Theoretical and Empirical Aspects of Matching Markets

Alfred Galichon
(Ecole polytechnique and Columbia University)

Columbia University, Economics Departement
Spring 2011
PRELIMINARY AND INCOMPLETE.
(Last updated April 27, 2011)

Contents

Introduction

Preliminary Stuff:

My office is SIPA 1113.
Email address is ag2457@columbia.edu
Office hours are scheduled by appointment (please email).
Course location: SIPA 1027.
Course webpage:
http://sites.google.com/site/alfredgalichon/teaching/matching.

Course material. The lecture slides and the syllabus are posted on the course webpage.

No text is required. For game-theoretic aspects, a worthwhile reading is Roth and Sotomayor, *Two-Sided Matching A study in Game-Theoretic Modeling and Analysis*, Cambridge.

For empirical aspects, see Bryan Graham's survey chapter in the *Handbook of Social Economics*:
https://files.nyu.edu/bsg1/public/
HandbookOfSocialEconChapter_Final_June10.pdf.

For the more mathematical aspects, the monograph by Cédric Villani, *Topics in Optimal Transportation*, AMS, is recommended.

Assessment. Students taking this course for credit should write a paper relevant to an aspect of the course, to be discussed with the instructor. This course will be graded on a pass/fail basis.

This course

The course will focus on the economic theory of matching from a number of points of view. It is intended to give an overview of the fundamental theory of the optimal assignment problem, as well as its application to various fields such as labor, family and transportation economics, focusing first on models with transferable utility. A particular emphasis is put on the empirical aspects and identification issues, and the main matching algorithms will also be discussed. The last part of the course tries to make a link with matching games with nontransferable (or partially transferable) utility and attempts to provide a unified treatment.

 These lecture notes were prepared while the author was visiting Columbia University in the Spring 2011. He thanks the Economics department for their hospitality and the Alliance program for its support. The author also thanks the course participants as well as Arnaud Dupuy for useful comments which helped improving these notes.

Course outline

Part I. Matching with Transferable Utility (TU) Feb 1,8,15, 22. 11am-1pm.

General introduction to matching. Optimal matching and duality. Optimal transportation theory. Transportation on networks.

Part II. Empirical issues in TU models. March 1, 8, 15, 22, 29. April 5, 12, 19, 26. 11am-1pm.

Identification and estimation issues. Economics of the family. Hedonic models. Labour economics. Revealed preferences

There will be two make-up classes which are informal "crash courses" on more technical aspects of the subject. Attendance is encouraged, but not necessary for following the rest of the course.

Feb 8, 9–11am. **Crash course on convex analysis and linear programming.**

Feb 22, 9-11am. **Crash course on algorithms and computational issues.**

Part I. Matching with transferable utility

1 General introduction

1.1 Models

In this section we shall review a variety of models that model various applied situations. The goal here is to give a feel of different classes of matching models, their similarities, differences, and their relevance in different settings. We start with arguably the simplest matching model: the housing problem of Shapley and Scarf (1974).

Example 1.1 (The housing problem without money). *Consider a situation where there are N individuals (or families) labeled by $i = \{1, ..., N\}$, and N houses labeled by $j = \{1, ..., N\}$. Each individual needs exactly one house; each house can only accommodate one individual. Individuals form preferences over houses:*

$j_1 \leq_i j_2$ (resp $<$) means "i weakly (resp. strictly) prefers house j_2 over house j_1".

We shall assume that these preferences are bona fide *preferences, that is they satisfy completeness, transitivity, and reflexivity.*

Assignments, permutations. An *assignment* of houses to individuals is a map $\sigma : \{1, ..., N\}$ to $\{1, ..., N\}$ such that individual i will be matched with house j. Because no house can accommodate two individuals, σ should be one-to-one; because there are as many houses as there are individuals, σ should be onto. Thus σ is invertible; $\sigma^{-1}(j)$ is the individual living in house j. σ is called a *permutation* of $\{1, ..., N\}$. The set of permutations of $\{1, ..., N\}$ is denoted $\mathfrak{S}_N$.

For now on we shall assume that when houses are initially assigned, initial assignment consists in assigning house i to individual i, hence initial assignment is $\sigma_0(i) = i$. This is obviously w.l.o.g. (without loss of generality), as one can always relabel houses.

Swaps, cycles. Now, given an initial assignment of houses, a pair of individuals have the possibility to swap houses, if they each decide to do so.

Formally, a *swap* is a situation in which i_1 gets i_2's house and i_2 gets i_1's house. One says that assignment σ contains the swap (i_1, i_2) if

$$i_2 = \sigma(i_1) \text{ and } i_1 = \sigma(i_2).$$

Swaps are a possible scheme for the agents, but there may be more involved ones. There can be a situation where i_1 gets i_2's house, i_2 get i_3's house, and so on until someone, call her i_p, gets i_1's house. Such an assignment is called a *cycle*. One says that σ contains the cycle $(i_1, ..., i_p)$ if

$$i_2 = \sigma(i_1), \; i_3 = \sigma(i_2), \; ..., i_p = \sigma(i_{p-1}), \; i_1 = \sigma(i_p)$$

$i_2 = \sigma(i_1)$, $i_3 = \sigma(i_2)$,..., $i_p = \sigma(i_{p-1})$ and $i_1 = \sigma(i_p)$. For convenient reasons that will appear later on, we shall take the convention to identify i_{p+1} and i_1. We shall denote $\mathfrak{C}_N$ the set of cycles over $\{1, ..., N\}$; clearly $\mathfrak{C}_N \subset \mathfrak{S}_N$.

In principle (that is, without taking individual preferences into consideration), any assignment can be obtained by a sequence of house swaps, irrespective of the initial assignment. In fact, we have the following mathematical result:

Proposition 1.1. *One has:*
(i) Any permutation can be obtained by a composition of disjoint cycles.
(ii) Any cycle can be obtained by a composition of swaps.
(iii) Hence, any permutation can be obtained by a composition of swaps.

Proof. Consider some permutation σ. Take any individual i_1, and let $i_2 = \sigma(i_1)$, and so on until $i_1 = \sigma(i_p)$ at which point we have got a cycle. Iterate with an element not in the cycle until there is no such element. σ is the composition of the cycles obtained so, which proves (i). (ii) follows from the fact that the cycle $(i_1, ..., i_p)$ coincides with the composition of swaps (i_1, i_{p-1}), (i_{p-1}, i_{p-2}),..., (i_3, i_2) (the iterated application from right to left). □

Again, in the statement of this result we have ignored the fact that agents have preferences, so some agents may not agree on some of these swaps. Hence not any assignment will be a reasonable outcome, even if it is feasible. We shall come back on this crucial point when we discuss solution concepts.

Example. *Consider the permutation*

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 1 & 3 \end{pmatrix}$$

meaning that $\sigma(1) = 4$, $\sigma(2) = 5$, *etc. One sees that this permutation is the product of cycles* $(1\ 4)$ *and* $(2\ 5\ 3)$*; and cycle* $(2\ 5\ 3)$ *is the composition (iterated application from right to left) of swaps* $(5\ 3)$ *and* $(2\ 3)$*.*

Cardinal reformulation. Note that the problem stated as such in the example above was purely ordinal: the input of the problem is whether Mrs. Smith prefers house A to house B, not *by how much* she prefers it. In an economy without money, this does not matter. One cardinal way to rephrase it is by introducing

$$U(i,j) = \text{utility brought to } i \text{ by house } j.$$

Of course, there is too much information in this setting: it is equivalent to any reformulation replacing $U(i,j)$ by $\tilde{U}(i,j)$ where $\tilde{U}(i,j) = \varphi_i(U(i,j))$, $\varphi_i(.)$ increasing. We now turn to an example where the intensity of preferences matters because there is a second good which serves as a numéraire.

Example 1.2 (The housing problem with money). *Consider the previous problem, where we introduce a divisible exchangeable good (money), and assume that utilities are semi-linear w.r.t. money. Letting* π *be the payment made by individual* i*, assume that the utility of individual* i *that has been assigned house* j *and has to make payment* π *is*

$$U(i,j;\pi) = U(i,j) - \pi.$$

In this economy, house j *has price* p_j*; and if agent* i *is initially matched to house* j_1 *but trades it for house* j_2*, his net payment will be* $p_{j_2} - p_{j_1}$ *and its resulting utility shall be*

$$U(i,j_2) - p_{j_2} + p_{j_1}$$

Here, the problem is obviously a cardinal problem, as the intensity of preferences will be somehow reflected in the prices at which agents are willing to trade houses.

Here again, the trading possibilities are not limited to swaps. In general, there may be a trading cycle $(i_1, ..., i_p)$ *where individual* i_k *buys* i_{k+1}*'s house and sells her house to* i_{k-1}*. In that case, the resulting utility of* i_k *will be*

$$U\left(i_k, i_{k+1}\right) - p_{i_{k+1}} + p_{i_k}.$$

Again, we shall study later on in this course what will be economically plausible outcomes of this economy.

Transferable Utility vs. Non Transferable Utility. There is a very important distinction between the models in Example 1 and Example 2. In Example 1, agents have no way to compensate each other by utility transfers. The model in Example 1 is thus said to belong to *Non-transferable utility (NTU) models* , while that of Example 2 is one instance of the class of *Transferable Utility (TU) models.* This distinction is fundamental and will shape this course. Transferable utility models usually have less problematic, "more unique" solutions, are more and more often identified. Non-transferable utility models are more discrete in nature, have usually more solutions, are more pervasive to strategic considerations, and are more prone to partial identification. The techniques used to provide solutions are quite different too. NTU models are usually dealt with using discrete methods such as graph theory or lattice theory. TU models are often solved using convex analysis or linear programming methods.

Although the link between these two classes of models remains for the most part an open problem in Economic theory, there are very interesting links between these two classes of models. We shall discuss these links at the end of the course.

Bilateral preferences. In the previous two example, houses do not have their word to say about who is moving in: they do not form preferences over individuals. This may not be so problematic in the case of houses, but may be more so in other matching situations, marriage for instance, where one should take the preferences of both sides into consideration.

Example 1.3 (The marriage problem without transfers). *We take the same setting as in Example 1 where we now have a population of* N *men and* N *women, instead of individuals and houses.* $i \in \{1, ..., N\}$ *now represents men and* $j \in \{1, ..., N\}$ *now represents women. The novelty is that now both*

sides form preferences over the other side:
$j_1 \leq_i j_2$ *(resp* $<$*) means "man* i *weakly (resp. strictly) prefers woman* j_2 *over woman* j_1*"; and*
$i_1 \leq_j i_2$ *(resp* $<$*) means "woman* j *weakly (resp. strictly) prefers man* i_2 *over man* i_1*".*

As before – and although it is of course unnecessary at this stage – we introduce a cardinal representation of these preferences:
$U(i,j)$ is the utility that Mr. i gets if matched to Mrs. j; and $V(i,j)$ is the utility that Mrs. j gets if matched to Mr. i.

Unsurprisingly, the marriage problem has also a version with transfers. There may be several types of transfers which we may consider: within genders only (i.e. the net balance of utility transfer among men should be zero and similarly among the women), across the population (i.e. the net balance of utility transfers among the whole population of men and women alike should be zero), or within household. Consistently with the literature on economics of marriage (starting with Becker 1973), we shall consider the latter problem, where if Mr. i and Mrs. j marry they may decide on a transfer which will decrease one of the partner's utility and increase the other's.

Example 1.4 (The marriage problem with transfers). *We take the same setting as in Example 3 where we now allow for utility transfers between partners: if Mr.* i *and Mrs.* j *are matched and decide on transfer* t *from Mr.* i *to Mrs.* j *(the direction of the transfer is a pure convention as the sign of* t *is unrestricted), then:*
Mr. i *gets utility* $U(i,j) - t$*; and*
Mrs. j *gets utility* $V(i,j) + t$*.*

We can give a first typology in Figure 1.1 below.

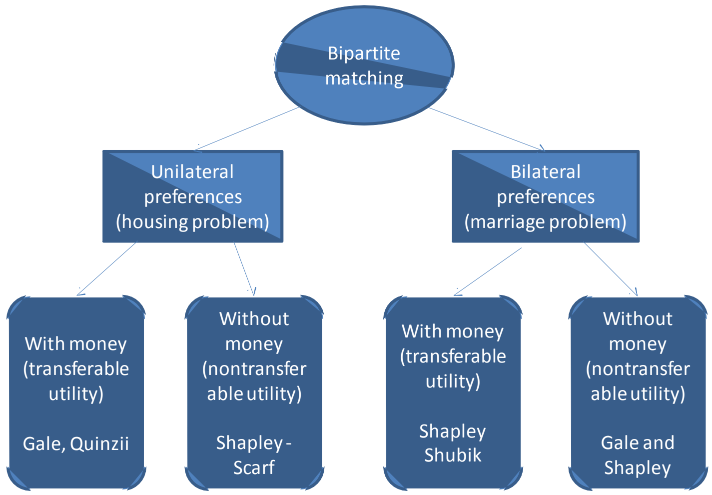


Figure 1.1: A typology of some matching problems

Besides the four basic problems of matching we just reviewed, there are a number of other matching problems of interest.

First, we may relax the assumption that there are as many players on each side of the market.

Example 1.5 (The marriage problem with singles). *Same as the marriage problem with or without transfers, but we no longer assume that there is an equal number of men and women, but we assume that man i has a reservation utility $U_0(i)$ for remaining single, and $V_0(j)$ for woman j.*

The reservation utilities will induce a participation constraint, which determines endogenously the population that chooses to match. Note that one recovers the classical cases by setting reservation utilities equal to $-\infty$ when there is the same number of men and women.

Bipartite matching, So far in all the matching we saw there were two populations: individuals and houses, men and women, etc. This class of problems is called *bipartite matching*. We may relax the assumption.

Example 1.6 (The marriage problem with single-sex households (or the roommate problem)). *Same as the marriage problem with or without transfers and with single, but we no longer assume that there are men and women: individual i will get utility $U(i,i') - t_{ii'}$ when matched with individual i', if $t_{ii'}$ is the transfer decided from individual i to i'.*

Note that one recovers the previous cases by setting $U(i,i')$ equal to minus infinity when i and i' have the same gender.

Pure vs. fractional matching. So far all the matchings we saw were *pure*, in the sense that we ruled out the possibility of splitting either side of the market. E.g. in the marriage problem we forced one man to marry one woman. But assume we have agents (workers) on one side of the market and machines on the other side of the market, and assume that each agent can split the time she spends on a given machine. Then we are led to a class of problem called fractional matching problems. Let us give the most famous of them.

Example 1.7 (The assignment problem). *Consider N agents $i \in \{1,...,N\}$ and N machines $j \in \{1,...,N\}$. Assume that when i works on machine j they produce together quantity Φ_{ij} of output per amount of time. The daily amount of time (of agents and machines alike) is normalized to one. Let π_{ij} be the amount of time that agent i spends on machine j. The constraints on the production plan π are that both agents and machines are fully employed: $\pi_{ij} \geq 0$ and*

$$\begin{aligned}\sum_{i=1}^{n} \pi_{ij} &= 1 \text{ for each } j\text{, and} \\ \sum_{j=1}^{n} \pi_{ij} &= 1 \text{ for each } i.\end{aligned}$$

Given a production plan π satisfying the constraints above, the total output is

$$\sum_{ij} \pi_{ij}\Phi_{ij}.$$

One-to-many matching. Until now the matching situations we discussed were one-to-one matching: we did not allow one single agent to match with more than one agent. But friendship, for instance, is a situation of many-to-many matching. There are many other interesting situations, as in the study of labour markets, where this assumption should be removed.Our next and final two examples illustrate situations of many-to-one matching.

