Matching In Closed Forms

Ulysse Lawogni
THEMA - CY Cergy Paris University

November 5, 2021

Abstract

In this paper, we consider an equally weighted bipartite matching market with transferability of the utility. We analyze the equilibrium of the market in the case of nearly full matching and in the case of nearly scarce matching. Considering a particular case of full matching with a quadratic specification and Gaussian distributions of the characteristics, we find entirely the analytical expression the optimal matching and of the individual surpluses gained by the partners from the sharing of the joint surplus at the equilibrium. We provide two ways to estimate the model.

1 Introduction

There exists a rich literature on matching models. Gale and Shapley (1962) defined theoretically an equilibrium for two-sided matching markets. The mathematical exploitation of this definition of equilibrium will be one the main later researches in this field. Shapely and Shubik (1971) proved an important result that is the maximization of the total social welfare by the optimal matching. But the resolution of their equation requires too strong assumptions. The first econometric framework to estimate matching models has been proposed by Choo and Siow (2006) with transferability of the utility. They considered a discrete logit framework to model the gain from marriage. They have been able to identify individual gains from marriage and the matching matrix in function of the primitives of the model. They obtain closed forms that allow easily a nonparametric estimation of the model. Chiappori, Salanié and Weiss (2008) used an heteroskedastic version of the Choo and Siow model to study the assortative matching on the marriage market. Lindenlaub (2017), considering multidimensional setting, studied the equilibrium properties and particularly on a matching problem with quadratic surplus and Gaussian distributions of the populations, and such that there is no unobserved heterogeneity and bivariate observed characteristics. Galichon and Salanié (2012) give a general discrete framework with
transferable utility and with equilibrium in the sense of Gale and Shapley. From their framework we can essentially derive as particular cases, the results of Choo and Siow and the heteroskedastic model of Chiappori, Salanié and Weiss. One of the main results of their study is to prove that the optimal matching maximizes the total social welfare and this social welfare can be decomposed as the sum of the averaged social surplus and an entropy which quantifies somehow the statistical disorder in the population. Dupuy and Galichon (2014) have proposed a continuous extension of the econometric matching model of Choo and Siow (2006). They found that the individual surpluses can be identified up to two existing, unique but undetermined functions; the individual surplus of a man in a union with a woman is identified up to a function \( a(.) \) that depends on the characteristics of the man, and analogously, the individual surplus of woman in a union with a man is identified up to a function \( b(.) \) that depends on the characteristics of the woman. Bojilov and Galichon (2016) with the setting of Dupuy and Galichon show that the assignment problem can be solved analytically in quadratic specification with Gaussian distributions of the observable characteristics, i.e., they derive analytically the optimal matching function by identifying the affinity matrix, and they also derive the conditional distributions \( X|Y \) and \( Y|X \) where \( X \) and \( Y \) are respectively the characteristics of men and of women. But what they do not do it to provide analytical expression of the individual surpluses and the functions \( a(.) \) and \( b(.) \) are then not precised analytically. The interest to have analytical expression of the individual is the possibility to compare the surplus gap for instance within the unions. Actually the surplus gap between partners is a separable function in the characteristics of the partners and it involves directly the functions \( a(x) \) and \( b(y) \). Finding analytically these functions implies automatically the analytical expressions of the individual surplus and reversely. Our main goal in this work is to solve entirely analytically the matching problem.

We adopt the Dupuy and Galichon continuous framework, and so we consider a bipartite matching market with transferable utility, possibly infinite and where the two sides have the same weight and in addition, we assume that the deterministic part of the utility of single individuals is a constant parameter. We rewrite the equilibrium and this rewriting has the advantage to allow directly the generalization of the model to the case \( \alpha \in \{0, 1\} \) where \( \alpha \) is the probability of matching on each of the two sides of the market. We show briefly that the framework we suggest here finds analytically the equilibrium of the market when the matching tends to be scarce. We then go on the fundamental purpose of this work. We consider in addition to the framework, as Bojilov and Galichon (2016), a quadratic specification of the joint surplus, with Gaussian distributions of the observable characteristics. We provide analytical expressions for the matching function and we retrieve the theorem of Bojilov and Galichon (2016) identifying the affinity matrix. Furthermore we find entirely the analytical expressions of individual surpluses. We finally present two ways to estimate parametrically the
model, by maximum likelihood eased by the fact that it is completely expressed analytically, and the estimation by moments based on the identification of the affinity matrix.

2 The model

2.1 Matching Market Description

We assume the market to be bipartite, one-to-one with transferable utility. One part of the market is \( \mathcal{H} \), and the second one is \( \mathcal{F} \). The whole population will be denoted by \( \mathcal{P} \). We will use the generic terms \( m \) and \( w \) respectively for individuals in \( \mathcal{H} \) and for individuals in \( \mathcal{F} \). Without losing in generality, the set \( \mathcal{H} \) will be considered as the set of men and the set \( \mathcal{F} \) the set of women. This framework can perfectly also be applied to any matching market such as employment market, Real Estate market, etc... The market we consider here is possibly infinite and its two sides are assumed to be in equal proportions.

\[
P(\mathcal{H}) = P(\mathcal{F}) = \frac{1}{2}
\]

Individuals of each side decide either to match with another individual of the opposite side or to remain single. The men are characterized by a continuous vector \( X \) and observable on the market. Analogously, women are characterized by a continuous vector \( Y \) observable on the market too. The support spaces of \( X \) and \( Y \) are respectively \( \mathbb{R}^p \) and \( \mathbb{R}^q \). We denote respectively by \( P \) and \( Q \) the probability distributions of \( X \) and \( Y \). These probability distributions are assumed to have respectively a probability density \( f \) and a probability density \( g \). These functions are assumed to be exogenous to the model. On this market, each individual aims to match with another individual of the opposite side. He or she remains unmatched in the case he or her does not find the "ideal" partner. We will bring more precision about the mechanism of matching on the market. To go further, we need to define the term of matching.

**Definition 1** (Matching). A matching is any density function \( \pi \) associated to \( P \) and \( Q \) and defined on the Cartesian space \( X \times Y \).

A matching describes precisely the manner any particular union of attributes \((x, y)\) is formed relatively to the set of couples. For more precision, it is important to know that all the density functions on \( X \times Y \) are not realistic. Before refining the definition of matching, we need to adopt some useful notations. We define the variable \( M \) that is the indication of the matching status. As we can guess it, this variable is endogenous because it is completely determined by the market.

\[
\forall k \in \mathcal{P}, \ M_k = 1\{ k \text{ is matched } \}
\]
We also define the variables \( \tilde{X} \) and \( \tilde{Y} \) respectively observable on the set of women \( F \) and on the set of men \( H \). For any particular man \( m \), the variable \( \tilde{Y}_m \) represents the attributes of his partner. And for any particular woman \( w \), \( \tilde{X}_w \) represents the attributes of her partner. In the case the partner does not exist, the attributes are \( \emptyset \), i.e the individual is unmatched. So \( \tilde{X} \) and \( \tilde{Y} \) have respectively as support spaces \( X \cup \emptyset \) and \( Y \cup \emptyset \). We can then define for any man \( m \), the variable \( \tilde{Y}_m \) and for any woman \( w \), \( \tilde{X}_w \) respectively as the attributes of his partner and of her partner. In the case the partner does not exist, the attributes are \( \emptyset \), i.e the individual is unmatched. So \( \tilde{X} \) and \( \tilde{Y} \) have respectively as support spaces \( X \cup \emptyset \) and \( Y \cup \emptyset \). We can then define for any man \( m \), the variable \( (X_m, \tilde{Y}_m) \) and for any woman \( w \) the variable \( (\tilde{X}_w, Y_w) \). We will denote by \( \alpha \) the probability to choose at the equilibrium a matched man in the set of the men \( H \). This probability is also equal to the probability to choose at the equilibrium a matched woman in the set of women \( F \).

\[
\alpha = \Pr(M = 1|H) = \Pr(M = 1|F)
\]

Let \( f_1 \) and \( f_0 \) be respectively the conditional probability density of \( X \) with respect to the event \( \{M = 1, H\} \) and to the event \( \{M = 0, H\} \). Analogously, we define respectively \( g_1 \) and \( g_0 \) the conditional probability density of \( Y \) with respect to the event \( \{M = 1, F\} \) and to the event \( \{M = 0, F\} \).

\[
f_1(x) = f(x|M = 1, H) \quad \text{and} \quad f_0(x) = f(x|M = 0, H)
\]

and

\[
g_1(y) = g(y|M = 1, F) \quad \text{and} \quad g_0(y) = g(y|M = 0, F)
\]

The density functions \( f \) and \( g \) are exogenous. Since \( f_1(x) \) and \( f_0(x) \) are conditional probability density functions, we have:

\[
f(x) = \alpha \cdot f_1(x) + (1 - \alpha) \cdot f_0(x).
\]

Similarly, we have:

\[
g(y) = \beta \cdot g_1(y) + (1 - \beta) \cdot g_0(y)
\]

The model will aim to identify \( \alpha, f_0, f_1, g_0 \) and \( g_1 \). In some particular cases, the functions \( f_1 \) and \( g_1 \) can be exogenous; for instance when the matching is assumed to be full i.e all the individuals are matched, then the conditional densities to marriage \( f_1 \) and \( g_1 \) are exactly equal respectively to \( f \) and \( g \). We will treat later that particular case.

All the matchings are not feasible. We refine the definition as follows.

**Definition 2 (Feasible Matching).** A feasible matching is a joint density function \( \pi(x, y) \) defined on \( X \times Y \) and a probability \( \alpha \).

Note that \( \pi(x, y) \) is the conditional joint density of \((X, \tilde{Y})\) with respect to \((M = 1, H)\) and equivalently, it is also the conditional joint density of \((\tilde{X}, Y)\) with respect to \((M = 1, F)\). We have:

\[
\pi(x, y) = \pi(x, y|M = 1, H) = \pi(x, y|M = 1, F)
\]

Let’s denote by \( \Pi(P, Q) \) the set of feasible matchings. In the following section, we will introduce the notion of matching surplus and we will add heterogeneity to the model.
2.2 Matching Surplus and Heterogeneity

The population is infinite. We make the separability assumption (previously suggested by Choo and Siow (2006) and formalized by Galichon and Salanié (2012) and Chiappori, Salanié and Weiss (2017)): that is the joint utility generated when a man $m$ with characteristics $x_m$ matches with a woman $w$ with characteristics $y_w$ does not depend on interactions between their unobserved characteristics, conditional on $(x_m, y_w)$.

**Assumption 1 (Separability).** The joint utility of a couple $(m, w)$ with characteristics $(x, y)$ is written:

$$S(x, y) + \sigma \varepsilon_m(y) + \sigma \eta_w(x)$$

where $S(x, y)$ is the deterministic part of the joint utility, $\sigma$ is a positive constant, $\varepsilon_m(y)$ and $\eta_w(x)$ are two stochastic terms.

The function $S(x, y)$ can be interpreted as the joint systematic surplus generated by the matching between a man of characteristics $x$ and a woman of characteristics $y$. The term $\varepsilon_m(y)$ is a stochastic process that represents for the man $m$ his utility from sympathy shock with a woman with attributes $y$. The term $\eta_w(x)$ is a stochastic process that represents for the woman $w$ her utility from sympathy shock with a man with attributes $x$, and $\sigma$ is a positive parameter. The joint utility from marriage represents the utility that partners will share between them, according to a rule that depends on the competition force in the marriage market.

The individuals compare their utility from marriage with what they would have if they remain single when they are single. We model in the following assumption, the utility of single individuals.

**Assumption 2.** The utility of a single man $m$ is:

$$u^0_m = A^0 + \sigma \varepsilon^0_m$$

and the utility of a single woman $w$ is:

$$v^0_w = B^0 + \sigma \eta^0_w$$

where $A^0$ is a general parameter common to all single men and $B^0$ a general parameter common to all single women; $\sigma$ is a positive constant, $\varepsilon^0_m$ is a stochastic term specific to man $m$ and $\eta^0_w$ also a stochastic term specific to woman $w$.

Remark that this assumption is equivalent to the general case in which the deterministic parts of the utility of the single are functions $A^0(x)$ and $B^0(y)$. In that case, it suffices to redefine a new joint surplus as $S(x, y) = A^0(x) - B^0(y) + E(A^0(X)) + E(B^0(Y))$ and new utilities of singles $u^0 = E(A^0(X)) + \sigma \varepsilon^0$ for single men and $v^0 = E(B^0(X)) + \sigma \eta^0$ for single women and this redefinition makes us come back to verify Assumption 2.
The following specification is in the light of Dupuy and Galichon (2014). We describe it in the following lines. Each man $m$ of attributes $x$ meets a subset, possibly infinite, of the set of women to make his choice. We assume $\varepsilon^0_m$ to have a standard Gumbel distribution. We index his acquaintances by $N^*$. From a particular acquaintance $k$ with a woman of attributes $y^m_k$, the man gets the utility:

$$u^k_m = U(x, y^m_k) + \sigma \varepsilon^m_k$$

We assume the process $\{(y^m_k, \varepsilon^m_k) : k \in N^*\}$ to be a Poisson process on the space $Y \times \mathbb{R}$ with intensity $e^{-\varepsilon - \gamma}$, where $\gamma$ is the Euler-Mascheroni constant.

Analogously, for each woman $w$ of attributes $y$, we assume $\eta^0_w$ to have have standard Gumbel distribution; the woman $w$ meets a possibly infinite subset of the men to make her choice. Her acquaintances are indexed by $N^*$. From a particular acquaintance $k$ with a man of attributes $x^w_k$, she gets the utility:

$$v^k_w = V(x^w_k, y) + \sigma \eta^w_k$$

The process $\{(x^w_k, \eta^w_k) : k \in N^*\}$ is assumed to be a Poisson process on the space $X \times \mathbb{R}$ with intensity $e^{-\eta - \gamma}$.

We assume that for any man $m$ of attributes $x_m = x$, the term $\varepsilon^0_m$ is independent from any sympathy shock $\varepsilon^m_k$ from an acquaintance $k$. And analogously for women, the term $\eta^0_w$ is independent from any sympathy shock $\eta^w_k$ from an acquaintance $k$.

