Cupid's Invisible Hand:

Social Surplus and Identification in Matching Models

Alfred Galichon[1] **Bernard Salanié**[2]

April 6, 2011[3]

[1]Economics Department, École polytechnique; e-mail: alfred.galichon@polytechnique.edu

[2]Department of Economics, Columbia University; e-mail: bsalanie@columbia.edu.

[3]The authors are grateful to Pierre-André Chiappori for useful comments and discussions. Galichon gratefully acknowledges support from Chaire EDF-Calyon "Finance and Développement Durable," Chaire Axa "Assurance et Risques Majeurs" and FiME, Laboratoire de Finance des Marchés de l'Energie. Part of the research underlying this paper was done when Galichon was visiting the University of Chicago Booth School of Business and Columbia University. Galichon thanks the Alliance program for its support.

Abstract

We investigate a matching game with transferable utility when some of the characteristics of the players are unobservable to the analyst. We allow for a wide class of distributions of unobserved heterogeneity, subject only to a separability assumption that generalizes Choo and Siow (2006). We first show that the stable matching maximizes a social gain function that trades off two terms. The first term is simply the average surplus due to the observable characteristics; and the second one can be interpreted as a generalized entropy function that reflects the impact of the unobserved characteristics. We use this result to derive simple closed-form formulæ that identify the joint surplus in every possible match and the equilibrium utilities of all participants, given any known distribution of unobserved heterogeneity. Moreover, we show that if transfers are observed, then the pre-transfer utilities of both partners are also identified. We conclude by discussing some empirical approaches suggested by these results for the study of marriage markets, hedonic prices, and the market for CEOs.

Keywords: matching, marriage, assignment, hedonic prices.

JEL codes: C78, D61, C13.

Introduction

Since the seminal contribution of Becker (1973), economists have modeled marriage markets as a matching problem in which each potential match generates a marital surplus. Given transferable utilities, the distributions of tastes and of desirable characteristics determine equilibrium shadow prices, which in turn explain how partners share the marital surplus in any realized match. This insight is not specific of the marriage market: it characterizes the "assignment game" (Shapley and Shubik (1972)), i.e. models of matching with transferable utilities. These models have also been applied to competitive equilibrium with hedonic pricing (Chiappori, McCann, and Nesheim (2008)) and the market for CEOs (Gabaix and Landier (2008)). We will show how our results can be used in these three contexts; but for concreteness, we often refer to partners as men and women in the exposition of the main results.

While Becker presented the general theory, he focused on the special case in which the types of the partners are one-dimensional and are complementary in producing surplus. As is well-known, the socially optimal matches then exhibit positive assortative matching. Moreover, the resulting configuration is stable, it is in the core of the corresponding matching game, and it can be efficiently implemented by classical optimal assignment algorithms.

This result is both simple and powerful; but its implications are also quite unrealistic and at variance with the data, in which matches are observed between partners with quite different characteristics. To account for this wider variety of matching patterns, one could introduce search frictions, as in Shimer and Smith (2000). But the resulting model is hard to handle, and under some additional conditions it still implies assortative matching. A simpler solution consists in allowing the joint surplus of a match to incorporate latent characteristics—heterogeneity that is unobserved by the analyst. Choo and Siow (2006) showed that it can be done in a way that yields a highly tractable model in large populations, provided that the unobserved heterogeneities enter the marital surplus quasi-additively and that they are distributed as standard type-I extreme value terms. Then the usual apparatus of multinomial logit discrete choice models applies, linking marriage patterns to marital surplus in a very simple manner[1]. Choo and Siow (2006) used this model to link the changes in gains to marriage and abortion laws; Siow and Choo (2006) applied it to Canadian data to measure the impact of demographic changes. It has also been used to study increasing returns in marriage markets (Botticini and Siow (2008)) and to test for complementarities across partner educations (Siow (2009)).

We revisit here the theory of matching with transferable utilities in the light of Choo and Siow's insights. Our contribution is threefold. First, we extend this framework to more general distributions of utility shocks. Chiappori, Salanié, and Weiss (2010) showed that quasi-additivity by itself reduces the complexity of the matching model to a series of discrete choice problems. We prove that with quasi-additive surplus, the market equilibrium maximizes a social surplus function that consists of two terms: a term that describes assortativeness on the observed characteristics; and a generalized entropic term that describes the random character of matching conditional on observed characteristics. While the first term tends to match partners with complementary observed characteristics, the second one pulls towards randomly assigning partners to each other. The social gain from any matching patterns trades off these two terms. In particular, when unobserved heterogeneity is distributed as in Choo and Siow (2006), the generalized entropy is simply the usual entropy measure.

Our second contribution is to show that the maximization of the social surplus function described above has very straightforward consequences in terms of identification, both when equilibrium transfers are observed and when they and are not. In fact, most quantities of interest can be obtained from derivatives of the terms that constitute generalized entropy. We show in particular that the joint surplus from matching is (minus) a derivative of the generalized entropy, computed at the observed matching. The expected and realized utilities of all types of men and women follow just as directly. If moreover equilibrium transfers are observed, then we also identify the pre-transfer utilities on both sides of the market.

[1]Fox (2010) relies instead on a rank-order property to identify the surplus function from the matching patterns. The handbook chapter by Graham (2011) discusses these and other approaches.

These results suggest various empirical strategies that can be used to estimate the parameters of models of matching with transferable utilities. We show how they fit within the framework of minimum distance estimation, and we discuss their applicability to the three classes of markets: marriage markets, where transfers between spouses not observed; the market for CEOs and competitive market with hedonic prices, where transfers (CEO compensation, the equilibrium prices of different varieties of products) may be observed.

Section 1 sets up the model and the notation. We prove our main results in section 2, and we specialize them to leading examples in section 3. Our results very significantly extend the Choo and Siow framework: they allow for general error distributions with heteroskedasticity and correlation across alternatives, as in generalized extreme values models or mixed logit models for instance. They open the way to new and richer specifications; section 4 explains how various restrictions can be imposed to identify and estimate the underlying parameters[2].

1 The Assignment Problem with Unobserved Heterogeneity

Throughout the paper, we maintain the basic assumptions of the transferable utility model of Choo and Siow: utility transfers between partners are unconstrained, matching is frictionless, and there is no asymmetric information. We also try to stay as close as possible to the notation Choo and Siow used. Men can belong to I groups, indexed by i; and women can belong to J groups, indexed by j. Groups can for instance be defined by education, race, and other characteristics which are observed by all men and women and also by the analyst. On the other hand, men and women of a given group differ along some dimensions that they all observe, but which do not figure in the analyst's dataset.

