{r}) = min_{\boldsymbol{U}, \boldsymbol{V} \in \mathbb{R}^{\mathcal{X} \times \mathcal{Y}}} \big(G(\boldsymbol{U}, \boldsymbol{n}) + H(\boldsymbol{V}, \boldsymbol{m})\big) \quad (4.6)$$
$$s.t. \quad U_{xy} + V_{xy} \geq \Phi_{xy} \; \forall x \in \mathcal{X}, y \in \mathcal{Y}$$

- Optimal solutions relationship

$$\mu_{xy} = \frac{\partial G}{\partial U_{xy}}(\boldsymbol{U}, \boldsymbol{n}) = \frac{\partial H}{\partial V_{xy}}(\boldsymbol{V}, \boldsymbol{m}) \quad (4.7)$$

Remarks of Theorem 3

1. **The components of social surplus and their meanings (4.5)**
   - The first term reflects “systematic preferences”
     - If it dominates, it is the linear programming problem of Shapley and Shubik (1972)
   - The second term reflects “idiosyncratic preferences”
     - If it dominates ($\Phi \cong 0$), it looks like random matching
2. **Dual problem (4.6)**
   - The dual problem (4.6): The destination of the surplus shared at equilibrium between men and women
     - $n_x G_x(\boldsymbol{U}_x)$: the total amount of utility going to men of group x
     - $m_y H_y(\boldsymbol{V}_y)$: the total amount of utility going to women of group y
   - The primary problem (4.7): The origin of surplus
     - Φ_{xy}: The part of the surplus that comes from the interaction between observable characteristics in pair xy
     - $\varepsilon(\boldsymbol{\mu}, \boldsymbol{r})$: unobservable heterogeneities in tastes
3. **The first-order conditions and the equality between the demand (4.7)**
   - (4.7) is the first-order conditions of (4.6)
   - The right-hand side is the demand of women of group y for men of group x and vice versa
   - In equilibrium, these numbers must both equal μ_{xy}
4. A wealth of comparative statics results and testable predictions

Individual and group surplus

Proposition 2 (Individual and group surplus)

Let $(\boldsymbol{U}, \boldsymbol{V})$ solve (4.6), and $U_{x0} = V_{0y} = 0$. Under Assumptions 1 and 2,

- A man i of group x who marries a woman of group y^* obtains utility

$$U_{xy^*} + \varepsilon_{iy^*} = max_{y \in \mathcal{Y}_0}\left(U_{xy} + \varepsilon_{iy}\right)$$

- The average utility of men of group x is

$$u_x = G_x(\boldsymbol{U_x}) = \frac{\partial \mathcal{W}}{\partial n_x}(\boldsymbol{\Phi}, \boldsymbol{r})$$

- These can also be applied to women's side

Identification

- Focus on the case when **the distributions of the error terms are known**

Theorem 4.

Under Assumptions 1 and 2:

1. $\boldsymbol{U}$ and $\boldsymbol{V}$ are identified from $\boldsymbol{\mu}$ by
$$\boldsymbol{U} = \frac{\partial G^*}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu}) \text{ and } \boldsymbol{V} = \frac{\partial H^*}{\partial \boldsymbol{\mu}}(\boldsymbol{\mu})$$
2. $U_{xy} + V_{xy} = \Phi_{xy}$ for every $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The matching surplus $\boldsymbol{\Phi}$ is identified by
$$\Phi_{xy} = -\frac{\partial \varepsilon}{\partial \mu_{xy}}(\boldsymbol{\mu}, \boldsymbol{r}), \qquad (4.9)$$
$$\Phi_{xy} = \frac{\partial G_x^*}{\partial \boldsymbol{\mu}_{y|x}}(\boldsymbol{\mu}_{.|x}) + \frac{\partial H_y^*}{\partial \boldsymbol{\mu}_{x|y}}(\boldsymbol{\mu}_{.|y}),$$

where $\mu_{xy} = \mu_{y|x} n_x = \mu_{x|y} m_y$

- Combining Theorem 2 and 4 shows that all of the quantities in Theorem 3 can be computed by solving **simple convex optimization problems**

Example 4.1 (The Choo and Siow specification)