We summarize that in the following assumption.

**Assumption 3.** (Dupuy and Galichon, 2014)

(i) The stochastic processes $\varepsilon^m(y)$ and $\eta^w(x)$ are max-stable process of the form:

$$\varepsilon^m(y) = \max_k (\varepsilon^m_k : y_k = y) \text{ if the set } \{k : y_k = y\} \text{ is non-empty}$$

$$= -\infty \text{ otherwise,}$$

where $\{(y^m_k, \varepsilon^m_k) : k \in K\}$ follows a Poisson process on the space $Y \times \mathbb{R}$ of intensity $e^{-\varepsilon - \gamma} dy dz$, and

$$\eta^w(x) = (\eta^w_l : x_l = x) \text{ if the set } \{l : x_l = x\} \text{ is non-empty}$$

$$= -\infty \text{ otherwise.}$$

where $\{(x^w_k, \eta^w_k) : l \in L\}$ follows a Poisson process on $X \times \mathbb{R}$ with intensity $e^{-\eta - \gamma} dx d\eta$.

(ii) The stochastic terms $\varepsilon^0_m$ and $\eta^0_w$ follow a standard Gumbel distribution, independent of $\varepsilon^m_k$ for $k \in K$ and independent of $\eta^w_l$ for $l \in L$, respectively.
We will denote respectively by $\bar{U}(x, y)$ and $\bar{V}(x, y)$ the net individual surpluses of men of attributes $x$ and women of attributes $y$ from their matching and we denote by $\bar{S}(x, y)$ the net joint surplus.

\[
\bar{U}(x, y) = U(x, y) - A^0, \quad \bar{V}(x, y) = V(x, y) - B^0 \quad \text{and} \quad \bar{S}(x, y) = S(x, y) - A^0 - B^0.
\]

### 2.3 Equilibrium and Identification

We suppose that the market equilibrium is stable in the sense of Gale and Shapley (1962), that is, there is no married person who would rather be single, and there is no pair of (married or unmarried) persons who prefer to form a new union. The equilibrium is the solution of the following maximization problems. Each man must choose in the set of his acquaintances the partner that brings him the optimal utility; in the case this utility is higher than his proper utility he chooses her. And reversely, each woman must choose in the set of her acquaintances the partner that brings her the optimal utility; she chooses this man in the case this utility she gets from the union is higher than her proper utility. A marriage between a man and a woman occurs when they choose each other. Let's denote respectively by $u_m(x)$ and $v_w(y)$ the utility at equilibrium of a man $m$ of attributes $x$ and the utility at equilibrium of a woman $w$ of attributes $y$. We also denote:

\[
\bar{u}_m = u_m - A^0 \quad \text{and} \quad \bar{v}_w = v_w - B^0.
\]

A man $m$ of attributes $x$ solves:

\[
\bar{u}_m(x) = \max \{\sigma_{z_m}^0, \max_{k \in \mathbb{N}} \{\bar{U}(x, y_k) + \sigma_{z_m}^k\}\} \tag{3}
\]

And a woman $w$ of attributes $y$ solves:

\[
\bar{v}_w(y) = \max \{\sigma_{z_w}^0, \max_{k \in \mathbb{N}} \{\bar{V}(x_k, y) + \sigma_{z_w}^k\}\} \tag{4}
\]

At the equilibrium, each individual does his choice in the set of his or her acquaintances. A match occurs between a man $m$ and a woman $w$ if the man is an acquaintance of the man $m$ and $w$ brings to him the maximum utility in his set of acquaintances and this utility is higher to his utility of remaining single. And reversely, man $m$ is an acquaintance of the woman $w$ and brings to her the maximum of utility in the set of her acquaintances and this utility is higher to her utility of being single.

To simplify notation, we write

\[
G_x(\bar{U}) = \mathbb{E}_P(\bar{u} | x)
\]

and

\[
H_y(\bar{V}) = \mathbb{E}_Q(\bar{v} | y)
\]
the net expected value of the utility of man with attributes $x$ and the net expected value of the utility of woman with attributes $y$, respectively. This notation emphasizes the fact that the net expected values depend on the net utility functions $\bar{U}(x, y)$ and $\bar{V}(x, y)$.

The problems (3) and (4) can be seen as the primal problems. The corresponding dual problems, which will be used hereafter, are defined as:

$$\max_{\bar{U}(x, \cdot)} \left( \int_y \frac{\alpha \pi(x, y)}{f(x)} \cdot \bar{U}(x, y) \cdot dy - G_x(\bar{U}) \right)$$  \hspace{1cm} (5)

where $\alpha \pi(x, y)/f(x)$ is density of men of attributes $x$ married to a woman of attributes $y$ in the whole population of men, and

$$\max_{\bar{V}(\cdot, y)} \left( \int_x \frac{\alpha \pi(x, y)}{g(y)} \cdot \bar{V}(x, y) \cdot dx - H_y(\bar{V}) \right)$$  \hspace{1cm} (6)

where $\alpha \pi(x, y)/g(y)$ is the density of women of attributes $y$ married to a man of attributes $x$ in the whole population of women. As pointed out by Galichon and Salanié (2012) and Dupuy and Galichon (2014) from the first order conditions of the optimization problems (5) and (6), we have:

$$\frac{\partial G_x(\bar{U})}{\partial \bar{U}}(x, y) = \frac{\alpha \pi(x, y)}{f(x)}$$  \hspace{1cm} (7)

and

$$\frac{\partial H_y(\bar{V})}{\partial \bar{V}}(x, y) = \frac{\alpha \pi(x, y)}{g(y)}$$  \hspace{1cm} (8)

from the envelop theorem, where $\alpha \pi(x, y)/f(x)$ is the probability for man of attributes $x$ of being married to a woman of attributes $y$ and $\alpha \pi(x, y)/g(y)$ is the probability for woman of attributes $y$ of being married to a man of attributes $x$.

**Proposition 1.** Assume a stable matching and Assumption 3. Then, for any man $m$ of attributes $x$, whose utility at the equilibrium is $u_m(x)$, and for any woman $w$ of attributes $y$ whose utility at the equilibrium is $v_w(y)$ we have:

$$u_m(x) \sim \text{Gumbel} \left( U^0(x) + \sigma \ln \left( 1 + \int_y e^{\bar{U}(x, y) / \sigma} \cdot dy \right), \sigma \right)$$

and

$$v_w(y) \sim \text{Gumbel} \left( V^0(y) + \sigma \ln \left( 1 + \int_x e^{\bar{V}(x, y) / \sigma} \cdot dx \right), \sigma \right)$$
Proof. We consider the case of the distribution of men’ utility and start from the preceding result:

\[
u_m(x_m) = \max \left( \max_{w \in W} (U(x_m, y_w) + \sigma \varepsilon_m(y_w)), U^0(x_m) + \sigma \varepsilon_m^0 \right).
\]

Using assumption S, the problem can be written as:

\[
u_m(x_m) = \max \left( \max_{k \in K} (U(x_m, y_{m_k}) + \sigma \varepsilon_k^m), U^0(x_m) + \sigma \varepsilon_m^0 \right)
\]

where \((y_{m_k}, \varepsilon_k^m)\) is following a Poisson process of intensity \(e^{-\varepsilon} d\varepsilon dy\). If the cumulative distribution of \(u_m(x)\) is denoted by \(F_{u_m}\), then:

\[F_{u}(t) = \Pr \left( \varepsilon_m^0 \leq \frac{t - U^0(x)}{\sigma} \right) \cdot \Pr \left( \varepsilon_k^m \leq \frac{t - U(x_m, y_{m_k})}{\sigma}, \forall k \right)
\]

because \(\varepsilon_m^0\) and \(\varepsilon_k^m\) are independent. Since \(\varepsilon_m^0\) is following a standard Gumbel distribution and \((\varepsilon_k^m, y_k)\) is following a Poisson distribution of intensity \(e^{-\varepsilon} d\varepsilon dy\), we have:

\[F_{u}(t) = \exp \left( - \exp \left( - \frac{t - U^0(x)}{\sigma} \right) \right) \times \exp \left( - \int_Y \int_{\frac{t - U(x_m, y)}{\sigma}}^{+\infty} \exp (-\varepsilon) d\varepsilon dy \right)
\]

In this expression, the second term of the right-hand-side represents the probability of not observing \((y_{m_k}, \varepsilon_k^m)\) in the set \(Y \times [(t - U(x_m, y))/\sigma, +\infty]\). Integrating the second term of this expression with respect ot \(\varepsilon\) gives:

\[F_{u}(t) = \exp \left( - \exp \left( - \frac{t - U^0(x_m)}{\sigma} \right) \right) \times \exp \left( - \int_Y \exp \left( \frac{t - U(x_m, y)}{\sigma} \right) dy \right)
\]

Simplifying gives:

\[F_{u}(t) = \exp \left( - \exp \left( - \frac{t - U^0(x_m)}{\sigma} \right) \right) \times \exp \left( - \int_Y \exp \left( \frac{U(x_m, y) - U^0(x_m)}{\sigma} \right) dy \right)
\]

That is, \(u_m\) has a Gumbel distribution of parameters

\[\left\{ U^0(x_m) + \sigma \log \left( 1 + \int_Y \exp \left( \frac{U(x_m, y) - U^0(x_m)}{\sigma} \right) dy \right) \right\}, \sigma \right\}
\]

We do the same demonstration for \(v_w(y)\) to prove its distribution. \(\square\)
From well-known results regarding the Gumbel distribution, we directly get:

\[ G_x(\bar{U}) = \sigma \ln \left( 1 + \int_y \exp \left( \frac{\bar{U}(x,y)}{\sigma} \right) dy \right) + \sigma \gamma \tag{9} \]

and

\[ H_y(\bar{V}) = \sigma \ln \left( 1 + \int_x \exp \left( \frac{\bar{V}(x,y)}{\sigma} \right) dx \right) + \sigma \gamma, \tag{10} \]

where \( \gamma = \ln(\ln(2)) \) is the Euler-Mascheroni constant.

From Proposition (1), we can also derive an expression for the systematic utility of each individual. If we apply the identity (7) with the specification for men’s conditional utility (9), we obtain:

\[ \frac{\exp \left( \frac{\bar{U}(x,y)}{\sigma} \right)}{1 + \int_y \exp \left( \frac{\bar{U}(x,y)}{\sigma} \right) dy} = \frac{\alpha \pi(x,y)}{f(x)} \tag{11} \]

Similarly, if we apply the identity (8) with the specification for women’ conditional utility (10), we obtain:

\[ \frac{\exp \left( \frac{\bar{V}(x,y)}{\sigma} \right)}{1 + \int_x \exp \left( \frac{\bar{V}(x,y)}{\sigma} \right) dx} = \frac{\alpha \pi(x,y)}{g(y)} \tag{12} \]

Together with identity (1), the right-hand-side of expression (11) can be written as:

\[ \frac{\alpha \pi(x,y)}{1 - \alpha} \frac{1}{f_0(x)} \left( 1 + \frac{\alpha}{1 - \alpha} \frac{f_1(x)}{f_0(x)} \right)^{-1}. \]

In addition, the feasibility of matchings implies that

\[ f_1(x) = \int_y \pi(x,y)dy. \]

Thus, we have

\[ \frac{\exp \left( \frac{\bar{U}(x,y)}{\sigma} \right)}{1 + \int_y \exp \left( \frac{\bar{U}(x,y)}{\sigma} \right) dy} = \frac{\alpha \pi(x,y)}{1 - \alpha} \frac{1}{f_0(x)} \left( 1 + \int_y \frac{\alpha \pi(x,y)}{1 - \alpha} \frac{dy}{f_0(x)} \right)^{-1}, \]

which implies that:

\[ \exp \left( \frac{\bar{U}(x,y)}{\sigma} \right) = \frac{\alpha \pi(x,y)}{(1 - \alpha) f_0(x)}. \tag{13} \]
In other words, the systematic utility obtained by a man of attributes \(x\) married to a woman of attributes \(y\) is equal to the ratio of the joint density of observing matches of man of attributes \(x\) and woman of attributes \(y\) and the density of observing a single of attributes \(x\). Similarly, we have:

\[
\exp\left(\frac{\bar{V}(x, y)}{\sigma}\right) = \frac{\alpha}{1 - \alpha} \frac{\pi(x, y)}{g_0(y)}. \tag{14}
\]

In other terms, we have:

\[
U(x, y) = A^0 + \sigma \ln \left(\frac{\alpha}{1 - \alpha} \frac{\pi(x, y)}{f_0(x)}\right) \tag{15}
\]

and

\[
V(x, y) = B^0 + \sigma \ln \left(\frac{\alpha}{1 - \alpha} \frac{\pi(x, y)}{g_0(y)}\right) \tag{16}
\]

These expressions show how the individual surpluses are linked to the joint matching density function. From these expressions of \(U(x, y)\) and \(V(x, y)\), we have, by rewriting:

\[
\frac{1 - \alpha}{\alpha} f_0(x) \exp \left(\frac{U(x, y)}{\sigma}\right) = \exp \left(\frac{A^0}{\sigma}\right) \pi(x, y)
\]

and

\[
\frac{1 - \alpha}{\alpha} g_0(y) \exp \left(\frac{V(x, y)}{\sigma}\right) = \exp \left(\frac{B^0}{\sigma}\right) \pi(x, y)
\]

The function \(\pi\) is a density function on \(\mathcal{X} \times \mathcal{Y}\) so:

\[
\int_{\mathcal{X}} \int_{\mathcal{Y}} \pi(x, y) dx \, dy = 1
\]

This leads to the following lines:

\[
\exp \left(\frac{A^0}{\sigma}\right) = \frac{1 - \alpha}{\alpha} \int_{\mathcal{X}} \int_{\mathcal{Y}} f_0(x) \exp \left(\frac{U(x, y)}{\sigma}\right) dx \, dy
\]

and

\[
\exp \left(\frac{B^0}{\sigma}\right) = \frac{1 - \alpha}{\alpha} \int_{\mathcal{X}} \int_{\mathcal{Y}} g_0(y) \exp \left(\frac{V(x, y)}{\sigma}\right) dx \, dy
\]