Choo and Siow assumed that the utility of a man m of group i who marries a woman

[2] This paper builds on and significantly extends our earlier discussion paper (Galichon and Salanié (2010)), which is now obsolete.

of group j can be written as

$$\tilde{\alpha}_{ij} - \tau_{ij} + \varepsilon_{ijm},$$

where τ_{ij} represents the utility that the man has to transfer to his partner in equilibrium, and ε_{ijm} is a standard type-I extreme value disturbance. If such a man remains single, he gets utility

$$\tilde{\alpha}_{i0} + \varepsilon_{i0m}.$$

Similarly, the utility of a woman w of group j who marries a man of group i can be written as

$$\tilde{\gamma}_{ij} + \tau_{ij} + \eta_{ijw},$$

and she gets utility

$$\tilde{\gamma}_{0j} + \eta_{0jw}.$$

is she is single.

Only utility differences matter in this model; we denote

$$\alpha_{ij} = \tilde{\alpha}_{ij} - \tilde{\alpha}_{i0} \text{ and } \gamma_{ij} = \tilde{\gamma}_{ij} - \tilde{\gamma}_{0j}.$$

The key assumption here is that the utility of a man m of group i who marries a woman w of group j does not depend on who this woman is—with a similar assumption for women. We will return to the interpretation of this assumption, which we will call "separability". When there are very large numbers of men and women within each group, Choo and Siow showed that there is a simple equilibrium relationship between group preferences, as defined by α and γ, and equilibrium marriage patterns. Denote μ_{ij} the number of marriages between men of group i and women of group j; μ_{i0} the number of single men of group i; and μ_{0j} the number of single women of group j. Denote

$$\pi_{ij} = \frac{\alpha_{ij} + \gamma_{ij}}{2}.$$

the *total systematic net gains to marriage*; and note that by construction, π_{i0} and π_{0j} are zero. Choo and Siow proved the following result:

Theorem 1 (Choo and Siow) *In equilibrium, for all $i, j \geq 1$*

$$\exp(\pi_{ij}) = \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}}.$$

Therefore marriage patterns μ directly identify the gains to marriage π in such a model.

It turns out that the assumption on the distribution of the utility shocks ε and η is not crucial. As shown in Chiappori, Salanié and Weiss (2010), some of the structure of the problem is preserved if this assumption is relaxed. The crucial assumption is what they call "separability". To state it, let Φ_{mw} denote the joint surplus created by a match between a man m and a women w.

Assumption 1 (Separability) *If men m and m' belong to the same group i and women w and w' belong to the same group j, then*

$$\Phi_{mw} + \Phi_{m'w'} = \Phi_{mw'} + \Phi_{m'w}.$$

It is easy to see that under Assumption 1, the surplus from a match between a man m of group i and a woman w of group j must decompose into

$$\Phi_{mw} = 2\pi_{ij} + \varepsilon_{ijm} + \eta_{ijw},$$

where the ε and η can be normalized to have zero mean. Again, $\pi_{i0} = \pi_{0j} = 0$: without loss of generality, singles get zero mean utility.

This assumption rules out interactions between unobserved characteristics in the marital output from a match, given the observed characteristics of both partners. On the other hand, it does not restrict group preferences in any way; and it also allows for variation in marital output within groups, as long as they do not interact across partners. For instance, men of a given group may differ in the marital outputs they can form, but only as relates to the group of their partner.

To take an analogy with discrete choice models of consumer purchases, take the following standard specification for the utility a buyer b derives from a variety v:

$$U_{bv} = \pi(X_b, X_v) + X_v\varepsilon_b + X_b\varepsilon_v + \varepsilon_{bv}.$$

In this context, separability would allow for variation in tastes over observed characteristics of products (through ε_v), and for group-dependent tastes for unobserved product characteristics ε_b. On the other hand, it would rule out the interaction term ε_{bv}.

We denote p_i the number of men of group i, and q_j the number of women of group j; then

$$\forall i \geq 1,\ \sum_{j=0}^{J} \mu_{ij} = p_i\ ;\ \forall j \geq 1,\ \sum_{i=0}^{I} \mu_{ij} = q_j. \tag{1.1}$$

For future reference, we denote $\mathcal{M}$ the set of $(IJ + I + J)$ non-negative numbers (μ_{ij}) that satisfy these $(I + J)$ equalities. Each element of $\mathcal{M}$ is called a "matching" as it defines a feasible set of matches (and singles).

Like Choo and Siow, we assume that the p_i's and q_j's are "large": there are a large number of men in any group i, and of women in any group j. More precisely, our statements in the following are exactly true when the number of individuals goes to infinity and the proportions of genders and types converge. To simplify the exposition, we consider the limit of a sequence of large economies where the proportion of each type remains constant:

Assumption 2 (Large Market) *The number of individuals on the market $N = \sum_{i=1}^{I} p_i + \sum_{j=1}^{J} q_j$ goes to infinity; and the ratios (p_i/N) and (q_j/N) are constant.*

With finite N we would need to introduce corrective terms; we leave this for further research.

2 Social Surplus, Utilities, and Identification

As Choo and Siow (2006) remind us (p. 177): "A well-known property of transferable utility models of the marriage market is that they maximize the sum of marital output in the society". This is true when marital output is defined as it is evaluated by the participants: the market equilibrium in fact maximizes

$$\sum_{mw} \delta_{mw} \Phi_{mw}$$

over the set of feasible matchings (δ_{mw}). On the other hand, this is not very useful to the analyst: she does not observe some of the characteristics of the players, and she can only compute quantities that depend on the observed groups of the partners in a match. A very naive evaluation of the sum of marital output, computed from the groups of partners only, would be

$$2 \sum_{ij} \mu_{ij} \pi_{ij}; \tag{2.1}$$

but this is clearly misleading. Realized matches by nature have a value of the unobserved marital surplus $(\varepsilon_{ijm} + \eta_{ijw})$ that is more favorable than an unconditional draw; and as a consequence, the equilibrium marriage patterns (μ) do not maximize the value in (2.1) over $\mathcal{M}$.

In order to find the expression of the value function that (μ) maximizes, we need to account for terms that reflect the conditional expectation of the unobserved parts of the surplus, given a match on observable types. To make this more precise, we need to introduce some notation. We continue to assume separability (Assumption 1) and a large market (Assumption 2); but we allow for quite general distributions of unobserved heterogeneity:

Assumption 3 (Distribution of Unobserved Variation in Surplus)

a) For any man $m \in i$, the ε_{ijm} are drawn from a $(J+1)$-dimensional distribution $\mathbf{P}_i$;

b) For any woman $w \in j$, the η_{ijw} are drawn from an $(I+1)$-dimensional distribution $\mathbf{Q}_j$;

c) These draws are independent across men and women.

