Migration in China: to Work or to Wed?

Arnaud Dupuy, CEPS/INSTEAD
Joint work with Alfred Galichon and Liping Zhao

September 19, 2014

1 Motivation

Since the late 1990s, rural migration in China went up by more than 100 million from 45 to 145 million. (Meng, 2012).

China is about to witness mankind's largest population movement: more than 300 million rural individuals moving to cities. (Meng, 2012)

Why do people migrate?...

.......To improve their welfare.

.......But there are many different margins involved, corresponding to many different markets.

For instance in China:

- 1/2 migrants moved for work and,
- 1/6 for marriage (Liang and Ma, 2004).

Mobility induces interesting trade-offs and in particular between marriage and work.

- Some men/women may accept to marry down in exchange of better labor market prospects elsewhere.

A simplistic example to fix ideas:

Consider Mr x, born and raised in a rural area, that obtained a college degree and has a taste for “rural” women.

Given his skills, Mr x has better labor market perspectives in the city than in his local rural area.

However, assuming that “rural” women prefer living in a rural area, his marriage market prospects would be better if he stayed in a rural area.

Mr x faces a trade-off between moving for work and staying for love.

To study these trade-offs one needs to unify:

- the matching model a la Becker (1973) and Shapley and Shubik (1972).
  - men care about type $y \in \mathcal{Y}$ of women and women care about type $x \in \mathcal{X}$ of men, interaction "x" $\times$ "y",

- with the hedonic model a la Rosen (1974)
  - men $x \in \mathcal{X}$ care about where women want to live, i.e. $z \in \mathcal{Z}$, and women $y \in \mathcal{Y}$ care about where men want to live $z \in \mathcal{Z}$, interactions "x" $\times$ "z" and "z" $\times$ "y".

The unified problem studies interactions "x" $\times$ "z" $\times$ "y". It has received very little attention in the literature: Hwang et. al, 1992, Dupuy, 2009 and Quintana-Domeque, 2011.

Note that this goes beyond the equivalence results by Chiappori et al. 2010.

2 Related literature

There is a growing literature analyzing the effects of sex ratios differences

- on the marriage market: Angrist, 2002, Porter, 2007a and 2007b, Abramitzky et al., 2011, Weiss, Yi and Zhang, 2013,

- on the labor market Angrist, 2002, Chang and Zhang, 2012, Grossbard and Amuedo-Dorantes, 2008 and Wei and Zhang, 2011.

We contribute to this literature by considering how variations in local marriage market conditions (sex ratios but also the distribution of types of men and women) and labor market conditions affect migration and matching on the marriage market.

Our work is also related to McCann, Shi, Siow and Wolthoff, 2014 and Zhang 2014.

3 Outline

1. This paper unifies the "classical" matching model a la Becker (1973) and Shapley and Shubik (1972) with the hedonic model a la Rosen (1974).

2. We show that any stable outcome maximizes the utilitarian social welfare (Pareto optimal) and is equivalent to a Walrasian equilibrium.

3. Building on Choo and Siow (2006) and Galichon and Salanié (2013), we introduce unobserved heterogeneity into the model and provide identification results for the costs/benefits of migration.

4. We present an application to the marriage market in China.

4 Encompassing matching and hedonic models

Sets:

Let $z \in \mathcal{Z}$ be the set of (geographic) locations.

Let $x \in \mathcal{X}$ denote the type of men, let there be n_x men of type x.

Let $y \in \mathcal{Y}$ denote the type of women, let there be m_y women of type y.

Let $x^1 \in \mathcal{Z}$ and $y^1 \in \mathcal{Z}$ denote the location at origin, and x^- and y^- denote the non geographic attributes $x = \left(x^1, x^-\right)$, and similarly $y^- = \left(y^1, y^-\right)$.

Utilities:

We consider a transferable utility model.

Let α_{xyz} be the pre-transfer utility of a man x married with a woman y living at z.

Let γ_{xyz} be the pre-transfer utility of a woman y married with a man x living at z.

The joint utility is therefore: $\Phi_{xyz} := \alpha_{xyz} + \gamma_{xyz}$.

We allow individuals to remain single, i.e. Φ_{x0z} and Φ_{0yz} are the utilities of singles.