In other terms, we can write:

\[
A^0 = \sigma \ln \left(\frac{1 - \alpha}{\alpha}\right) + \sigma \ln \left(\int_{\mathcal{X}} \int_{\mathcal{Y}} f_0(x) \exp \left(\frac{U(x, y)}{\sigma}\right) dx \, dy\right)
\]
and
\[ B^0 = \sigma \ln \left( \frac{1 - \alpha}{\alpha} \right) + \sigma \ln \left( \int_{x} \int_{y} g_0(y) \exp \left( \frac{V(x, y)}{\sigma} \right) d x \, d y \right) \]

We denote:
\[ \delta_1 := \ln \left( \int_{x} \int_{y} f_0(x) \exp \left( \frac{U(x, y)}{\sigma} \right) d x \, d y \right) \tag{17} \]
and
\[ \delta_2 := \ln \left( \int_{x} \int_{y} g_0(y) \exp \left( \frac{V(x, y)}{\sigma} \right) d x \, d y \right). \tag{18} \]

The parameters \( A^0 \) and \( B^0 \) can then be written:
\[ A^0 = \sigma \ln \left( \frac{1 - \alpha}{\alpha} \right) + \sigma \delta_1 \tag{19} \]
and
\[ B^0 = \sigma \ln \left( \frac{1 - \alpha}{\alpha} \right) + \sigma \delta_2 \tag{20} \]

This is not an identification of \( A^0 \) and \( B^0 \). Actually the equations (17) and (19) are equivalent and the equations (18) and (20) are equivalent as well. To prove that, using the relations \( U = \bar{U} + A^0 \) and \( V = \bar{V} + B^0 \), we can rewrite \( \delta_1 \) and \( \delta_2 \) at the forms:
\[ \delta_1 = \ln \left( \int_{x} \int_{y} f_0(x) \exp \left( \frac{U(x, y) + A^0}{\sigma} \right) d x \, d y \right) = \frac{A^0}{\sigma} + \ln \left( \int_{x} \int_{y} f_0(x) \exp \left( \frac{U(x, y)}{\sigma} \right) d x \, d y \right) \]
and
\[ \delta_2 = \ln \left( \int_{x} \int_{y} g_0(y) \exp \left( \frac{V(x, y) + B^0}{\sigma} \right) d x \, d y \right) = \frac{B^0}{\sigma} + \ln \left( \int_{x} \int_{y} g_0(y) \exp \left( \frac{V(x, y)}{\sigma} \right) d x \, d y \right) \]

Combining (1), (13) and the feasibility constraint \( f_1(x) = \int_{y} \pi(x, y) d y \), we show that:
\[ \alpha = \int_{x} \int_{y} \frac{e^{\bar{U}(x, y)}}{1 + \int_{y} e^{\bar{U}(x, y)} d y} f(x) d x \, d y \tag{21} \]
and
\[ f_0(x) = \frac{1}{(1 - \alpha) \left( 1 + \int_{y} e^{\bar{U}(x, y)} d y \right)} f(x) \tag{22} \]
Analogously, exploiting (2), (14) and the feasibility constraint $g_1(y) = \int_X \pi(x, y) dx$, we show also that:

$$\alpha = \int_X \int_Y \frac{\bar{V}(x, y)}{\sigma} \frac{e^{-\frac{\bar{V}(x, y)}{\sigma}} g(y) dx dy}{1 + \int_X e^{-\frac{\bar{V}(x, y)}{\sigma}} dx}$$

(23)

and

$$g_0(y) = \frac{1}{(1 - \alpha) \left(1 + \int_X e^{-\frac{\bar{V}(x, y)}{\sigma}} dx\right)} g(y)$$

(24)

By replacing (22) in the rewritten expression of $\delta_1$, we find that:

$$\delta_1 = \frac{A^0}{\sigma} - \ln(1 - \alpha) + \ln \left(\int_X \int_Y \frac{e^{-\frac{\bar{U}(x, y)}{\sigma}} f(x)}{1 + \int_Y e^{-\frac{\bar{U}(x, y)}{\sigma}} dy} dx\right)$$

and analogously, by replacing (24) in the rewritten expression of $\delta_2$, we have:

$$\delta_2 = \frac{B^0}{\sigma} - \ln(1 - \alpha) + \ln \left(\int_X \int_Y \frac{e^{-\frac{\bar{V}(x, y)}{\sigma}} g(y)}{1 + \int_Y e^{-\frac{\bar{V}(x, y)}{\sigma}} dy} dx\right)$$

The integration of (21) and (23) respectively in these expressions gives:

$$\delta_1 = \frac{A^0}{\sigma} - \ln \left(\frac{1 - \alpha}{\alpha}\right) \quad \text{and} \quad \delta_2 = \frac{B^0}{\sigma} - \ln \left(\frac{1 - \alpha}{\alpha}\right)$$

And these expressions are respectively equivalent to (19) and (20). What do the equivalence between (19) and (17) and the equivalence between (20) and (18) mean? These equivalences mean that one cannot exploit separately (19) and (17) nor exploit separately (20) and (18) since the exploitation of one is equivalent to the exploitation of the second. Concretely, it means that we know $\delta_1$ from (17), then we can no more use (19) to derive the value of $A^0$ and reversely and same of $\delta_2$ and $B^0$. But we can fix $A^0$ and $B^0$ and then derive $\delta_1$ and $\delta_2$. The value of $\delta_1$ and the value of $\delta_2$ cannot be chosen arbitrarily because they depend actually on $f_0$ and $\bar{U}$ and on $g_0$ and $\bar{V}$ given by the equilibrium of the market.

To ensure that (19) and (20) are not an identification, let’s verify if it is possible to identify $\sigma$. As we know that only the individual net surpluses $\bar{U}$ and $\bar{V}$ are identified, we can, without losing in generality, assume $A^0 = B^0 = 0$. we have then $U = \bar{U}$ and $V = \bar{V}$. As we have shown that:

$$\delta_1 = \frac{A^0}{\sigma} - \ln \left(\frac{1 - \alpha}{\alpha}\right) \quad \text{and} \quad \delta_2 = \frac{B^0}{\sigma} - \ln \left(\frac{1 - \alpha}{\alpha}\right)$$
we deduce automatically that:

\[ \delta_1 = \delta_2 = -\ln \left( \frac{1 - \alpha}{\alpha} \right) \]

Then by replacing these expressions respectively in (19) and (20), we obtain:

\[ 0 = \sigma \times 0 \]

The term \( \sigma \) can then be set to any positive and finite value. In conclusion, when \( A^0 = B^0 = 0 \), we cannot identify \( \sigma \). In other terms, we cannot identify \( \sigma \). Even if (19) and (20) are not an identification, we can nevertheless give an interpretation for these relations. Consider two different stable markets 1 and 2 with the same joint function, the same distributions \( P \) and \( Q \), and the same \( \sigma \) and on which the number of men and the number of women are equal and such that at the equilibrium the conditional probability of marriage \( \alpha_1 \) on the male population in the market 1 is lower than the conditional probability of marriage \( \alpha_2 \) on the male population in the market 2. Single men on the market 1 are expected to be happier than single men on the market 2 i.e \( A_1 \geq A_2 \) where \( A_1 \) is the deterministic part of the utility of single men on the market 1 and \( A_2 \) is the deterministic part of the utility of single men on the market 2. The intuition is that the utility of singles is expected to be relatively higher when marriage is rare and reversely, it is lower when marriage is very frequent. And we can remark that:

\[ \lim_{\alpha \to 1^-} \ln \left( \frac{1 - \alpha}{\alpha} \right) = -\infty \quad \text{and} \quad \lim_{\alpha \to 0^+} \ln \left( \frac{1 - \alpha}{\alpha} \right) = +\infty \]

Before going further, we can make the following assumption without loosing in generality, to simplify the setting.

**Assumption 4.** The utility of a single man \( m \) is:

\[ u^0_m = \Phi^0 + \sigma \varepsilon^0_m \]

and the utility of a single woman \( w \) is:

\[ v^0_w = \Phi^0 + \sigma \eta^0_w, \]

where \( \Phi^0 \) is general parameter common to all single individuals, \( \sigma \) is a positive constant, \( \varepsilon^0_m \) a stochastic term specific to man \( m \) and \( \eta^0_w \) a stochastic term specific to woman \( w \).

This assumption states that:

\[ A^0 = B^0 = \Phi^0 \]
i.e all the single individuals have the same deterministic part denoted $\Phi^0$. It is legitimate to assume that since we have explained above that actually we do not identify $A^0$ and $B^0$. As only the individual net surpluses $\bar{U}$ and $\bar{V}$ are identified, the parameters $A^0$ and $B^0$ are actually simply respectively a reference point for married men and a reference point for married women. We can then assume without loosing in generality the equality between these two reference points $A^0$ and $B^0$. One advantage of this assumption is the equivalence of the comparison between $U$ and $V$ and the comparison between $\bar{U}$ and $\bar{V}$ because we will have then $U - V = \bar{U} - \bar{V}$. By now, we fill consider then single men and single women have the same constant deterministic part of utility that is denoted $\Phi^0$. The reader may wonder why we do not simply set this parameter to 0. Of course, we could, but for technical usefulness, we prefer to keep it as a parameter even if we are not going to identify it. The interesting trick will be to vanish it and that will turn out to be very simplifying.

From the relations (19) and (20) combined to the assumption stated above, we have $\delta_1 = \delta_2$. We will then denote

$$\delta = \delta_1 = \delta_2$$

The relations (19) and (20) take then both the form:

$$\Phi^0 = \sigma \ln \left( \frac{1 - \alpha}{\alpha} \right) + \sigma \delta$$

The relations (15) and (15) become respectively:

$$U(x, y) = \Phi^0 + \sigma \ln \left( \frac{\alpha}{1 - \alpha} \frac{\pi(x, y)}{f_0(x)} \right)$$

and

$$V(x, y) = \Phi^0 + \sigma \ln \left( \frac{\alpha}{1 - \alpha} \frac{\pi(x, y)}{g_0(y)} \right)$$

The combination of (26) with (25) and the combination of (27) with (25) give respectively:

$$U(x, y) = \sigma \delta + \sigma \ln \left( \frac{\pi(x, y)}{f_0(x)} \right)$$

and

$$V(x, y) = \sigma \delta + \sigma \ln \left( \frac{\pi(x, y)}{g_0(y)} \right)$$

From (28) and (29), and using the fact that $U(x, y) + V(x, y) = S(x, y)$, we can state the following proposition.
Proposition 2. Assume a stable matching market. Under Assumption 1, Assumption 3 and Assumption 4, we have:

1. for any $x \in X$, $y \in Y$,

$$S(x,y) = 2\sigma \delta + \sigma \ln \left( \frac{(\pi(x,y))^2}{f_0(x)g_0(y)} \right)$$

2. The systematic surplus of a man of attributes $x$ from a matching with a woman of attributes $y$ is such as:

$$U(x,y) = \frac{1}{2} (S(x,y) - \sigma \ln (f_0(x)) + \sigma \ln (g_0(y)))$$

3. The systematic surplus of a woman of attributes $y$ from a matching with a man of attributes $x$ is such as:

$$V(x,y) = \frac{1}{2} (S(x,y) + \sigma \ln (f_0(x)) - \sigma \ln (g_0(y)))$$

4. for any $x \in X$,

$$f_0(x) = \frac{e^\delta}{(1-\alpha)e^\delta + \alpha \int_Y e^{\frac{U(x,y)}{\sigma}} dy} f(x) \quad \text{and} \quad f_1(x) = \frac{\int_Y e^{\frac{U(x,y)}{\sigma}} dy}{(1-\alpha)e^\delta + \alpha \int_Y e^{\frac{U(x,y)}{\sigma}} dy} f(x)$$

5. for any $y \in Y$,

$$g_0(y) = \frac{e^\delta}{(1-\alpha)e^\delta + \alpha \int_X e^{\frac{V(x,y)}{\sigma}} dx} g(y) \quad \text{and} \quad g_1(y) = \frac{\int_X e^{\frac{V(x,y)}{\sigma}} dx}{(1-\alpha)e^\delta + \alpha \int_X e^{\frac{V(x,y)}{\sigma}} dx} g(y)$$

6. \[ \delta = \ln \left( \int_X \int_Y (f_0(x))^{\frac{1}{2}} (g_0(y))^{\frac{1}{2}} \exp \left( \frac{S(x,y)}{2\sigma} \right) dx dy \right) \]

Proof. 1. The combination of (28) and (29) with the relation $U(x,y) + V(x,y) = S(x,y)$ leads directly to:

$$S(x,y) = 2\sigma \delta + \sigma \ln \left( \frac{(\pi(x,y))^2}{f_0(x)g_0(y)} \right)$$

2. From the relation above, we can deduce:

$$\ln(\pi(x,y)) = \frac{S(x,y) + \sigma \ln(f_0(x)) + \sigma \ln(g_0(y)) - (\delta_1 + \delta_2)\sigma}{2\sigma}$$

We use this expression of $\ln(\pi(x,y))$ in (28) and we deduce the stated expression in the proposition for $U(x,y)$. 

16
3. Analogously, we replace the expression of $\ln(\pi(x,y))$ in (29) and we obtain the expression of $V(x,y)$.