Example 1.8 (The labour market problem). *Assume there are $\mathcal{I} = \{1, ..., I\}$ workers and $\mathcal{J} = \{1, ..., J\}$ firms. Let $u_i(j, w)$ be the utility of workers i working for firm j with salary w. Let C^j be the set of workers hired by firm j. Let $v_j(C^j)$ be the valuation by firm j of set of workers C^j. Its net profit is*

$$\pi_j\left(C^j, w_j\right) = v_j\left(C_j\right) - \sum_{i \in C_j} w_j^i.$$

Example 1.9 (The fountain problem). *Consider a city, which we shall assume for simplicity is modeled as the square $[0,1]^2$. This city is uniformly populated. We assume that there are N fountains on the surface of the city, located at point $Y_1, ..., Y_N$. Fountains charge different prices for access to water: π_k is the price charged by fountain k. Assume that inhabitant located at $x \in [0,1]^2$ has a utility for using fountain k of the form*

$$U(x, Y_k) - \pi_k$$

for instance, it may be the case that

$$U(x, Y_k) = -|x - Y_k|^2,$$

in which case the utility of a given inhabitant depends on the squared Euclidian distance to the fountain (transportation cost), and on the price charged.

We are going to encounter many of these problems throughout this course. So far we have seen several important classes of models: TU/NTU models; models with or without singles; models with unilateral or bilateral preferences; bipartite or non-bipartite models; one-to-one, one-to-many, many-to-many; fractional or pure matching. Each situation to be modeled calls for a different setting. We are now going to review important solution concepts.

1.2 Properties of solutions

Pareto efficiency. The most obvious solution one may think of is Pareto efficiency. A matching is Pareto efficient if no other matching offers a better welfare to anyone, weakly for all and strictly for some. In some cases it is the right concept, in some others it is not.

When preferences are unilateral it is natural to look for Pareto efficient solutions. In the NTU housing problem of Example 1, for instance, an assignment σ is Pareto efficient if there is no other assignment $\tilde{\sigma}$ such that $\tilde{\sigma}(i) \geq_i \sigma(i)$ for all i with at least one of these preferences being strict. For any assignment σ that is not Pareto efficient there will be an improving cycle, namely a cycle $(i_1, ..., i_p)$ such that $\sigma(i_{k+1}) \geq_{i_k} \sigma(i_k)$, with at least one strict inequality. Therefore, if an assignment is not Pareto efficient, agents will have an incentive to form a coalition as an improving cycle.

Stability. It turns out, however, that this is not always the most natural property in many situations, mainly in situations with bilateral preferences. Consider for example, a NTU marriage example where there are two men i_1 and i_2 and two women j_1 and j_2, and each men prefer j_1 to j_2 and each women prefer i_1 to i_2. A matching where i_1 is matched to j_2 and i_2 is matched to j_1 is Pareto optimal: indeed, the only alternative is to marry i_1 to j_1 and i_2 to j_2; but moving to that outcome would of course hurt i_2's welfare, hence it is not a Pareto improvement. But this matching is, however, quite implausible: indeed, both i_1 and j_1 would be willing to leave their current partner and form a match – at least in a situation where divorce is allowed.

One sees that in many settings with bilateral preferences, the operational concept is stability. A matching is *stable* if no two unmatched agents have an incentive in leaving their current partner and forming a pair. Such a pair is called a *blocking pair*.

Of course, this is a stronger concept than Pareto efficiency.

Proposition 1.2. *In the marriage model (TU or NTU), stability implies Pareto efficiency.*

Proof. We will show the reverse statement. Assume a matching is not Pareto efficient. Then there is a Pareto improvement, which is strictly preferred by at least one agent. Then this agent and his/her match under the Pareto improving matching form a blocking pair, hence the matching is not stable. □

The previous concepts were normative, but it is of interest to understand what may take place when starting from an initial allocation, typically not Pareto efficient, either in a decentralized economy (equilibrium) or in the presence of a mechanism.

Equilibrium. In many actual examples of matching (such as marriage, labor market, housing market, etc.) the economy is decentralized, and there is no central matching mechanism. This leads to the question of existence of an equilibrium, mostly in TU models. In the TU housing problem, for instance, one may wonder whether the houses have an equilibrium price, which would adjust supply to demand. In the fountain example, one may wonder whether there is a set of prices that adjust supply to demand. We will get back to these questions later.

Mechanism. There are significant examples where there is a centralized mechanism that performs the matching. In the US, the National Intern Matching Program which matches the interns to hospitals, is an example of such mechanisms. Many school districts also have a centralized matching procedure.

1.3 Questions

When facing a matching situation, a number of questions arise:

1. **What model?** what model best describes the economic situation at hand: TU or NTU? fractional or pure? should one make participation endogenous or not?

2. **What solution concept?** it will be tightly linked to the rules of the game that agents play: legal environment, market structure, etc. For instance, in the context of the marriage market, in societies where divorce is permitted, stability is a plausible assumptions. In others, we may have to focus on Pareto optimality.

3. **Strategic issues:** are player's types fully observable by other agents? is truth-telling a dominant strategy?

4. **Empirical implications:** what are the empirical implications of the model? is it identified?

1.4 References and notes

Example 1 is Shapley and Scarf (1974) [99]. Example 2 is Gale (1984) [50] and Quinzii (1987) [85]. Example 3, as example 5 is Gale and Shapley (1962) [51], and Example 4 is Becker (1973) [7] and Shapley and Shubik (1971) [100]. Example 6 is in Gale and Shapley and has been further studied by Granot (1984) [56], Gusfield (1988) [59], and Irving (1985) [69]. Example 7 is Koopmans and Beckmann (1957) [72], building on Dantzig's pioneering work on Linear Programming in the 1930's, see [35]. Example 8 is Kelso and Crawford (1982) [70].

2 Two important examples

We are now going to fully work out two examples that were previously discussed: the optimal assignment problem of machines to agents (Example 1.7), and the fountain problem (Example 1.9). Although these two problems will appear quite different in nature, and in the techniques used, our goal in the next chapter will be to relate them through the use of the Monge-Kantorovich theorem and optimal transport theory.

2.1 The fountain problem

In the fountain problem (Example 1.9) we considered a city, which was modeled as the square $\Omega = [0,1]^2$, on which dwells a uniformly distributed population. Let λ be the Lebesgue measure which is the distribution of the inhabitants; we normalize the total number of inhabitants to one. There are N fountains on the surface of the city, located at point $Y_1, ..., Y_N$. Fountains charge different prices for access to water: π_k is the price charged by fountain k. Assume that inhabitant located at $x \in [0,1]^2$ has a utility for using fountain k of the form

$$U(x, Y_k) - \pi_k$$

where we will choose

$$U(x, Y_k) = -|x - Y_k|^2,$$

in which case the utility of a given inhabitant depends on the squared Euclidian distance to the fountain (transportation cost), and on the price charged.

The capacity of a fountain is measured in terms of the quantity of individuals it can serve. We assume that fountain k has capacity q_k; and that $\sum_{k=1}^{N} q_k = 1$, that is aggregate supply equal aggregate demand.

If there is a central planner, there does not need to be prices. The central planner will assign to each inhabitant x a fountain $\tau(x) \in \{1, ..., N\}$, in such a way that each fountain k is used to full capacity, hence

$$\lambda\left(\tau^{-1}(\{k\})\right) = q_k, \ \forall k. \tag{2.1}$$

While there are many such choices, the central planner may want to choose to optimize some welfare criterion under the balance constraints (2.1).

One natural welfare criterion to think of is the average utility of inhabitants. Hence, one may look for

$$\begin{aligned} W &= \max \int_{\Omega} U(x, \tau(x))\, d\lambda(x) \\ & \quad s.t.\ (2.1) \end{aligned} \tag{2.2}$$

and the first question is:

1. How to determine the optimal assignment τ of fountains to individuals?

A second question we may ask is about the decentralized market version of this economy. As we saw, aggregate supply equals the aggregate demand. But it does not mean that the market clears. Indeed, there may be regions where there are a higher concentration of fountains, and regions where there will be a lower concentration of them. Without prices, or if prices cannot perfectly signal scarcity and adjust demand and supply, some of these fountains will be in excess supply, some of them in excess demand. On the fountains where there is excess demand, there will be shortages, congestions etc. and the demand will be imperfectly served. We thus ask:

2. Is there a system of prices such that the market clears, i.e. such that the demand for fountain k is exactly q_k?

As it will turn out, the two questions are intimately related. To study them, we need to have a better understanding of how the demand responses to the evolution of prices. Assume for a moment that fountains do not charge for access to water. Then the inhabitants will choose the fountain that maximizes their utility, that is each inhabitant will choose the fountain that is closest. Hence there is a partition of the surface of the city Ω into regions Ω_k such that each x in Ω_k is closest to fountain Y_k than any other fountain. (In fact it is not quite a partition as there are inhabitants indifferent between two fountains, sometimes more – but this does not really matter).

The graph it leads to is called a *Voronoi tesselation*. Of course, the demand for fountain Y_k is $\lambda(\Omega_k)$ the area of Ω_k which does sum to 1, but has no reason to be equal to supply q_k. Hence some fountains will be over-demanded ($\lambda(\Omega_k) > q_k$) while some will be under-demanded ($\lambda(\Omega_k) < q_k$).

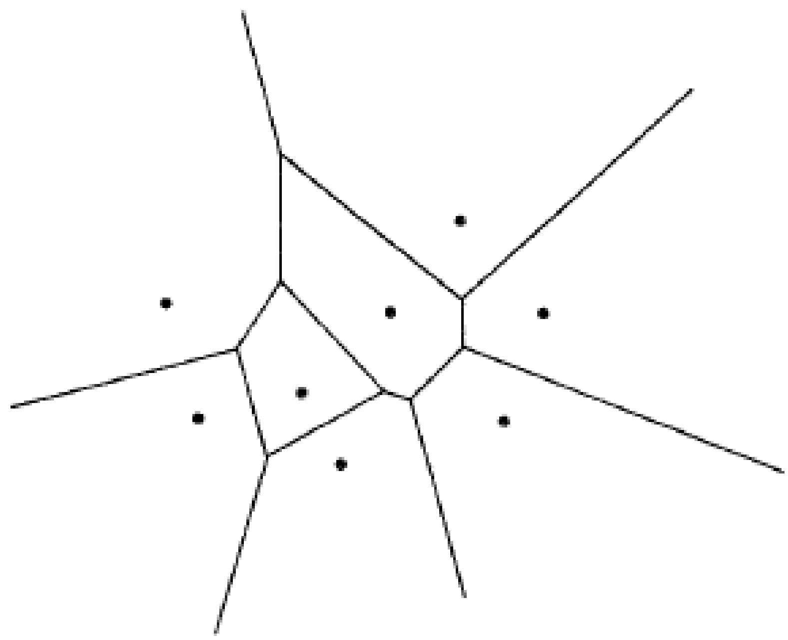


Figure 1: Source: Aurenhammer (1991).

Let us now introduce a system of prices π, where π_k is the price of fountain k. Intuitively, we would like to raise the price of those fountains that are over demanded, and lower the price of the under-demanded fountains. Let us now determine the area of consumers who will choose fountain k. We have

$$\Omega_k(\pi) = \left\{ x : |x - Y_k|^2 - \pi_k = \min_{i=1\dots N} |x - Y_i|^2 - \pi_i \right\}.$$

This graph is called a *power diagram*, and it is a straightforward generalization of a Voronoi tesselation (which can be obtained by taking $\pi = 0$). This graph has the following properties.

Proposition 2.1. *One has:*

(i) Each Ω_k is a convex polyhedron.

(ii) The intersection of two Ω_k's is either empty, or is a line segment.

(iii) The sets $\Omega_k(\pi)$ are decreasing with respect to π_k, all other prices being fixed.

(iv) The sets $\Omega_k(\pi)$ are left unchanged when the same constant c is added to all the prices π_k.

Proof. Left as an exercise. ■

We can introduce the indirect utility of inhabitant x as the utility given to her by her most favorite fountain.

$$U^*(x, \pi) = \max_{i=1\dots N} \left(U(x, Y_i) - \pi_i \right).$$

We now recall the optimal welfare problem (2.2)

$$\begin{aligned} W &= \max_{\tau(.)} \int_\Omega U(x, \tau(x))\, d\lambda(x) \\ & \quad s.t.\ \lambda\left(\tau^{-1}(\{k\})\right) = q_k,\ \forall k. \end{aligned}$$

Theorem 2.1. *One has:*

(i) the social welfare W is equal to

$$W = \min_\pi \int U^*(x, \pi)\, dx + \sum_k \pi_k q_k$$

(ii) There is a set of prices π that maximize W; for these prices

$$\lambda\left(\Omega_k\left(\pi\right)\right) = q_k$$

hence markets clear.

(iii) The optimal assignment in the optimal welfare problem (2.2) is given by

$$\tau\left(x\right) = \arg\min_{i=1\ldots N}\left(\left|x - Y_i\right|^2 - \pi_i\right).$$

Proof. (i) The full proof of this statement will be given in the next chapter; here we shall only give an informal justification of the proof. Start from

$$\begin{aligned} W &= \max_{\tau(.)}\int_{\Omega} U\left(x, \tau\left(x\right)\right) d\lambda\left(x\right) \\ & \quad s.t.\ \lambda\left(\tau^{-1}\left(\{k\}\right)\right) = q_k,\ \forall k. \end{aligned}$$

One has

$$\begin{aligned} W = \max_{\tau(.)}\min_{\pi} \quad & \int_{\Omega} U\left(x, \tau\left(x\right)\right) d\lambda\left(x\right) + \sum_k \pi_k q_k \\ & - \int \pi_{\tau(x)} d\lambda\left(x\right) \end{aligned}$$

where the maximization is now unconstrained. Hence

$$W = \max_{\tau(.)}\min_{\pi}\int_{\Omega} U\left(x, \tau\left(x\right)\right) - \pi_{\tau(x)} d\lambda\left(x\right) + \sum_k \pi_k q_k.$$

Assuming that we can exchange the min and the max, which is not at all obvious and will be only justified later on in the course, we get

$$W = \min_{\pi}\max_{\tau(.)}\int_{\Omega} U\left(x, \tau\left(x\right)\right) - \pi_{\tau(x)} d\lambda\left(x\right) + \sum_k \pi_k q_k,$$

hence

$$W = \min_{\pi}\int_{\Omega}\max_k\left(U\left(x, k\right) - \pi_k\right) d\lambda\left(x\right) + \sum_k \pi_k q_k$$

which proves (i).