Post-transfer utilities: $\alpha_{xyz} + w_{xyz}$ and $\gamma_{xyz} - w_{xyz}$.

Relation to existing models:

1. Matching model (a unique z): $\alpha_{xyz} = \alpha_{xy}$ and $\gamma_{xyz} = \gamma_{xy}$ s.t. $\Phi_{xyz} = \Phi_{xy}$.

2. Hedonic model: $\alpha_{xyz} = \alpha_{xz}$ and $\gamma_{xyz} = \gamma_{yz}$.

3. Matching model on perfectly segmented markets: $\alpha_{xyz} = \gamma_{xyz} = -\infty$ for all x and y s.t. $x^1 \neq z$ and $y^1 \neq z$. Matching model on each segmented market.

5 Definitions

Definition 1. *A matching μ is feasible if and only if $\mu \in \mathcal{M}$ where $\mathcal{M}$ is defined as follows:*

$$\mathcal{M} = \left\{ \mu_{xyz} | \sum_{y \in \mathcal{Y}, z \in \mathcal{Z}} \mu_{xyz} \leq n_x; \sum_{x \in \mathcal{X}, z \in \mathcal{Z}} \mu_{xyz} \leq m_y \right\}$$

Single men of type x at z,i.e. μ_{x0z} and hence single men of type x: $\mu_{x0} = \sum_{z \in \mathcal{Z}} \mu_{x0z} = n_x - \sum_{y \in \mathcal{Y}, z \in \mathcal{Z}} \mu_{xyz}$.

Single women of type y at z, i.e. μ_{0yz} and hence single women of type y: $\mu_{0y} = \sum_{z \in \mathcal{Z}} \mu_{0yz} = m_y - \sum_{x \in \mathcal{X}, z \in \mathcal{Z}} \mu_{xyz}$.

Definition 2. *The outcome* (μ, u, v) *is stable if and only if*

$$u_x \geq \Phi_{x0z},$$

$$v_y \geq \Phi_{0yz},$$

and

$$u_x + v_y \geq \Phi_{xyz},$$

with equality if respectively $\mu_{x0z} > 0$, $\mu_{0yz} > 0$ *and* $\mu_{xyz} > 0$.

Proposition 5.1. *The outcome* (μ, u, v) *is stable if and only if* $\mu \in \mathcal{M}$ *and* (u, v) *is solution to*

$$\min_{u,v} \sum_x u_x n_x + \sum_y v_y m_y,$$

$$s.t. \left\{ \begin{array}{c} u_x + v_y \geq \Phi_{xyz} \ \left[\mu_{xyz}\right] \\ u_x \geq \Phi_{x0z} \ [\mu_{x0z}] \\ u_y \geq \Phi_{0yz} \ \left[\mu_{0yz}\right] \end{array} \right. .$$

Definition 3. *A Walrasian equilibrium is a pair* (μ, w) *such that* $\mu \in \mathcal{M}$ *and men and women maximize their utilities:*

$$
\begin{aligned}
u_x &= \max_{y,z} \left(\alpha_{xyz} + w_{xyz}, \Phi_{x0z}\right) \\
v_y &= \max_{x,z} \left(\gamma_{xyz} - w_{xyz}, \Phi_{0yz}\right)
\end{aligned}
$$

and such that:

$\mu_{xyz} > 0$ *implies that* (y, z) *and* (x, z) *are respectively solutions to the first and second problem,*

$\mu_{x0z} > 0$ *implies that* $(0, z)$ *is solution to the first and,*

$\mu_{0yz} > 0$ *implies that* $(0, z)$ *is solution to the second.*

6 Existence and properties of equilibrium

Theorem 1. *Any stable outcome maximizes the utilitarian social welfare.*

Proof. The utilitarian social welfare is

$$\begin{aligned}\max_{\mu\geq 0} \quad & \sum_{xyz} \mu_{xyz}\Phi_{xyz} + \sum_{xz} \mu_{x0z}\Phi_{x0z} + \sum_{yz} \mu_{0yz}\Phi_{0yz} \\ s.t. \quad & \sum_{y\in\mathcal{Y},z\in\mathcal{Z}} \mu_{xyz} \leq n_x \ \ [u_x] \\ & \sum_{x\in\mathcal{X},z\in\mathcal{Z}} \mu_{xyz} \leq m_y \ \ [v_y]\end{aligned}$$

whose dual is

$$\begin{aligned}&\min_{u_x\geq\Phi_{x0z},v_y\geq\Phi_{0yz}} \sum_x u_x n_x + \sum_y v_y m_y \\ &s.t.\ u_x + v_y \geq \Phi_{xyz} \ \ \left[\mu_{xyz}\right]\end{aligned}$$