4. From the relation 

$$U(x,y) = \sigma \delta + \sigma \ln \left( \frac{\pi(x,y)}{f_0(x)} \right)$$

we can deduce:

$$e^{\frac{U(x,y)}{\sigma}} f_0(x) = e^\delta \pi(x,y)$$

The feasibility of the optimal matching implies that:

$$\int_Y \pi(x,y) dy = f_1(x)$$

So, we have:

$$f_0(x) \int_Y e^{\frac{U(x,y)}{\sigma}} dy = e^\delta f_1(x)$$

We then use the relation:

$$f(x) = \alpha f_1(x) + (1 - \alpha) f_0(x)$$

and we get the equation for $f_0(x)$:

$$f_0(x) \int_Y e^{\frac{U(x,y)}{\sigma}} dy = e^\delta \left( \frac{f(x) - (1 - \alpha) f_0(x)}{\alpha} \right)$$

and the solution is given by:

$$f_0(x) = \frac{e^\delta}{\left(1 - \alpha\right)e^\delta + \alpha \int_Y e^{\frac{U(x,y)}{\sigma}} dy} f(x)$$

We deduce the expression of $f_1(x)$ by replacing the found expression of $f_0(x)$ in:

$$f_1(x) = \frac{f(x) - (1 - \alpha) f_0(x)}{\alpha}$$

5. We do analogous reasoning for $g_0(y)$ and $g_1(y)$ as above.

6. We use the relation:

$$\ln(\pi(x,y)) = \frac{S(x,y) + \sigma \ln(f_0(x)) + \sigma \ln(g_0(y)) - 2\sigma \delta}{2\sigma}$$

i.e

$$\pi(x,y)e^\delta = e^{\frac{S(x,y)}{2\sigma}} (f_0(x))^{\frac{1}{2}} (g_0(y))^{\frac{1}{2}}$$

We then use the relation:

$$\int_X \int_Y \pi(x,y) dx \ dy = 1$$

and we deduce directly $\delta$. 

\[\square\]
Remark: We treat in Appendix, the case in which the deterministic part of the utility of single men $A_0$ and the deterministic part of the utility of single women $B_0$ are still constant but may be different. That is traduced by Assumption 2. In the following development, we will make a quick remark about the case of nearly scarce matching i.e $\alpha = 0^+$ and then we will entirely and particularly focus on the interesting case of nearly full matching i.e $\alpha = 1^-$. Remark that the probability of matching $\alpha$ at the equilibrium is:

$$\alpha = \int_x \int_y \frac{e^{\theta(x,y)}}{1 + \int_y e^{\theta(x,y)}} f(x) \, dx \, dy = \int_x \int_y \frac{e^{\triangledown(x,y)}}{1 + \int_x e^{\triangledown(x,y)}} g(y) \, dx \, dy$$

That implies $0 < \alpha < 1$. This is why we will prefer the terminology of nearly scarce matching ($\alpha = 0^+$) and of nearly full matching ($\alpha = 1^-$). And as we have:

$$f(x) = \alpha f_1(x) + (1 - \alpha) f_0(x) \quad \text{and} \quad g(y) = \alpha g_1(y) + (1 - \alpha) g_0(y)$$

where $f(x)$ and $g(y)$ are respectively the density functions of $X$ and $Y$, then the feasibility constraints will always be satisfied and in particular:

$$\int_x f_0(x) \, dx = \int_x f_1(x) \, dx = \int_y g_0(y) \, dy = \int_y g_1(y) \, dy = 1$$

This remark will be crucial particularly in the case of nearly scarce matching and nearly full matching.

**Equilibrium in Nearly Scarce Matching: $\alpha = 0^+$**

In the case at the equilibrium the probability of matching $\alpha$ is close to 0, the deterministic part of the utility of singles $\Phi^0$ tends to $+\infty$ and so the individuals prefer to remain single and they do not participate to the marriage market. The direct consequence is the absolute lack of matching. In other words, in this situation, the social planer cannot increase anymore the social surplus by matching. It is optimal for all the individuals to remain single. So there is no participation to the marriage market and it occurs then on the market a general celibacy. But note that, as the joint surplus is assumed to be exogenous to the model, then, even if there is no formation of couples, it is possible to derive the theoretical surpluses $U(x, y)$ and $V(x, y)$. These functions represent the eventual surpluses partners in isolated cases of matching have at the equilibrium of this market. We may observe actually some isolated cases of matching but the measure of their set will be 0. To precise our idea, considering the market is infinite, then the number of matchings in this particular case of scarce matching will be finite, i.e the probability of marriage is 0. If the market is finite then there will be no marriage. As we assume the possibility of infinitude of the market, then marriage can occur on this market discontinuously and finitely. The optimal matching function is
simply the density function limit to which $\pi(x, y)$ tends to when $\alpha$ is close 0. It will still be a density function in this case of scarce matching respecting feasibility constraints. From Proposition 2, we can deduce for $\alpha = 0^+$ at the equilibrium, we have:

1. For any $x \in \mathcal{X}$, $y \in \mathcal{Y}$,
$$\pi(x, y) = e^{-\delta(f(x))} \frac{1}{2} (g(y)) \frac{1}{2} e^{\frac{S(x,y)}{2\sigma}}$$

2. The systematic surplus of a man of attributes $x$ from a matching with a woman of attributes $y$ is such as:
$$U(x, y) = \frac{1}{2} (S(x, y) - \sigma \ln (f(x)) + \sigma \ln (g(y)))$$

3. The systematic surplus of a woman of attributes $y$ from a matching with a man of attributes $x$ is such as:
$$V(x, y) = \frac{1}{2} (S(x, y) + \sigma \ln (f(x)) - \sigma \ln (g(y)))$$

4. for any $x \in \mathcal{X}$,
$$f_0(x) = f(x) \text{ and } f_1(x) = e^{-\delta(f(x))} \frac{1}{2} \int_{\mathcal{Y}} (g(y)) \frac{1}{2} e^{\frac{S(x,y)}{2\sigma}} dy$$

5. for any $y \in \mathcal{Y}$,
$$g_0(y) = g(y) \text{ and } g_1(y) = e^{-\delta(f(x))} \frac{1}{2} \int_{\mathcal{X}} (f(x)) \frac{1}{2} e^{\frac{S(x,y)}{2\sigma}} dx$$

6. 
$$\delta = \ln \left( \int_{\mathcal{X}} \int_{\mathcal{Y}} (f(x)) \frac{1}{2} (g(y)) \frac{1}{2} e^{\frac{S(x,y)}{2\sigma}} dx dy \right)$$

*Proof.* We simply replace $\alpha = 0^+$ in Proposition 2.

In this particular state of the market all the unknown of the model are completely analytically determined. We can also remark that even there is no matching observed, the matching function $\pi(x, y)$ still satisfies the feasibility constraints. In fact we have:

$$\int_{\mathcal{Y}} \pi(x, y) dy = e^{-\delta(f(x))} \frac{1}{2} \int_{\mathcal{Y}} (g(y)) \frac{1}{2} e^{\frac{S(x,y)}{2\sigma}} dy = f_1(x)$$

and

$$\int_{\mathcal{X}} \pi(x, y) dx = e^{-\delta(g(y))} \frac{1}{2} \int_{\mathcal{X}} (f(x)) \frac{1}{2} e^{\frac{S(x,y)}{2\sigma}} dx = g_1(y)$$

The expression of $\delta$ guarantees the satisfaction of the constraint

$$\int_{\mathcal{X}} \int_{\mathcal{Y}} \pi(x, y) dx dy = 1.$$ 

In what follows, we will focus on a particular matching market in which nearly all the individuals are matched.
3 Nearly Full Matching and Quadratic Specification

3.1 General Equilibrium and Comparison with the Dupuy-Galichon Model

We consider an equally bipartite matching market verifying Assumptions 1, 3 and 4. In addition, we assume that nearly all the individuals are matched at the equilibrium. So here, the probability $\alpha$ is equal to $1^-$. This has been treated by Dupuy and Galichon (2014). The equilibrium is given by the maximization of the social surplus. Proposition 2 gives the equilibrium and moreover we have:

1. for any $x \in \mathcal{X}$,
   \[ f_0(x) = \frac{e^\delta}{\int_\mathcal{Y} e^{-\delta \frac{U(x,y)}{\sigma}} dy} f(x) \quad \text{and} \quad f_1(x) = f(x) \]

2. for any $y \in \mathcal{Y}$,
   \[ g_0(y) = \frac{e^\delta}{\int_\mathcal{X} e^{-\delta \frac{V(x,y)}{\sigma}} dx} g(y) \quad \text{and} \quad g_1(y) = g(y) \]

*Proof.* We obtain these results by replacing $\alpha = 1^-$ in Proposition 2.

The expressions of the surpluses $\pi(x, y)$, $U(x, y)$, $V(x, y)$ and $\delta$ are given by Proposition 2.

3.2 Parametric Inference

We specify parametrically the joint surplus $S(x, y)$ with a parameter $\theta$ belonging to a subset $\Theta$ of a real vectors subspace. We denote by $\mathcal{C} \subset \mathcal{H} \times \mathcal{F}$ the set of the couples. There exists no single on this market. We denote by $\pi^\theta$ the equilibrium matching density function corresponding to each particular $\theta$. Considering a man $m$ with characteristics $x_m$ and a woman $w$ with characteristics $y_w$, the likelihood of the union $(m, w)$ is the product of the conditional probability the man $m$ chooses the woman $w$ with respect to $X = x_m$ with the conditional probability the woman $w$ chooses the man $m$ with respect to $Y = y_w$. That is:

\[ \pi^\theta(y_w|x_m) \cdot \pi^\theta(x_m|y_w) \]

with

\[ \pi^\theta(y_w|x_m) = \frac{\pi^\theta(x_m, y_w)}{f(x_m)} \quad \text{and} \quad \pi^\theta(x_m|y_w) = \frac{\pi^\theta(x_m, y_w)}{f(y_w)} \]
The likelihood $L$ of all the unions is the product of all the product all the likelihoods of the unions. So we have:

$$\forall \theta \in \Theta, \quad \log L(\theta) = \sum_{(m,w) \in C} \left[ 2 \ln(\pi^\theta(x_m, y_w)) - \ln(f(x_m)) - \ln(g(y_w)) \right]$$

We maximize the log-likelihood over the set $\Theta$. Note the density functions $f$ and $g$ are assumed to be completely known. So we can withdraw them from the log-likelihood. And we then obtain a reduced log-likelihood which is:

$$\forall \theta \in \Theta, \quad \log L(\theta) = \sum_{(m,w) \in C} \ln \pi^\theta(x_m, y_w)$$

In the following, we will try to give some closed forms at the equilibrium of the model and the likelihood for the inference, by considering a quadratic specification for the joint surplus.

### 3.3 Analytical Solution of the Matching Equilibrium in Case of Quadratic Specification

In this section, we provide the analytical solution of the model in this particular case of full marriage and under some assumptions on the distributions of the characteristics. We assume $X$ and $Y$ to be gaussian and we specify a quadratic form for the net joint surplus $S(X,Y)$.

**Assumption 5.**

1. 

$$X = \begin{pmatrix} X^1 \\ \vdots \\ X^p \end{pmatrix} \sim \mathcal{N}_p(0, \Sigma_X) \quad \text{and} \quad Y = \begin{pmatrix} Y^1 \\ \vdots \\ Y^q \end{pmatrix} \sim \mathcal{N}_q(0, \Sigma_Y)$$

2. 

$$S(X,Y) = (X' Y') K \begin{pmatrix} X \\ Y \end{pmatrix}$$

where:

$$K = \begin{pmatrix} A & C' \\ C & B \end{pmatrix} \in \mathcal{M}_{n,n}(\mathbb{R})$$

with: $n = p + q$, $A \in \mathcal{M}_{p,p}(\mathbb{R})$, $B \in \mathcal{M}_{q,q}(\mathbb{R})$ and $C \in \mathcal{M}_{q,p}(\mathbb{R})$.

The following theorem describes analytically the equilibrium of this market.

**Theorem 1.** Assume a stable nearly full matching verifying Assumption 1, Assumption 3, Assumption 4 and Assumption 5. Then,
1. the optimal matching $\pi$ is the density function of a $N_{p+q}(0, \Sigma)$ such that:

(a) 

$$
\Sigma^{-1} = \begin{pmatrix}
\Sigma_X^{-1} (\Sigma_X - \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{XY})^{-1} & -\Sigma_X^{-1} \Sigma_{XY} (\Sigma_Y - \Sigma_{YX} \Sigma_Y^{-1} \Sigma_{XY})^{-1} \\
-\Sigma_Y^{-1} \Sigma_{YX} (\Sigma_X - \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{XY})^{-1} & (\Sigma_Y - \Sigma_{YX} \Sigma_Y^{-1} \Sigma_{XY})^{-1}
\end{pmatrix}
$$

(b) we also have:

$$
\Sigma^{-1} = \begin{pmatrix}
\Gamma_X & -\frac{1}{\sigma} C' \\
-\frac{1}{\sigma} C & \Gamma_Y
\end{pmatrix}
$$

$$
= \begin{pmatrix}
\frac{1}{2} \Sigma_X^{-1} + \left(\frac{1}{4} \Sigma_X^{-1} + \frac{1}{\sigma^2} C' \Sigma_Y C\right) \Sigma_X^{-1} & -\frac{1}{\sigma} C' \\
-\frac{1}{\sigma} C & \frac{1}{2} \Sigma_Y^{-1} + \left(\frac{1}{4} \Sigma_Y^{-1} + \frac{1}{\sigma^2} C' \Sigma_X C\right) \Sigma_Y^{-1}
\end{pmatrix}
$$

(c) and moreover

$$
\Sigma = \begin{pmatrix}
\Sigma_X & \Sigma_{XY} \\
\Sigma_{YX} & \Sigma_Y
\end{pmatrix} = \begin{pmatrix}
\Sigma_X & \frac{1}{\sigma} \Gamma^{-1}_X C' \Sigma_Y \\
\frac{1}{\sigma} \Sigma_Y C \Gamma^{-1}_X & \Sigma_Y
\end{pmatrix}
$$

2.

$$f_0 \text{ is the density function of a Gaussian } N_p\left(0, \left(\frac{2A + 2\sigma \Gamma_X}{\sigma}\right)^{-1}\right)$$

and

$$g_0 \text{ is the density function of a Gaussian } N_q\left(0, \left(\frac{2B + 2\sigma \Gamma_Y}{\sigma}\right)^{-1}\right)$$

3. $\forall (X, Y) \in \mathbb{R}^p \times \mathbb{R}^q$,

$$U(X, Y) = \frac{1}{2} X'(2A + \sigma \Gamma_X) X - \frac{\sigma}{2} Y' \Gamma_Y Y + X' C' Y - \frac{\sigma}{4} \ln \left|\frac{2A + 2\sigma \Gamma_X}{\sigma}\right| + \frac{\sigma}{4} \ln \left|\frac{2B + 2\sigma \Gamma_Y}{\sigma}\right| - \frac{\sigma (p - q)}{4} \ln(2\pi)$$

and

$$V(X, Y) = \frac{1}{2} Y'(2B + \sigma \Gamma_Y) Y - \frac{\sigma}{2} X' \Gamma_X X + Y' C X - \frac{\sigma}{4} \ln \left|\frac{2B + 2\sigma \Gamma_Y}{\sigma}\right| + \frac{\sigma}{4} \ln \left|\frac{2A + 2\sigma \Gamma_X}{\sigma}\right| + \frac{\sigma (p - q)}{4} \ln(2\pi)$$
4. \[\delta = \frac{1}{4} \left( \ln \left| \frac{2A + 2\sigma \Gamma_X}{\sigma} \right| + \ln \left| \frac{2B + 2\sigma \Gamma_Y}{\sigma} \right| - \ln \left( \frac{\Gamma_X ||\Gamma_Y|}{\Sigma_X ||\Sigma_Y|} \right) \right)\]

Proof. The proof is given in the appendix. \(\square\)

Remark: We treat in Appendix the case the joint surplus has linear terms.