- Assume that $\boldsymbol{P}_x$ and $\boldsymbol{Q}_y$ are the distributions of centred i.i.d standard type I extreme value random variables
- Generalized entropy:
  $\varepsilon = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}_0}\mu_{xy}\log\mu_{y|x} - \sum_{y\in\mathcal{Y},x\in\mathcal{X}_0}\mu_{xy}\log\mu_{x|y}$
- Averaged utilities with matching patterns:
  $$u_x = -\log\mu_{0|x},\ v_y = -\log\mu_{0|y}$$
- Surplus with matching patterns:
  $\Phi_{xy} = 2\log\mu_{xy} - \log\mu_{x0} - \log\mu_{0y}$
  $$\mu_{xy} = \sqrt{\mu_{x0}\mu_{0y}}\exp(\frac{\Phi_{xy}}{2})$$
- Define:
  $F(\boldsymbol{u},\boldsymbol{v};\boldsymbol{\Phi},\boldsymbol{r}) \coloneqq \sum_{x\in\mathcal{X}} n_x(u_x + e^{-u_x} - 1) + \sum_{y\in\mathcal{Y}} m_y\left(v_y + e^{-v_y} - 1\right) + 2\sum_{x\in\mathcal{X},y\in\mathcal{Y}}\sqrt{n_x m_y}e^{\frac{\Phi_{xy}-u_x-v_y}{2}}$
  - Sum of exponentials and of linear functions
  - Globally strictly convex function of $(\boldsymbol{u},\boldsymbol{v})$
  - Social welfare $\mathcal{W}(\boldsymbol{\Phi};\boldsymbol{r})$ equals F's minimum value and at the minimum,
    $$\mu_{x0} = n_x\exp(-u_x)$$
    $$\mu_{0y} = m_y\exp\left(-v_y\right)$$
    $$\mu_{xy} = \sqrt{n_x m_y}\exp(\frac{\Phi_{xy} - u_x - v_y}{2})$$

5. Computation

Min-Emax method (based on gradient descent)

- Two expressions for the social surplus (Theorem 3)
  - (4.5) solves for the matching patterns $\boldsymbol{\mu}$:
    The globally concave unconstrained maximization problem (4.5)
  - (4.6) solves for the $\boldsymbol{U}$ and $\boldsymbol{V}$ utility components:
    The globally convex unconstrained minimization problem (dual)
    $$min_{\boldsymbol{U}\in\mathbb{R}^{xy}}(G(\boldsymbol{U},\boldsymbol{n})+H(\boldsymbol{\Phi}-\mathbf{U},\mathbf{m})) \qquad (5.1)$$
- Min-Emax method based on (5.1)
  - (5.1) has dimension $|\mathcal{X}|\times|\mathcal{Y}|$, unconstrained, very sparse structure
  - The Hessian of the objective function contains many zeroes → easy
  - Closed form → only require evaluating the G_x and H_y
  - Not closed form → use simulation and linear programming
- (5.1) is globally convex → a descent algorithm converges nicely under weak conditions
- In the Choo and Siow specification, the sparse structure of the problem can be used very easily to reduce the dimensionality
  - Only $|\mathcal{X}|+|\mathcal{Y}|$ arguments

Iterative projection fitting procedure (IPFP)

- In some cases, the number of groups $|\mathcal{X}|$ and $|\mathcal{Y}|$ is too large and min-Emax method is not practical option
  → Extended the IPFP, if the generalized entropy ε is easy to evaluate

IPFP

- The average utilities (u_x) and (v_y) of the groups of men and women play the role of prices that equate demand and supply
- Adjust the prices alternatively on each side of the market

1. Fix the prices (v_y) and find the prices (u_x)
$$\sum_{y\in\mathcal{Y}} \mu_{xy} + \mu_{x0} = n_x \qquad \text{for each } x \in \mathcal{X}$$
2. Fix the prices (u_x) and find the prices (v_y)
$$\sum_{x\in\mathcal{X}} \mu_{xy} + \mu_{0y} = m_y \qquad \text{for each } y \in \mathcal{Y}$$
3. Iterate these procedures (coordinate descent procedure)

Theorem 5
Under Assumption 1 and 2, the IPFP algorithm converges to the solution of (4.5) and to the corresponding average utilities $\boldsymbol{u}$ and $\boldsymbol{v}$

The performance of the proposed algorithms

- Test on the Choo and Siow model
- The IPFP algorithm
  - Extremely fast compared to standard optimization or equation-solving methods
- The min-Emax method of (5.1)
  - Slower but it still works very well for medium-size problems
  - Applicable to all separable models