Let us now turn to (ii). Note that

$$F(\pi) = \int U^*(x,\pi)\,d\lambda(x) + \sum_k \pi_k q_k$$

is a convex function of π. Hence there is an optimal price system π. By first order conditions, at the optimal π

$$\frac{\partial F(\pi)}{\partial \pi_k} = 0$$

thus

$$\int \frac{\partial U^*(x,\pi)}{\partial \pi_k} d\lambda(x) = q_k.$$

But by the envelope theorem.

$$\frac{\partial U^*(x,\pi)}{\partial \pi_k} = 1\{\tau(x) = k\}$$

hence the first order constraints rewrite

$$\int 1\{\tau(x) = k\}\,d\lambda(x) = \lambda\left(\tau^{-1}(\{k\})\right) = q_k,$$

which proves (ii) and (iii) follows. ■

Numerical computation of the prices. Lastly, we turn to a discussion on the numerical determination of the prices. The function F to minimize being convex, it is standard to use a descent algorithm: starting at a given point π^0, update at each step by $\nabla F(\pi^t)$

$$\pi^{t+1} = \pi^t - \varepsilon \nabla F\left(\pi^t\right),$$

where $\nabla F(\pi)$ is the vector $\left(\frac{\partial F(\pi)}{\partial \pi_1}, ..., \frac{\partial F(\pi)}{\partial \pi_N}\right)$. But

$$\frac{\partial F(\pi^t)}{\partial \pi_k} = q_k - \lambda\left(\Omega_k\left(\pi^t\right)\right)$$

hence the descent algorithm consists in

$$\pi^{t+1} - \pi^t = \varepsilon\left(\lambda\left(\Omega_k\left(\pi^t\right)\right) - q_k\right),$$

which takes immediately an economic interpretation: the fountains that are over-demanded *raise* their prices, while the fountains that are under-demanded *lower* their prices. This is exactly the mechanism of the *tâtonnement process*, first imagined by Walras and formalized by Samuelson.

2.2 The optimal assignment problem

Recall the assignment problem (Example 1.7). There are N agents $i \in \{1, ..., N\}$ and N machines $j \in \{1, ..., N\}$. Assume that when i works on machine j they produce together quantity Φ_{ij} of output per amount of time. Φ_{ij} will be referred to as the *production function*. The daily amount of time (of agents and machines alike) is normalized to one. Let π_{ij} be the amount of time that agent i spends on machine j. The constraints on the production plan π are that both agents and machines are fully employed: $\pi_{ij} \geq 0$ and

$$\sum_{i=1}^{N} \pi_{ij} = 1 \text{ for each } j\text{, and} \tag{2.3a}$$

$$\sum_{j=1}^{N} \pi_{ij} = 1 \text{ for each } i. \tag{2.3b}$$

Note that these constrains just mean that each agent and machine should be fully employed. Such a π is called a *bistochastic matrix*.

Definition 2.1. *A bistochastic matrix π_{ij} is a matrix satisfying $\pi_{ij} \geq 0$ and equations (2.3a) and (2.3b) above. We shall denote by $\mathcal{B}$ the set of bistochastic matrices.*

Given a production plan π satisfying the constraints above, the total output to be maximized is

$$\sum_{ij} \pi_{ij} \Phi_{ij}. \tag{2.4}$$

Hence, the firm's program is to maximize (2.4) under constraints (2.3a) and (2.3a). Note that the set of bistochastic matrices is compact, hence an optimal production plan π always exist.

2.2.1 Linear programming formulation

The firm's problem is a linear programming problem, and it is feasible. Hence it is equal to its dual formulation.

Theorem 2.2. *The linear program*

$$V = \max_{\pi_{ij} \geq 0} \sum_{1 \leq i,j \leq N} \pi_{ij} \Phi_{ij} \tag{2.5}$$

$$s.t. \quad \sum_{i=1}^{N} \pi_{ij} = 1 \textit{ for each } j$$

$$\sum_{j=1}^{N} \pi_{ij} = 1 \textit{ for each } i.$$

is dual to the program

$$V = \min_{u_i, v_j} \sum_{i} u_i + \sum_{j} v_j \tag{2.6}$$

$$s.t. \ u_i + v_j \geq \Phi_{ij}$$

and the values of these programs coincide. Further, if π is a solution of (2.5) and (u, v) is a solution of (2.6), we have by complementary slackness that

$$\pi_{ij} > 0 \textit{ implies } u_i + v_j = \Phi_{ij}.$$

Proof. Start with the primal problem

$$\begin{aligned} V &= \max_{\pi_{ij} \geq 0} \sum_{1 \leq i,j \leq N} \pi_{ij} \Phi_{ij} \\ s.t. & \quad \sum_{i=1}^{N} \pi_{ij} = 1 \text{ for each } j \\ & \quad \sum_{j=1}^{N} \pi_{ij} = 1 \text{ for each } i. \end{aligned}$$

and introduce u_i and v_j the Lagrange multipliers of the constraints, we can rewrite the problem as a min-max problem

$$
\begin{aligned}
V &= \max_{\pi_{ij}\geq 0}\min_{u_i,v_j} \sum_{1\leq i,j\leq N} \pi_{ij}\Phi_{ij} + \sum_i u_i - \sum_{ij} u_i\pi_{ij} \\
&\quad + \sum_j v_j - \sum_{ij} v_j\pi_{ij} \\
&= \max_{\pi_{ij}\geq 0}\min_{u_i,v_j} \sum_i u_i + \sum_j v_j \\
&\quad + \sum_{1\leq i,j\leq N} \pi_{ij}\left(\Phi_{ij} - u_i - v_j\right)
\end{aligned}
$$

Hence, as the primal equals the dual, we get

$$
\begin{aligned}
V &= \min_{u_i,v_j}\max_{\pi_{ij}\geq 0} \sum_i u_i + \sum_j v_j \\
&\quad + \sum_{1\leq i,j\leq N} \pi_{ij}\left(\Phi_{ij} - u_i - v_j\right)
\end{aligned}
$$

and we now see that the π_{ij} become the Lagrange multipliers associated to constraint $u_i + v_j \geq \Phi_{ij}$. Hence, we arrive at the dual formulation

$$
\begin{aligned}
V &= \min_{u_i,v_j} \sum_i u_i + \sum_j v_j \\
s.t. &\quad u_i + v_j \geq \Phi_{ij}.
\end{aligned}
$$

■

Let us now give an economic interpretation of this result. Assume that, instead of having a planning bureau decide on the optimal production plan, the firm decides to externalize its operations, and hire external agents to work on its machines on an external, competitive market. Let v_j be the market price of output of machine j per hour. The firm now has to pay an hourly market wage w_i to agent i it hires. Then if agent i worked time π_{ij} on machine j, the profit of the firm out of this combination is $\pi_{ij}\left(v_j - w_i\right)$, hence total profit of the firm is

$$
\sum_{ij} \pi_{ij}\left(v_j - w_i\right) = \sum_j v_j - \sum_i w_i.
$$

Now, note that $v_j - w_i \geq \Phi_{ij}$ – otherwise the firm would choose to operate by itself. Hence

$$\sum_j v_j - \sum_i w_i \geq \sum_{ij} \pi_{ij} \Phi_{ij}$$

but if the markets are perfectly competitive, the firm should be indifferent between hiring external agents or operate on its own – otherwise external agents would have a rent. Hence the above inequality is actually an equality. Setting $w_i = -u_i$, we recover the duality of Theorem 2.2.

Also, note that we do not change the solution to the dual problem if we add a constant c to the u_i's and subtract it to the v_j's. Hence, the salaries and the prices are not perfectly determined in this model; in order to determine them one would need an alternative technology which would provide agents (or machines) with a reservation utility, and make participation endogenous. We will turn to this issue at section 2.2.4 below.

2.2.2 Purity

It turns out one can say more about the solution. We shall in fact show that the solution may be chosen pure, that is that the optimal π can be taken such that $\pi_{ij} \in \{0, 1\}$. In other words, there is an optimal solution in (2.5) such that each agent i works for one and only one machine $\sigma(i)$, and such that each machine j is operated by one and only one agent $\sigma^{-1}(j)$. Of course, σ is a permutation matrix. In terms of the production plan π, this means that it is a *permutation matrix* in the following sense.

Definition 2.2. *A matrix π is a permutation matrix if there is a permutation $\sigma \in \mathfrak{S}_N$ such that $\pi_{ij} = 1_{\{j=\sigma(i)\}}$. We will denote π^σ this matrix.*

In a more economical parlance, we shall refer to permutations as *deterministic assignments*, bistochastic matrices as *fractional assignments*.

Let us give some examples of permutation matrices. Matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

are permutation matrices.

Obviously, a permutation matrix is a bistochastic matrix. But the converse does not hold. For instance,

$$\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

is bistochastic but is not a permutation matrix.

Now consider the problem of the firm with the additional constraint that agents cannot "multitask", that it, one agent should work for one and only one machine. This may be due to switching costs, etc. The firm's objective is then

$$\max_{\sigma \in \mathsf{S}_N} \sum_{i=1}^{N} \Phi_{i\sigma(i)}.$$

Quite obviously, as permutation matrices are a subset of bistochastic matrices, we have

$$\max_{\sigma \in \mathsf{S}_N} \sum_{i=1}^{N} \Phi_{i\sigma(i)} \leq V.$$

Amazingly, it turns out that the inequality is actually an equality. In other words, the constraint that agents cannot multitask does not hurt the firm. This is expressed in the following theorem.

Theorem 2.3. *In the program (2.5) above, the solution may be chosen so that π is a permutation matrix, hence*

$$\max_{\sigma \in \mathsf{S}_N} \sum_{i=1}^{N} \Phi_{i\sigma(i)} = V.$$

Note that this theorem does not imply that every solution to the assignment problem is pure; it says that there exists a pure solution (among a set of possibly many solutions). The proof is extremely important, as it is based on one of the most important ideas in bilateral matching.

Proof. Let π be a solution of program (2.5), and assume that π is not a permutation matrix. Then there are indices i_0 and j_1 such that $\pi_{i_0 j_1} \in (0,1)$. We can then construct a sequence of indices $i_1, j_2, i_2...$ such that $\pi_{i_1 j_1}$, $\pi_{i_1 j_2}$, $\pi_{i_2 j_2}$ all are in $(0,1)$. Repeating the process, one index, either among the i's or the j's, is going to be repeated twice: doing so we have isolated a cycle. This is illustrated on Figure 2. Let x be the smallest value of π_{ij} along the cycle. We remove x from the corresponding entry, and alternatively add and remove x to π along this cycle. On Figure 2 for example, $x = 0.3$ and we remove 0.3 to the squares and add 0.3 to the circles. The variation to the objective function is therefore

$$0.3\left(\Phi_{3A} - \Phi_{1A} + \Phi_{1B} - \Phi_{4B} + \Phi_{4D} - \Phi_{3D}\right).$$

The quantity in parenthesis must be zero; otherwise it would be possible to improve the objective function while preserving the constraints, which would violate the fact that one is at optimum. Hence alternatively adding and removing x along the cycle does not change the value of the objective $\sum_{ij} \pi_{ij} \Phi_{ij}$. Doing so, we have decreased the number of fractional entries while preserving the bistochasticity constraint, and leaving the objective function unchanged. We can iterate the process until the resulting matrix is a permutation matrix. ■

2.2.3 Geometric interpretation and consequences for identification

The above results can be reinterpreted geometrically as properties of the set $\mathcal{B}$ of bistochastic matrices. We begin by a series of obvious observations we state as lemmas. The proofs are left as exercises.

Lemma 2.1. *The set of bistochastic matrices is convex and compact.*

As a consequence an optimal solution to program (2.5) always exists, and is on the boundary of the set of bistochastic matrices $\mathcal{B}$.

Lemma 2.2. *The set of production plans π that are optimal for some production function Φ_{ij} is the boundary of $\mathcal{B}$.*

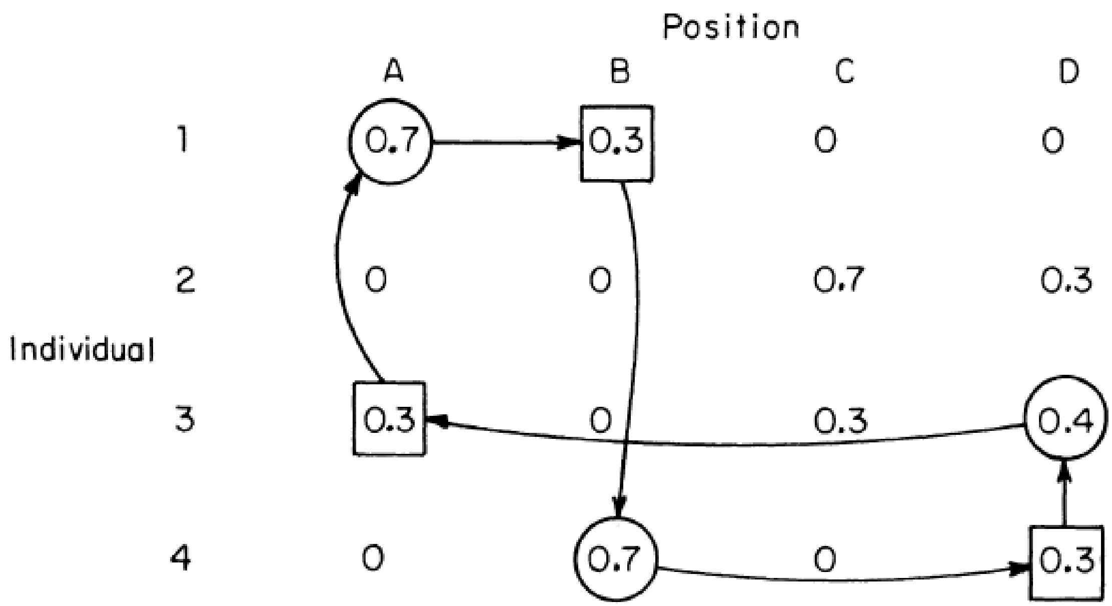


FIG. 1.—A probability cycle in the assignment procedure

Figure 2: Source: Hylland and Zeckhauser (1979).

Hence, the set $\mathcal{B}$ is a *polyhedron* (i.e. a compact and convex set defined by a finite number of linear inequalities) in the vector space of $N \times N$ matrices whose facets are the solutions of an optimal assignment problem. Matrix Φ_{ij}, seen as a vector, is normal to these facets. It turns out that one can say much more, and give a characterization of the vertices of this polyhedron.

Theorem 2.4 (Birkhoff-von Neumann). *The set of extreme points (vertices) of $\mathcal{B}$ is the set of permutation matrices.*

The proof of this result if very similar to the proof of Theorem 2.3.

Proof. First, it is easy to see that every permutation matrix is an extreme point. We are going to show that every bistochastic matrix that is not a permutation matrix is not an extremal point. Let π be a bistochastic matrix that is not a permutation matrix. We can then construct a sequence of indices $i_1, j_2, i_2 ... j_{p+1} = j_1$ such that $\pi_{i_1 j_1}$, $\pi_{i_1 j_2}$, $\pi_{i_2 j_2}$, ..., $\pi_{i_p j_{p+1}}$ all are in $(0, 1)$ (the repeated index may also be a line index; this is without loss of generality). Let x resp. y be the smallest resp. largest value of π_{ij} along the cycle. Taking $\varepsilon = \min(x, 1 - y)$, we see that π can be seen as the barycenter of one matrix where we have alternatively added and subtracted ε and $-\varepsilon$, respectively, to the entries of the cycle. This shows that π is not an extremal point of $\mathcal{B}$. ■

As a consequence, we get the very important result that any production plan may be interpreted as the outcome of a lottery over permutations.