From the statements 1.a and 1.b of Theorem 1., we can remark that we obtain an interesting result of identification:

\[ C = \sigma \Sigma_Y^{-1} \Sigma_X \left( \Sigma_X - \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{XY} \right)^{-1} \]

and equivalently,

\[ C' = \sigma \Sigma_X^{-1} \Sigma_{XY} \left( \Sigma_Y - \Sigma_{XY} \Sigma_X^{-1} \Sigma_{XY} \right)^{-1} \]

With this result, we find exactly the Theorem 2 of Bojilov and Galichon (2016) identifying the affinity matrix of their model. Actually, the matrix 2C corresponds to the affinity matrix of the framework of Bojilov and Galichon.

The statement 1.c of Theorem 1. precises the matrix of covariance between X and Y:

\[ \Sigma_{XY} = \frac{1}{\sigma} \Sigma_Y \Gamma^{-1}_X = \frac{1}{\sigma} \Gamma^{-1}_Y C \Sigma_X \text{ and } \Sigma_{XY} = \frac{1}{\sigma} \Gamma^{-1}_X C' \Sigma_Y = \frac{1}{\sigma} \Sigma_X C' \Gamma^{-1}_Y \]

Note that we have considered here for simplicity that the characteristics such that \(E(X) = E(Y) = 0\). This can of course be extended to a generalized case.

The 'Dupuy-Galichon Constant' In the Dupuy and Galichon’s (2014) model, they have proved the following expressions for the individual surpluses:

\[ U(x, y) = \frac{\Phi(x, y) + a(x) - b(y)}{2} \text{ and } V(x, y) = \frac{\Phi(x, y) - a(x) + b(y)}{2} \]

where \(\Phi(x, y)\) is the joint surplus function, and \(a(x)\) and \(b(y)\) are potentials guaranteeing the feasibility of \(\pi\); these potentials \(a(x)\) and \(b(y)\) exist, are unique and can be determined respectively up to two constants \(c_1\) and \(c_2\). We have:

\[ a(x) = -\sigma \ln (f_0(x)) + c_1 \text{ and } b(y) = -\sigma \ln (g_0(y)) + c_2 \]

where the functions \(f_0(x)\) and \(g_0(y)\) are given analytically by Theorem 1. We can tell more about the constants \(c_1\) and \(c_2\). Actually, we can identify the sum of the constants \(c_1 + c_2\). Actually, \(c_1 + c_2 = 2\sigma \delta\) and \(\delta\) is given by Theorem 1. So we have:

\[ c_1 + c_2 = \frac{\sigma}{2} \left( \ln \left| \frac{2A + 2\sigma \Gamma_X}{\sigma} \right| + \ln \left| \frac{2B + 2\sigma \Gamma_Y}{\sigma} \right| - \ln \left( \frac{\Gamma_X ||\Gamma_Y|}{\Sigma_X ||\Sigma_Y|} \right) \right) \]

23
where the matrices $\Gamma_X$ and $\Gamma_Y$ are given by Theorem 1. The exact determination of the constants $c_1$ and $c_2$ remains then on the value of $c_1 - c_2$. We will denote:

$$Cst^{DG} = \frac{c_1 - c_2}{2}$$

and we will call it The Dupuy-Galichon Constant for the reason that, as we know the value of the sum $c_1 + c_2 = 2\sigma\delta$, one has:

$$c_1 = \frac{(c_1 + c_2)}{2} + \frac{(c_1 - c_2)}{2} = \sigma\delta + Cst^{DG}$$

and

$$c_2 = \frac{(c_1 + c_2)}{2} - \frac{(c_1 - c_2)}{2} = \sigma\delta - Cst^{DG}.$$ 

So the constants $c_1$ and $c_2$ are actually determined up to a constant that is the Dupuy-Galichon Constant $Cst^{DG}$. The value of this constant remains basically on the specification of the deterministic part of the utility of single individuals at the beginning of the market even if at the equilibrium no one has remained unmatched. More precisely, we have:

$$Cst^{DG} = \sigma \left( \int_X \int_Y f_0(x) \exp \left( \frac{U(x, y)}{\sigma} \right) dx dy - \int_X \int_Y g_0(y) \exp \left( \frac{V(x, y)}{\sigma} \right) dx dy \right)$$

This can also be expressed as follows:

$$Cst^{DG} = \sigma \left( \ln \left( \int_X \int_Y \exp \left( \frac{A^0(x)}{\sigma} \right) \pi(x, y) dx dy \right) - \ln \left( \int_X \int_Y \exp \left( \frac{B^0(y)}{\sigma} \right) \pi(x, y) dx dy \right) \right)$$

or equivalently here in the case of full matching, we have:

$$\int_Y \pi(x, y) dy = f_1(x) = f(x) \quad \text{and} \quad \int_X \pi(x, y) dx = g_1(y) = g(y)$$

then we can rewrite $Cst^{DG}$ completely with deterministic terms:

$$Cst^{DG} = \sigma \left( \ln \left( \int_X \exp \left( \frac{A^0(x)}{\sigma} \right) f(x) dx \right) - \ln \left( \int_Y \exp \left( \frac{B^0(y)}{\sigma} \right) g(y) dy \right) \right)$$

where $A^0(x)$ and $B^0(y)$ represent respectively the deterministic part of the utility of a single man with characteristic $x$ and the deterministic part of the utility of a single woman with characteristic $y$. One implication of this identity is that, the Dupuy-Galichon Constant $Cst^{DG}$ is actually exogenous. The specification of the deterministic parts of the utility of singles determines in advance the value this constant. In the case, we do not make any specification on the deterministic part of the utility of singles, $Cst^{DG}$ becomes undetermined. It is the case in the Dupuy and Galichon’s model (2014). In the case the deterministic parts of the utility
of singles are constant functions $A^0$ and $B^0$ as stated in Assumption 2, then we simply have $\text{Cst}^{DG} = A^0 - B^0$. Assumption 4 eases even things by assuming $A^0 = B^0 = \Phi^0$, it implies $\text{Cst}^{DG} = 0$ and therefore:

$$c_1 = c_2 = \sigma \delta = \frac{\sigma}{4} \left( \ln \left| \frac{2A + 2\sigma \Gamma_X}{\sigma} \right| + \ln \left| \frac{2B + 2\sigma \Gamma_Y}{\sigma} \right| - \ln \left( \frac{||\Gamma_X||}{||\Gamma_Y||} \right) \right)$$

where the matrices $\Gamma_X$ and $\Gamma_Y$ are given by Theorem 1.

As we can remark, the optimal distribution $\pi$ depends only on the sub-matrix $C$ of the matrix $K$. In what follows, we propose two ways to estimate it.

**Parametric Inference**

1. **Maximum Likelihood Estimation**
   We maximize over the set of the sub-matrices $C \in \mathcal{M}_{q,p}(\mathbb{R})$, the log-likelihood:

   $$\text{Log } L(C) = - \sum_{(m,w) \in C} \left[ (X'_m, Y'_w) \Sigma^{-1} \left( \begin{array}{c} X_m \\ Y_w \end{array} \right) + \ln |\Sigma| \right]$$

   with

   $$\Sigma^{-1} = \begin{pmatrix} \frac{1}{2} \Sigma^{-1}_X + \left( \frac{1}{4} \Sigma^{-1}_X + \frac{1}{\sigma^2} C'' \Sigma_X C \right)^{\frac{1}{2}} \left( \Sigma^{-1}_X \right)^{-\frac{1}{2}} & -\frac{1}{\sigma} C' \\ -\frac{1}{\sigma} C & \frac{1}{2} \Sigma^{-1}_Y + \left( \frac{1}{4} \Sigma^{-1}_Y + \frac{1}{\sigma^2} \Sigma_X C' \right)^{\frac{1}{2}} \left( \Sigma^{-1}_Y \right)^{-\frac{1}{2}} \end{pmatrix}$$

   To ease the computation of the likelihood, one may simply use:

   $$\Sigma^{-1} = \begin{pmatrix} \left( \Sigma_X - \Sigma_{XY} \Sigma^{-1}_Y \Sigma_{YX} \right)^{-1} & -\frac{1}{\sigma} C' \\ -\frac{1}{\sigma} C & \left( \Sigma_Y - \Sigma_{YX} \Sigma^{-1}_X \Sigma_{XY} \right)^{-1} \end{pmatrix}$$

   by replacing the matrix $\Sigma_{XY}$ by the empirical matrix of covariance:

   $$\hat{\Sigma}_{XY} = \frac{1}{|C|} \sum_{(m,w) \in C} X^i_m Y^j_w$$

2. **Moment-based Estimation**

   From the identification of $C$, by a first approach, we can have as estimator:

   $$\hat{C} = \sigma \Sigma^{-1} \hat{\Sigma}_{YX} \left( \Sigma_X - \hat{\Sigma}_{XY} \Sigma^{-1}_Y \hat{\Sigma}_{YX} \right)^{-1}$$

   or

   $$\hat{C'} = \sigma \Sigma^{-1} \hat{\Sigma}_{XY} \left( \Sigma_Y - \hat{\Sigma}_{YX} \Sigma^{-1}_X \hat{\Sigma}_{XY} \right)^{-1}$$
But note the matrix Σ_{XY} or Σ_{YX} depend actually on the parameter C. In fact, from the lemma, we have the relation:

\[ \Sigma_{YX} = \frac{1}{\sigma} \Gamma^{-1}_Y C \Sigma_X \quad \text{and} \quad \Sigma_{XY} = \frac{1}{\sigma} \Gamma^{-1}_X C' \Sigma_Y \]

and the proposition precises:

\[ \Gamma_X = \frac{1}{2} \Sigma^{-1}_X + \left( \frac{1}{4} \Sigma^{-1}_X + \frac{1}{\sigma^2} C' \Sigma_Y C \right)^{\frac{1}{2}} \left( \Sigma^{-\frac{1}{2}}_X \right)^{-1} \]

and

\[ \Gamma_Y = \frac{1}{2} \Sigma^{-1}_Y + \left( \frac{1}{4} \Sigma^{-1}_Y + \frac{1}{\sigma^2} C \Sigma_X C' \right)^{\frac{1}{2}} \left( \Sigma^{-\frac{1}{2}}_Y \right)^{-1} \]

So we have:

\[ \Sigma_{YX} = \frac{1}{\sigma} \left[ \frac{1}{2} \Sigma^{-1}_Y + \left( \frac{1}{4} \Sigma^{-1}_Y + \frac{1}{\sigma^2} C \Sigma_X C' \right)^{\frac{1}{2}} \left( \Sigma^{-\frac{1}{2}}_Y \right) \right]^{-1} C \Sigma_X \]

and

\[ \Sigma_{XY} = \frac{1}{\sigma} \left[ \frac{1}{2} \Sigma^{-1}_X + \left( \frac{1}{4} \Sigma^{-1}_X + \frac{1}{\sigma^2} C' \Sigma_Y C \right)^{\frac{1}{2}} \left( \Sigma^{-\frac{1}{2}}_X \right) \right]^{-1} C' \Sigma_Y \]

We then estimate the matrix C by solving the equation:

\[ \hat{\Sigma}_{YX} = \frac{1}{\sigma} \left[ \frac{1}{2} \Sigma^{-1}_Y + \left( \frac{1}{4} \Sigma^{-1}_Y + \frac{1}{\sigma^2} \hat{C} \Sigma_X \hat{C} \right)^{\frac{1}{2}} \left( \Sigma^{-\frac{1}{2}}_Y \right) \right]^{-1} \hat{C} \Sigma_X \]

or

\[ \hat{\Sigma}_{XY} = \frac{1}{\sigma} \left[ \frac{1}{2} \Sigma^{-1}_X + \left( \frac{1}{4} \Sigma^{-1}_X + \frac{1}{\sigma^2} \hat{C} \Sigma_Y \hat{C} \right)^{\frac{1}{2}} \left( \Sigma^{-\frac{1}{2}}_X \right) \right]^{-1} \hat{C}' \Sigma_Y \]

A Particular but General Case of Specification

We assume that the joint surplus function is specified as follows:

\[ S(X, Y) = X'C'Y' = \frac{1}{2} (X'Y') \begin{pmatrix} 0 & C' \\ C & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \quad \text{where} \quad C \in \mathcal{M}_{q,p}(\mathbb{R}) \]

where

\[
X = \begin{pmatrix}
X^1 \\
\vdots \\
X^p
\end{pmatrix} \sim \mathcal{N}_p(0, \Sigma_X) \quad \text{and} \quad Y = \begin{pmatrix}
Y^1 \\
\vdots \\
Y^q
\end{pmatrix} \sim \mathcal{N}_q(0, \Sigma_Y)
\]
We can include in the components of $X$ polynomials of the attributes of men, and in the components of $Y$ polynomials of the attributes of women. In a general case where the attributes $X$ and $Y$ are not centered, i.e $E(X)$ and $E(Y)$ can possibly be different from 0, we can consider for instance $X = (1 \ x \ x^2)'$ and $Y = (1 \ y \ y^2)'$, and then we have:

$$S(x, y) = a \ x^2 + b \ y^2 + c \ xy + a_1 \ x + b_1 \ y$$

We can write $S(x, y)$ at the form:

$$S(x, y) = (1 \ x \ x^2) \begin{pmatrix} 0 & b_1 & b \\ a_1 & c & 0 \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ y \\ y^2 \end{pmatrix} = X' \begin{pmatrix} 0 & 0 & b \\ 0 & c & 0 \\ a & 0 & 0 \end{pmatrix} Y$$

So the specification we are considering here is a particular case of the general case we have presented above because it is its restriction with $A = 0$ and $B = 0$, but nevertheless this particular restriction is equivalent to the initial specification considered.