We note that by complementary slackness one has $\Phi_{xyz} - (u_x + v_y) = 0$ whenever $\mu_{xyz} > 0$. □

Theorem 2. *Any stable outcome is equivalent to a Walrasian equilibrium.*

Proof. A stable outcome must satisfy the incentive constraint by definition, hence u and v must be such that

$$u_x - \alpha_{xyz} \geq \gamma_{xyz} - v_y.$$

Using the definition of a Walrasian equilibrium, one has

$$\begin{aligned} u_x &\geq \alpha_{xyz} + w_{xyz} \\ v_y &\geq \gamma_{xyz} - w_{xyz}. \end{aligned}$$

Rewriting these inequalities, one obtains

$$u_x - \alpha_{xyz} \geq w_{xyz} \geq \gamma_{xyz} - v_y.$$

Therefore, there always exist transfers w supporting a stable equilibrium (u, v). In particular, from the duality one has

$$\begin{aligned} \mu_{xyz}\boldsymbol{\Phi}_{xyz} &= \mu_{xyz}(\alpha_{xyz} + \gamma_{xyz}) \\ &= \mu_{xyz}(u_x + v_y). \end{aligned}$$

This implies that $\alpha_{xyz}+\gamma_{xyz} = u_x+v_y$ for all $\mu_{xyz} > 0$. We conclude that for all $\mu_{xyz} > 0$ one has

$$u_x - \alpha_{xyz} = w_{xyz} = \gamma_{xyz} - v_y.$$

□

Two extreme cases:

1. Matching: $w_{xyz} = w_{xy}$, since one has

$$u_x - \alpha_{xy} \geq w_{xyz} \geq \gamma_{xy} - v_y.$$

2. Hedonic: $w_{xyz} = w_z$, since one has

$$\begin{aligned} u_x - \alpha_{xz} &\geq \min_x \{u_x - \alpha_{xz}\} \geq w_z \\ &\geq \max_y \left\{\gamma_{yz} - v_y\right\} \geq \gamma_{yz} - v_y. \end{aligned}$$

7 Adding heterogeneity

Drawing insights from Choo and Siow (2006) and Galichon and Salanié (2013), we introduce unobserved heterogeneity into the model.

Notation:

Let i index men and j index women.

x_i observed type of man i

ε_{iyz} unobserved taste of man i for *any* women of observed type y when living at z.

(for women, y_j and η_{jxz}).

Let $\left(\varepsilon_{i0z_1}, ..., \varepsilon_{i|\mathcal{Y}||\mathcal{Z}|}\right) \sim P_x$ and $\left(\eta_{j0z_1}, ..., \eta_{j|\mathcal{X}||\mathcal{Z}|}\right) \sim Q_y$.

Assumption on the joint utility:

Let the joint utility of a man i of type $x_i = x$ and a woman j of type $y_j = y$ be

$$\widetilde{\Phi}_{ijz} = \Phi_{xyz} + \varepsilon_{iyz} + \eta_{jxz},$$

and for singles

$$\begin{aligned} \widetilde{\Phi}_{i0z} &= \Phi_{x0z} + \varepsilon_{i0z}, \\ \widetilde{\Phi}_{0jz} &= \Phi_{0yz} + \eta_{0jz}. \end{aligned}$$

Proposition 7.1. *The outcome* $(\mu, \widetilde{u}, \widetilde{v})$ *is stable if and only if* $\mu \in \mathcal{M}$ *and* $(\widetilde{u}, \widetilde{v})$ *is solution to*

$$\min_{(\widetilde{u},\widetilde{v}) \in \widetilde{\mathcal{S}}} \sum_i \widetilde{u}_i + \sum_j \widetilde{v}_j$$

where

$$\widetilde{\mathcal{S}} = \left\{ (u, v) | u_i \geq \widetilde{\Phi}_{i0z} \forall i; v_j \geq \widetilde{\Phi}_{0jz} \forall j; u_i + v_j \geq \widetilde{\Phi}_{ijz} \forall i, j, z \right\}$$

This problem can be recast as a two-sided discrete choice problem.

