Lecture 5: Estimation of a One-to-One Transferable Utility Matching Model

Bryan S. Graham, UC - Berkeley & NBER

September 4, 2015

Consider a single matching market composed of two large populations of, for concreteness, women and men. For each woman and man we respectively observe the discretely-valued characteristics $W_i \in \mathbb{W} = \{w_{1,}, \ldots, w_K\}$ and $X^j \in \mathbb{X} = \{x_1, \ldots, x_L\}$.[1] The K types of women and L types of men may encode, for example, different unique combinations of years-of-schooling and age. While K and L are assumed finite, they may be very large in practice.

Observationally identical women have heterogeneous preferences over different types of men, but are indifferent between men of the same type. Specifically female i's utility from matching with male j is given by

$$U\left(W_i, X^j, \underline{\varepsilon}_i\right) = \alpha\left(W_i, X^j\right) + \tau\left(W_i, X^j\right) + \varepsilon_i\left(X^j\right),$$

where $\alpha(w_k, x_l)$ is the systematic utility a type $W_i = w_k$ women derives from matching with a type $X^j = x_l$ man, $\varepsilon_i(X^j) = \sum_{l=1}^{L} \mathbf{1}(X^j = x_l)\,\varepsilon_{il}$ captures unobserved heterogeneity in women's preferences over alternative *types* of men, and $\tau(w_k, x_l)$ is the *equilibrium* transfer that a type $X^j = x_l$ man must pay a type $W_i = w_k$ women in order to match. Transfers may be negative and their determination is discussed below. Here $\mathbf{1}(\bullet)$ denotes the indicator function. Since the stochastic component of female match utility, $\varepsilon_i(X^j)$, varies with male type alone (i.e., his specific identify does not matter), women are indifferent amongst observationally identical men. A similar restriction on male preferences ensures that the equilibrium transfer, $\tau(w_k, x_l)$, depends on agent types alone.

A women may also choose to remained unmatched, or 'single', in which case her utility is given by

$$\underline{U}(W_i, \underline{\varepsilon}_i) = \underline{\alpha}(W_i) + \varepsilon_{i0}.$$

[1] The subscript 'i' denotes a generic random draw from the population of women, while the superscript 'j' one from men. See Shapley and Shubik (1971) for the theory of one-to-one matching with transfers.

Table 1: Feasible matchings

W\M	Single(x_0)	x_1	$\cdots$	x_L	
Single(w_0)	-	$q_1 - \sum_{k=1}^{K} r_{k1}$	$\cdots$	$q_L - \sum_{k=1}^{K} r_{kL}$	-
w_1	$p_1 - \sum_{l=1}^{L} r_{1l}$	r_{11}	$\cdots$	r_{1L}	p_1
$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\vdots$
w_K	$p_K - \sum_{l=1}^{L} r_{Kl}$	r_{K1}	$\cdots$	r_{KL}	p_K
	-	q_1	$\cdots$	q_L	

Notes: Let $r_{kl} \geq 0$ denote the number of k-to-l matches. Feasibility of a matching imposes the $K+L$ adding up constraints $\sum_{l=1}^{L} r_{kl} \leq p_k$ for $k = 1, \ldots, K$ and $\sum_{k=1}^{K} r_{kl} \leq q_l$ for $l = 1, \ldots, L$.

Men also have heterogenous preferences. Man j's utility from matching with woman i is given by

$$V\left(W_i, X^j, \underline{\upsilon}^j\right) = \beta\left(W_i, X^j\right) - \tau\left(W_i, X^j\right) + \upsilon^j\left(W_i\right),$$

where $\beta\left(w_k, x_l\right)$ is the systematic utility a type $X^j = x_l$ men derives from matching with a type $W_i = w_k$ woman and $\upsilon^j\left(W_i\right) = \sum_{k=1}^{K} \mathbf{1}\left(W_i = w_k\right) \upsilon_k^j$ is a heterogenous component of match utility. Here $\tau\left(w, x\right)$ enters with a negative sign as we conceptually imagine men 'paying' women (recall that transfers may be negative). The utility from remaining unmatched is

$$\underline{V}\left(X^j, \underline{\upsilon}^j\right) = \underline{\beta}\left(X^j\right) + \upsilon_0^j.$$

Preference heterogeneity ensures that, for any given transfer function $\tau\left(w, x\right)$, observationally identical women will match with different types of men. If the support of the heterogeneity distribution is rich enough all types of matches will be observed in equilibrium.