Theorem 2.5. *Any fractional assignment can be seen as the outcome of lotteries on deterministic assignments. More precisely, for any $\pi \in \mathcal{B}$, there exist a probability distribution defined over the space of permutations $\alpha(\sigma) \geq 0$, $\sum_{\sigma \in \mathsf{S}_N} \alpha(\sigma) = 1$, such that*

$$\pi_{ij} = \sum_{\sigma \in \mathsf{S}_N} \alpha(\sigma) \pi_{ij}^{\sigma}.$$

In other words, the set of bistochastic matrices is the convex hull of the set of permutation matrices.

Proof. By the Birkhoff-von Neumann theorem, the set of extreme points is the set of permutation matrices; by the Krein-Millman theorem, a convex set is the convex hull of its extreme points. Hence the set of bistochastic matrices is the convex hull of the set of permutation matrices. ■

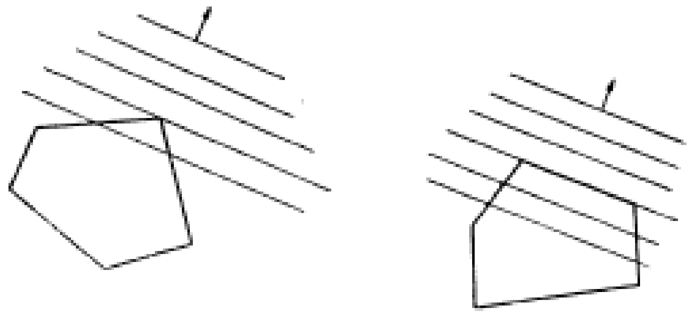


FIGURE 1.—Maximum of a Linear Function on a Polyhedron.

Figure 3: Source: Koopmans and Beckmann (1957)

In fact, it is easy to see that Theorem 2.5 is a direct consequence of the previous theorem. Consider a bistochastic matrix $\pi \in \mathcal{B}$ solution to program (2.5), and let V be the value of that program. Then there exists a probability distribution α defined over $\mathfrak{S}_N$ such that $\pi_{ij} = \sum_{\sigma \in \mathsf{S}_N} \alpha(\sigma) \pi^{\sigma}_{ij}$. It is now obvious that

$$\alpha(\sigma) > 0 \text{ implies } \sum_{i=1}^{N} \Phi_{i\sigma(i)} = V.$$

Consequences for identification. The previous geometric picture has very important consequences for identification. We are going to study this issue in much detail later on in this course, but it is useful to start a discussion at this point. Assume that one observe the fraction of time π_{ij} that each agent i has worked on machine j, and one wonders if there is a production function Φ_{ij} such that the fractional assignment π_{ij} is optimal. Two questions arise.

1. Rationalizability: Is π_{ij} rationalizable by some production function Φ?

2. Identification: if yes, what is the set of production functions that rationalize the data.

Of course, we would like to rule out trivial functions that would yield to theories void of testable implications. Hence Φ cannot be chosen constant. For similar reasons, Φ should not be separatively additive in i and j: indeed, if $\Phi_{ij} = a_i + b_j$ then for any fractional assignment plan $\pi \in \mathcal{B}$, one would have $\sum_{ij} \pi_{ij} \Phi_{ij} = \sum_i a_i + \sum_j b_j$, and the theory, again, cannot deliver testable implications. We shall see later on that if Φ cannot be written in a separatively additive way, then the theory has testable implications. We shall assume from now on that this is the case.

The answer to the previous two questions above is then the following.
1. π is rationalizable if and only if it is on the boundary of $\mathcal{B}$.
2. The set of production functions that rationalize the data coincides with the subdifferential of $\mathcal{B}$ at π.

We shall explain in detail these results later in the course, and elaborate on them. But let us state at this stage that the important consequence of this result is that in practice, empirically observed matchings are *almost never rationalizable*, and that some form of heterogeneity is needed both for identification and inference.

Example 2.1. *Consider the matrix*

$$\pi = \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix}.$$

This matrix is equal to $0.7 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 0.3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. *Hence for a production function* Φ, *we get*

$$\sum_{ij} \pi_{ij} \Phi_{ij} = 0.7 \left(\Phi_{11} + \Phi_{22} \right) + 0.3 \left(\Phi_{12} + \Phi_{21} \right)$$

Hence it cannot be rationalized by a production function Φ *unless* $\Phi_{11} + \Phi_{22} = \Phi_{12} + \Phi_{21}$. *But in that case, set* $a_1 = \Phi_{11}$, $b_1 = 0$, $a_2 = \Phi_{21}$, *and* $b_2 = \Phi_{12} - \Phi_{11}$, *and we can verify that* $\Phi_{ij} = a_i + b_j$ *– a case we precisely ruled out. Therefore* π *cannot be rationalized.*

2.2.4 Endogenous participation

We are now going to modify the optimal assignment problem by allowing for unaffected agents and machines, and assuming there is an alternative technology to employ agents: if agent i is not affected to any machine, she produces output Φ_{i0} per amount of time; similarly, a machine j not operated by any agent is rented out and produces Φ_{0j} per amount of time.

The firm's problem is then

$$\begin{aligned} V = \max_{\pi_{ij} \geq 0} & \sum_{1 \leq i,j \leq N} \pi_{ij} \Phi_{ij} \\ & + \sum_{i=1}^{N} \pi_{i0} \Phi_{i0} + \sum_{j=1}^{N} \pi_{0j} \Phi_{0j} \\ s.t. \quad & \sum_{i=1}^{N} \pi_{ij} + \pi_{0j} = 1 \text{ for each } j \geq 1 \\ & \sum_{j=1}^{N} \pi_{ij} + \pi_{i0} = 1 \text{ for each } i \geq 1. \end{aligned} \tag{2.7}$$

where we have introduced two variables π_{0i}, the fraction of time that agent i spends on its own, and a similar quantity π_{0j} for machine j.

Note that in the limit where Φ_{0i} and Φ_{0j} tend to $-\infty$, we recovered a standard optimal assignment problem. Hence problem (2.7) encompasses standard optimal assignment problems. But although this problem looks more general, we are going to show that problem (2.7) is *formally equivalent to a standard optimal assignment problem.* This is a surprising result, which we now explain. Add one individual "0" both to agents and machines, and set $\Phi_{00} = 0$. Call π_{00} the total time worked by agents on machines; this should coincide with the total time where machines have been affected to agents, as

$$\pi_{00} = \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_{ij}$$

hence

$$\begin{aligned} \sum_{i=0}^{N} \pi_{i0} &= \pi_{00} + \sum_{i=1}^{N} \pi_{i0} = n, \\ \sum_{j=0}^{N} \pi_{0j} &= \pi_{00} + \sum_{j=1}^{N} \pi_{0j} = n. \end{aligned}$$

Therefore one may rewrite problem (2.7) as a standard assignment problem

$$V = \max_{\pi_{ij} \geq 0} \sum_{0 \leq i,j \leq N} \pi_{ij} \Phi_{ij} \tag{2.8}$$

$$\begin{aligned} s.t. \quad & \sum_{i=0}^{N} \pi_{ij} = 1\{j \geq 1\} + n1\{j = 0\} \\ & \sum_{j=0}^{N} \pi_{ij} = 1\{i \geq 1\} + n1\{i = 0\} \end{aligned}$$

In this case, participation is endogenously determined by the level of Φ_{i0} and Φ_{0j}. This can be seen by looking at the dual of problem (2.8). One has

$$\begin{aligned} & V = \min_{u_i, v_j} \sum_{i=1}^{N} u_i + \sum_{j=1}^{N} v_j + n(u_0 + v_0) \\ & s.t.\ u_i + v_j \geq \Phi_{ij},\ i,j \geq 0 \end{aligned}$$

The constraint of the dual problem can be rewritten

$$\begin{aligned} u_i + v_j &\geq \Phi_{ij},\ i,j \geq 1, \\ u_i + v_0 &\geq \Phi_{i0},\ i \geq 1, \\ u_0 + v_j &\geq \Phi_{0j},\ j \geq 1, \\ u_0 + v_0 &\geq 0 \end{aligned}$$

and recall that the Lagrange multiplier associated to the last constraint is $\pi_{00} > 0$ unless everybody chooses not to participate; in that case, by complementary slackness the constraint is saturated, so $u_0 + v_0 = 0$, and as there is one degree of freedom in the choice of the variables in the dual program, one can set $u_0 = v_0 = 0$. Therefore, the constraints become

$$\begin{aligned} u_i + v_j &\geq \Phi_{ij},\ i,j \geq 1, \\ u_i &\geq \Phi_{i0},\ i \geq 1, \\ v_j &\geq \Phi_{0j},\ j \geq 1. \end{aligned}$$

We can therefore state the following result.

Theorem 2.6. *The dual of problem (2.7) is*

$$V = \min_{u_i, v_j} \sum_{i=1}^{N} u_i + \sum_{j=1}^{N} v_j \tag{2.9}$$
$$\textit{s.t. for } i, j \geq 1 :$$
$$u_i + v_j \geq \Phi_{ij},$$
$$u_i \geq \Phi_{i0},$$
$$v_j \geq \Phi_{0j}.$$

2.2.5 The Hide-and-Seek game

In 1953, von Neumann described the following two-person, zero-sum game. Let (α_{ij}) be a $N \times N$ matrix with positive terms. There are two players, the "Hider" and the "Seeker".

- First, Hider hides in a cell (i, j),
- Then, Seeker highlights either a row or a column where he believes Hider rests hidden. If Seeker is correct, then Hider pays him α_{ij}, otherwise 0. Thus the value of this game for Seeker is

  $$A = \min_{x \geq 0, \sum_{ij} x_{ij} = 1} \max_{i', j'} \left(\sum_j \alpha_{i'j} x_{i'j}, \sum_i \alpha_{ij'} x_{ij'} \right)$$

  where the maximization is over the vector of mixed strategies which verify $\sum_{ij} x_{ij} = 1$, x_{ij} being interpreted as the probability that the seeker hides in cell (i, j).

Interestingly, this game is related to an assignment problem in a nontrivial way. The following result was proved by von Neumann.

Theorem 2.7. *The value A of the game for Seeker is equal to the inverse of the assignment problem with production function equal to the inverse of the payoff α_{ij}, that is*

$$A = \frac{1}{\max_{\sigma \in \mathsf{S}_N} \sum_{i=1}^{N} \frac{1}{\alpha_{i\sigma(i)}}}.$$

Denote $\Phi_{ij} = 1/\alpha_{ij}$*. Let* $\sigma^* \in \mathfrak{S}_N$ *be an optimal assignment for the problem*

$$V = A^{-1} = \max_{\sigma \in \mathsf{S}_N} \sum_{i=1}^{N} \Phi_{i\sigma(i)},$$

and let u_i *and* v_j *be the solution to the dual of that problem, so that* $u_i, v_j \geq 0$ *and*

$$u_i + v_j \geq \Phi_{ij}.$$

Then,

– The optimal strategy of the Hider is to hide in cell (i, j) *with probability* $x_{ij} = A\Phi_{ij} 1\{j = \sigma(i)\}$*, and*

– The optimal strategy for the Seeker is to highlight cell i *with probability* Au_i*, or cell* j *with probability* Av_j*.*

Proof. Start from

$$\begin{aligned} V &= \min_{u_i, v_j} \sum_i u_i + \sum_j v_j \\ &\quad s.t. \quad u_i + v_j \geq \Phi_{ij}. \end{aligned}$$

One can equivalently rewrite this problem as

$$\begin{aligned} V &= \min_{u_i, v_j} \sum_i u_i + \sum_j v_j \\ &\quad s.t. \quad \frac{u_i + v_j}{\Phi_{ij}} \geq 1, \end{aligned}$$

hence

$$\begin{aligned} V &= \min_{u_i, v_j} \max_{\pi_{ij} \geq 0} \sum_i u_i + \sum_j v_j + \sum_{ij} \pi_{ij} \left(1 - \frac{u_i + v_j}{\Phi_{ij}}\right) \\ &= \max_{\pi_{ij} \geq 0} \sum_{ij} \pi_{ij} + \min_{u_i, v_j} \sum_i u_i + \sum_j v_j - \sum_{ij} \pi_{ij} \frac{u_i + v_j}{\Phi_{ij}} \\ &= \max_{\pi_{ij} \geq 0} \sum_{ij} \pi_{ij} \\ &\quad + \min_{\gamma \geq 0} \gamma \min_{\sum_i u_i + \sum_j v_j = 1} \left(1 - \sum_{ij} \pi_{ij} \frac{u_i + v_j}{\Phi_{ij}}\right) \end{aligned}$$

so, introducing

$$f(\pi) = \max_{\sum_i u_i + \sum_j v_j = 1} \left(\sum_{ij} \pi_{ij} \frac{u_i + v_j}{\Phi_{ij}} \right),$$

one has $f(t\pi) = tf(\pi)$ for $t > 0$, and

$$\begin{aligned} V &= \max_{\pi_{ij} \geq 0} \min_{\gamma \geq 0} \sum_{ij} \pi_{ij} + \gamma (1 - f(\pi)) \\ &= \max_{\pi_{ij} \geq 0} \left\{ \sum_{ij} \pi_{ij} : f(\pi) \leq 1 \right\} \end{aligned}$$

thus

$$\begin{aligned} V &= \max_{\pi \geq 0, \sum_{ij} \pi_{ij} = 1} \max_{t \geq 0} \{ t : f(t\pi) \leq 1 \} \\ &= \max_{\pi \geq 0, \sum_{ij} \pi_{ij} = 1} \frac{1}{f(\pi)}, \end{aligned}$$

and therefore

$$V = \frac{1}{\min_{\pi \geq 0, \sum_{ij} \pi_{ij} = 1} f(\pi)}$$

hence,

$$\min_{\substack{\pi \geq 0, \\ \sum_{ij} \pi_{ij} = 1}} \max_{\sum_i u_i + \sum_j v_j = 1} \left(\sum_{ij} \pi_{ij} \frac{u_i + v_j}{\Phi_{ij}} \right) = \frac{1}{V},$$

QED. The optimal strategies are left as an exercise. ■

Let us give an example.

Example 2.2. *Consider payoff matrix*

$$\alpha_{ij} = \begin{pmatrix} \overline{1/5} & 3 & 1/2 \\ 1/2 & 1 & \overline{1/2} \\ 1/4 & \underline{1/3} & 1 \end{pmatrix}$$

hence,

$$\Phi = \begin{pmatrix} \underline{5} & 3 & 2 \\ 2 & 1 & \underline{2} \\ 4 & \underline{3} & 1 \end{pmatrix}$$

Clearly, the solution σ of the optimal assignment problem is such that $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$, corresponding to the underlined entries. Thus $V = 10$, and the strategy for the hider will be to hide in cell $(1,1)$ with probability $\frac{5}{10} = 0.5$, in cell $(2,3)$ with probability 0.2, and in cell $(3,2)$ with probability 0.3.

It is clear that whichever pure strategy the Seeker may play, her value will be $A = 1/10$: say it plays line i, then the probability that the Hider hides there is $A\Phi_{ij}$, hence expected payoff to the Seeker is $A\Phi_{ij}\alpha_{ij} = A$.

The value of the game for the Seeker is $A = 1/10$.

2.3 References and notes

The ideas in Section 2.1 can be found in several different literatures. For power diagrams, see Aurenhammer (1991) [5] and references therein. In an industrial organization context, see the monograph by Anderson, de Palma and Thisse (1992) [2], and the paper by Feentra and Levinsohn (1995) [44]. For the numerical computation of the prices and the link with the tâtonnement process, and a completely different application, see Ekeland, Galichon and Henry (2010) [42].