Indeed, incentive constraint can be reformulated as

$$\widetilde{u}_i - \varepsilon_{iyz} + \widetilde{v}_j - \eta_{jxz} \geq \Phi_{xyz} \forall i, j, z.$$

In particular, this constraint must be true for

$$\min_{i|x_i=x} \left\{ \widetilde{u}_i - \varepsilon_{iyz} \right\} + \min_{j|y_j=y} \left\{ \widetilde{v}_j - \eta_{jxz} \right\} \geq \Phi_{xyz} \forall x, y, z.$$

Using the notation $U_{xyz} := \min_{i|x_i=x} \left\{ \widetilde{u}_i - \varepsilon_{iyz} \right\}$ and $V_{xyz} = \min_{j|y_j=y} \left\{ \widetilde{v}_j - \eta_{jxz} \right\}$ this rewrites as

$$U_{xyz} + V_{xyz} \geq \Phi_{xyz} \forall x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}.$$

By definition one has

$$\widetilde{u}_i \geq U_{xyz} + \varepsilon_{iyz} \forall x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}.$$

$$\widetilde{v}_j \geq V_{xyz} + \eta_{jxz} \forall x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}.$$

Men therefore solve

$$\widetilde{u}_i = \max_{y \in \mathcal{Y}_0, z \in \mathcal{Z}} \left\{ U_{xyz} + \varepsilon_{iyz} \right\},$$

using the notation $U_{x0z} = \Phi_{x0z}$ and $\mathcal{Y}_0 = \mathcal{Y} \cup \{\emptyset\}$ and,

women solve

$$\widetilde{v}_j = \max_{x \in \mathcal{X}_0, z \in \mathcal{Z}} \left\{ V_{xyz} + \eta_{jxz} \right\},$$

where $V_{0yz} = \Phi_{0yz}$ and $\mathcal{X}_0 = \mathcal{X} \cup \{\emptyset\}$.

The expected utility of a man of type x reads as

$$G_x(U_{x.}) = \mathbb{E}_{P_x}\left[\max_{y\in\mathcal{Y}_0, z\in\mathcal{Z}}\left\{U_{xyz} + \varepsilon_{iyz}\right\}\right]$$

where $U_{x.} = (U_{xyz})_{y\in\mathcal{Y}_0, z\in\mathcal{Z}}$. (With a similar expression for women).

It follows that

$$\begin{aligned} & \frac{\partial G_x(U_{x.})}{\partial U_{xyz}} \\ = \; & \Pr\left[U_{xyz} + \varepsilon_{iyz} \geq U_{xy'z'} + \varepsilon_{iy'z'} \forall y' \in \mathcal{Y}_0, z' \in \mathcal{Z}\right] \\ = \; & \frac{\mu_{xyz}}{n_x} = \mu_{yz|x}. \end{aligned}$$

Assuming P_x and Q_y are i.i.d. Gumbel-type I distributions with parameter $\sigma = 1$, obtains

$$\frac{\mu_{xyz}}{n_x} = \frac{\exp U_{xyz}}{\sum_{y' \in \mathcal{Y}_0, z' \in \mathcal{Z}} \exp U_{xy'z'}}.$$

Using the IIA property of the logit structure, one can easily derive the following expressions

$$\frac{\mu_{xyz}}{n_x - \mu_{x0}} = \frac{\exp U_{xyz}}{\sum_{y' \in \mathcal{Y}, z' \in \mathcal{Z}} \exp U_{xy'z'}}, \tag{7.1}$$

$$\frac{\mu_{x0z}}{\mu_{x0}} = \frac{\exp U_{x0z}}{\sum_{z' \in \mathcal{Z}} \exp U_{x0z'}}. \tag{7.2}$$