Assumption 3 clearly is a substantial generalization with respect to Choo and Siow

(2006), who assume that $\mathbf{P}_i$ and $\mathbf{Q}_j$ are independent products of standard type-I extreme values distributions:

Assumption 4 (Type-I extreme values distribution)

a) For any man $m \in i$, the $(\varepsilon_{ijm})_{j=0,\ldots,J}$ are drawn independently from a standard type-I extreme value distribution;

b) For any woman $w \in j$, the $(\eta_{ijm})_{i=0,\ldots,I}$ are drawn independently from a standard type-I extreme value distribution;

c) These draws are independent across men and women.

Assumption 3 generalizes assumption 4 in three important ways: it allows for different families of distributions, with any form of heteroskedasticity, and with any pattern of correlation across partner groups.

2.1 A Heuristic Derivation

Now suppose that the men of group i expect to get mean utilities w_j from marrying partner type j, for $j = (0, \ldots, J)$. A given man of this group, characterized by a draw (ε_j) from $\mathbf{P}_i$, would then choose the partner type j that makes $(w_j + \varepsilon_j)$ largest. Therefore the sum of the expected utilities of these men would be

$$G_i(w) = p_i \mathbb{E}_{\mathbf{P}_i} \left[\max_{j=0,\ldots,J} (w_j + \varepsilon_j) \right],$$

where the expectation is taken over a random vector $(\varepsilon_0, \ldots, \varepsilon_J) \sim \mathbf{P}_i$.

Similarly, the sum of the expected utilities of the women of group j is

$$H_j(z) = q_j \mathbb{E}_{\mathbf{Q}_j} \left[\max_{i=0,\ldots,I} (z_i + \eta_i) \right].$$

The social surplus is simply the sum of the expected utilities of all types of men and women. Thus if we denote U_{ij} and V_{ij} the mean utilities of a man of type i and of a woman of type j when they are matched, the social surplus is

$$\sum_{i=1}^{I} G_i(U_{i.}) + \sum_{j=1}^{J} H_j(V_{.j}),$$

denoting $U_{i.} = (U_{i0} = 0, U_{i1}, \dots, U_{iJ})$ and $V_{.j} = (V_{0j} = 0, V_{1j}, \dots, V_{Ij})$. Of course these mean utilities are unobserved, and we must find a way to write them in terms of the matching patterns μ. We will give here a heuristic explanation of how we obtain such a formula[3].

Let us focus on the function G_i. By construction,

$$G_i(w) = p_i \sum_{j=0}^{J} \Pr(j|i;w) \left(w_j + e(j|i;w)\right), \tag{2.2}$$

where we denote $\Pr(j|i;w)$ the probability that the maximum is achieved for a choice of partner in group j when mean utilities are w, and $e(j|i;w)$ the conditional expectation of ε_j in this case. In particular, if $w = U_{i.}$ then $\Pr(j|i;w) = \mu_{ij}/p_i$; and we obtain

$$G_i(U_{i.}) = \sum_{j=0}^{J} \mu_{ij} U_{ij} + p_i \sum_{j=0}^{J} \Pr(j|i;U_{i.}) e(j|i;U_{i.}). \tag{2.3}$$

Now $U_{ij} + V_{ij} = 2\pi_{ij}$ since the mean utilities of the partners must add up to the marital surplus; and when we add the first term in (2.3) to the corresponding term for women, we will find

$$\sum_{i=1}^{I}\sum_{j=0}^{J} \mu_{ij} U_{ij} + \sum_{j=1}^{J}\sum_{i=0}^{I} \mu_{ij} V_{ij} = 2 \sum_{i,j \geq 1} \mu_{ij}\pi_{ij}$$

which is just the "naive" formula in (2.1). We still have to evaluate the conditional terms. To do this, note that G_i is convex since it is a linear combination of the maxima of linear functions; as such it is almost everywhere differentiable, with derivatives

$$\frac{\partial G_i}{\partial w_j}(w) = p_i \Pr(j|i;w). \tag{2.4}$$

But this means that we can rewrite the second term in (2.2) as

$$p_i \sum_{j=0}^{J} \Pr(j|i;w) e(j|i;w) = G_i(w) - \sum_{j=0}^{J} w_j \frac{\partial G_i}{\partial w_j}(w) \equiv \mathcal{G}(w). \tag{2.5}$$

[3] Appendix A gives rigorous proofs of all of our results.

If G_i were a homogeneous function of degree 1 then the right-hand side $\mathcal{G}(w)$ would be zero. But G_i is convex, and therefore the right-hand side is positive; in fact, it is is easily computed from the Legendre-Fenchel transform of G_i. This associates to any $(a_0, \ldots, a_J)$ the number

$$\max_{w=(w_0,\ldots,w_J)} \left(\sum_{j=0}^{J} a_j w_j - G_i(w) \right).$$

Note that if t is any scalar and $w'_j = w_j + t$, then $G_i(w') = G_i(w) + p_i t$: the function G_i is convex but not strictly convex. As a consequence, the value of its Legendre-Fenchel transform in $(a_0, \ldots, a_J)$ is infinite if $\sum_{j=0}^{J} a_j \neq p_i$. Accordingly, we focus on its restriction to the hyperplane $\sum_{j=0}^{J} a_j = p_i$, which always takes finite values; and we use a slightly different argument list:

$$G_i^*(p_i, a_1, \ldots, a_J) = \max_{w=(w_0,\ldots,w_J)} \left(\left(p_i - \sum_{j=1}^{J} a_j \right) w_0 + \sum_{j=1}^{J} a_j w_j - G_i(w) \right),$$

extended to $G^*(p_i; a_1, \ldots, a_J) = +\infty$ if $\sum_{j=1}^{J} a_j \geq p_i$. We define $H_j^*(q_i, b_1, \ldots, b_I)$ similarly for women of type j.

Denoting $a_0 = p_i - \sum_{j=1}^{J} a_j$, the first-order conditions in $G_i^*(a)$ can be written

$$a_j = \frac{\partial G_i}{\partial w_j}(w); \tag{2.6}$$

so that if w achieves the maximum in $G_i^*(p_i, a)$,

$$G_i^*(p_i, a) = -\mathcal{G}(w).$$

While this may seem like replacing an unknown quantity with another, combining equations (2.4) and (2.6) implies that if $a_j = \mu_{ij}$ for all j, then the solution w is simply $w_j \equiv U_{ij}$, the vector of mean utilities for which men of type i split across partner types with the probabilities given by $\mu_{i.}/p_i$. Going back to (2.2), we finally obtain

$$G_i(U_{i.}) = \sum_{j=0}^{J} \mu_{ij} U_{ij} - G_i^*(p_i, \mu_{i.}), \tag{2.7}$$

and

$$\sum_{i=1}^{I} G_i(U_{i.}) + \sum_{j=1}^{J} H_j(V_{j.}) = 2 \sum_{i,j\geq 1} \mu_{ij}\pi_{ij} - \sum_{i=1}^{I} G_i^*(p_i, \mu_{i.}) - \sum_{j=1}^{J} H_j^*(q_j, \mu_{.j}).$$

The right-hand side of this equation gives the value of the social surplus when the matching patterns are (μ_{ij}). The first term $2\sum_{ij}\mu_{ij}\pi_{ij}$ reflects "group preferences": if groups i and j generate more surplus when matched, then they should be matched with higher probability. In the one-dimensional Beckerian example, an increasing i or j could reflect higher education. If the marital surplus is complementary in the educations of the two partners, π_{ij} is supermodular and this first term is maximized when matching partners with similar education levels (as far as feasibility constraints allow.) On the other hand, the second and the third term reflect the effect of the dispersion of individual affinities, conditional on observed characteristics: those men m in a group i that have more affinity to women of group j should be matched to women of group j.