The optimal matching $\pi$ is the density function of a Gaussian $\mathcal{N}_{p+q}(0, \Sigma)$ where

$$\Sigma^{-1} = \begin{pmatrix} \Gamma_X & -\frac{1}{2\sigma} C' \\ -\frac{1}{2\sigma} C & \Gamma_Y \end{pmatrix}$$

The individual surpluses are given by:

$$U(X, Y) = \frac{\sigma}{2} X' \Gamma_X X - \frac{\sigma}{2} Y' \Gamma_Y Y + \frac{1}{2} X' C' Y - \frac{\sigma}{4} \ln \left[ \frac{\det \Gamma_X}{\det \Gamma_Y} \right] - \frac{\sigma(p - q)}{4} \ln(4\pi)$$

and

$$V(X, Y) = \frac{\sigma}{2} Y' \Gamma_Y Y - \frac{\sigma}{2} X' \Gamma_X X + \frac{1}{2} Y' C X + \frac{\sigma}{4} \ln \left[ \frac{\det \Gamma_X}{\det \Gamma_Y} \right] + \frac{\sigma(p - q)}{4} \ln(4\pi)$$

where,

$$\Gamma_X = \frac{1}{2} \Sigma_X^{-1} + \left( \frac{1}{4} \Sigma_X^{-1} + \frac{1}{4\sigma^2} C' \Sigma_Y C \right)^{\frac{1}{2}} \left( \Sigma_X^{-\frac{1}{2}} \right)^{-1} = (\Sigma_X - \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{YX})^{-1}$$

and

$$\Gamma_Y = \frac{1}{2} \Sigma_Y^{-1} + \left( \frac{1}{4} \Sigma_Y^{-1} + \frac{1}{4\sigma^2} C \Sigma_X C' \right)^{\frac{1}{2}} \left( \Sigma_Y^{-\frac{1}{2}} \right)^{-1} = (\Sigma_Y - \Sigma_{YX} \Sigma_X^{-1} \Sigma_{XY})^{-1}$$

The matrix $C$ is called the affinity matrix by Dupuy-Galichon. We can estimate this matrix by maximization of the log-likelihood and we have:
\[ \hat{C}_{MLE} = \text{Arg} \max_{C \in \mathcal{M}_{\rho}(\mathbb{R})} - \sum_{(m,w) \in C} (X_m' \ Y_w') \Sigma^{-1} \left( \begin{array}{c} X_m \\ Y_w \end{array} \right) + \ln |\Sigma| \]

with
\[ \Sigma^{-1} = \begin{pmatrix} (\Sigma - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YY})^{-1} & -\frac{1}{2\sigma} C' \\ -\frac{1}{2\sigma} C & (\Sigma - \Sigma_{XY} \Sigma_{XX}^{-1} \Sigma_{XY})^{-1} \end{pmatrix} \]

We can also estimate by:
\[ \hat{C} = \sigma \Sigma_{YY}^{-1} \Sigma_{XY} \left( \Sigma - \hat{\Sigma}_{XY} \Sigma_{YX}^{-1} \hat{\Sigma}_{XY} \right)^{-1} \]

or
\[ \hat{C}' = \sigma \Sigma_{XX}^{-1} \Sigma_{XY} \left( \Sigma - \hat{\Sigma}_{XY} \Sigma_{X}^{-1} \hat{\Sigma}_{XY} \right)^{-1} \]

**Uni-dimensional Case**

We consider here that the characteristics follow standard gaussian distributions. Also for simplicity, we assume \( \sigma = 1 \). We obtain as results in this particular case the following results.

\[ S(x, y) = (x, y) \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \]

The optimal matching \( \pi \) that is the density function of \((X, Y)\) and:

\[ \pi(x, y) = \frac{1}{2\pi} \sqrt{\frac{1 + \sqrt{1 + 4c^2}}{2}} \exp \left\{ -\frac{1}{2} (x \ y) \begin{pmatrix} \frac{1 + \sqrt{1 + 4c^2}}{2} & -c \\ -c & \frac{1 + \sqrt{1 + 4c^2}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\} \]

The individual surpluses are given by:

\[ U(x, y) = \frac{1 + 4a + \sqrt{1 + 4c^2}}{4} x^2 - \frac{1 + \sqrt{1 + 4c^2}}{4} y^2 + c x y - \frac{1}{4} \ln \left( \frac{1 + 2a + \sqrt{1 + 4c^2}}{1 + 2b + \sqrt{1 + 4c^2}} \right) \]

and

\[ V(x, y) = \frac{1 + 4b + \sqrt{1 + 4c^2}}{4} y^2 - \frac{1 + \sqrt{1 + 4c^2}}{4} x^2 + c x y + \frac{1}{4} \ln \left( \frac{1 + 2a + \sqrt{1 + 4c^2}}{1 + 2b + \sqrt{1 + 4c^2}} \right) \]

The functions \( f_0(x) \) and \( g_0(y) \) are given by:

\[ f_0(x) = \sqrt{\frac{1 + 2a + \sqrt{1 + 4c^2}}{2\pi}} \exp \left\{ -\frac{1 + 2a + \sqrt{1 + 4c^2}}{2} x^2 \right\} \]
and
\[ g_0(y) = \sqrt{\frac{1 + 2b + \sqrt{1 + 4c^2}}{2\pi}} \exp \left\{ -\frac{1 + 2b + \sqrt{1 + 4c^2}}{2} y^2 \right\} \]
The constant \( \delta \) is given by:
\[ \delta = \frac{1}{4} \ln \left( \left| 1 + 2a + \sqrt{1 + 4c^2} \right| \right) + \frac{1}{4} \ln \left( \left| 1 + 2b + \sqrt{1 + 4c^2} \right| \right) - \frac{1}{2} \ln \left( \frac{1 + \sqrt{1 + 4c^2}}{2} \right) \]

**A Particular Application**

We consider the following simple case. We assume \( \sigma = 1 \), and \( X \) and \( Y \) are assumed to be standard Gaussian variables, and the joint surplus is defined as:
\[ S(x, y) = -(x - y)^2 = (x \ y) \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \]

Then the optimal matching joint density function \( \pi(x, y) \) is the density function of the following two-dimensional gaussian:
\[ \mathcal{N} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \varphi^{-1} \\ \varphi^{-1} & 1 \end{pmatrix} \right) \]
where
\[ \varphi = \frac{1 + \sqrt{5}}{2} \]
\( \varphi \) is called the golden section. This natural constant was considered by Greeks as the absolute perfect proportion. The optimal matching is given by:
\[ \pi(x, y) = \frac{\sqrt{\varphi}}{2\pi} \exp \left\{ -\frac{1}{2}(x \ y) \begin{pmatrix} \varphi & -1 \\ -1 & \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\} \]

Moreover, the functions \( f_0(x) \) and \( g_0(y) \) are equal to the density function of a gaussian:
\[ \mathcal{N} \left( 0, \frac{1}{2} \varphi \right) \]
In other words:
\[ f_0(x) = \frac{1}{\sqrt{\pi \varphi}} \exp \left\{ -\frac{1}{\varphi} x^2 \right\} = \sqrt{\frac{-1 + \sqrt{5}}{2\pi}} \exp \left\{ \frac{1 - \sqrt{5}}{2} x^2 \right\} \]
and
\[ g_0(y) = \frac{1}{\sqrt{\pi \varphi}} \exp \left\{ -\frac{1}{\varphi} y^2 \right\} = \sqrt{\frac{-1 + \sqrt{5}}{2\pi}} \exp \left\{ \frac{1 - \sqrt{5}}{2} y^2 \right\} \]
The individual surpluses functions $U(x, y)$ and $V(x, y)$ are:

$$U(x, y) = -\frac{1}{2}(2 - \varphi)x^2 - \frac{1}{2}\varphi y^2 + xy = -\frac{1}{4}(3 - \sqrt{5})x^2 - \frac{1}{4}(1 + \sqrt{5})y^2 + xy$$

and

$$V(x, y) = -\frac{1}{2}(2 - \varphi)y^2 - \frac{1}{2}\varphi x^2 + xy = -\frac{1}{4}(3 - \sqrt{5})y^2 - \frac{1}{4}(1 + \sqrt{5})x^2 + xy$$

And finally, the constant $\delta$ from its expression in Theorem 1. can be computed and is equal here to:

$$\delta = \frac{1}{2} \ln \left(3 - \sqrt{5}\right)$$

4 Empirical Application

We use the PSID data on the period 1978-2007 to estimate the affinity between men and women relatively to the normalized hourly incomes of the partners:

$$x_m = \frac{\text{Inc}_m - \text{E}[\text{Inc}_{\text{Men}}]}{\text{V}[\text{Inc}_{\text{Men}}]} \quad \text{and} \quad y_w = \frac{\text{Inc}_w - \text{E}[\text{Inc}_{\text{Women}}]}{\text{V}[\text{Inc}_{\text{Women}}]},$$

where $\text{Inc}_{\text{Men}}$ is the hourly income of men, and $\text{Inc}_{\text{Women}}$ is the hourly income of women. We assume that the characteristics $x_m$ and $y_w$ have standard Gaussian distributions. We assume $\sigma = 0$ and we specify the joint surplus from a matching of partners with characteristics $x$ and $y$ as follows:

$$S(x, y) = \left(\begin{array}{c} x_m \\ y_w \end{array}\right) \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

The affinity between $x$ and $y$ can be measured by $\frac{\partial^2 S(x,y)}{\partial x \partial y} = 2c$.

The optimal matching $\pi$ is given by:

$$\pi(x, y) = \frac{1}{2\pi} \sqrt{\frac{1 + \sqrt{1 + 4c^2}}{2}} \exp \left\{ -\frac{1}{2}(x, y) \begin{pmatrix} 1 + \sqrt{1 + 4c^2} \\ 2 \\ -c \\ 1 + \sqrt{1 + 4c^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\}$$

We maximize the log-likelihood over the set of the parameter $c$:

$$\log L(c) = \sum_{(m,w) \in C} \ln \pi^c(x_m, y_w)$$

So the parameter we can estimate here is $c$. 
Data

We use the PSID data on the period 1978-2005. We exclude from the data under 17 years old individuals. Our main variable is the income per hour. To deal with missing data on the income, we used the Heckman method to estimate the unobserved income on some individuals as the decision of unemployment can be partially explained by some observable determinants such as the age, the education or the sex. We estimate by the Heckman method of selection the income per hour for all the individuals by using as variables explaining the selection, the age, the square of age, the sex, and the education and these variables are present in the data basis. The hourly income will be the variable on which we will compute the gain from marriage. This variable is somehow the income of full worked time (by normalizing the worked time to 1 for all the individuals). We keep only married individuals.

Results

We estimate the model. The following graph shows the evolution of the affinity $c$ over the period 1978-2005 and we compare it to the evolution of the correlation between the incomes of the partners.

The affinity and the correlation between the incomes of the partners have both similar trend of variations. They have both generally increased on the period 1978-1989 but the intensity of the increase of the affinity has been higher than the intensity of the increase of the correlation between the incomes of the partners. We observe a decrease for both from 1989 to 1990 followed by an increase on the period 1990-1993. The affinity has decreased from 1993 to 1999 whereas the correlation between the partners’ incomes has generally increased on the same period 1993-1999. Then both decreased from 1999 to 2001 and increased from 2001 to 2005 and the intensity of the increase of the affinity is the higher of the two. So this comparison of evolution between the estimated affinity and the
correlation in the partners’ incomes shows that even the correlation between the incomes of partners has increased on the period 1990-1999, the affinity between the partners seems to have decreased on that period.