The standard reference for Section 2.2, more precisely 2.2.1 and 2.2.4 is Koopmans and Beckmann (1957) [72]; see also the treatise by Dantzig on linear programming [35], Chapters 14 and 15. The classical results in Sections 2.2.2 and 2.2.3 trace back to Birkhoff (1946) [13] and von Neumann (1953) [116]. There are innumerable applications to economics, see [67] just to cite one. The result in Section 2.2.5 first appeared in [116]; however the presentation given here is original.

3 Optimal transport theory

In this chapter we set out the main theoretic basis to handle bipartite matching models in a TU setting with a sufficient level of generality. Economic applications, some of which we shall review in subsequent chapters, include labour economics, hedonic models, transportation economics, the economics of marriage...

The apparatus we shall develop here is part of the so-called theory of Optimal Transport. We can see it as a generalization of the assignment model introduced in Section 2.2 to handle situations where there is possibly a continuum of types; thus, the theory is essentially linear programming in infinite dimension. As we shall see these problems have a lot of structure, and even when no closed-form solutions are available there is much to say about qualitative properties of the solutions.

3.1 The Monge-Kantorovich theorem

We are about to state an abstract result general enough to include as special cases both the fountain problem and the assignment problem discussed above, and several others. Consider $\mathcal{X}$ and $\mathcal{Y}$ two probability spaces, which in full generality are assumed to be Polish spaces (i.e. complete and separable metric spaces), but for all practical applications in this course will be finite dimensional vector spaces, or subsets of the latter. Let μ and ν be Borel probability measures on $\mathcal{X}$ and $\mathcal{Y}$ respectively. [Remember, a Borel probability measure is defined on the σ-algebra generated by open sets of the space.]

Introduce $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_+ \cup +\infty$ a lower semi-continous cost function. [Remember, φ is lower semi-continuous whenever $\varphi(\lim x_n) \leq \lim \varphi(x_n)$ for each converging sequence x_n.]

Set of couplings of μ and ν. Define $\mathcal{M}(\mu, \nu)$ as the set of *couplings* of probability measures μ and ν; this is the set of Borel probability measures π on $\mathcal{X} \times \mathcal{Y}$ with marginals μ and ν, respectively, that is such that for all measurable sets $A \subset \mathcal{X}$ and $B \subset \mathcal{Y}$, we have

$$\begin{aligned} \pi(A \times \mathcal{Y}) &= \mu(A) \\ \pi(\mathcal{X} \times B) &= \nu(B) \end{aligned}$$

In other words, one should have

$$\begin{aligned}\int_{\mathcal{Y}} d\pi\left(x,y\right) &= d\mu\left(x\right)\\ \int_{\mathcal{X}} d\pi\left(x,y\right) &= d\nu\left(y\right)\end{aligned}$$

Yet in other terms, $\mathcal{M}\left(\mu,\nu\right)$ is the set of distributions of random pairs $\left(X,Y\right)$ such that the distributions of X and Y are fixed: $X\sim\mu$ and $Y\sim\nu$.

Stable set. Define the *stable set* as the set S_c of pairs of functions $\left(\varphi,\psi\right)$ such that $\varphi\in L^1\left(d\mu\right)$, $\psi\in L^1\left(d\nu\right)$ satisfying

$$\varphi\left(x\right)+\psi\left(y\right)\leq c\left(x,y\right). \tag{3.1}$$

The optimal transport problem. Assume that we would like to move a pile of sand to fill a hole. The quantity of sand available in a small volume around location x is $d\mu\left(x\right)$, while the quantity of sand needed in the hole around location y is $d\nu\left(y\right)$. The solution to this problem is how much sand is moved from location around x to location around y, call that quantity $d\pi\left(x,y\right)$. Of course, the total quantity of sand taken away around location x should be given by $d\mu\left(x\right)$, while the total quantity of sand brought around location y should be $d\nu\left(y\right)$, thus imposing $\pi\in\mathcal{M}\left(\mu,\nu\right)$.

One assume there is a cost $c\left(x,y\right)$ of moving a unit amount of mass of sand from x to y, so the cost of the whole operation is

$$\int_{\mathcal{X}\times\mathcal{Y}} c\left(x,y\right)d\pi\left(x,y\right),$$

so if we would like to operate at minimal cost, our problem is thus the *optimal transport problem*:

$$W=\inf_{\pi\in\mathcal{M}\left(\mu,\nu\right)}\int_{\mathcal{X}\times\mathcal{Y}} c\left(x,y\right)d\pi\left(x,y\right).$$

W is sometimes called the *Wasserstein distance* associated to cost c.

We have the following result.

Theorem 3.1 (Monge-Kantorovich). *Under the assumptions above, one has:*

(i) The primal problem

$$W = \inf_{\pi \in \mathcal{M}(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi(x,y) \tag{3.2}$$

is equal to its dual problem

$$W = \sup_{(\varphi,\psi)\in S_c} \int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\nu. \tag{3.3}$$

where the stable set S_c has been defined by Eq. (3.1).

(ii) The infimum in the primal problem (3.2) is attained.

(iii) Assume that $c(x,y)$ is bounded above by $\varphi(x)+\psi(y)$ such that $\varphi \in L^1(d\mu)$, $\psi \in L^1(d\mu)$. Then the supremum in the dual problem (3.2) is attained by a pair (φ,ψ) valued in $\mathbb{R}\cup +\infty$ such that one has

$$\begin{aligned}\varphi(x) &= \inf_{y\in\mathcal{Y}} \{c(x,y) - \psi(y)\} \\ \psi(y) &= \inf_{x\in\mathcal{X}} \{c(x,y) - \varphi(x)\}.\end{aligned}$$

A coupling π which is optimal for problem (3.2) is called an *optimal coupling*, or an *optimal transportation plan.* φ and ψ are sometimes called *transportation potentials*, or *Kantorovich potentials.*

Informal proof. It is out of the scope of these lecture notes to provide a rigorous proof of the Monge-Kantorovich theorem in its full generality, and we refer to [110], Chapters 1 and 2 for a full proof. We will only give an idea of how the duality in part (i) works. Note that one inequality is easy to establish, as integrating (3.1) over any $\pi \in \mathcal{M}(\mu,\nu)$ yields

$$\int (\varphi(x)+\psi(y))\, d\pi(x,y) \leq \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi(x,y)$$

hence

$$\int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\nu \leq \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, d\pi(x,y)$$

thus, taking the infimum for $\pi \in \mathcal{M}(\mu,\nu)$ on the right hand-side and the supremum for $(\varphi,\psi) \in S_c$ on the left hand-side, one has the weak duality inequality

$$
\begin{aligned}
& \sup_{(\varphi,\psi)\in S_c} \int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\nu \\
\leq & \inf_{\pi\in\mathcal{M}(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} c\left(x,y\right) d\pi\left(x,y\right).
\end{aligned}
$$

As for proving that equality actually holds, we will not establish formally the absence of duality gap, but we will give an informal argument. Fist, note that the function

$$
\begin{aligned}
L_P\left(\pi\right) &= \sup_{\varphi,\psi} \int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\nu \\
&\quad - \int_{\mathcal{X}\times\mathcal{Y}} \left(\varphi\left(x\right) + \psi\left(y\right)\right) d\pi\left(x,y\right)
\end{aligned}
$$

takes values

$$
\begin{aligned}
L_P\left(\pi\right) &= 0 \text{ if } \pi \in \mathcal{M}\left(\mu,\nu\right) \\
&= +\infty \text{ otherwise}
\end{aligned}
$$

hence,

$$
\begin{aligned}
W &= \inf_{\pi\geq 0} \int_{\mathcal{X}\times\mathcal{Y}} c\left(x,y\right) d\pi\left(x,y\right) + L_P\left(\pi\right) \\
&= \inf_{\pi\geq 0} \sup_{\varphi,\psi} \int_{\mathcal{X}\times\mathcal{Y}} c\left(x,y\right) d\pi\left(x,y\right) \\
&\quad + \int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\mu \\
&\quad - \int_{\mathcal{X}\times\mathcal{Y}} \left(\varphi\left(x\right) + \psi\left(y\right)\right) d\pi\left(x,y\right)
\end{aligned}
$$

thus, admitting without a proof that the inf and the sup can be interchanged (this is not as obvious as it is in the finite dimensional case)

$$
\begin{aligned}
W &= \sup_{\varphi,\psi} \int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\nu \\
&\quad + \inf_{\pi\geq 0} \int_{\mathcal{X}\times\mathcal{Y}} c\left(x,y\right) - \left(\varphi\left(x\right) + \psi\left(y\right)\right) d\pi\left(x,y\right)
\end{aligned}
$$

and in turn, we note that, defining

$$L_D(\varphi,\psi) = \inf_{\pi\geq 0} \int_{\mathcal{X}\times\mathcal{Y}} c(x,y) - (\varphi(x) + \psi(y))\, d\pi(x,y)$$

we get

$$\begin{aligned} L_D(\varphi,\psi) &= 0 \text{ if } (\varphi,\psi) \in S_c \\ &= -\infty \text{ otherwise} \end{aligned}$$

so W eventually rewrites as

$$W = \sup_{(\varphi,\psi)\in S_c} \int_{\mathcal{X}} \varphi d\mu + \int_{\mathcal{Y}} \psi d\nu.$$

■

Finite case. We now show that the duality results on the assignment problem of Section 2.2 are just a specialization of the above results in the case where the sets $\mathcal{X}$ and $\mathcal{Y}$ are finite. In the optimal assignment example, both the sets $\mathcal{X}$ and $\mathcal{Y}$ are $\{1,...,N\}$, and $c_{ij} = -\Phi_{ij}$ for $x = i$ and $y = j$. Measure μ and ν coincide with the uniform probability distribution on $\{1,...,N\}$, which we write

$$\mu = \nu = \frac{1}{N}\sum_{i=1}^{N} \delta_i.$$

In this case, $\pi \in \mathcal{M}(\mu,\nu)$ becomes

$$\begin{aligned} \sum_{i=1}^{N} \pi_{ij} &= 1/N \\ \sum_{j=1}^{N} \pi_{ij} &= 1/N \\ \pi &\geq 0 \end{aligned}$$

that is, matrix $(N\pi_{ij})_{ij}$ is bistochastic: $\pi N \in \mathcal{B}$.

Then formula (3.2) becomes

$$\min_{N\pi\in\mathcal{B}} \sum_{ij} \pi_{ij} c_{ij} = \max_{u_i+v_j\leq c_{ij}} \sum_{i=1}^{N} u_i + \sum_{j=1}^{N} v_j,$$

which is (up to a sign inversion and a scaling by a factor N), the formula in Theorem 2.2.

Actually, slightly more general version of the theorem holds in the discrete case, where we consider general probability measures on the finite sets $\mathcal{X} = \{1, ..., N\}$ and $\mathcal{Y} = \{1, ..., M\}$. The following theorem is nothing else than the specialization of Theorem 3.1 in the case where the sets $\mathcal{X}$ and $\mathcal{Y}$ are finite.

Theorem 3.2. *Let* (p_i) *and* (q_j) *be two probability vectors on* $\{1, ..., N\}$ *and* $\{1, ..., M\}$ *respectively. Then the value of primal program*

$$\begin{aligned} &\min_{\pi_{ij} \geq 0} \sum_{ij} \pi_{ij} c_{ij} \\ s.t. \quad &\sum_{j=1}^{M} \pi_{ij} = p_i \ , \ \sum_{i=1}^{N} \pi_{ij} = q_j \end{aligned}$$

coincides with the value of the dual

$$\begin{aligned} &\max \sum_{i=1}^{N} u_i p_i + \sum_{j=1}^{M} v_j q_j \\ s.t. \quad &u_i + v_j \leq c_{ij} \end{aligned}$$

and $\pi_{ij} > 0$ *for some solution* π *of the primal problem implies* $u_i + v_j = c_{ij}$.

Purity. In the theory of optimal transportation, important efforts are devoted to understand when the solution is *pure*, namely when all the sand in a small volume around location x is send to the same place, or whether it is split between several locations. In other words, the solution is pure whenever the distribution of y conditional on x, $\pi(y|x)$, is concentrated at one point $T(x)$. This motivates the following definition.

Definition 3.1. *A coupling* $\pi \in \mathcal{M}(\mu, \nu)$ *is called pure when there exists a map* $T : \mathcal{X} \to \mathcal{Y}$ *such that*

$$\pi(x, y) = \mu(x) \delta(y = T(x)),$$

or equivalently, when π is the distribution of a random pair (X,Y) such that

$$\begin{aligned} X &\sim \mu\ ,\ Y \sim \nu \\ Y &= T(X)\,. \end{aligned}$$

In that case T is called an optimal transportation map.

We saw already an example of such a result in the case where the marginal distributions coincide and are uniform probability distributions on the same finite set: in Theorem 2.3 it was shown that the assignment problem has at least one pure solution (even though all solutions may not be pure). We shall next see two particularly important examples where the solution is still generically pure: the case of submodular cost, with sets $\mathcal{X}$ and $\mathcal{Y}$ univariate; and the case of quadratic costs, with sets $\mathcal{X}$ and $\mathcal{Y}$ multivariate.

3.2 Dimension one, submodular cost

The univariate case is one of the few cases where we can construct an explicit formula for the optimal transportation map, which is shown to exist. Many early models of matching in labour economics, family economics and industrial organization are univariate models: agents' types, or products characteristics are represented by a scalar (single-dimensional) parameter. While these models are often tractable and simple it turns out, however, that many of the intuitions that work in the univariate case do not extend to the multivariate case, making multi-dimensional models significantly more difficult to apprehend.

In this paragraph, we set $\mathcal{X} = \mathcal{Y} = \mathbb{R}$. μ and ν are two probability distribution on the real line, with respective cumulative distribution function (cdf) F_μ and F_ν

$$F_\mu(x) = \mu((-\infty, x])\,,\ F_\nu(x) = \nu((-\infty, x])\,.$$

Let us start by the following definition.

Definition 3.2 (Submodularity). *A function $c : \mathbb{R}^2 \to \mathbb{R}$ is* submodular *if for all quadruple of real numbers x, x', y, y' such that $x \leq x'$ and $y \leq y'$, one has*

$$c(x,y) + c(x',y') \leq c(x,y') + c(x',y)\,.$$

It is strictly submodular if the inequality above is strict as soon as $x < x'$ and $y < y'$ hold.

The notion of submodularity is quite intuitively explained in terms of matching. It means that when the x's and the y's are ranked in the same way, then the cost of the matching that respects the relative ranks is less than the cost of the matching that interchanges the ranks.

When c is twice continuously differentiable, there is the following differential characterization of submodularity, which will be very useful.

Proposition 3.1. *When $c \in C^2$, then c is submodular if and only if*

$$\frac{\partial^2 c(x,y)}{\partial x \partial y} \leq 0;$$

and it is strictly submodular if the inequality above is strict:

$$\frac{\partial^2 c(x,y)}{\partial x \partial y} < 0. \tag{3.4}$$

We will come back on this condition and its economic interpretation in Section 3.4.