The formula for the social surplus incorporates these two considerations. To take the education example again, a marriage between a man with a college degree and a woman who is a high-school dropout generates less marital surplus on average than a marriage between college graduates; but because of the dispersion of marital surplus that comes from the ε and η terms, it will be optimal to have some marriages between dissimilar partners.

2.2 Main Results

We now give formal statements of our results. As Legendre-Fenchel transforms of convex functions, the functions G_i^* and H_j^* are also convex; as such, they are differentiable almost everywhere—and very mild assumptions on the distributions $\mathbf{P}_i$ and $\mathbf{Q}_j$ would make them differentiable everywhere. We will use their derivatives in stating our results; they should be replaced with subgradients at hypothetical points of non-differentiability.

We denote $\mu_{i.}$ the vector $(\mu_{i1}, \ldots, \mu_{iJ})$, and similarly for $\mu_{.j}$. For simplicity, we will assume that all matching patterns are possible at the optimal matching:

Assumption 5 (Interior Solution) *For every $i \geq 0$ and $j \geq 0$, μ_{ij} is positive.*

Assumption 5 must hold in large markets if the functions G_i and H_j are increasing in all of their arguments; as they will be if the distributions $\mathbf{P}_i$ and $\mathbf{Q}_j$ all have unbounded support, or if their supports are wide enough relative to the variation in π.

We prove two results in Appendix A. The first one characterizes the social surplus function that the stable matching maximizes. As explained in section 2.1, the social surplus trades off matching on observables and on unobservables:

Theorem 2 ***(Social Surplus)*** *Under assumptions 1, 2, 3, and 5, the market equilibrium* $\mu = (\mu_{ij})_{i,j \geq 1}$ *maximizes the social gain*

$$\mathcal{W}(\mu) = 2 \sum_{i,j \geq 1} \mu_{ij} \pi_{ij} + \mathcal{E}(p, q, \mu),$$

where $\mathcal{E}$ *is defined by*

$$\mathcal{E}(p, q, \mu) = -\sum_{i=1}^{I} G_i^* \left(p_i, \mu_{i.}\right) - \sum_{j=1}^{J} H_j^* \left(q_j, \mu_{.j}\right);$$

and the probabilities of singlehood are given by

$$\mu_{i0} = p_i - \sum_{j=1}^{J} \mu_{ij} \text{ and } \mu_{0j} = q_j - \sum_{i=1}^{I} \mu_{ij}.$$

We call $\mathcal{E}$ the *generalized entropy function* of the distribution of characteristics—as we will see, in the simple case analyzed by Choo and Siow (2006) it is just the usual notion of entropy. Theorem 2 has several important consequences. In particular, it yields a remarkably simple formula for the utilities participants of any type obtain in equilibrium. We state the result for men—the one for women follows with the obvious change in notation.

Theorem 3 ***(Participant Utilities)*** *Under assumptions 1, 2, 3, and 5,*

a) In equilibrium, a man $m \in i$ *who marries a woman of group* j *obtains utility*

$$U_{ij} + \varepsilon_{ijm}$$

where

$$U_{ij} = \frac{\partial G_i^*}{\partial \mu_{ij}} \left(p_i, \mu_{i.}\right)$$

can also be computed by solving the system of equations

$$\frac{\partial G_i}{\partial w_{ij}}(U_{i.}) = \mu_{ij} \text{ for } j = 0, \ldots, J,$$

given the normalization $U_{i0} = 0$.

b) The average expected utility of the men of group i *is*

$$u_i = \frac{G_i(U_{i.})}{p_i} = -\frac{\partial G_i^*}{\partial p_i}(p_i, \mu_{i.}). \tag{2.8}$$

Part b) of Theorem 3, in particular, makes it extremely easy to evaluate the participant utilities. The data directly yield the number of participants of this type (p_i) and their matching patterns ($\mu_{i.}$); and the specification of the distribution of unobserved heterogeneity determines the function G_i^*, thus allowing for the computation of u_i.

Remember that $2\pi_{ij} = U_{ij} + V_{ij}$; then Theorem 3 implies directly the following relationship between the matching patterns and the underlying surplus function:

Theorem 4 ***(Identification)*** *Under assumptions 1, 2, 3, and 5,*

a) In equilibrium, for any $i, j \geq 1$

$$2\pi_{ij} = -\frac{\partial \mathcal{E}(p, q, \mu)}{\partial \mu_{ij}} = \frac{\partial G_i^*}{\partial \mu_{ij}}(p_i, \mu_{i.}) + \frac{\partial H_j^*}{\partial \mu_{ij}}\left(q_j, \mu_{.j}\right); \tag{2.9}$$

b) Denote the systematic part of pre-transfer utilities (α, γ) *and of transfers* τ *as in section 1. Then*

$$U_{ij} = \alpha_{ij} - \tau_{ij} \text{ and } V_{ij} = \gamma_{ij} + \tau_{ij}.$$

Therefore if transfers are observed, both pre-transfer utilities α_{ij} *and* γ_{ij} *are also identified.*

Equation (2.9) identifies the marital surplus matrix π from the observed matching patterns μ, given the distribution of unobserved heterogeneities. It can also be used to solve for the optimal matching, given full knowledge of the surplus function; it generates a system of IJ (typically nonlinear) equations with IJ independent unknowns : the values of μ_{ij} for $i, j \geq 1$. Once the μ's are computed, they can for instance be injected into formula (2.8) to compute the expected utilities of each type of participant.

3 Examples

The functions G_i^* and H_j^*, and hence $\mathcal{E}$ can often be found in closed form. Appendix B gives the resulting formulæ for McFadden's Generalized Extreme Value (GEV) framework. This comprises some very useful special cases. We now study a couple of examples.