5 Conclusion

In this paper, we have considered a bipartite matching market with transferable utility, on which the two sides have the same weight. We use the framework of Dupuy and Galichon (2014) with some additional assumptions. The model allows to solve analytically the equilibrium of the market in the case of scarce matching. But the main goal of the paper consists of investigating on closed for the equilibrium in a particular case of full matching. For that we have restricted our framework to a matching market without singles, with a quadratic specification of the joint surplus and Gaussian distributions of the observable characteristics. This question has been treated by Bojilov and Galichon (2016); on the basis of Dupuy and Galichon setting, they have proposed the analytical expression of the optimal matching and have identified the affinity matrix of their model for this particular matching market. We find exactly the same identification of the affinity matrix with Bojilov and Galichon; we also provide closed forms for the optimal matching that are as well similar to the ones found by Bojilov and Galichon. But at the difference of Bojilov and Galichon, in addition we provide the analytical expressions of the individual surpluses. Dupuy and Galichon (2014) have actually proved that the individual surpluses $U(x,y)$ and $V(x,y)$ are identified respectively up to a function $a(x)$ and up to a function $b(y)$ and these functions exist, are unique and can be determined up to a constant. In the approach we suggest here, we can entirely retrieve the functions $U(x,y)$ and $V(x,y)$. This result is an interesting novelty in the sense that for such models, until now no closed forms are proposed for the individual surpluses to our knowledge. In fact, Bojilov and Galichon (2016) have given the conditional distributions $X|Y$ and $Y|X$ but have not given the expressions of the functional constants $a(x)$ and $b(y)$ crucial to recover the entirely the expressions of the individual surpluses in their framework. We finally provide two approaches to estimate the model: by maximum likelihood and by moments based estimation. The framework we propose here is based on transferable utility; but these results can be applied in the case of full matching with quadratic specification of the surplus in case of non transferable utility since following Menzel (2015), matching equilibrium in such circumstances of Gaussian distributions and quadratic specification and with transferable utility is identical to the equilibrium in these same circumstances with non transferable utility.
References


6 Appendix

6.1 Proof of Theorem 1.

Let:

\[ X = \begin{pmatrix} X^1 \\ \vdots \\ X^p \end{pmatrix} \sim N(0, \Sigma_X) \text{ and } Y = \begin{pmatrix} Y^1 \\ \vdots \\ Y^q \end{pmatrix} \sim N(0, \Sigma_Y) \]
and

\[ S(X, Y) = (X', Y') \begin{pmatrix} A & C' \\ C & B \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \]

where

\[ A \in \mathcal{M}_{p,p}(\mathbb{R}), \ B \in \mathcal{M}_{q,q}(\mathbb{R}) \text{ and } C \in \mathcal{M}_{q,p}(\mathbb{R}) \]

Let us denote:

\[ K = \begin{pmatrix} A & C' \\ C & B \end{pmatrix} \]

From Proposition 2., we have:

\[ S(X, Y) = 2\sigma\delta + 2\sigma \ln (\pi(X, Y)) - \sigma \ln (f_0(X)) - \sigma \ln (g_0(Y)) \]

Then we deduce:

\[ \text{Hess}(S)(X, Y) = 2K = 2\sigma \text{Hess}(\ln(\pi))(X, Y) - \begin{pmatrix} \sigma \text{Hess}(\ln(f_0))(X) & 0 \\ 0 & \sigma \text{Hess}(\ln(g_0))(Y) \end{pmatrix} \]

i.e.

\[ 2 \begin{pmatrix} A & C' \\ C & B \end{pmatrix} = 2\sigma \text{Hess}(\ln(\pi))(X, Y) - \begin{pmatrix} \sigma \text{Hess}(\ln(f_0))(X) & 0 \\ 0 & \sigma \text{Hess}(\ln(g_0))(Y) \end{pmatrix} \]

As the Hessian matrix of \( S \) is symmetric so is the Hessian matrix of \( \ln(\pi) \). Moreover, \( X \) and \( Y \) are Gaussian and as the matching is assumed to be full, then the distributions of \( X \) and \( Y \) are exactly their conditional distributions with respect to marriage \( M = 1 \). So the feasibility constraints are traduced by:

\[ \int_{\mathbb{R}^q} \pi(X, Y) dY = f(X) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_X|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}X'\Sigma_X^{-1}X \right\} \]

and

\[ \int_{\mathbb{R}^p} \pi(X, Y) dX = g(Y) = \frac{1}{(2\pi)^{\frac{q}{2}}|\Sigma_Y|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2}Y'\Sigma_Y^{-1}Y \right\} \]

These feasibility constraints require the Hessian matrix of \( \ln(\pi) \) to be constant.

The Hessian matrix of \( \ln(\pi) \) is symmetric and constant so we can let it be at the form:

\[ \text{Hess}(\ln(\pi))(X, Y) = -\begin{pmatrix} \Gamma_X & \Lambda' \\ \Lambda & \Gamma_Y \end{pmatrix} \]

where \( \Gamma_X \in \mathcal{M}_{p,p}(\mathbb{R}), \ \Gamma_Y \in \mathcal{M}_{q,q}(\mathbb{R}) \) and \( \Lambda \in \mathcal{M}_{q,p}(\mathbb{R}) \). We then have:

\[ 2 \begin{pmatrix} A & C' \\ C & B \end{pmatrix} = -2\sigma \begin{pmatrix} \Gamma_X & \Lambda' \\ \Lambda & \Gamma_Y \end{pmatrix} - \begin{pmatrix} \sigma \text{Hess}(\ln(f_0))(X) & 0 \\ 0 & \sigma \text{Hess}(\ln(g_0))(Y) \end{pmatrix} \]

From this equation, we identify immediately:

\[ \Lambda = -\frac{1}{\sigma}C' \]
As the Hessian matrix $\ln(\pi)$ is constant, then $\pi$ is a multidimensional Gaussian density function. We can then denote:

$$\text{Hess}(\ln(\pi))(X, Y) = -\Sigma^{-1}$$

We deduce:

$$\pi(X, Y) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (X' Y' \Sigma^{-1} \begin{pmatrix} X \\ Y \end{pmatrix}) \right\}$$

We then exploit the feasibility constraints and we have:

$$\int_{\mathbb{R}^q} \pi(X, Y) dY = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_X|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} X' \Sigma^{-1}_X X \right\}$$

and

$$\int_{\mathbb{R}^p} \pi(X, Y) dX = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_Y|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} Y' \Sigma^{-1}_Y Y \right\}$$

Remark that:

$$(X' Y') \Sigma^{-1} \begin{pmatrix} X \\ Y \end{pmatrix} = X' \Gamma_X X + 2X'\Lambda' Y + Y'\Gamma_Y Y$$

$$= X' (\Gamma_X - \Lambda \Gamma_Y^{-1} \Lambda') X + (Y + \Gamma_Y^{-1} \Lambda X)' \Gamma_Y (Y + \Gamma_Y^{-1} \Lambda X)$$

$$= Y' (\Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda') Y + (X + \Gamma_X^{-1} \Lambda' Y)' \Gamma_X (X + \Gamma_X^{-1} \Lambda' Y)$$

We then have by computation,

$$\int_{\mathbb{R}^q} \pi(X, Y) dY = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^\frac{1}{2} |\Gamma_X|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} X' \left( \Gamma_X - \Lambda \Gamma_Y^{-1} \Lambda' \right) X \right\}$$

and

$$\int_{\mathbb{R}^p} \pi(X, Y) dX = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^\frac{1}{2} |\Gamma_Y|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} Y' \left( \Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda' \right) Y \right\}$$

Thus the feasibility constraints are traduced by:

$$\frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_X|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} X' \Sigma^{-1}_X X \right\} = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^\frac{1}{2} |\Gamma_X|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} X' \left( \Gamma_X - \Lambda \Gamma_Y^{-1} \Lambda \right) X \right\}$$

and

$$\frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_Y|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} Y' \Sigma^{-1}_Y Y \right\} = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^\frac{1}{2} |\Gamma_Y|^\frac{1}{2}} \exp\left\{ -\frac{1}{2} Y' \left( \Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda' \right) Y \right\}$$

This is equivalent to:
\[
\begin{cases}
|\Sigma||\Gamma_Y| = |\Sigma_X| \\
\Gamma_X - \Lambda^t \Gamma_Y^{-1} \Lambda = \Sigma_X^{-1} \\
|\Sigma||\Gamma_X| = |\Sigma_Y| \\
\Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda' = \Sigma_Y^{-1}
\end{cases}
\]

We can then deduce from this system the following equations:

\[
\left(\begin{array}{cc}
\Gamma_X - \Lambda^t \Gamma_Y^{-1} \Lambda & -\Gamma_X^{-1} \Lambda^t (\Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda')^{-1} \\
-(\Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda')^{-1} \Lambda \Gamma_X^{-1} & (\Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda')^{-1}
\end{array}\right) = \left(\begin{array}{cc}
\Sigma_X & -\Gamma_X^{-1} \Lambda^t \Sigma_Y \\
-\Sigma_Y \Lambda \Gamma_X^{-1} & \Sigma_Y
\end{array}\right)
\]

and

\[
\left(\begin{array}{cc}
(\Gamma_X - \Lambda \Gamma_Y^{-1} \Lambda)^{-1} & -(\Gamma_X - \Lambda \Gamma_Y^{-1} \Lambda)^{-1} \Lambda \Gamma_Y^{-1} \\
-\Gamma_Y^{-1} \Lambda (\Gamma_X - \Lambda \Gamma_Y^{-1} \Lambda)^{-1} & (\Gamma_Y - \Lambda \Gamma_X^{-1} \Lambda')^{-1}
\end{array}\right) = \left(\begin{array}{cc}
\Sigma_X & -\Sigma_X \Lambda \Gamma_Y^{-1} \\
-\Gamma_Y^{-1} \Lambda \Sigma_X & \Sigma_Y
\end{array}\right)
\]

And the two matrices in left members of the equations are exactly the inverse of the matrix \(\Sigma^{-1}\) in the matrix inversion algorithm developed by Strassen (1969).

In other words, all these matrices are equal to \(\Sigma\). This is traduced by:

\[
\Sigma = \left(\begin{array}{cc}
\Gamma_X & \Lambda' \\
\Lambda & \Gamma_Y
\end{array}\right)^{-1} = \left(\begin{array}{cc}
\Sigma_X & -\Gamma_X^{-1} \Lambda^t \Sigma_Y \\
-\Sigma_Y \Lambda \Gamma_X^{-1} & \Sigma_Y
\end{array}\right) = \left(\begin{array}{cc}
\Sigma_X & -\Sigma_X \Lambda \Gamma^{-1}_Y \\
-\Gamma_Y^{-1} \Lambda \Sigma_X & \Sigma_Y
\end{array}\right)
\]

This proves the statement 1.c of Theorem 1., and this is equivalent to the result:

\[
\Sigma_Y \Lambda \Gamma^{-1}_X = \Gamma_Y^{-1} \Lambda \Sigma_X
\]

As we have:

\[
\Lambda = -\frac{1}{\sigma} C
\]

we obtain the equation:

\[
\Sigma_Y C \Gamma^{-1}_X = \Gamma_Y^{-1} C \Sigma_X
\]

We summary the main elementary results necessary for the proof in the following lemma.

**Lemma 1.** Assume a stable full matching verifying Assumption 1, Assumption 3, Assumption 4 and Assumption 5. Then there exist an invertible \(p \times p\) squared matrix \(\Gamma_X\) and an invertible \(q \times q\) squared matrix \(\Gamma_Y\) such that the optimal matching joint density function \(\pi(X, Y)\) verifies:

\[
\pi(X, Y) = \frac{1}{(2\pi)^{\frac{p+q}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} (X' Y') \Sigma^{-1} \left(\begin{array}{c} X \\ Y \end{array}\right)\right\}
\]

where:

\[
\Sigma = \left(\begin{array}{cc}
\frac{\Gamma_X}{\frac{1}{\sigma} C} & -\frac{1}{\sigma} C' \\
-\frac{1}{\sigma} C & \Gamma_Y
\end{array}\right)^{-1}
\]
and
\[
\begin{align*}
\Gamma_X - \frac{1}{\sigma^2} C\Gamma_Y^{-1} C &= \Sigma_X^{-1} \\
\Gamma_Y - \frac{1}{\sigma^2} CT_X^{-1} C' &= \Sigma_Y^{-1} \\
\Sigma_Y CT_X^{-1} &= \Gamma_Y^{-1} C\Sigma_X
\end{align*}
\]

Proof. The proof is immediate from the development and the following remark. We have:

\[
\begin{align*}
\Gamma_X - \frac{1}{\sigma^2} C\Gamma_Y^{-1} C &= \Sigma_X^{-1} \\
\Gamma_Y - \frac{1}{\sigma^2} CT_X^{-1} C' &= \Sigma_Y^{-1} \\
\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \pi(X, Y) dX dY &= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \pi(X, Y) dX dY = 1 \implies \left\{ \begin{array}{l}
|\Sigma| |\Gamma_Y| = |\Sigma_X| \\
|\Sigma| |\Gamma_X| = |\Sigma_Y|
\end{array} \right.
\end{align*}
\]

By equivalence, we have:

\[
\begin{align*}
\Gamma_X - \frac{1}{\sigma^2} C\Gamma_Y^{-1} C &= \Sigma_X^{-1} \\
\Gamma_Y - \frac{1}{\sigma^2} CT_X^{-1} C' &= \Sigma_Y^{-1} \\
\implies \left\{ \begin{array}{l}
\Gamma_X \Sigma_X \Gamma_X - \Gamma_X = \frac{1}{\sigma^2} C\Sigma_Y C' \\
\Gamma_Y \Sigma_Y \Gamma_Y - \Gamma_Y = \frac{1}{\sigma^2} C\Sigma_X C'
\end{array} \right.
\end{align*}
\]

\[
\begin{align*}
\implies \left\{ \begin{array}{l}
(\Gamma_X - \frac{1}{2} \Sigma_X^{-1}) \Sigma_X \left( \Gamma_X - \frac{1}{2} \Sigma_X^{-1} \right) = \frac{1}{4} \Sigma_X^{-1} + \frac{1}{\sigma^2} C\Sigma_Y C' \\
(\Gamma_Y - \frac{1}{2} \Sigma_Y^{-1}) \Sigma_Y \left( \Gamma_Y - \frac{1}{2} \Sigma_Y^{-1} \right) = \frac{1}{4} \Sigma_Y^{-1} + \frac{1}{\sigma^2} C\Sigma_X C'
\end{array} \right.
\end{align*}
\]

The matrices $\Gamma_X$ and $\Gamma_Y$ are defined positive. We have:

\[
\Gamma_X = \frac{1}{2} \Sigma_X^{-1} + \left( \frac{1}{4} \Sigma_X^{-1} + \frac{1}{\sigma^2} C\Sigma_Y C \right)^{\frac{1}{2}} \left( \Sigma_X^{-1} \right)^{-1}
\]

and

\[
\Gamma_Y = \frac{1}{2} \Sigma_Y^{-1} + \left( \frac{1}{4} \Sigma_Y^{-1} + \frac{1}{\sigma^2} C\Sigma_X C' \right)^{\frac{1}{2}} \left( \Sigma_Y^{-1} \right)^{-1}
\]

This proves the first point of the theorem as:

\[
\Sigma^{-1} = \begin{pmatrix}
\Gamma_X & -\frac{1}{\sigma^2} C' \\
-\frac{1}{\sigma^2} C & \Gamma_Y
\end{pmatrix}
\]

and with the Strassen algorithm, we have:

\[
\Sigma^{-1} = \begin{pmatrix}
\Sigma_X & \Sigma_{XY} \\
\Sigma_{YX} & \Sigma_Y
\end{pmatrix}^{-1}
\]

\[
= \begin{pmatrix}
\left( \Sigma_X - \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{YX} \right)^{-1} & -\Sigma_X^{-1} \Sigma_{XY} \left( \Sigma_Y - \Sigma_{YX} \Sigma_Y^{-1} \Sigma_{XY} \right)^{-1} \\
-\Sigma_Y^{-1} \Sigma_{YX} \left( \Sigma_X - \Sigma_{XY} \Sigma_Y^{-1} \Sigma_{YX} \right)^{-1} & \left( \Sigma_Y - \Sigma_{YX} \Sigma_Y^{-1} \Sigma_{XY} \right)^{-1}
\end{pmatrix}
\]