Let $\underline{\varepsilon} = \left(\varepsilon_{i0}, \varepsilon_{i1}, \ldots, \varepsilon_{iL}\right)'$ and $\underline{\upsilon} = \left(\upsilon_0^j, \upsilon_1^j, \ldots, \upsilon_K^j\right)'$. I assume that the components of these vectors are independently and identically distributed Type I extreme value random variables

$$F_{\underline{\varepsilon}|W}\left(\underline{e} \middle| W = w_k\right) = \prod_{l=0}^{L} \exp\left(-\exp\left(-\frac{e_l}{\sigma_\varepsilon}\right)\right) \tag{1}$$

$$F_{\underline{\upsilon}|X}\left(\underline{\upsilon} \middle| X = x_l\right) = \prod_{k=0}^{K} \exp\left(-\exp\left(-\frac{\upsilon_k}{\sigma_\upsilon}\right)\right).$$

Assumption (1) is slightly more general than that maintained by Choo and Siow (2006a,b) who additionally impose the restriction $\sigma_\varepsilon = \sigma_\upsilon$. [2]

[2]Graham (2011) considers the case where $F_{\underline{\varepsilon}|W}$ and $F_{\underline{\upsilon}|X}$ are left nonparametric.

Equilibrium

Let $\alpha_{kl} = \alpha(w_k, x_l)$, $\underline{\alpha}_{k0} = \underline{\alpha}(w_k)$, $\beta_{kl} = \beta(w_k, x_l)$, $\underline{\beta}_{0l} = \underline{\beta}(x_l)$ and $\tau_{kl} = \tau(w_k, x_l)$. Let θ be a vector of model parameters – to be more precisely specified below – and $\underline{\tau}$ a $KL \times 1$ vector of transfers. The total number of type k women is given by p_k, that of type l men by q_l. Let $\mathbf{p} = (p_1, \ldots, p_K)'$ and $\mathbf{q} = (q_1, \ldots, q_L)'$.

Denote the probability, given a hypothetical transfer vector $\underline{\tau}$, that a type k woman matches with a type l man by $e_{kl}^{D}(\theta; \underline{\tau})$. The probability of remaining unmatched is $e_{k0}^{D}(\theta, \underline{\tau}) = 1 - \sum_{l=1}^{L} e_{kl}^{D}(\theta, \underline{\tau})$. Under the Type I extreme value assumption we have for $k = 1, \ldots, K$ (McFadden, 1974):

$$
\begin{aligned}
e_{k0}^{D}(\theta, \underline{\tau}) &= \frac{1}{1 + \sum_{n=1}^{L} \exp\left(\sigma_{\varepsilon}^{-1}\left[\alpha_{kn} - \underline{\alpha}_{k0} + \tau_{kn}\right]\right)} \\
e_{kl}^{D}(\theta, \underline{\tau}) &= \frac{\exp\left(\sigma_{\varepsilon}^{-1}\left[\alpha_{kl} - \underline{\alpha}_{k0} + \tau_{kl}\right]\right)}{1 + \sum_{n=1}^{L} \exp\left(\sigma_{\varepsilon}^{-1}\left[\alpha_{kn} - \underline{\alpha}_{k0} + \tau_{kn}\right]\right)}, \; l = 1, \ldots, L.
\end{aligned}
$$

Total ‘demand’ for type l men by type k women is therefore

$$
\begin{aligned}
r_{k0}^{D}(\theta, \underline{\tau}, \mathbf{p}, \mathbf{q}) &\stackrel{def}{\equiv} p_k e_{k0}^{D}(\theta, \underline{\tau}) \\
r_{kl}^{D}(\theta, \underline{\tau}, \mathbf{p}, \mathbf{q}) &\stackrel{def}{\equiv} p_k e_{kl}^{D}(\theta, \underline{\tau}),
\end{aligned} \tag{2}
$$

which, after some manipulation, gives

$$
\sigma_{\varepsilon} \ln\left(\frac{r_{kl}^{D}(\theta, \underline{\tau}, \mathbf{p}, \mathbf{q})}{r_{k0}^{D}(\theta, \underline{\tau}, \mathbf{p}, \mathbf{q})}\right) = \alpha_{kl} - \underline{\alpha}_{k0} + \tau_{kl}. \tag{3}
$$