6. Parametric Inference

Parametric model

Single matching market assumption

- Focus on observations from a single matching market

Need for parametric model

- Joint surplus functions Φ_{xy}^{λ} and distributions $\boldsymbol{P}_x^{\lambda}$ and $\boldsymbol{Q}_y^{\lambda}$ with parameters $\boldsymbol{\lambda}$

Sampling Assumption

- At household level
- Consist of H households, including couples and singles
- Number of individuals $\hat{S} = \sum_x \widehat{N}_x + \sum_y \widehat{M}_y$, where $\widehat{N}_x$and $\widehat{M}_y$ are the number of men and women in the sample
- Empirical frequencies $\hat{n}_x = \widehat{N}_x/\hat{S}$ and $\widehat{m}_y = \widehat{M}_y/\hat{S}$

Estimation method

Matching patterns and Margins

- Observed matches $\hat{\mu}_{xy}$ satisfy:

$$\sum_{y\in\mathcal{Y}} \mu^{\boldsymbol{\lambda}}_{xy} + \mu^{\boldsymbol{\lambda}}_{x0} = \hat{n}_x \qquad \forall x \in \mathcal{X}$$
$$\sum_{x\in\mathcal{X}} \mu^{\boldsymbol{\lambda}}_{xy} + \mu^{\boldsymbol{\lambda}}_{0y} = \hat{m}_y \qquad \forall y \in \mathcal{Y} \qquad (6.1)$$

- Data assumed from a population with true parameter $\boldsymbol{\lambda}_0$

Social Surplus and stable matching

- Social surplus:

$$\mathcal{W}\left(\boldsymbol{\Phi}^{\lambda}, \hat{\boldsymbol{r}}\right) = max_{\boldsymbol{\mu}\in\mathcal{M}(\hat{r})}\left(\sum_{x,y} \mu_{xy}\Phi^{\lambda}_{xy} + \varepsilon^{\lambda}(\boldsymbol{\mu}, \hat{\boldsymbol{r}})\right)$$

- Stable matching $\mu^{\lambda}(\hat{\boldsymbol{r}})$ computed efficiently

Estimation methods for $\boldsymbol{\lambda}$

1. Maximum likelihood estimation
2. Moment matching method
3. Minimum distance estimator

Maximum likelihood estimation (MLE)

1. **Compute the optimal matching with parameters λ for given populations of men and women**
   - Fix $\hat{n}_x$ and $\hat{m}_y$, impose constraints (6.1)
2. **Simulated number of households:**
   $H^\lambda \equiv \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \mu_{xy}^\lambda + \sum_{x\in\mathcal{X}} \mu_{x0}^\lambda + \sum_{y\in\mathcal{Y}} \mu_{0y}^\lambda = \sum_{x\in\mathcal{X}} \hat{n}_x + \sum_{y\in\mathcal{Y}} \hat{m}_y - \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \mu_{xy}^\lambda$
3. **Observed matches:**
   - μ_{x0}^λ: Number of single men with characteristic x
   - μ_{0y}^λ : Number of single women with characteristic y
   - μ_{xy}^λ: Number of (x, y) couples
4. **Log-likelihood function:**

$$\log L(\lambda) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} \hat{\mu}_{xy} \log \frac{\mu_{xy}^\lambda}{H^\lambda} + \sum_{x\in\mathcal{X}} \hat{\mu}_{x0} \log \frac{\mu_{x0}^\lambda}{H^\lambda} + \sum_{y\in\mathcal{Y}} \hat{\mu}_{0y} \log \frac{\mu_{0y}^\lambda}{H^\lambda}$$

- **Maximum likelihood estimator $\hat{\lambda}^{MLE}$:**
  - Consistent, asymptotically normal, and asymptotically efficient under the usual set of assumptions

Moment-based estimation in semi-linear models

Alternative to MLE:

- MLE is powerful but often difficult to maximize due to several local extrema

Conditions for moment-based method:

1. Distribution of the unobserved heterogeneities must be parameter-free (e.g. Choo and Siow, 2006)
2. Parametrization of the $\boldsymbol{\Phi}$ matrix must be linear in the parameter vector
   - $\Phi_{xy}^{\boldsymbol{\lambda}} = \sum_{k=1}^{K} \lambda_k \phi_{xy}^{k}$
   - $\boldsymbol{\lambda} \in \mathbb{R}^K$ and $\widetilde{\boldsymbol{\phi}} \coloneqq (\boldsymbol{\phi}^{\mathbf{1}}, \dots, \boldsymbol{\phi}^{\boldsymbol{K}})$ are K known linearly independent basis surplus vectors

Moment-matching estimator:

- Matches predicted moments with empirical moments:

$$\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \hat{\mu}_{xy} \phi_{xy}^{k} = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mu_{xy}^{\lambda} \phi_{xy}^{k} \quad \forall k$$

- The moment-matching estimator:

$$\hat{\boldsymbol{\lambda}}^{MM} \coloneqq \arg max_{\boldsymbol{\lambda} \in \mathbb{R}^K} \left(\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \hat{\mu}_{xy} \Phi_{xy}^{\boldsymbol{\lambda}} - \mathcal{W}(\Phi^{\lambda}, \hat{r}) \right)$$

Minimum distance estimation

Mixed hypothesis:

$$\exists \boldsymbol{\lambda}, \boldsymbol{D}^{\lambda} \equiv \boldsymbol{\Phi}^{\lambda} + \frac{\partial \varepsilon^{\boldsymbol{\lambda}}}{\partial \boldsymbol{\mu}}$$

Estimation process:

- Choose $\hat{\boldsymbol{\lambda}}$ to minimize $\left\|\boldsymbol{D}^{\lambda}\right\|_{\boldsymbol{\Omega}}^{2}$ for some positive definite matrix $\boldsymbol{\Omega}$
- Particularly appealing when distributions P_x and Q_y are parameter-free and surplus matrix $\boldsymbol{\Phi}^{\lambda}$ is linear in the parameters

7. Empirical Application

Testing methods on Choo and Siow's dataset

Objective:

- Testing Choo and Siow's specification against alternative models

Selected sub-sample:

- Time period: 1970s wave (younger marriage age)
- Age range: 16-40 years
- Sub-sample: "non-reform states"
  - 75,265 observations representing 13.3m individuals

Analysis approach:

- Non-parametric surplus models fit all separable models so it's hard to choose between models
- Two steps process
  1. Fit parametric surplus models:
     - Use semi-linear model and select basis functions (ϕ_{xy}^{k}) using Bayesian Information Criterion (BIC)
  2. Fit alternative specifications
     - Utilize chosen basis functions and test different error term distributions

Heteroskedastic logit models

Method:

- Add heteroskedasticity to benchmark model while maintaining scale normalization
- To determine the best fit, use BIC

Findings:

- Gender heteroskedasticity: minimal improvement in fit, worsens BIC
- Gender and age heteroskedasticity: significant improvement in both fit and BIC
  - Preferred model: replace $\varepsilon_{iy} + \eta_{xj}$ with $\sigma_x \varepsilon_{iy} + \tau_y \eta_{xj}$
    - $\sigma_x = \exp(\sigma_1 x + \sigma_2 x^2)$, $\tau_y = \exp(\tau_0)$
  - Results:
    - +29.4 points of log-likelihood and +25.1 points on BIC
    - Estimated parameters: $\tau_y = 0.16, \sigma_x$ from 0.40 at age 16 to 2.49 at age 40

Impact on surplus share

- Heteroskedasticity affects surplus share in matches:

$$\frac{u_x}{u_x + v_y} = \frac{\sigma_x \log \mu_{0|x}}{\sigma_x \log \mu_{0|x} + \tau_y \log \mu_{0|y}}$$

- Figure 1: Surplus share ratio for same-age couples in 3 models (homoscedastic, gender-heteroskedastic, and gender- and age-heteroskedastic)
- Men's surplus share increases with age in gender- and age-heteroskedastic model