Example. *For $f \in C^2$ and $c(x,y) = f(x-y)$, c is submodular if and only if f is convex. Thus, all the functions $c(x,y) = |x-y|^p$, $p \geq 1$ are submodular.*

Pseudo-inverses, quantiles. Let $\varphi : \mathbb{R} \to [0,1]$ a nondecreasing function; we define the *generalized inverse* of φ as a function φ^{-1} defined on $[0,1]$ by

$$\varphi^{-1}(t) = \inf\{x \in \mathbb{R} : \varphi(x) > t\}.$$

The quantile function associated to μ, denoted $Q_\mu(t)$ is the inverse function of the cdf of μ

$$Q_\mu(t) = F_\mu^{-1}(t).$$

One has the following property.

Proposition 3.2. *For $U \sim \mathcal{U}([0,1])$ a random variable distributed uniformly on $[0,1]$, the random variable $Q_\mu(U)$ has distribution μ.*

Proof. One has

$$\Pr\left(Q_{\mu}\left(U\right) \leq x\right) = \Pr\left(U \leq F_{\mu}\left(x\right)\right) = F_{\mu}\left(x\right).$$

■

We have the following theorem, which characterizes the solution of the univariate optimal transport problem, in the case where the cost function is submodular.

Theorem 3.3 (Hardy-Littlewood). *Let μ and ν be two probability measures on $\mathbb{R}$, and c a submodular cost function. Then an optimal coupling $\pi \in \mathcal{M}\left(\mu, \nu\right)$ for (3.2) has a c.d.f. given by*

$$F_{\pi}\left(x, y\right) = \min\left(F_{\mu}\left(x\right), F_{\nu}\left(y\right)\right);$$

and π it is the probability distribution of the joint random variable $\left(X, Y\right)$, where

$$\begin{aligned} X &= F_{\mu}^{-1}\left(U\right) \\ Y &= F_{\nu}^{-1}\left(U\right) \end{aligned} \tag{3.5a}$$

and $U \sim \mathcal{U}\left(\left[0,1\right]\right)$. As a result, if F_X is continuous (which is the case when μ does not give mass to atoms), one has $U = F_X\left(X\right)$, and hence

$$Y = T\left(X\right)$$

where $T\left(x\right) = F_{\nu}^{-1}\left(F_{\mu}\left(x\right)\right)$ is a nondecreasing function.

Further, when c is strictly submodular, the solution above is the only solution.

Vocabulary. One says that random variables X and Y satisfying (3.5) are *comonotone*, or have the *Positive Assortative Matching* (PAM) property.

Proof. Let us first see this for the case of uniform measures over sets of size N. Assume μ^N is the uniform distribution over a set $\{X_1, ..., X_N\}$ and ν^N the uniform distribution over a set $\{Y_1, ..., Y_N\}$. In that case, (and because of Theorem 2.3) problem (3.2) boils down to

$$\min_{\sigma \in \mathsf{S}_N} \sum_{i=1}^{N} c\left(X_i, Y_{\sigma(i)}\right) \tag{3.6}$$

We claim that there is a σ_N solution of (3.6) such that $X_i \leq X_j$ and $Y_{\sigma(i)} \geq Y_{\sigma(j)}$ cannot hold with at least one strict inequality for two indices i and j. Indeed, assume there are two such indices, and let σ be a solution of (3.6). Then, c being submodular, one has by the very definition of submodularity

$$\begin{aligned} & c\left(X_i, Y_{\sigma(j)}\right) + c\left(X_j, Y_{\sigma(i)}\right) \\ \leq \quad & c\left(X_i, Y_{\sigma(i)}\right) + c\left(X_j, Y_{\sigma(j)}\right) \end{aligned}$$

thus, defining $\hat{\sigma}(k) = \sigma(k)$ for $k \notin \{i, j\}$, $\hat{\sigma}(i) = \sigma(j)$ and $\hat{\sigma}(j) = \sigma(i)$, one has $\hat{\sigma} \in \mathfrak{S}_N$, and

$$\sum_{i=1}^{N} c\left(X_i, Y_{\hat{\sigma}(i)}\right) \leq \sum_{i=1}^{N} c\left(X_i, Y_{\sigma(i)}\right).$$

thus $\hat{\sigma}$ is still a solution of (3.6), and we can iterate this, strictly decreasing the number of such pairs (i, j), until no such pair exists. As a result

$$F_{\pi^N}(x, y) = \min\left(F_{\mu^N}(x), F_{\nu^N}(y)\right).$$

One extends this to the general case by taking i.i.d.samples of size N of probability distributions μ and ν for increasingly large N, and passing to the limit. The proof of the unicity of the solution in the case of strict submodularity is left as an exercise. ■

3.3 Higher dimension, quadratic cost

In this paragraph, we assume that $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$, μ and ν are two probability distributions with finite second order moments, and $c(x, y) = |x - y|^2 / 2$ is the quadratic distance cost function (up to a factor a half). We shall assume that μ is absolutely continuous with respect to the Lebesgue measure, and we shall denote f_μ its pdf.

Notations. In our notations, $|z|^2$ denotes the squared Euclidian norm $\sum_{i=1}^{d} z_i^2$, while $x.y$ denotes the scalar product of x and y $\sum_{i=1}^{d} x_i y_i$. The operator ∇ denotes the gradient; for a function $u : \mathbb{R}^d \to \mathbb{R}$, $\nabla u(x)$ is the vector $\left(\frac{\partial u(x)}{\partial x_i}\right)_i$, and u^* denotes its convex conjugate (Legendre-Fenchel) function.

Given the assumption made on μ, ν and the quadratic cost, it is easy to verify that the assumptions of Theorem 3.1 are met. In particular, $c(x, y) =$

$|x-y|^2/2 \leq |x|^2+|y|^2$ which is an integrable function by hypothesis. Hence, by the first part of Theorem 3.1, one has

$$\begin{aligned} W &= \inf_{\pi\in\mathcal{M}(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|x-y|^2}{2} d\pi(x,y) \\ &= \sup_{(\varphi,\psi)\in S_c} \int_{\mathbb{R}^d} \varphi(x)\, d\mu(x) + \int_{\mathbb{R}^d} \psi(y)\, d\nu(y). \end{aligned}$$

In this case, $2W$ is called the *quadratic Wasserstein distance.*

Note that $(\varphi,\psi)\in S_c$ means $\varphi(x)+\psi(y) \leq |x-y|^2/2$, thus

$$\underbrace{|x|^2/2-\varphi(x)}_{u(x)} + \underbrace{|y|^2/2-\psi(y)}_{v(y)} \geq x.y$$

where we have introduced $u(x) = |x|^2/2-\varphi(x)$ and $v(y) = |y|^2/2-\psi(y)$. We can let

$$M = \int_{\mathbb{R}^d} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^d} \frac{|y|^2}{2} d\nu(y) - W$$

and we have

$$\begin{aligned} M &= \sup_{\pi\in\mathcal{M}(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} x.y\ d\pi(x,y) \\ &= \inf_{u(x)+v(y)\geq x.y} \int_{\mathbb{R}^d} u(x)\, d\mu(x) + \int_{\mathbb{R}^d} v(y)\, d\nu(y). \end{aligned}$$

Now by the second part and third part of Theorem 3.1, we know respectively that a maximizer π and a minimizer (u,v) exist. Also, still by the third part, we can choose

$$\begin{aligned} v(y) &= \sup_{x\in\mathbb{R}^d} (x.y-u(x)) \\ u(x) &= \sup_{v\in\mathbb{R}^d} (x.y-v(y)) \end{aligned}$$

hence both functions u and v can be chosen convex and such that $v=u^*$, where u^* denotes the convex conjugate function to u.

We summarize this in the following proposition, which is a mere restatement of Theorem 3.1.

Proposition 3.3. *Under the maintained assumptions on μ and ν, consider the primal problem*

$$M = \sup_{\pi \in \mathcal{M}(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} x.y \; d\pi\left(x,y\right). \tag{3.7}$$

Then one has:

(i) The value of the primal problem is equal to the value of the dual problem

$$M = \inf_{u(x)+v(y)\geq x.y} \int_{\mathbb{R}^d} u\left(x\right) d\mu\left(x\right) + \int_{\mathbb{R}^d} v\left(y\right) d\nu\left(y\right) \tag{3.8}$$

(ii) Both the primal and the dual problem admit optimizers.

(iii) The minimizers in dual problem (3.8), u and v can be chosen convex and such that

$$v\left(y\right) = u^{*}\left(y\right).$$

It turns out that one can say much more about this problem. One can actually show that the solution is pure. Further, one can relate the solutions of the primal and the dual problem in a very insightful way. This is provided in the following result.

Theorem 3.4. *Under the maintained assumptions on μ and ν, one has:*

(i) the optimal coupling π solution of (3.7) is pure: there is a map $T : \mathcal{X} \rightarrow \mathcal{Y}$ such that

$$\pi\left(x,y\right) = \mu\left(x\right)\delta\left(y = T\left(x\right)\right),$$

or in other words that π is the distribution of a random pair $\left(X,Y\right)$ such that

$$Y = T\left(X\right) \;\; , \;\; X \sim \mu \;\; , \;\; Y \sim \nu.$$

(ii) $T\left(x\right)$ coincides with $\nabla u\left(x\right)$, where u is a solution of the dual problem (3.8) and u is μ-almost surely differentiable, so $T\left(x\right)$ exists μ-almost surely.

We will not provide a proof for this result, referring instead to [110], Th. 2.12. Let us just try to provide intuition for the result, by gaining insight from the discrete case. In the discrete case, if u_i and v_j are solutions of the dual problem and π_{ij} is a solution of the primal problem, one has

$$\pi_{ij} > 0 \text{ implies } u_i + v_j = c_{ij},$$

that is, if there is a positive probability of matching i with j at optimum, then the dual feasibility constraint should be saturated. Here this fact still holds: x will have a positive probability of being assigned to y at optimum if and only if the dual feasibility constraint is saturated, hence if and only if

$$u(x) + v(y) = x.y$$

thus

$$u(x) + u^*(y) = x.y$$

which is equivalent to $y \in \partial u(x)$. And when u is differentiable at x, this is equivalent to

$$y = \nabla u(x).$$

By a result known as Rademacher's theorem, the set of points x where a convex function u is not differentiable has zero Lebesgue measure; but, as we have assumed that μ is absolutely continuous with respect to the Lebesgue measure, this set will be of μ-measure zero. This implies that (except for a zero-probability set of value of x where u is not differentiable), x will be assigned to the unique y which is such that

$$y = \nabla u(x),$$

hence under quadratic costs the solution is pure.

In general there is no closed-form formula for the transportation potential between μ and ν in dimension greater than one. However there is one example of interest where there is actually a closed-form, when both measures μ and ν are Gaussian.

Example 3.1. *Let Σ_μ and Σ_ν be two symmetric positive matrices, and Σ_μ definite positive. Consider $\mu \sim \mathcal{N}(0, \Sigma_\mu)$ and $\nu \sim \mathcal{N}(0, \Sigma_\nu)$, then the transportation potential u is given by*

$$u(x) = \frac{1}{2} x' A x$$

where

$$A = \Sigma_\mu^{-1/2} (\Sigma_\mu^{1/2} \Sigma_\nu \Sigma_\mu^{1/2})^{1/2} \Sigma_\mu^{-1/2},$$

hence the transportation plans is

$$T(x) = Ax$$

and

$$M = \mathrm{Tr}\left[\left(\Sigma_\mu^{1/2}\Sigma_\nu\Sigma_\mu^{1/2}\right)^{1/2}\right]$$

thus

$$W = \mathrm{Tr}\left[\frac{1}{2}\left(\Sigma_\mu + \Sigma_\nu\right) - \left(\Sigma_\mu^{1/2}\Sigma_\nu\Sigma_\mu^{1/2}\right)^{1/2}\right].$$

Indeed, $u(x)$ is convex because A is positive symmetric and $T(x) = \nabla u(x) = Ax$ sends the probability measure $N(0,\Sigma_\mu)$ to the probability measure $N(0,\Sigma_\nu)$; the expressions for M and W follow.

In particular, in the case $\mu = N(0, I_d)$, $M = \mathrm{Tr}\left[\Sigma_\nu^{1/2}\right]$ which in matrix analysis is called the trace norm *of Σ_ν. When $d = 2$, and*

$$\Sigma_\nu = \begin{pmatrix} \sigma_1^2 & \varrho\sigma_1\sigma_2 \\ \varrho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

we have the formula tr $\left(\sqrt{S}\right) = \sqrt{\mathrm{Tr}\,(S) + 2\sqrt{\det S}}$, so we get an explicit expression for M

$$M = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2\sqrt{1-\varrho^2}}.$$

Let us finally briefly come back on the fountain problem. Recall that there both the spaces $\mathcal{X}$ and $\mathcal{Y}$ are the square $[0,1]^2$ which models the surface of the city; the measure μ is the measure of the inhabitants, it is the uniform (Lebesgue) probability measure on $[0,1]^2$; while the measure ν is the measure whose N atoms are the N fountains, hence

$$\nu = \frac{1}{N}\sum_{i=1}^{N}\delta_{Y_i}.$$

Finally, the cost is the square Euclidian distance, so $c(x,y) = |x-y|^2/2$.

Because of Theorem 3.4, we know that there is a function $T(x)$ such that $T(x)$ is the fountain chosen by inhabitant x, where if $X \sim \mu$, we have

$$T(X) \sim \nu$$

which is exactly the capacity constraint on each fountain. Further, we know that T verifies

$$T(x) = \nabla u(x)$$

where u solves the dual Kantorovich problem

$$\inf_{u(x)+v_j \geq x.Y_j} \int u(x)\, d\mu(x) + \frac{1}{N} \sum_{j=1}^{N} v_j$$

but u may be chosen so that $u(x) = \max(x.Y_j - v_j)$, which boils down to

$$\inf_{v \in \mathbb{R}^N} \int \max(x.Y_j - v_j)\, d\mu(x) + \frac{1}{N} \sum_{j=1}^{N} v_j.$$

3.4 Higher dimension, more general costs

We now want to investigate existence and qualitative properties of solutions of the optimal transport problem still in the multivariate case, but with costs that are more general than the case of quadratic cost that we have considered thus far. It turns out that, under assumptions which generalize the univariate submodularity assumption from Section 3.2, we are still able to say much, and the properties of the solutions resemble the ones derived in Section 3.3 for quadratic costs – we just need to replace the standard notion of convexity by a notion of generalized convexity, which is explained in Appendix E for interested readers. In this section and in the next one we shall state the results without any proofs; see Section 3.6 for references to texts where these results are proven.

As in Section 3.3, we assume that $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$. Let us make some assumptions on the probability measures μ and ν.

Assumption A0. μ and ν are two Borel probability measures, and μ is absolutely continuous with respect to the Lebesgue measure, with a pdf denoted f_μ.