37
In the next step of the proof, we will identify \( f_0(X) \) and \( g_0(Y) \). We have:
\[
\begin{align*}
2A &= -2\sigma \Gamma_X - \sigma \text{Hess} \ln (f_0)(X) \\
2B &= -2\sigma \Gamma_Y - \sigma \text{Hess} \ln (g_0)(Y)
\end{align*}
\]
\[
\iff \begin{align*}
\text{Hess} \ln (f_0)(X) &= -\frac{2A - 2\sigma \Gamma_X}{\sigma} \\
\text{Hess} \ln (g_0)(Y) &= -\frac{2B - 2\sigma \Gamma_Y}{\sigma}
\end{align*}
\]

As we are in nearly full matching, we still have:
\[
\int_{\mathbb{R}^p} f_0(X) dX = 1 \quad \text{and} \quad \int_{\mathbb{R}^q} g_0(Y) dY = 1
\]
and we deduce:
\[
f_0(X) = \left| \frac{A + \sigma \Gamma_X}{\pi \sigma} \right|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} X' \left( \frac{2A + 2\sigma \Gamma_X}{\sigma} \right) X \right\}
\]
and
\[
g_0(Y) = \left| \frac{B + \sigma \Gamma_Y}{\pi \sigma} \right|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} Y' \left( \frac{2B + 2\sigma \Gamma_Y}{\sigma} \right) Y \right\}
\]
We then have:
\[
\ln (f_0(X)) = -\frac{1}{2} \left( X' \left( \frac{2A + 2\sigma \Gamma_X}{\sigma} \right) X - \ln \left| \frac{A + \sigma \Gamma_X}{\pi \sigma} \right| \right)
\]
and
\[
\ln (g_0(Y)) = -\frac{1}{2} \left( Y' \left( \frac{2B + 2\sigma \Gamma_Y}{\sigma} \right) Y - \ln \left| \frac{B + \sigma \Gamma_Y}{\pi \sigma} \right| \right)
\]
From Proposition 2., we have:
\[
U(X,Y) = \frac{1}{2} \left( S(X,Y) - \sigma \ln(f_0(X)) + \sigma \ln(g_0(Y)) \right)
\]
and
\[
V(X,Y) = \frac{1}{2} \left( S(X,Y) + \sigma \ln(f_0(X)) - \sigma \ln(g_0(Y)) \right)
\]
The constant \( \delta \) is given by:
\[
\delta = \ln \left( \int_{\mathbb{R}^p} \int_{\mathbb{R}^q} e^{\frac{S(X,Y)}{2\sigma}} \left( f_0(X) \right)^{\frac{1}{2}} \left( g_0(Y) \right)^{\frac{1}{2}} dX dY \right)
\]
With these formulas we derive the expressions of \( U(X,Y) \), \( V(X,Y) \) and \( \delta \).

6.2 General Case with \( A^0 \neq B^0 \)

We will consider here that the deterministic part of the utility of single men is a constant \( A^0 \) and the deterministic part of utility of single women is a constant \( B^0 \). We assume that \( A^0 \) may be different from \( B^0 \).
Proposition 3. Assume a stable matching market. Under Assumption 1, Assumption 2, and Assumption 3 we have:

1. for any $x \in X$, $y \in Y$,

$$S(x, y) = \sigma(\delta_1 + \delta_2) + \sigma \ln \left( \frac{(\pi(x, y))^2}{f_0(x)g_0(y)} \right)$$

2. The systematic surplus of a man of attributes $x$ from a matching with a woman of attributes $y$ is such as:

$$U(x, y) = \frac{1}{2} (S(x, y) - \sigma \ln (f_0(x)) + \sigma \ln (g_0(y)) + \sigma(\delta_1 - \delta_2))$$

3. The systematic surplus of a woman of attributes $y$ from a matching with a man of attributes $x$ is such as:

$$V(x, y) = \frac{1}{2} (S(x, y) + \sigma \ln (f_0(x)) - \sigma \ln (g_0(y) - \sigma(\delta_1 - \delta_2)))$$

4. for any $x \in X$,

$$f_0(x) = \frac{e^{\delta}}{(1 - \alpha)e^{\delta} + \alpha \int_Y e^{\frac{U(x, y)}{\sigma}} dy} f(x) \quad \text{and} \quad f_1(x) = \frac{\int_Y e^{\frac{U(x, y)}{\sigma}} dy}{(1 - \alpha)e^{\delta} + \alpha \int_Y e^{\frac{U(x, y)}{\sigma}} dy} f(x)$$

5. for any $y \in Y$,

$$g_0(y) = \frac{e^{\delta}}{(1 - \alpha)e^{\delta} + \alpha \int_X e^{\frac{V(x, y)}{\sigma}} dx} g(y) \quad \text{and} \quad g_1(y) = \frac{\int_X e^{\frac{V(x, y)}{\sigma}} dx}{(1 - \alpha)e^{\delta} + \alpha \int_X e^{\frac{V(x, y)}{\sigma}} dx} g(y)$$

6. $\delta_1 + \delta_2 = 2 \ln \left( \int_X \int_Y (f_0(x))^\frac{1}{2} (g_0(y))^\frac{1}{2} \exp \left( \frac{S(x, y)}{2\sigma} \right) dx dy \right)$

The proof is analogous to the proof of Proposition 2. with:

$$\delta_1 = \ln \left( \int_X \int_Y f_0(x) \exp \left( \frac{U(x, y)}{\sigma} \right) dx dy \right)$$

and

$$\delta_2 = \ln \left( \int_X \int_Y g_0(y) \exp \left( \frac{V(x, y)}{\sigma} \right) dx dy \right)$$
6.3 Nearly Full Matching and Quadratic Specification of the joint surplus including linear terms

We make the following assumption.

Assumption 6. 1.

\[
X = \begin{pmatrix} X^1 \\ \cdot \\ \cdot \\ X^p \end{pmatrix} \sim N_p(0, \Sigma_X) \quad \text{and} \quad Y = \begin{pmatrix} Y^1 \\ \cdot \\ \cdot \\ Y^q \end{pmatrix} \sim N_q(0, \Sigma_Y)
\]

2.

\[
S(X, Y) = (X' Y') K \begin{pmatrix} X \\ Y \end{pmatrix} + X' a + Y' b
\]

where:

\[
K = \begin{pmatrix} A & C' \\ C & B \end{pmatrix} \in \mathcal{M}_{n,n}(\mathbb{R})
\]

with: \( n = p + q \), \( A \in \mathcal{M}_{p,p}(\mathbb{R}) \), \( B \in \mathcal{M}_{q,q}(\mathbb{R}) \) and \( C \in \mathcal{M}_{q,p}(\mathbb{R}) \), and \( a \in \mathbb{R}^p \), \( b \in \mathbb{R}^q \).

Proposition 4. Assume a stable nearly full matching. Under the Assumption 1, Assumption 3, Assumption 4 and Assumption 6, we have:

1. the optimal matching \( \pi \) is the density function of a \( N_{p+q}(0, \Sigma) \) such that:

   (a)

   \[
   \Sigma^{-1} = \begin{pmatrix} \Sigma_X - \Sigma_XY^{-1}\Sigma_Y^{-1}\Sigma_XY & \Sigma_X^{-1}\Sigma_XY \Sigma_Y^{-1}\Sigma_XY^{-1} \\ \Sigma_X^{-1}\Sigma_XY \Sigma_Y^{-1}\Sigma_XY^{-1} & \Sigma_Y^{-1}\Sigma_XY^{-1}\Sigma_Y^{-1} \end{pmatrix}
   \]

   (b) we also have:

   \[
   \Sigma^{-1} = \begin{pmatrix} \Gamma_X & -\frac{1}{\sigma} C' \\ -\frac{1}{\sigma} C & \Gamma_Y \end{pmatrix}
   \]

   \[
   = \begin{pmatrix} \frac{1}{2} \Sigma_X^{-1} + \frac{1}{4} \Sigma_X^{-1} + \frac{1}{\sigma^2} C' \Sigma_Y C' \frac{1}{2} \left( \Sigma_X \right)^{-1} \\ -\frac{1}{\sigma} C & \frac{1}{2} \Sigma_Y^{-1} + \frac{1}{4} \Sigma_Y^{-1} + \frac{1}{\sigma^2} C' \Sigma_X C' \frac{1}{2} \left( \Sigma_Y \right)^{-1} \end{pmatrix}
   \]

   (c) and moreover

   \[
   \Sigma = \begin{pmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma_{XY} & \Sigma_Y \end{pmatrix} = \begin{pmatrix} \Sigma_X & \frac{1}{\sigma} \Gamma_X^{-1} C \Sigma_Y \\ \frac{1}{\sigma} \Sigma_Y C \Gamma_Y^{-1} \Sigma_X & \Sigma_Y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma} \Gamma_X^{-1} C \Sigma_Y \\ \frac{1}{\sigma} \Sigma_Y C \Gamma_Y^{-1} \Sigma_X \end{pmatrix}
   \]
2. 

\( f_0 \) is the density function of a Gaussian \( \mathcal{N}_p \left( \mu_0, \left( \frac{2A + 2\Gamma X}{\sigma} \right)^{-1} \right) \) with 

\[ \mu_0 = (2A + 2\Gamma X)^{-1} a \]

and

\( g_0 \) is the density function of a Gaussian \( \mathcal{N}_q \left( \mu_0, \left( \frac{2B + 2\Gamma Y}{\sigma} \right)^{-1} \right) \) with 

\[ \mu_0 = (2B + 2\Gamma Y)^{-1} b \]

3. \( \forall (X, Y) \in \mathbb{R}^p \times \mathbb{R}^q, \)

\[
U(X, Y) = \frac{1}{2} X' \left( 2A + \sigma \Gamma X \right) X - \frac{\sigma}{2} Y' \Gamma Y Y + X' C' Y + X' a - \frac{\sigma}{4} \ln \left| \frac{2A + 2\sigma \Gamma X}{\sigma} \right| \\
- \frac{1}{8} a' (A + \sigma \Gamma X)^{-1} a + \frac{1}{8} b' (B + \sigma \Gamma Y)^{-1} b + \frac{\sigma}{4} \ln \left| \frac{2B + 2\sigma \Gamma Y}{\sigma} \right| - \frac{\sigma (p - q)}{4} \ln (2\pi)
\]

and

\[
V(X, Y) = \frac{1}{2} Y' \left( 2B + \sigma \Gamma Y \right) Y - \frac{\sigma}{2} X' \Gamma X X + Y' C X + Y' b - \frac{\sigma}{4} \ln \left| \frac{2B + 2\sigma \Gamma Y}{\sigma} \right| \\
+ \frac{1}{8} a' (A + \sigma \Gamma X)^{-1} a - \frac{1}{8} b' (B + \sigma \Gamma Y)^{-1} b + \frac{\sigma}{4} \ln \left| \frac{2A + 2\sigma \Gamma X}{\sigma} \right| + \frac{\sigma (p - q)}{4} \ln (2\pi)
\]

4. 

\[ \delta = \frac{1}{4} \left( \ln \left| \frac{2A + 2\sigma \Gamma X}{\sigma} \right| + \ln \left| \frac{2B + 2\sigma \Gamma Y}{\sigma} \right| - \ln \left( \frac{||\Gamma X|| \Gamma Y ||}{||\Sigma X|| \Sigma Y ||} \right) \right) \]

Proof. The proof is then analogous to the proof of Theorem 1. and in addition we use the following equations:

\[
\text{Grad}(S)(X, Y) = \begin{pmatrix}
2AX + 2C'Y + a \\
2BY + 2C'X + b
\end{pmatrix} = 2\sigma \text{Grad} \ln(\pi)(X, Y) - \begin{pmatrix}
\sigma \text{Grad}(\ln(f_0)(X)) \\
\sigma \text{Grad}(\ln(g_0)(Y))
\end{pmatrix}
\]

where

\[ \text{Grad}(\ln(\pi))(X, Y) = - \begin{pmatrix}
\Gamma X X + \Lambda' Y \\
\Gamma Y Y + \Lambda X
\end{pmatrix} \]

and:

\[ \Lambda = - \frac{1}{\sigma} C \]
and we deduce that:

\[ \sigma \text{Grad}(\ln(f_0))(X) = -2(A + \sigma \Gamma_X)X - a \]

and

\[ \sigma \text{Grad}(\ln(g_0))(Y) = -2(B + \sigma \Gamma_Y)Y - b \]

Furthermore, we have:

\[ \text{Grad}(\ln(f_0))(X) = -\left(\frac{2A + 2\sigma \Gamma_X}{\sigma}\right)X + \left(\frac{2A + 2\sigma \Gamma_X}{\sigma}\right)\mu^0_X \]

and

\[ \text{Grad}(\ln(g_0))(Y) = -\left(\frac{2B + 2\sigma \Gamma_Y}{\sigma}\right)Y + \left(\frac{2B + 2\sigma \Gamma_Y}{\sigma}\right)\mu^0_Y \]

with

\[ \mu^0_X = -(2A + 2\sigma \Gamma_X)^{-1}a \]

and

\[ \mu^0_Y = -(2B + 2\sigma \Gamma_Y)^{-1}b \]

As we are in nearly full matching, we still have:

\[ \int_{\mathbb{R}^p} f_0(X)\,dX = 1 \quad \text{and} \quad \int_{\mathbb{R}^q} g_0(Y)\,dY = 1 \]

And we deduce immediately the expressions of \( f_0(X) \) and \( g_0(Y) \). We then use Proposition 2. to deduce the expressions of the individual surpluses \( U(X,Y) \), \( V(X,y) \) and of \( \delta \).

\[ \square \]