Example 1 (Heteroskedastic logit) *Assume that ε_{ijm} and η_{ijw} are type-I extreme value random variables with scaling factors σ_i^m and σ_j^w respectively. Then (focusing on men)*

$$G_i(w) = p_i \sigma_i^m \log \sum_{j=0}^{J} \exp\left(\frac{w_j}{\sigma_i^m}\right).$$

Take numbers of marriages $(a_1, \ldots, a_J)$ for men of type i, and denote $a_0 = p_i - \sum_{j=1}^{J} a_j$. These marriage patterns can be rationalized by the mean utilities

$$w_i^j(p_i, a) = \sigma_i^m \log \frac{a_j}{p_i} + t_i(a),$$

where $t_i(a)$ is an arbitrary scalar function. As a result,

$$G_i^*(p_i, a_1, \ldots, a_J) = \sigma_i^m \sum_{j=0}^{J} a_j \ln \frac{a_j}{p_i};$$

and

$$\mathcal{E}(p, q, \mu) = -\sum_{i=1}^{I} \sigma_i^m \sum_{j=0}^{J} \mu_{ij} \ln \frac{\mu_{ij}}{p_i} - \sum_{j=1}^{J} \sigma_j^w \sum_{i=0}^{I} \mu_{ij} \ln \frac{\mu_{ij}}{q_j}.$$

Hence (2.9) simplifies to

$$2\pi_{ij} = \left(\sigma_i^m + \sigma_j^w\right) \ln \mu_{ij} - \sigma_i^m \ln \mu_{i0} - \sigma_j^w \ln \mu_{0j}; \tag{3.1}$$

men of type i get an average expected utility

$$u_i = -\sigma_i^m \ln \frac{\mu_{i0}}{p_i},$$

and women of type j get an average expected utility

$$v_j = -\sigma_j^w \ln \frac{\mu_{0j}}{q_j}.$$

As a particular case of the above example when $\sigma_i^m = \sigma_j^w = 1$, we get Choo and Siow's model:

Proposition 1 *Under assumptions 1, 2 and 4 (which implies 5), the function $\mathcal{E}$ is simply*

$$\mathcal{E}(\mu) = -\sum_{i=1}^{I}\sum_{j=0}^{J}\mu_{ij}\ln\frac{\mu_{ij}}{p_i} - \sum_{j=1}^{J}\sum_{i=0}^{I}\mu_{ij}\ln\frac{\mu_{ij}}{q_j},$$

so that

$$\mathcal{W}(\mu) = 2\sum_{i,j\geq 1}\mu_{ij}\pi_{ij} - \sum_{i=1}^{I}\sum_{j=0}^{J}\mu_{ij}\ln\frac{\mu_{ij}}{p_i} - \sum_{j=1}^{J}\sum_{i=0}^{I}\mu_{ij}\ln\frac{\mu_{ij}}{q_j}. \tag{3.2}$$

Moreover, surplus and matching patterns are linked by

$$2\pi_{ij} = 2\ln\mu_{ij} - \ln\mu_{i0} - \ln\mu_{0j},$$

which is Choo and Siow's result (Theorem 1 above.)

To interpret formula (3.2), start with the case when unobserved heterogeneity is dwarfed by variation due to observable characteristics: $\Phi_{mw} \simeq 2\pi_{ij}$ if $m \in i$ and $w \in j$. Then we know that the observed matching μ must maximize the value in (2.1); but this is precisely what the more complicated expression $\mathcal{W}(\mu)$ above boils down to if we scale up the values of π to infinity. On the other hand, if data is so poor that unobserved heterogeneity dominates ($\pi \simeq 0$), then the analyst should observe something that, to her, looks like completely random matching. But information theory tells us that entropy is a natural measure of statistical disorder; and the entropy of a discrete probability distribution $(x_1, \ldots, x_N)$ is simply

$$-\sum_{i=1}^{N} x_i \ln x_i,$$

which appears in the two terms that constitute $\mathcal{W}$ in the framework of Proposition 1 when π is zero. In the intermediate case in which some of the variation in marital surplus is driven by group characteristics (through the π_{ij}) and some is carried by the unobserved heterogeneity terms ε_{ijm} and η_{ijw}, the market equilibrium trades off matching on group characteristics (as in (N)) and randomness, as measured by the entropy terms in $\mathcal{W}(\mu)$.

In the more general cases that Assumption 3 allows for, the function $\mathcal{E}(\mu)$ cannot be interpreted as the entropy of a probability distribution; we call it a *generalized entropy* since it plays a similar role.

As a more complex example of a GEV distribution, consider a nested logit.

Example 2 (Nested logit) *Suppose for instance that men of type i choose among "nests" A_l^i for $l = 1, \ldots, m_i$, and that the scale parameter is σ_{il}^m in nest l, and s_i^m overall. Then the system of equations that defines the U_{ij}:*

$$\frac{\partial G_i}{\partial w_{ij}}(U_{i.}) = \mu_{ij} \text{ for } j = 0, \ldots, J,$$

can be rewritten as

$$\frac{\mu_{ij}}{p_i} = \frac{\left(\sum_{j' \in A_l^i} \exp\left(\frac{U_{ij'}}{\sigma_{il}^m}\right)\right)^{\sigma_{il}^m / s_i^m}}{\sum_{k=1}^{m_i}\left(\sum_{j' \in A_k^i} \exp\left(\frac{U_{ij'}}{\sigma_{ik}^m}\right)\right)^{\sigma_{ik}^m / s_i^m}} \frac{\exp\left(U_{ij}/\sigma_{il}^m\right)}{\sum_{j' \in A_l^i} \exp\left(U_{ij'}/\sigma_{il}^m\right)} \tag{3.3}$$

where l is the index of the nest such that $j \in A_l^i$. There is no general closed-form expression for U_{ij}; however, note that within a nest A_l^i,

$$U_{ij} = \sigma_{il}^m \log \frac{\mu_{ij}}{p_i} + t_l^i$$

and that in (3.3) only the constants t_l^i remain to be determined numerically.

While the GEV framework is convenient, the mixed logit model has also become quite popular in the applied literature; it is our last example[4].

Example 3 (Mixed logit) *Take nonnegative numbers α_{ik} such that $\sum_{k=1}^K \alpha_{ik} = 1$ for each i.. Consider the mixture model in which for any type i of men, with probability α_{ik} the distribution $\mathbf{P}_i$ is iid type-I extreme value with standard error σ_{ik}^m.*

Then the U_{ij} solve

$$\frac{\mu_{ij}}{p_i} = \sum_{k=1}^K \alpha_{ik} \frac{e^{U_{ij}/\sigma_{ik}^m}}{\sum_{j'=0}^J e^{U_{ij'}/\sigma_{ik}^m}}.$$

[4]Our framework allows for more general specifications, e.g. mixed GEV models.