Taking the logarithm of equations 7.1 and 7.2 yields

$$\ln \mu_{xyz} = U_{xyz} + a_x, \tag{7.3}$$

$$\ln \mu_{x0z} = \Phi_{x0z} + a_x. \tag{7.4}$$

One obtains similar expressions for women

$$\ln \mu_{xyz} = V_{xyz} + b_y, \tag{7.5}$$

$$\ln \mu_{0yz} = \Phi_{0yz} + b_y. \tag{7.6}$$

8 Nonparametric Identification

Combining equations 7.3 and 7.5, obtains

$$\mu_{xyz} = \exp \frac{\Phi_{xyz} + a_x + b_y}{2}.$$

Similarly, from equations 7.4 and 7.6 one obtains

$$\begin{aligned} \mu_{x0z} &= \exp\left(\Phi_{x0z} + a_x\right), \\ \mu_{0yz} &= \exp\left(\Phi_{0yz} + b_y\right). \end{aligned}$$

Note that a_x and b_y are determined such that $\mu \in \mathcal{M}$.

1. μ_{x0z} and μ_{0yz} identify Φ_{x0z} and Φ_{0yz}, up to additive vectors a_x and b_y.

2. μ_{xyz} identifies Φ_{xyz} up to additive vectors a_x and b_y.

Several objects of interest can be identified:

1. The moving surplus of singles (similar object for women): one has

$$\Phi_{(x^1,x^-)0z}-\Phi_{(x^1,x^-)0x^1}=\ln\mu_{(x^1,x^-)0z}-\ln\mu_{(x^1,x^-)0x^1}$$

2. The marriage surplus of non migrant couples:

$$\begin{aligned}&\Phi_{(z,x^-)(z,y^-)z}-\Phi_{(z,x^-)0z}-\Phi_{0(z,y^-)0z}\\=\ &2\ln\mu_{(z,x^-)(z,y^-)z}-\ln\mu_{(z,x^-)0z}-\ln\mu_{0(z,y^-)z},\end{aligned}$$

3. The moving surplus of couples:

$$\begin{aligned}&\Phi_{(x^1,x^-)(x^1,y^-)z}-\Phi_{(x^1,x^-)(x^1,y^-)x^1}\\=\ &2\left(\ln\mu_{(x^1,x^-)(x^1,y^-)z}-\ln\mu_{(x^1,x^-)(x^1,y^-)x^1}\right).\end{aligned}$$

9 Additive separable specification

$$\Phi_{xyz} = \varphi_{xy} + c_{xz} + d_{yz}.$$

Complementarities in joint utility between the types of men and women are market invariant.

Three extreme cases:

1. Matching model on an integrated market: this case is met when $c_{xz} = d_{yz} = 0$ for all x, y, z.

2. Matching model on perfectly segmented markets: this case is met when $c_{xz} = d_{yz} = -\infty$ if $x^1 \neq y^1 \neq z$.

3. Hedonic model: this case is met when $\varphi_{xy} = 0$ for all x and y.

10 Estimation

Remember that

$$\frac{\partial G_x\left(U_{x.}\right)}{\partial U_{xyz}} = \mu_{yz|x}.$$

Inference requires inverting the problem. Write the Legendre-Fenchel transform of G_x

$$G_x^*(\mu_{.|x}) = \max_{\widetilde{U}_{x.}} \left(\sum_{yz} \mu_{yz|x} \widetilde{U}_{xyz} - G_x(\widetilde{U}_{x.}) \right).$$

Convex duality yields

$$G_x(U_{x.}) = \sum_{yz} \mu_{yz|x} U_{xyz} - G_x^*(\mu_{.|x}).$$

A similar expression yields for women

$$H_y(V_{.y}) = \sum_{xz} \mu_{xz|y} V_{xyz} - H_y^*(\mu_{.|y}).$$

Note that total utility of men is

$$\sum_i \widetilde{u}_i = \sum_x n_x G_x(U_{x.}),$$

and similar for women

$$\sum_j \widetilde{v}_j = \sum_y m_y H_y(V_{.y}).$$

We therefore have by duality

$$\min_{(\widetilde{u},\widetilde{v}) \in \widetilde{\mathcal{S}}} \sum_i \widetilde{u}_i + \sum_j \widetilde{v}_j =$$

$$\max_{\mu \in \mathcal{M}} \sum_{xyz} \mu_{xyz} \Phi_{xyz} + \sum_{xz} \mu_{x0z} \Phi_{x0z} + \sum_{yz} \mu_{0yz} \Phi_{0yz}$$
$$- \sum_x n_x G_x^*(\mu_{.|x}) - \sum_y m_y H_y^*(\mu_{.|y}).$$