REVIEW OF ECONOMIC STUDIES

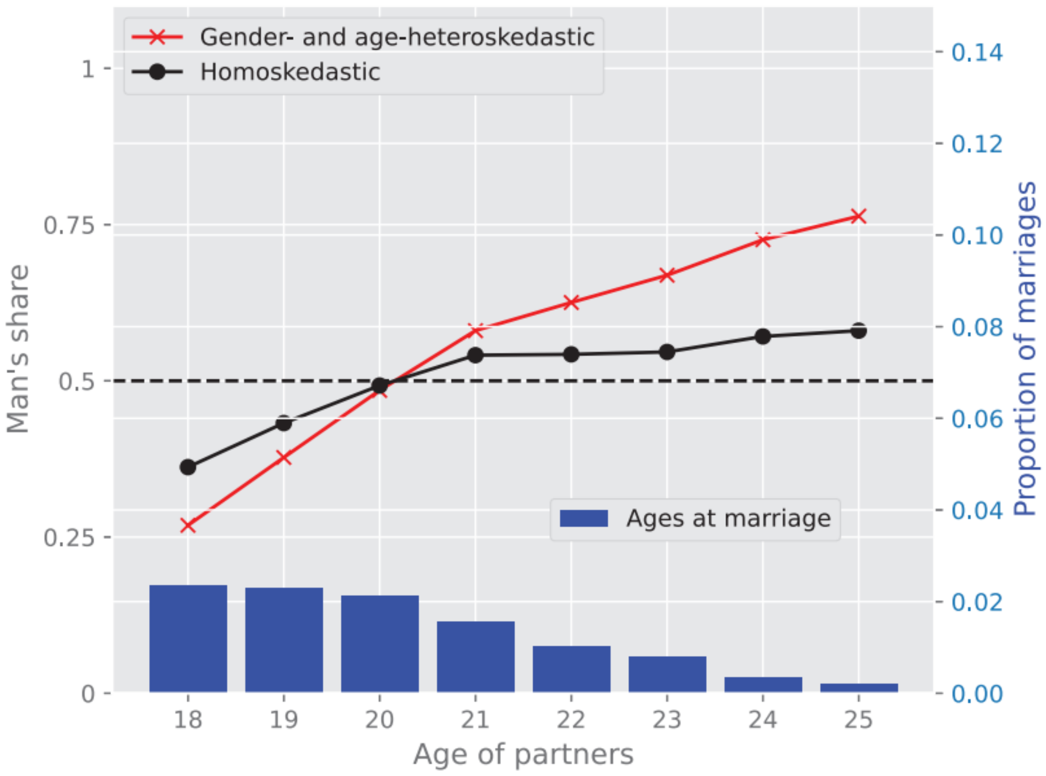

FIGURE 1

Men's share of the marriage surplus in the logit models

Flexible MNL models

Background:

- Nested logit model limitation:
  - Equal correlation between all alternatives in a nest
  - Not suit for capturing age-local correlations

Model choice:

- Flexible coefficient multinomial logit (FC-MNL) model (Davis and Schiraldi, 2014)

Method:

- Reformulate as an MPEC
- Maximize log-likelihood for parameters and $\boldsymbol{U}$ under constraint $\nabla G(\boldsymbol{U}) = \nabla H(\boldsymbol{\Phi} - \boldsymbol{U})$

Model specification:

- Substitution patterns matrix:

$$b_{y,y'}^{x} = \begin{cases} \dfrac{b_m(x)}{|y - y'|} & if\ y \neq y' \\ 1 & if\ y = y' \end{cases}$$

- Similar for women's side with $b_w(y)$ divided by $|x - x'|$

8. Concluding Remarks

Concluding remarks

Validation of assumptions:

- Separability and large market assumptions are tested and supported by the simulations (Chiappori et al., 2019b)

Potential extensions:

- Continuous characteristics
  - Dupuy and Galichon (2014) address this issue for the Choo and Siow model using the theory of extreme value processes and propose testing the number of relevant dimensions

Broader applications:

- Beyond bipartite matching
  - “Roommate” problem (Chiappori et al., 2019a)
  - Trade on networks with transfers (Hatfield and Kominers, 2012) & (Hatfield et al., 2013)

Relaxing utility assumptions:

- Imperfectly transferable utility and separable logit heterogeneity
- Non-transferable utility and a similar form of heterogeneity

所感

- 数学的なテクニカルな手法（双対性等）が何回も用いられていたので、その展開を追うのが難しかった
- 一つ一つの話題に関しては理解することができた気がするが、論文自体の長さ（おそらく書かれ方が丁寧？）もあり、全体の流れを掴むのが大変だった
- 初めてまともに読む論文ということもあり、読み進めるときに命題や定理などについて、どこまでそう言うものだと想定して（なぜそうなるのかわからなくても）読み進めていけるのかの判断が難しかった