4 Empirical Approaches

The assumptions in Choo and Siow (2006) imply a stark trade-off in the specification of the model: in order to keep the joint surplus π entirely non-parametric, the distribution of the unobserved heterogeneity is very tightly specified—in fact, imposing Assumption 4 leaves it with no free parameter at all. Our results open the way to a wider range of empirical strategies, in which the analyst can leverage on restrictions on the joint surplus in order to allow for more general distributions of unobserved heterogeneity.

4.1 Estimation

Suppose for instance that the joint surplus π and the distributions of unobserved heterogeneity $(\mathbf{P}_i), (\mathbf{Q}_j)$ are specified up to a parameter vector θ, so that $\pi_{ij} \equiv \pi_{ij}(\theta)$ and the generalized entropy $\mathcal{E}_\theta$ also depends on the unknown parameters. Then given observed matching patterns (μ_{ij}) we could estimate θ by using minimum distance methods. To see this, take a hypothetical matching (ν_{ij}) that satisfies the feasibility constraints in (1.1):

$$\forall i \geq 1,\ \sum_{j=0}^{J} \nu_{ij} = p_i\ ;\ \forall j \geq 1,\ \sum_{i=0}^{I} \nu_{ij} = q_j. \tag{4.1}$$

Then (up to constant terms) the log-likelihood of the matching ν is

$$L(\nu;\mu) = 2 \sum_{i\geq 1, j\geq 1} \mu_{ij} \log \nu_{ij} + \sum_{i\geq 1} \mu_{i0} \log \nu_{i0} + \sum_{j\geq 1} \mu_{0j} \log \nu_{0j}.$$

Since this is concave in ν and the feasibility constraints are linear, maximizing the log-likelihood over the set of all feasible matchings ν simply gives $\nu_{ij} = \mu_{ij}$. But the parameterization imposes constraints, in the form of equation (2.9): given parameter values θ, for any $i, j \geq 1$,

$$\pi_{ij}(\theta) = -\frac{\partial \mathcal{E}_\theta(p, q, \nu)}{\partial \mu_{ij}}. \tag{4.2}$$

These additional constraints exhaust the restrictions from the theoretical model; therefore if the model identifies θ, choosing θ and ν to maximize the log-likelihood under the constraints (4.1) and (4.2) yields a consistent and asymptotically efficient estimator of θ.

Conceptually, this is a minimum distance estimator: if for every θ equation (4.2) can be solved for

$$\nu = K(\pi(\theta), p, q, \theta),$$

then maximizing $L(K(\pi(\theta), p, q, \theta); \mu)$ over θ amounts to minimizing the Kullback-Leibler distance between the model and the data. Moreover, if the model is overidentified, then a specification test can be constructed in the usual way from the value of the distance at its minimum.

Rather than solving for ν for each current value of θ, it may be more efficient to maximize $L(\nu; \mu)$ under the constraints in (4.2). This empirical strategy is similar to that advocated by Su and Judd (2010) for discrete choice models. Modern software implementing the Mathematical Programming with Equilibrium Constraints approach can solve this very efficiently, and a specification test can be constructed from the multiplier of the constraint at the optimum.

This can easily be extended to the case when transfers τ_{ij} are observed. Suppose that the analyst has specified pre-transfer utilities $\alpha_{ij}(\theta)$ and $\gamma_{ij}(\theta)$. Then it follows from Theorem 3 that

$$\tau_{ij} = \alpha_{ij}(\theta) + \frac{\partial G_i^*}{\partial \mu_{ij}} (p_i, \mu_{i\cdot}, \theta) = -\gamma_{ij}(\theta) - \frac{\partial H_j^*}{\partial \mu_{ij}} \left(q_j, \mu_{\cdot j}, \theta\right), \tag{4.3}$$

where the argument θ in G_i^* and H_j^* reflects the possible dependence of the distributions $\mathbf{P}_i$ and $\mathbf{Q}_j$ on θ.

Like (4.2), the equations in (4.3) can simply be added as a constraint in the maximization of the likelihood of the model. Using information on transfers of course makes the estimator more efficient and the specification tests more powerful.

4.2 Identification

As the previous subsection shows, if the model is identified (or overidentified) then estimating and testing it is fairly straightforward. Identification, however, is not a foregone conclusion. Proposition 1 illustrates the underlying difficulty in the basic Choo and Siow

(2006) framework, when transfers are not observed. If the analyst leaves the joint surplus function unrestricted, then the parameters of interest θ consist in the marital surplus matrix π. We can pick ν to be the observed matching μ; then π obtains from equation (4.2),

$$2\pi_{ij} = \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}},$$

and the specification test has no bite: the model is just identified, and any observed matching pattern can be rationalized by choosing the marital surplus matrix π as above.

On marriage markets, even indirect estimates of transfers between partners are hard to come by; therefore identification requires additional restrictions. If the analyst can observe several markets and exclusion restrictions are imposed, this information can be used to generate specification tests. Decker, Lieb, McCann, and Stephens (2010) consider the case when the analyst observes several markets with the same function π, the same distribution of unobserved heterogeneity as in Choo and Siow (2006), and different distributions of men and women p_i and q_j. They derive comparative statics results that imply testable restrictions on the matching patterns across markets.

When the distributions $\mathbf{P}_i$ and $\mathbf{Q}_j$ are incompletely specified, identification can come from imposing exclusion restrictions across several markets, and/or from restrictions on the shape of the function π. Chiappori, Salanié, and Weiss (2010) rely on a mixture of both, within the heteroskedastic logit model of Example 1. The nature of the parametric restrictions that can be brought to bear is of course application-specific. In an earlier version (Galichon and Salanié (2010)) we used a semilinear parameterization of the joint surplus π and we showed that it is a fruitful way to explore such restrictions. We are currently working on an application of this approach.

When transfers are observed, the situation is typically less dire, as further restrictions it can also help in identification. The market for CEOs is a case in point: regulatory rules allow the analyst to observe compensation, and the pre-transfer utilities have a simple interpretation. Let CEOs be indexed by i and firms by j. Then $\alpha_{ij} = -d_{ij}$, with d_{ij} the disutility for CEO i of working in firm j; and $\gamma_{ij} = \rho_{ij}$, the profit of the firm (before it pays the CEO.) To take a simple example of identifying restrictions, if the work disutility of CEOs does not depend on the firm they work in, then in our notation

$$\alpha_{ij} = U_{ij} + \tau_{ij}$$

should not depend on the firm j. This generates additional constraints that can be used to estimate the model; and less blunt restrictions on the drivers of disutility and profit would also help, as illustrated by Dupuy (2010) in a rather different framework.