For Gumbel-type I distributions

$$G_x^*(\mu_{.|x}) = \sum_{yz} \mu_{yz|x} \ln \mu_{yz|x},$$

so that social welfare is

$$\begin{aligned}\mathcal{W}(\Phi) &= \max_{\mu \in \mathcal{M}} \sum_{xyz} \mu_{xyz} \Phi_{xyz} + \sum_{xz} \mu_{x0z} \Phi_{x0z} + \sum_{yz} \mu_{0yz} \Phi_{0yz} \\ &\quad - \sum_{xyz} \mu_{xyz} \ln \frac{\mu_{xyz}}{n_x} - \sum_{xyz} \mu_{xy} \ln \frac{\mu_{xyz}}{m_y}.\end{aligned}$$

Parametric specification:

Let $\Phi_{xyz} = \sum_{kl} A_{kl} \widetilde{x}^{(k)} \widetilde{y}^{(l)}$ where $\widetilde{x} = (x, z)$ and $\widetilde{y} = (y, z)$.

Denote the associated optimal matching by $\mu^{A}_{\widetilde{x}\widetilde{y}}$.

By the envelope theorem

$$\begin{aligned} \frac{\partial \mathcal{W}(A)}{\partial A_{kl}} &= \sum_{\widetilde{x}\widetilde{y}} \mu^{A}_{\widetilde{x}\widetilde{y}} \left[\widetilde{x}^{(k)} \widetilde{y}^{(l)} \right] \\ &= \mathbb{E}_{\mu^A} \left[\widetilde{x}^{(k)} \widetilde{y}^{(l)} \right]. \end{aligned}$$

This suggests using a matching moment estimator:

- find A such that $\mathbb{E}_{\mu^A}\left[\widetilde{x}^{(k)}\widetilde{y}^{(l)}\right] = \mathbb{E}\left[\widetilde{x}^{(k)}\widetilde{y}^{(l)}\right]$.

It requires to compute μ_{xyz}^A which is done using IPFP and

$$\mu_{xyz}^A = \exp \frac{\widetilde{x}' A \widetilde{y} + a_x + b_y}{2}.$$

11 Empirical analysis

11.1 Why China?

Migration used to be free in China up until the Great Famine in the late 1950s.

To prevent migration (rural-to-urban), the Chinese government implemented a permanent geographic registration of Chinese citizens, i.e. the "Hukou system".

Up until the end of the 1980s, migration was extremely difficult: changing one's Hukou was merely impossible and strict restrictions were imposed that inhibited rural households from migrating (urban food coupons for urban hukou only).

The booming economy and the entry of China into WTO (2001), led to a sharp increase in low-skilled jobs in urban areas.

China's policy on rural migration changed gradually allowing more people to migrate from rural to urban areas.

Rural-to-urban migration rose rapidly, doubling between 1989 and 1993. (Shi 2008)

Since the late 1990s the number of rural migrants increased by more than 100 million from 45 to 145 million. (Meng, 2012).

In the next few decades more than 300 million rural individuals may move to cities. (Meng, 2012).

However, even today, the employment of rural migrants is concentrated in jobs in the informal sector that local urban people do not want, the so-called four “Ds”: Dirt, Drain, Danger, and Disgrace.

Moreover, individuals with a rural hukou have only limited access to urban public goods such as health care(insurance) and education.

11.2 Data

We use data from the Rural Urban Migration in China (RUMiC) longitudinal dataset.

This dataset consists of three independent surveys in China,

1. the Urban Household Survey (UHS),
2. the Rural Household Survey (RHS), and
3. the Migrant Household Survey (MHS), collected since 2008.