For $l = 1, \ldots, L$ we get a conditional ‘supply’ of type l men to each of the $k = 1, \ldots, K$ types of women equal to

$$
\begin{aligned}
r_{0l}^{S}(\theta, \underline{\tau}, \mathbf{p}, \mathbf{q}) &\stackrel{def}{\equiv} q_l g_{0l}^{S}(\theta, \underline{\tau}) \\
r_{kl}^{S}(\theta, \underline{\tau}, \mathbf{p}, \mathbf{q}) &\stackrel{def}{\equiv} q_l g_{kl}^{S}(\theta, \underline{\tau}),
\end{aligned} \tag{4}
$$

where

$$
\begin{aligned}
g_{0l}^{S}(\theta, \underline{\tau}) &= \frac{1}{1 + \sum_{m=1}^{K} \exp\left(\sigma_{\upsilon}^{-1}\left[\beta_{ml} - \underline{\beta}_{0l} - \tau_{ml}\right]\right)} \\
g_{kl}^{S}(\theta, \underline{\tau}) &= \frac{\exp\left(\sigma_{\upsilon}^{-1}\left[\beta_{kl} - \underline{\beta}_{0l} - \tau_{kl}\right]\right)}{1 + \sum_{m=1}^{K} \exp\left(\sigma_{\upsilon}^{-1}\left[\beta_{ml} - \underline{\beta}_{0l} - \tau_{ml}\right]\right)}, \; k = 1, \ldots, K,
\end{aligned}
$$

so that

$$\sigma_\upsilon \ln\left(\frac{r^S_{kl}\left(\theta,\underline{\tau},\mathbf{p},\mathbf{q}\right)}{r^S_{0l}\left(\theta,\underline{\tau},\mathbf{p},\mathbf{q}\right)}\right)=\beta_{kl}-\underline{\beta}_{0l}-\tau_{kl}. \tag{5}$$

The transfer vector, $\underline{\tau}$, adjusts to equate the KL female 'demands' with the KL male 'supplies' so that in equilibrium

$$r^{\text{eq}}_{kl}\left(\theta,\underline{\tau}^{\text{eq}},\mathbf{p},\mathbf{q}\right)=r^D_{kl}\left(\theta,\underline{\tau}^{\text{eq}},\mathbf{p},\mathbf{q}\right)=r^S_{kl}\left(\theta,\underline{\tau}^{\text{eq}},\mathbf{p},\mathbf{q}\right),\ k=1,\ldots,K,\ l=1,\ldots,L, \tag{6}$$

with the 'eq' superscript denoting an equilibrium quantity.

Let

$$\gamma_{kl}=\frac{\alpha_{kl}+\beta_{kl}-\underline{\alpha}_{k0}-\underline{\beta}_{0l}}{\sigma_\varepsilon+\sigma_\upsilon},\ \lambda=\frac{\sigma_\upsilon}{\sigma_\varepsilon+\sigma_\upsilon}.$$

Imposing (6), adding (3) and (5), exponentiating and rearranging yields

$$r^{\text{eq}}_{kl}=\left(r^{\text{eq}}_{k0}\right)^{1-\lambda}\left(r^{\text{eq}}_{0l}\right)^{\lambda}\exp\left(\gamma_{kl}\right) \tag{7}$$

where I let $r^{\text{eq}}_{kl}=r^{\text{eq}}_{kl}\left(\theta,\underline{\tau}^{\text{eq}},\mathbf{p},\mathbf{q}\right)$ to economize on notation. See Graham (2013) for a fixed point representation of the matching market equilibrium. The set of feasible matchings is depicted in Table 1.

Estimation

Let i and j index two independent random draws from the equilibrium distribution of matches. Let

$$A^{klmn}_{ij}=1\left(W_i\in\{w_k,w_m\}\right)1\left(W_j\in\{w_k,w_m\}\right)1\left(X_i\in\{x_l,x_n\}\right)1\left(X_j\in\{x_l,x_n\}\right)$$

be an indicator for the event that this pair of matches belongs to the $klmn$ sub-allocation. That is the two women in the matches are of type w_k or w_m and the two men are of type x_l or x_n. We can assume that, without loss of generality, that the support points of female and male types are ordered so that $w_k<w_m$ and $x_l<x_n$. Now define

$$S_{ij}=\operatorname{sgn}\left\{\left(W_i-W_j\right)\left(X_i-X_j\right)\right\}.$$