To be consistent with our application to contract theory in the next section, we shall write the problem as to maximize a total production, with production function $\Phi(x, y)$, instead of minimizing a total cost with cost function $c(x, y)$. Of course, all the results we shall give readily apply to cost minimization by taking $c(x, y) = -\Phi(x, y)$. We shall need the following regularity assumption on Φ:

Assumption A1. Φ is twice continuously differentiable and for every compact set $\Omega \subset \mathcal{Y}$, there is $c_\Omega > 0$ such that

$$\sup_{y \in \Omega} |\Phi(x_1, y) - \Phi(x_2, y)| \leq c_\Omega |x_1 - x_2| .$$

We shall need another condition on Φ in order to replace the univariate submodularity assumption. Written on the production function $\Phi(x, y)$, the strict submodularity condition for c (3.4) translates into a *strict supermodularity condition*

$$\frac{\partial^2 \Phi(x, y)}{\partial x \partial y} > 0. \tag{3.9}$$

In mechanism design, condition (3.9) is called the *Spence-Mirrlees condition*, sometimes also called the *single-crossing condition.* If $\Phi(x, y)$ is the utility of an agent of type x assigned a good of characteristics y, it means that the indifference curve of two different types x_1 and x_2 can only cross once. It also implies that

$$\frac{\partial \Phi(x, y_1)}{\partial x} = \frac{\partial \Phi(x, y_2)}{\partial x} \Longrightarrow y_1 = y_2. \tag{3.10}$$

It turns out that we shall adopt this condition (3.10) as our multivariate generalization. Thus, we shall assume:

Assumption A2. Φ is such that:

$$\frac{\partial \Phi(x, y_1)}{\partial x} = \frac{\partial \Phi(x, y_2)}{\partial x} \Longrightarrow y_1 = y_2. \tag{3.11}$$

Throughout this section, we shall maintain Assumptions A0, A1 and A2 above.

One has the following theorem, which is a generalization of Theorem 3.4:

Theorem 3.5. *Under the maintained assumptions, consider the primal problem*

$$M = \sup_{\pi \in \mathcal{M}(\mu, \nu)} \int_{\mathcal{X} \times \mathcal{Y}} \Phi(x, y) \, d\pi(x, y) . \tag{3.12}$$

Then one has:

(i) The value of the primal problem is equal to the value of the dual problem

$$M = \inf_{u(x)+v(y)\geq\Phi(x,y)} \int_{\mathcal{X}} u(x)\, d\mu(x) + \int_{\mathcal{Y}} v(y)\, d\nu(y) \tag{3.13}$$

(ii) Both the primal and the dual problem admit optimizers.

(iii) the optimal coupling π solution of (3.12) unique and is pure: there is a map $T : \mathcal{X} \to \mathcal{Y}$ such that π is the distribution of a random pair (X, Y) such that

$$Y = T(X), \ X \sim \mu \ , \ Y \sim \nu. \tag{3.14}$$

(iv) The solutions (u, v) of the dual problem (3.13) are unique (up to a constant), and verify

$$\begin{aligned} v(y) &= \sup_{x\in\mathcal{X}} \Phi(x,y) - u(x) \\ u(x) &= \sup_{y\in\mathcal{Y}} \Phi(x,y) - v(y) \end{aligned}$$

(v) u is almost-everywhere differentiable, and the relation

$$\nabla u(x) = \frac{\partial\Phi}{\partial x}(x, T(x)) \tag{3.15}$$

holds μ-almost surely.

We refer to Appendix E for an interpretation of these results in terms of the theory of “Generalized convexity”.

Although proving this result is out of scope, we provide a very informal justification for it, hoping to convey a reasonable idea of what is going on. As in the quadratic case, by analogy with the complementary slackness relation in the finite case, it is reasonable to believe that x will be matched to y with some positive probability if the dual constraint is saturated, that is if

$$u(x) + v(y) = \Phi(x,y)$$

But it was seen that

$$v(y) = \sup_x (\Phi(x,y) - u(x))$$

hence at optimum, one gets by F.O.C.

$$\nabla u(x) = \frac{\partial\Phi(x,y)}{\partial x}$$

which, because of Assumption A2, is satisfied by at most one $y = T(x)$.

3.5 Application: multidimensional contracts

In this section we shall apply the apparatus developed in Section 3.4 to contract theory in the case both the space of types and of outcomes are multidimensional. The main result of this section will be to prove, that testing whether a given map from the space of types to the space of outcomes is implementable or not can be answered in terms of a simple Optimal Transport problem. Throughout this section we shall maintain Assumptions A1 and A2 of Section 3.4.

As before we assume that $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$. Suppose that the agents' preferences are given by the quasilinear utility function

$$U(x, y, w) = \Phi(x, y) + w$$

where $x \in \mathcal{X}$ is the agent's type, $y \in \mathcal{Y}$ is an outcome, and $w \in \mathbb{R}$ is some monetary transfer.

Define contracts in the following way:

Definition 3.3. *A contract curve is a pair of function $(T, w) : \mathcal{X} \to \mathcal{Y} \times \mathbb{R}$, where $(T(x), w(x))$ is the contract offered to individual announcing type x.*

Hence, for some announced type x, a contract specifies the action chosen $T(x)$ and the monetary transfer $w(x)$.

Next, one defines the potential associated with a contract as the utility of an agent who truthfully announces its type.

Definition 3.4. *The potential associated with a contract (T, w) is the function $U_{T,w}$ defined by*

$$U_{T,w}(x) = \Phi(x, T(x)) + w(x).$$

Of course, we need to worry about people having an incentive to fail to report their type truthfully. One thus says that a contract is incentive-compatible is telling the truth is a dominating strategy.

Definition 3.5. *The contract (T, w) is incentive-compatible if and only if*

$$\Phi(x, T(x)) + w(x) \geq \Phi(x, T(x')) + w(x')$$

holds for all $(x, x') \in \mathcal{X}^2$.

Definition 3.6. *A map $T : \mathcal{X} \to \mathcal{Y}$ is implementable if and only if there exists $w : \mathcal{X} \to \mathbb{R}$ such that (T, w) is incentive-compatible.*

Let us now relate this to our optimal transport. If (T, w) is incentive-compatible, then

$$U_{T,w}(x) \geq U_{T,w}(x') + \Phi(x, T(x')) - \Phi(x', T(x'))$$

holds for all pair (x, x'). Thus, by first order conditions at $x' = x$, assuming differentiability,

$$\nabla U_{T,w}(x) = \frac{\partial \Phi}{\partial x}(x, T(x))$$

which is exactly condition (3.15). This suggests that if T is implementable, then T is an optimal transportation map in an optimal transport problem with cost Φ.

In the one-dimensional case and when Φ satisfies the Spence-Mirrlees condition $\partial^2\Phi/\partial x\partial y > 0$, it is well-known that T is implementable if and only if T is nondecreasing, that is T is the gradient of a convex function, hence an optimal transport map in the problem with submodular costs. It turns out that this result is very general.

We state our main result for this section, which is a general characterization of implementable maps using the theory of optimal transport. Let μ be any Borel probability measure μ absolutely continuous with respect to the Lebesgue measure, and let ν be the image measure of μ by T, that is, the distribution of $T(X)$, where $X \sim \mu$. Then we have.

Theorem 3.6. *The following statements are equivalent:*
(i) The map $T : \mathcal{X} \to \mathcal{Y}$ is implementable
(ii) One has

$$\int_{\mathcal{X}} \Phi(x, T(x))\, d\mu(x) = \sup_{\pi \in \mathcal{M}(\mu,\nu)} \int_{\mathcal{X}\times\mathcal{Y}} \Phi(x, y)\, d\pi(x, y).$$

In that case, one has for μ-almost x

$$\nabla u(x) = \frac{\partial \Phi(x, T(x))}{\partial x}$$

where u appears in the solution of (3.13). In that case, the transfer schedule is given by

$$w(x) = u(x) - \Phi(x, T(x))$$

and u is the potential associated with the contract (T, w)

$$U_{T,w} = u.$$

From this equivalence we see that testing whether a given map T is implementable or not *has become a simple optimization problem!*

In particular, when $\Phi(x, y) = x \cdot y$, then the implementable functions are the gradient of convex functions. One example of this is the following:

Example 3.2 (Information elicitation over distributions). *Assume there are d states of the world $\{\omega_1, ..., \omega_d\}$, and the agent is an "expert" who has subjective probabilities, $x = p(\omega_1), ..., p(\omega_d)$ over these contingent states. The agent's type coincides with this vector of subjective probabilities, and the principal wants the agent to reveal her type. The outcome is a payoff associated to each $\{\omega_1, ..., \omega_d\}$, call it $y = (U(\omega_1), ..., U(\omega_d))$.*

Assuming the agent has von-Neumann utility, then

$$\Phi(x, y) = \sum_{\omega} p(\omega) U(\omega) = x \cdot y$$

Then the implementable functions are gradient of convex functions of $x = (p(\omega_1), ..., p(\omega_d))$. See [74] for more.

3.6 References and notes

Section 3.1 provides basic results in the theory of optimal transport. Our exposition follows from Villani (2003) [110], which we refer the reader for a mathematical yet intuitive introduction to the topic. Villani (2009) [111] is an up-to-date extensive treatment. The theory of optimal transport may be seen as an abstract theory of matching but it has also a wealth of applications in other sciences: meteorology, fluid mechanics, astrophysics, transportation of course, and even pure mathematics. Cédric Villani, a leading expert in this theory and the author of the two treatises [110] and [111], was one of the recipients of the Fields medal in 2010.

The optimal transport problem was first posed by Monge in his 1781 paper, *Mémoire sur la théorie des déblais et des remblais.* Although Monge provided geometrical insights, he failed to provide a full solution of the problem. This is due to two main reasons. First, costs considered by Monge were distance costs, which yields to considerable analytical complications compared to quadratic distance costs. Also, Monge only looked for pure solutions, which is a highly nonlinear problem, making it difficult to attack. It is one century and a half later that a next breakthrough occurred, when in the 1940's Soviet mathematician Leonid Kantorovich, as well as the Dutch economist Tjalling Koopmans independently studied the linear relaxation of the problem and its Linear Programming formulation. Both of them are considered to be the founders of Linear Programming – they shared the Sveriges Riksbank Prize in Economics in Memory of Alfred Nobel in 1975 for "their contribution to the theory of optimal allocation of resources".

The proof of Theorem 3.1 is in Villani (2003) [110], ch. 1 for points 1 and 2, and Ex. 2.36 for point 3. The Monge-Kantorovich has also been applied to identification in the presence of partial identification; see [52] and references therein.

Standard references for 3.2 and the Hardy-Littlewood inequality are the classical text by Hardy, Littlewood and Polya (1952) [61], and for the submodular extension, Lorentz (1953) [75]. See application to statistics in [24] and [25].

Section 3.3 deals with quadratic costs. Early results were obtained by Knott and Smith (1984 [71]). Theorem 3.4 is based on results of Brenier (1987 [18], 1991 [19]) and Rachev and Rüschendorf (1990 [86]). See a more complete treatment in [110], Chapters 2 and 3. Example 3.1 can be found in [87] I, Ex. 3.2.12; see reference therein.

The references for Sections 3.4 and 3.5 are the three papers by Carlier [20], [21] and [22], from which we have borrowed most of the exposition, and to which we refer for the proofs, as well as for an exposition of the theory of Φ-convex functions exposed in Appendix E. See also [77] for an early reference on multivariate contract theory. See Vohra (2008) [115] for results analogous to Theorem 3.6 and an interpretation in terms of cycles. Example 3.2 is taken from Lambert (2009) [74].

4 Transportation on networks

Our focus so far when we have considered "Optimal transport" was the cost of going from A to B, not really the way we went from A to B. In some cases, this is a trivial problem, as in the fountains problem, where individuals may go to fountains along a straight line; however there may be situations where this assumption is not realistic (e.g. urban transportation is done through a network). Hence, we shall focus more deeply on transportation on networks, when one may go from A to B only through a well-defined network, incurring some well-defined cost to move from a node of the network to an adjacent one.

4.1 Prelude: the Silk Road

As an introduction, we are going to take an ancient and fascinating example: the Silk road. The secret of silk manufacturing was one of the best-kept secrets of the ancient times. Sericulture, the culture of silk from the cocoons of the larvae of the mulberry silkworm *Bombyx mori*, was known to Chinese only for thousands of years; and during the Han dynasty (206 BC-220 AD) death penalty was the fate of anyone caught exporting the silkworm. Although the secret spread to Korea during that period, and then to India, it remained unknown to Europeans until the middle of the VIth century, when it reached the Byzantine empire during the reign of Justinian I. Needless to say, Europeans valued immensely the fabrics, which was sold at fantastic prices in Rome and Byzantium. The silk was shipped via a network of routes through Central Asia (and later, more and more by the sea).

Although one usually refers to "Silk Road", it would be more correct to use the term "silk roads", as there were in fact several routes, and yet this would omit the fact that through these routes also went various other goods, in order to balance the trade: silks, metals, woods, gems, spices, were shipped through these routes – but also, slaves... and philosophies, religions, and epidemics, such as plagues. The goods were shipped on caravans, which did not went along the whole route, but instead went back and forth on distances which rarely exceeded 200 or 300 miles.

It is a fascinating matter to think that silk was shipped from China to the Roman Empire, in spite of the absence of contact between these two far apart civilizations. An irresistible force moved the precious fabric over 6,000 miles, across deserts, mountains, all sorts of kingdoms and tribal areas, more

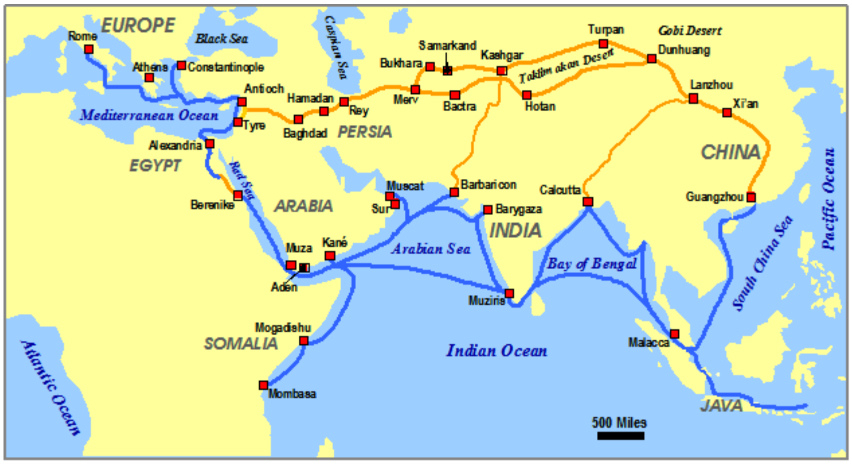


Figure 4: The silk roads. Source: themoderatevoice.com

surely than any planned expedition; and an unambiguous signal showed the way to Byzantium better than any compass. The force, of course, is greed – and the signal is prices: prices of silk (in its gold equivalent) steadily increased between Xi'an or Guangzhou all the way to Antioch or Rome; and it is the prospect of obtaining a portion of that value that traders would venture on caravans across the steps. An explorer fallen from Mars lost in a village in Central Asia would almost have been able to deduce its position in between China and Europe from the trading price of silk. Indeed, assuming that there were enough traders so to maintain competitive prices, that price would reflect the transportation cost along the cheapest path from producers.

Of course, the problem is more complicated than this, as there not a single location producing silk, but many, and not a single location producing silk, but many. Figure 4 is a sketch of these routes. We are now about to model this network and the transportation problem on it.

4.2 Network flows

Nodes, arcs. Consider a directed graph $G = (V, A)$ composed of a set of *nodes* (cities) $V = \{1, ..., n\}$ and a set of *arcs* A, which are pairs (i, j) where $i, j \in V$. The existence of an arc from i to j means that there is a direct way of transportation from i to j; note that these are *directed* arcs, as the existence of an (i, j) arc need not imply the existence of a (j, i) arc.