Urban (rural) residents are individuals who possess a urban (rural) hukou.

A migrant is an individual who has a rural hukou, but is living in a city at the time of the survey.

11.3 Couples and singles

We restrict our sample to women aged between 18 to 30 and men aged 18 to 35 years old.

Our definition of a couple is a man and a woman living in the same household and reporting being either head of the household or spouse of the head.

Unmatched individuals, are added to the sample of individuals reporting being neither the head nor the spouse of the head of the household $=>$ singles.

Table 1: Number of couples and singles in our working dataset by geographical groups (Urban, Rural and Migrants) and gender.

	Urban		Rural		Migrants	
	Men	Women	Men	Women	Men	Women
Singles	1,448	1,273	2,687	2,031	347	183
Couples	443	443	113	113	121	121

11.4 Selected Variables

Educational attainment: answers to highest level of education completed=> 3 standard educational levels: primary education, secondary education and tertiary education.

Height

Body Mass Index (BMI): body mass in kg divided by the square of its height in meters.

Subjective health: The respondents were also asked to report their general health. The phrasing of the question was: "What is your current health status (compared to people your age)".

Smoking: We use the answer to the question: "how many cigarettes on average do you smoke per day now?" to create a smoking variable that takes for value 0 if the answer is 0 and 1 else.

Our working dataset consists of those households with complete information on education, height, BMI, health, education, smoking behavior.

Table 2: Affinity matrix for young couples (N=677).

Men Women	**Height**	**Health**	**Education**	**BMI**	**Urban Hukou**	**Urban**
Height	**0.41**	-0.04	**0.31**	-0.10	**0.87**	**-0.90**
Health	0.08	**1.58**	0.10	-0.02	**1.09**	0.05
Smoking	-0.05	0.08	-0.10	-0.06	**0.91**	**-0.84**
Education	0.11	0.00	**0.84**	-0.05	**3.46**	**-3.69**
BMI	-0.01	**-0.11**	0.03	0.19	**-1.26**	**0.52**
Urban Hukou	-0.34	0.97	**1.59**	0.29	**10.38**	**-5.50**
Urban	**-0.41**	0.31	**-2.25**	-0.42	**-6.48**	**6.65**

Bold coefficients are significant at 5

12 Affinity matrix

13 Take away

1. This paper unifies the "classical" matching model a la Becker (1973) and Shapley and Shubik (1972) with the hedonic model a la Rosen (1974).

2. We show that any stable outcome maximizes the utilitarian social welfare (Pareto optimal) and is equivalent to a Walrasian equilibrium.

3. Building on Choo and Siow (2006) and Galichon and Salanié (2013), we introduce unobserved heterogeneity into the model and provide identification results for the costs/benefits of migration.

4. The unified model has many potential applications: Fair trade on product market, sorting on skills and preferences in labor market, etc.

5. We present an application to the marriage market in China.

14 Additional slides

Note on population sex-ratios:

Sex ratios (males/females):

R1015 (ages 10-15) 1.06 0.05 0.94 1.23

R1625 (ages 16-25) 1.02 0.03 0.95 1.15

R2645 (ages 26-45) 1.03 0.02 0.97 1.11

R4665 (ages 46-65) 1.04 0.05 0.82 1.14

SOURCES: Edlund et al. (2013).

However, several interesting objects are only identified up to additive terms:

1. The surplus of local singles over migrant singles (similar object for women):

$$\Phi_{(z,x^-)0z}-\Phi_{(x^1,x^-)0z}=\ln\mu_{(z,x^-)0z}-\ln\mu_{(x^1,x^-)0z}+a_{(x^1,\ldots}$$

2. The surplus of local couples over moving couples:

$$\begin{aligned}&\Phi_{(z,x^-)(z,y^-)z}-\Phi_{(x^1,x^-)(x^1,y^-)z}\\ =\;& 2\begin{pmatrix}\ln\mu_{(z,x^-)(z,y^-)z}-\ln\mu_{(x^1,x^-)(x^1,y^-)z}\\ +a_{(x^1,x^-)}-a_{(z,x^-)}+b_{(x^1,y^-)}-b_{(z,y^-)}\end{pmatrix},\end{aligned}$$