We have that $S_{ij}=1$ if the "higher" type women matches with the "higher" type man and $S_{ij}=-1$ in the configuration is anti-assortative. If either $W_i=W_j$ or $X_i=X_j$ or both, then $S_{ij}=0$ and the configuration is neither assortative or anti-assortative. Using (7) above we have

$$\begin{aligned}\Pr\left(S_{ij}=1\right|S_{ij}\in\{-1,1\},A_{ij}^{klmn}=1\big) &= \frac{r_{kl}^{\text{eq}}r_{mn}^{\text{eq}}}{r_{kl}^{\text{eq}}r_{mn}^{\text{eq}}+r_{kn}^{\text{eq}}r_{ml}^{\text{eq}}} \\ &= \frac{\exp\left(\gamma_{kl}\right)\exp\left(\gamma_{mn}\right)}{\exp\left(\gamma_{kl}\right)\exp\left(\gamma_{mn}\right)+\exp\left(\gamma_{kn}\right)\exp\left(\gamma_{ml}\right)} \\ &= \frac{\exp\left(\gamma_{mn}-\gamma_{ml}-\left(\gamma_{kn}-\gamma_{kl}\right)\right)}{1+\exp\left(\gamma_{mn}-\gamma_{ml}-\left(\gamma_{kn}-\gamma_{kl}\right)\right)}.\end{aligned}$$

If we let $\mathbf{A}_{ij}$ be the $J=\binom{K}{2}\binom{L}{2}$ vector of sub-allocation indicators and define ϕ to be the corresponding vector of complementarity parameters

$$\phi_{klmn}=\gamma_{mn}-\gamma_{ml}-\left(\gamma_{kn}-\gamma_{kl}\right),$$

we have

$$\Pr\left(S_{ij}=1\right|S_{ij}\in\{-1,1\},\mathbf{A}_{ij})=\frac{\exp\left(\mathbf{A}_{ij}'\phi\right)}{1+\exp\left(\mathbf{A}_{ij}'\phi\right)}.$$

In some settings it may be desirable to impose more structure on the complementarity parameters. Let

$$\gamma\left(w,x\right)=a\left(w\right)+b\left(x\right)+e\left(w,x\right)'\eta$$

with $e\left(w,x\right)$ a low dimensional vector of basis functions and $a\left(w\right)$ and $b\left(x\right)$ arbitrary. If we let $\mathbf{M}$ be a matrix composed of rows

$$m_{klmn}=\left\{e\left(w_m,x_n\right)-e\left(w_m,x_l\right)-\left[e\left(w_k,x_n\right)-e\left(w_k,x_l\right)\right]\right\}',$$

then we have $\phi=\mathbf{M}\eta$ and hence

$$\Pr\left(S_{ij}=1\right|S_{ij}\in\{-1,1\},\mathbf{A}_{ij})=\frac{\exp\left(\mathbf{A}_{ij}'\mathbf{M}\eta\right)}{1+\exp\left(\mathbf{A}_{ij}'\mathbf{M}\eta\right)}.$$

We can consistently estimate η but choosing $\hat{\eta}$ to minimize the U-process:

$$L_N\left(\eta\right)=\frac{2}{N\left(N-1\right)}\sum_{i=1}^{N}\sum_{j<i}\left|S_{ij}\right|\left\{S_{ij}'\mathbf{A}_{ij}'\mathbf{M}\eta-\ln\left[1+\exp\left(S_{ij}'\mathbf{A}_{ij}'\mathbf{M}\eta\right)\right]\right\}.$$

References

[1] Choo, Eugene and Aloysius Siow. (2006a). "Who marries whom and why?" *Journal of Political Economy* 114 (1): 175 - 201.

[2] Choo, Eugene and Aloysius Siow. (2006b). "Estimating a marriage matching model with spillover effects," *Demography* 43 (3): 464 - 490.

[3] Graham, Bryan S. (2011). "Econometric methods for the analysis of assignment problems in the presence of complementarity and social spillovers," *Handbook of Social Economics* 1B: 965 - 1052 (J. Benhabib, A. Bisin, & M. Jackson, Eds.). Amsterdam: North-Holland.

[4] Graham, Bryan S. (2013). "Comparative static and computational methods for an empirical one-to-one transferable utility matching model," *Advances in Econometrics* 31: 153 - 181.

[5] McFadden, Daniel. (1974). "Conditional logit analysis of qualitative choice behavior," *Frontiers in Econometrics*: 105 - 142 (P. Zarembka, Ed.). Academic Press: New York.

[6] Shapley, Lloyd S. and Martin Shubik. (1971) "The assignment game I: The core," *International Journal of Game Theory* 1 (1): 111 - 130.