Finally, consider competitive equilibrium in a market for differentiated products, with 0-1 demand and quasi-linear utilities. Chiappori, McCann, and Nesheim (2008) show how such a hedonic pricing model can be reinterpreted as an assignment game. Introducing unobserved heterogeneity, however, requires some care. Assume that buyer m of observed type i derives utility $\alpha_{ik} + \varepsilon_{ikm}$ from variety k, which seller w of observed type j produces at cost $-\gamma_{jk} + \eta_{jkw}$. Let P_k be the equilibrium price of variety k. Then buyer $m \in i$ chooses to buy the product that maximizes $(\alpha_{ik} - P_k + \varepsilon_{ikm})$, while seller $w \in j$ produces the variety that maximizes $(P_k - \gamma_{jk} + \eta_{jkw})$. The surplus from a match between a buyer m in group i and a seller w in group j is

$$\Phi_{mw} = \max_k (\alpha_{ik} - \gamma_{jk} + \varepsilon_{ikm} + \eta_{jkw}).$$

For the surplus to satisfy assumption 1, we need to impose that either ε_{ikm} or η_{jkw} is further separable. If for instance buyers of type i have the same preferences over varieties and only differ in their valuation for the good:

$$\varepsilon_{ikm} = \zeta_{ik} + \xi_{im},$$

then without further assumptions on η we can rewrite

$$\Phi_{mw} = \max_k (\alpha_{ik} - \gamma_{jk} + \eta_{jkw}) + \xi_{im},$$

which does not involve any interaction between m and w conditional on i and j and hence satisfies Assumption 1. Our other assumptions do not raise any new problem here; Theorems 2, 3 and 4 apply, and so does our discussion of identification. Note that the transfers here are just the prices of the different varieties, which are often available.

References

Becker, G. (1973): "A theory of marriage, part I," *Journal of Political Economy*, 81, 813–846.

Botticini, M., and A. Siow (2008): "Are there Increasing Returns in Marriage Markets?," University of Toronto mimeo.

Chiappori, P., R. McCann, and L. Nesheim (2008): "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," *Economic Theory*, pp. 1–49.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2010): "Partner Choice and the Marital College Premium," mimeo Columbia University.

Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114, 175–201.

Decker, C., E. Lieb, R. McCann, and B. Stephens (2010): "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market," mimeo.

Dupuy, A. (2010): "Sorting on Skills and Preferences: Tinbergen Meets Sattinger," Discussion paper, Maastricht University.

Fox, J. (2010): "Identification in Matching Games," *Quantitative Economics*, 1, 203–254.

Gabaix, X., and A. Landier (2008): "Why Has CEO Pay Increased So Much?," *Quarterly Journal of Economics*, 123, 49–100.

Galichon, A., and B. Salanié (2010): "Matching with Trade-offs: Revealed Preferences over Competing Characteristics," Discussion Paper 7858, CEPR.

Graham, B. (2011): "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. Jackson. Elsevier.

Shapley, L., and M. Shubik (1972): "The assignment game I: the core," *International Journal of Game Theory*, 1, 111–130.

Shimer, R., and L. Smith (2000): "Assortative matching and search," *Econometrica*, 68, 343–369.

Siow, A. (2009): "Testing Becker's Theory of Positive Assortative Matching," Discussion paper, University of Toronto.

Siow, A., and E. Choo (2006): "Estimating a Marriage Matching Model with Spillover Effects," *Demography*, 43(3), 463–490.

Su, C.-L., and K. Judd (2010): "Constrained Optimization Approaches to Estimation of Structural Models," mimeo Chicago Booth.

Appendix A:

Proofs of Theorems 2 and 3

Proof of Theorem 2

By the classical dual formulation of the matching problem, the market equilibrium assigns utilities u_{im} to man $m \in i$ and v_{jw} to woman $j \in w$ so as to solve

$$\mathcal{G} = \min\left(\sum_{i,m} u_{im} + \sum_{j,w} v_{jw}\right)$$

where the minimum is taken under the set of constraints

$$\begin{aligned} u_{im} + v_{jw} &\geq 2\pi_{ij} + \varepsilon_{ijm} + \eta_{ijw} \\ u_{im} &\geq \varepsilon_{i0m} \\ v_{jw} &\geq \eta_{0jw}. \end{aligned}$$

Denote

$$\begin{aligned} U_{ij} &= \min_m \{u_{im} - \varepsilon_{ijm}\},\ i \geq 1,\ j \geq 0 \\ V_{ij} &= \min_w \{v_{jw} - \eta_{ijw}\},\ i \geq 0,\ j \geq 1 \end{aligned}$$

so that

$$\begin{aligned} u_{im} &= \max_{j=0,\ldots,J} \{U_{ij} + \varepsilon_{ijm}\},\ i \geq 1 \\ v_{jw} &= \max_{i=0,\ldots,I} \{V_{ij} + \eta_{ijw}\},\ j \geq 1 \end{aligned}$$

Then

$$\mathcal{G} = \min\left(\sum_{i,m} \max_{j=0,\ldots,J} \{U_{ij} + \varepsilon_{ijm}\} + \sum_{j,w} \max_{i=0,\ldots,I} \{V_{ij} + \eta_{ijw}\}\right)$$

under the set of constraints

$$\begin{aligned} U_{ij} + V_{ij} &\geq 2\pi_{ij} \\ U_{i0} &\geq 0 \\ V_{0j} &\geq 0. \end{aligned}$$

Assign non-negative multipliers $\mu_{ij}, \mu_{i0}, \mu_{0j}$ to these constraints. By duality in Linear Programming, we can rewrite

$$\begin{aligned}\mathcal{G} &= \max_{\mu_{ij}\geq 0}\left(2\sum_{i,j\geq 1}\mu_{ij}\pi_{ij}\right.\\ &\quad -\max_{U_{ij}}\left\{\sum_{\substack{i\geq 1\\ j\geq 0}}\mu_{ij}U_{ij}-\sum_{i,m}\max_{j\geq 0}\{U_{ij}+\varepsilon_{ijm}\}\right\}\\ &\quad \left.-\max_{V_{ij}}\left\{\sum_{\substack{i\geq 0\\ j\geq 1}}\mu_{ij}V_{ij}-\sum_{j,w}\max_{i\geq 0}\{V_{ij}+\eta_{ijw}\}\right\}\right).\end{aligned}$$

Now

$$\sum_{i,m}\max_{j\geq 0}\{U_{ij}+\varepsilon_{ijm}\}=\sum_{i}p_i E_{m\in i}\max_{j\geq 0}\{U_{ij}+\varepsilon_{ijm}\}.$$

In this formula $E_{m\in i}$ denotes the empirical average over the population of men in group i. Now we invoke Assumption 2: if there is a large number of men in each group,

$$E_{m\in i}\max_{j\geq 0}\{U_{ij}+\varepsilon_{ijm}\}\approx E_{\mathbf{P}_i}\left[\max_{j=0,1,\ldots,J}\{w_j+\varepsilon_j\}\right].$$

Adding the similar expression for women, we get

$$\mathcal{G}=\max_{\mu_{ij}\geq 0}\left(2\sum_{i,j\geq 1}\mu_{ij}\pi_{ij}-A(\mu)-B(\mu)\right)$$

where

$$\begin{aligned}A(\mu) &= \max_{U_{ij}}\left\{\sum_{\substack{i\geq 1\\ j\geq 0}}\mu_{ij}U_{ij}-\sum_{i\geq 1}G_i(U_{ij})\right\}\\ B(\mu) &= \max_{V_{ij}}\left\{\sum_{\substack{i\geq 0\\ j\geq 1}}\mu_{ij}V_{ij}-\sum_{j\geq 1}H_j(V_{ij})\right\}\end{aligned}$$

Consider $A(\mu)$ for instance. It is the sum of Legendre-Fenchel transforms of the functions G_i; and as explained in the text, if $\sum_{j=0}^{J}\mu_{ij}\neq p_i$ for any i, then $A(\mu)$ is plus infinity.

Therefore at the maximum in $\mathcal{G}$, the feasibility constraints in (1.1) must hold. and we can rewrite $A(\mu)$ and $B(\mu)$ in terms of the restricted Legendre-Fenchel transforms:

$$A(\mu) = \sum_{i \geq 1} G_i^* (p_i, \mu_{i.}) \text{ and } B(\mu) = \sum_{j \geq 1} H_j^* \left(q_j, \mu_{.j}\right).$$

It follows that

$$\mathcal{G} = \max_{\mu_{ij} \geq 0} \left\{ 2 \sum_{i,j \geq 1} \mu_{ij} \pi_{ij} - \sum_{i=1}^{I} G_i^* (p_i, \mu_{i.}) - \sum_{j=1}^{J} H_j^* \left(q_j, \mu_{.j}\right) \right\}. \qquad \blacksquare$$

Proof of Theorem 3

From the proof of Theorem 2, and given $\mu_{ij} > 0$,

$$U_{ij} = \frac{\partial A}{\partial \mu_{ij}} (\mu);$$

but since

$$A(\mu) = \sum_{i \geq 1} G_i^* (p_i, \mu_{i.}),$$

part a) of Theorem 3 follows immediately.

For part b), note that the social surplus $\mathcal{G}$ is equal to the sum of expected utilities of all types $\sum_{i=1}^{I} p_i u_i + \sum_{j=1}^{J} q_j v_j$ at the optimum; and that the numbers of available men and women of each type p_i and q_j define the feasibility constraints in (1.1). As a consequence,

$$u_i = \frac{\partial \mathcal{G}}{\partial p_i};$$

But given Theorem 2,

$$\frac{\partial \mathcal{G}}{\partial p_i} = -\frac{\partial G_i^*}{\partial p_i} (p_i, \mu_{i.}). \qquad \blacksquare$$

Appendix B: The Generalized Extreme Values Framework

Consider functions $g_i : \mathbb{R}^{J+1} \longrightarrow \mathbb{R}$ and $h_j : \mathbb{R}^{I+1} \longrightarrow \mathbb{R}$ such that the following four conditions hold:

- each g_i or h_j is positive homogeneous of degree one
- they go to $+\infty$ whenever any of their arguments goes to $+\infty$
- their partial derivatives of order k exist outside of 0 and have sign $(-1)^k$
- the functions defined by

$$\begin{aligned} \mathbf{P}_i\left(w_0, ..., w_J\right) &= \exp\left(-g_i\left(e^{-w_0}, ..., e^{-w_J}\right)\right) \\ \mathbf{Q}_j\left(z_0, ..., z_I\right) &= \exp\left(-h_j\left(e^{-z_0}, ..., e^{-z_I}\right)\right) \end{aligned}$$

  are multivariate cumulative distribution functions.

Then introducing utility shocks $\varepsilon_i \sim \mathbf{P}_i$, and $\eta_j \sim \mathbf{Q}_j$, we have by a theorem of McFadden (1978):

$$\begin{aligned} \frac{G_i(w)}{p_i} &= \mathbb{E}_{\mathbf{P}_i}\left[\max_{j=0,1,\ldots,J}\{w_j + \varepsilon_j\}\right] = \log g_i\left(e^w\right) + \gamma \\ \frac{H_j(z)}{q_j} &= \mathbb{E}_{\mathbf{Q}_j}\left[\max_{i=0,1,\ldots,I}\{z_i + \eta_i\}\right] = \log h_j\left(e^z\right) + \gamma \end{aligned}$$

where γ is the Euler constant $\gamma \simeq 0.5772$.

Therefore,

$$G_i^*\left(p_i, a\right) = \left(p_i - \sum_{j=1}^{J} a_j\right) w_0^i(p_i, a) + \sum_{j=1}^{J} a_j w_j^i\left(p_i, a\right) - p_i\left(\log g_i\left(e^{w^i(p_i,a)}\right) + \gamma\right)$$

where for $i = 0, \ldots, I$, the vector $w^i\left(p_i, a\right)$ solves the system

$$\left(p_i - \sum_{j=1}^{J} a_j, a_1, \ldots, a_J\right) = p_i \frac{\partial}{\partial w} \log g_i\left(e^{w^i}\right). \tag{.4}$$

Similarly, if $\sum_{i=0}^{I} b_j = q_j$ then

$$H_j^*(b) = \sum_{i=0}^{I} b_i z_i^j(q_j, b) - q_j \left(\log h_j \left(e^{z^j(q_j, b)} \right) + \gamma \right) \tag{.5}$$

where the vectors $z^j(q_j, b)$ solve the systems

$$\left(q_j - \sum_{i=1}^{I} b_j, a_1, \ldots, b_I \right) = q_j \frac{\partial}{\partial z} \log h_j \left(e^{z^j} \right).$$

Hence,

$$\begin{aligned} \mathcal{E}(p, q, \mu) &= \sum_{i=1}^{I} \left(p_i \log g_i \left(e^{w^i(p_i, \mu_{i\cdot})} \right) - \sum_{j=0}^{J} \mu_{ij} w_j^i(p_i, \mu_{i\cdot}) \right) \\ &+ \sum_{j=1}^{J} \left(q_j \log h_j \left(e^{z^j(q_j, \mu_{\cdot j})} \right) - \sum_{i=0}^{I} \mu_{ij} z_i^j(q_j, \mu_{\cdot j}) \right) + C \end{aligned}$$

where $C = \gamma \left(\sum_{i=1}^{I} p_i + \sum_{j=1}^{J} q_j \right)$, and for $i, j \geq 1$

$$\frac{\partial \mathcal{E}}{\partial \mu_{ij}}(p, q, \mu) = -w_j^i(p_i, \mu_{i\cdot}) - z_i^j(q_j, \mu_{\cdot j}).$$