Paths, networks. There is a *path* from i to j if one can find arcs $(i_0, i_1), (i_1, i_2) ... (i_{p-1}, i_p)$ such that $i_0 = i$ and $i_p = j$. G is called *strongly connected* if for any pair (i, j) such that $i \neq j$, there is a path from i to j. For each pair (i, j) that is an arc, one defines the direct transportation cost of going from i to j as a real number k_{ij}. For each pair (i, j) that is not an arc, one can define $k_{ij} = +\infty$. A strongly connected directed graph G endowed with a direct transportation cost is called a *network*.

Flows. For each arc A, let π_{ij} be the *intermediate flow* of goods through arc (i, j). For each $i \in V$, let b_i be the *exiting flow* of goods disappearing from the network. If $b_i < 0$, there is an incoming flow of goods in the network, and one says i is a source node – modeling supply. If $b_i > 0$ there is an outgoing flow of goods and node i is called a terminal node – modeling demand. The set of source and terminal nodes are denoted S and T, respectively.

Capacity constraints. The network might have *capacity constraints*, which are upper and lower bounds on the intermediate flow

$$l_{ij} \leq \pi_{ij} \leq u_{ij} \tag{4.1}$$

In the sequel, we shall impose $l_{ij} = 0$ and $u_{ij} = +\infty$.

Conservation equation. The *conservation equation* requires that the total intermediate flow into a node minus the total intermediate flow out of a node must equal to the exiting flow. Namely,

$$b_i = \sum_{k:(k,i) \in A} \pi_{ki} - \sum_{k:(i,k) \in A} \pi_{ik}$$

should hold for all i. This equation is sometimes also called *Kirchoff's law*.

For a given node i and a given arc a, we define the *node-incidence matrix* $\mathcal{N}$ as

$$\begin{aligned} \mathcal{N}_{ia} &= 1 \text{ if } a \text{ is out of } i \\ &= -1 \text{ if } a \text{ is into } i \\ &= 0 \text{ otherwise} \end{aligned}$$

Hence, conservation equation can be rewritten as

$$\mathcal{N}\pi = b, \tag{4.2}$$

which is understood as

$$\sum_{a \in A} \mathcal{N}_{ia} \pi_a = b_i.$$

Feasible flows. A flow is called *feasible* if there exist a solution to conservation equation (4.2) which satisfies the capacity constraints (4.1).

Definition 4.1. *The set of feasible network flows $\mathcal{M}(b)$ is defined by*

$$\mathcal{M}(b) = \{\pi \geq 0 : \mathcal{N}\pi = b\}.$$

We shall often omit the adjective "feasible", implicitly meaning feasible network flows when mentioning only network flows.

We have obviously that for any arc a,

$$\sum_{i \in V} \mathcal{N}_{ia} = 0$$

thus if π satisfies (4.2), then $\sum_{i \in V} b_i = 0$, therefore for a feasible flow to exist, then total supply should necessarily equal total demand, i.e.

$$-\sum_{i \in S} b_i = \sum_{i \in T} b_i,$$

which we shall assume from now on.

Clearly, the notation $\mathcal{M}(b)$ is purposely chosen so to remind us of the notation $\mathcal{M}(p, q)$ in Chapter 3. Indeed, a (fractional) assignment is a particular case of network flow. (Note: there is a slight abuse of language here, as the network in the proposition to follow is not strongly connected; however it does not affect the results).

Proposition 4.1 (Matchings from network flows). *Consider the setting of Theorem 3.2 where $\mathcal{X} = \{1, ..., N\}$ and $\mathcal{Y} = \{1, ..., M\}$ on which are defined two probability vectors (p_i) and (q_j) respectively. Then one can setup* *the problem as a network flow problem by taking* $V = \mathcal{X} \cup \mathcal{Y}$*, and* $A = \{(i,j) : i \in \mathcal{X},\ j \in \mathcal{Y}\}$*, and*

$$\begin{aligned} b_k &= -p_k \text{ if } k \in \mathcal{X} \\ &= q_k \text{ if } k \in \mathcal{Y}. \end{aligned}$$

Then $\mathcal{M}(b)$ *and* $\mathcal{M}(p,q)$ *coincide.*

Proof. The feasibility constraint can be written in general

$$b_k = \sum_{i:(i,k)\in A} \pi_{ik} - \sum_{j:(k,j)\in A} \pi_{kj}$$

but for $k \in \mathcal{X}$, no arc arrives at k, thus this becomes

$$b_k = -\sum_{j:(k,j)\in A} \pi_{kj}$$

hence

$$\sum_{j:(k,j)\in A} \pi_{kj} = p_k \text{ for } k \in \mathcal{X}$$

while for $k \in \mathcal{Y}$, no arc leaves from k, thus the feasibility constraint becomes

$$b_k = \sum_{i:(i,k)\in A} \pi_{ik}$$

that is,

$$\sum_{i:(i,k)\in A} \pi_{ik} = q_k \text{ for } k \in \mathcal{Y}.$$

■

Just like the set of fractional matchings in convex, so it the set of (feasible) network flows. As it is defined by a finite number of linear inequalities, we have that:

Proposition 4.2. *The set of feasible network flows is a convex polyhedron.*

Just as it was the case for matchings, it will be of great importance to determine what are the extreme points of this polyhedron.

Flow decomposition. For any source node i and terminal node j, introduce $\mathcal{P}_{ij}$ the set of paths from i to j. Let $\mathcal{P}$ be the set of all paths from source nodes to terminal nodes. Finally, let $\mathcal{C}$ be the set of directed cycles in the network.

For any $\rho \in \mathcal{P}$, let h_ρ be the intensity of the flow along ρ, and for $\mu \in \mathcal{C}$, let g_μ be the intensity of the flow in the cycle μ. A flow on a path is called a *path flow*, and a flow on a cycle is called a *cycle flow*.

Clearly, (h, g) defines a network flow π (not necessarily feasible), which is defined by

$$\pi_a = \sum_{\rho \in \mathcal{P}} h_\rho 1\{a \in \rho\} + \sum_{\mu \in \mathcal{C}} g_\mu 1\{a \in \mu\}. \tag{4.3}$$

Now, if

$$\forall i \in S, \ \sum_{j \in T} \sum_{\rho \in \mathcal{P}_{ij}} h_\rho = -b_i \tag{4.4}$$

$$\forall j \in T, \ \sum_{i \in S} \sum_{\rho \in \mathcal{P}_{ij}} h_\rho = b_j \tag{4.5}$$

then π is feasible.

One can actually show that for any feasible network flow π, there exists (h, g) such that π can be decomposed into path flows and cycle flows of intensity h and g.

Theorem 4.1 (Flow decomposition). *Any feasible network flow π can be decomposed according to (4.3) into path flows and cycle flows of intensity h and g such that h satisfies Equations (4.4) and (4.5). Conversely, any distribution over path flows and cycle flows with intensity h and g satisfying Equations (4.4) and (4.5) can be associated to a feasible network flow by Equation (4.3).*

The proof of this results relies on the following lemma.

Lemma 4.1. *Let $\pi \in \mathcal{M}(b)$, such that $\pi_{ij} > 0$ for some arc (i, j). Then there is either a path from S to T, or a cycle passing through (i, j) along which $\pi > 0$.*

Proof. Step 1. Let us show that there is either a path from S to T passing through (i,j), or a cycle along which $\pi > 0$.

We extend (i,j) on the right to construct a path along which $\pi > 0$ by iterating forward from j on. Repeating this operation, either one forms a loop by repeating a node, or one forms a path from i to T. We then extend (i,j) on the left to construct a path along which $\pi > 0$ by iterating backward from i on. Repeating this operation, either one forms a loop by repeating a node, or one forms a path from S to j. Then we have either build a path from S to T passing through (i,j), or a loop somewhere.

Step 2. If the loop does not go through (i,j), we may remove the smallest value of the flow along it to the arcs involved and invoke Step 1 repeatedly until the loop has to pass through (i,j). ■

Proof of Theorem 4.1. Consider

$$\begin{aligned}
&\max \sum_{\rho\in\mathcal{P}} h_\rho + \sum_{\mu\in\mathcal{C}} g_\mu \\
&s.t. \\
&\sum_{\rho\in\mathcal{P}} h_\rho 1\{(i,j)\in\rho\} + \sum_{\mu\in\mathcal{C}} g_\mu 1\{(i,j)\in\mu\} \leq \pi_{ij}
\end{aligned}$$

Applying the Lemma to

$$\pi_{ij} - \sum_{\rho\in\mathcal{P}} h_\rho 1\{(i,j)\in\rho\} - \sum_{\mu\in\mathcal{C}} g_\mu 1\{(i,j)\in\mu\}$$

yields the desired result. ■

4.3 The shortest path problem

The shortest path problem is a very natural one. Given two nodes i and j, what is the path from i to j which minimizes the total cost incurred? in other words, what is the cost-optimal way to move one unit of mass from i to j?

For i and j two distinct nodes, define the *reduced cost* by

$$c_{ij} = \inf\left\{\sum_{k=0}^{p-1} k_{i_k i_{k+1}} : i_0 = i, ..., i_p = j\right\} \tag{4.6}$$

and if this infimum is attained by $i_0, ..., i_p \in V$, $i_0 = i, i_p = j$, such path is called is called a *shortest path* (sometimes also called *geodesic*) from i to j. Note that shortest paths are by no mean unique in general. Also note that the reduced cost satisfies the *triangle inequality*

$$c_{ij} \leq c_{ip} + c_{pj}$$

which the direct cost need not satisfy. Remember that we have assumed here that the network is strongly connected: there is a path from any node to another.

The following proposition is obvious. It says in particular that if k is a proper distance, then the reduce cost c coincides with the direct cost k.

Proposition 4.3. *One has*

$$\forall i, j: \; c_{ij} \leq k_{ij}$$

with equality for all i, j *if and only if* k *satisfies the triangle inequality*

$$\forall i, j, l: \; k_{ij} \leq k_{il} + k_{lj}.$$

Also, if k is symmetric, that is if $k_{ij} = k_{ji}$ for all pairs i,j, and if $k_{ii} = 0$ for all i, then c is a proper distance.

Proposition 4.4. *Assume there is no cycle of negative length in the network. Then for any two distinct nodes* i *and* j*, there is a shortest path from* i *to* j*, i.e.* $c_{ij} > -\infty$*.*

Conversely, if there is a cycle of negative length in the network, then for two distinct nodes i *and* j *there is no shortest path from* i *to* j*.*

Proof. Assume there is no cycle of negative length in the network. Then the infimum in (4.6) can be taken over cycles where there is at most a single occurrence of each node. There is a finite number of such cycles, hence the problem has a minimizer.

Conversely, assume there is a cycle of negative length, $c = (i_0, ..., i_p)$ such that $i_0 = i_p = p$ and

$$\sum_{k=0}^{p-1} k_{i_k i_{k+1}} = L < 0$$

Then for any i and j, one can attain p from i at finite cost, and then repeat the cycle c an arbitrary number of times, and then attain j from p at finite cost. See Figure 4.3. This allows to build paths of arbitrary negative length. Thus there is no shortest path from i to j. ■

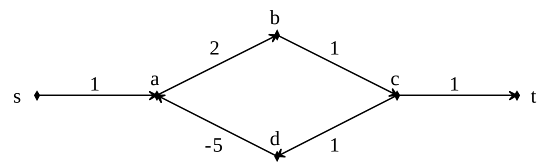


Figure 4.3. (a, b, c, d, a) is a cycle in the graph of length -1.

Consider now the shortest path from a source node $s \in S$ to a terminal node $t \in T$. Introduce $b^{s,t}$ accordingly as

$$\begin{aligned} b_i^{s,t} &= 0 \text{ if } i \in V \backslash \{s, t\} \\ b_s^{s,t} &= -1 \\ b_t^{s,t} &= 1 \end{aligned}$$

and the (s, t) *−shortest path polyhedron* is the set of solutions π of

$$\mathcal{N}\pi = b^{s,t}, \ \pi \geq 0$$

The name shortest path polyhedron is misleading: for π to be in the shortest path polyhedron does not implies it moves mass from s to t in a cost optimal for some cost. In fact, only boundary points of it will have that property.

The flow associated to the shortest path problem is a solution of

$$\begin{aligned} &\min \sum_{(i,j) \in A} \pi_{ij} k_{ij} \\ &s.t. \ \pi \geq 0 \\ &\mathcal{N}\pi = b^{s,t} \end{aligned}$$

Hence, this π can be chosen to be an extremal solution of the shortest path polyhedron.

We have the following result.

Proposition 4.5. *Every extreme point of the shortest path polyhedron is integral.*

Proof. Consider an extreme point of the shortest path polyhedron; it thus is a solution of the problem

$$\begin{aligned} c_{st} &= \min \sum_{(i,j)\in A} \pi_{ij} k_{ij} \\ &\quad s.t.\ \pi \geq 0 \\ &\quad \mathcal{N}\pi = b^{s,t} \end{aligned}$$

for some direct cost function k. From the Flow Decomposition theorem (4.1), π_a can be obtained as

$$\pi_a = \sum_{\rho\in\mathcal{P}_{st}} h_\rho 1\{a\in\rho\} + \sum_{\mu\in\mathcal{C}} g_\mu 1\{a\in\mu\}$$

where

$$\sum_{\rho\in\mathcal{P}_{st}} h_\rho = 1$$

Clearly, it is optimal to set $g_\mu = 0$ as there are no cycles of negative length for k (otherwise the optimum would be $-\infty$). Now one can weakly improve the objective by choosing ρ_0 among the $\rho \in \mathcal{P}_{st}$ such that $h_\rho > 0$ that minimizes the cost along ρ, and replace h by $1\{\rho = \rho_0\}$. ∎

Finally, the shortest path problem is a linear programming problem, and has an interesting formulation.

Theorem 4.2. *The value c_{st} of the primal of the shortest path problem*

$$\begin{aligned} &\min \sum_{(i,j)\in A} \pi_{ij} k_{ij} \\ &s.t.\ \pi \geq 0 \\ &\mathcal{N}\pi = b^{s,t} \end{aligned}$$

coincides with the value of its dual formulation

$$\begin{aligned} &\max w_t - w_s \\ &s.t.\ w_j - w_i \leq k_{ij} \end{aligned}$$

which is feasible if and only if the network has no cycle of negative length.

This result is a specialization of the so-called min-flow, max-cut theorem (Theorem 4.3 below) for which a proof will be given.

As an example, consider the network in Figure 4.3 below.

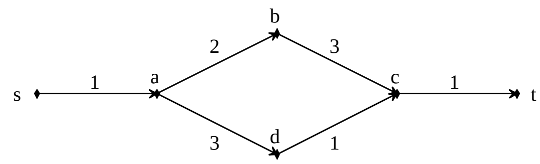


Figure 4.3. The shortest path from s to t, (s, a, d, c, t), has lenght 6.

Clearly, the shortest path has length 6. Also, we can set $w_s = 0$, then $w_a = 1$, $w_d = 4$, $w_c = 5$ and $w_t = 6$, saturating the dual constraint, while one may choose $w_b \in [2, 3]$ without saturating the two constraints relative to arcs (a, b) and (b, c).

4.4 Optimal transport on networks

In the previous section, we only dealt with the shortest path problem, which is the individual decision problem – or how to move mass from a given source point to a give terminal point. We are now concerned about how to move (still in a cost-optimal way) a *distribution* of mass over source nodes to a distribution of mass over terminal nodes. Throughout this section, we shall assume that the network has no cycle of negative length – hence the reduced cost function c_{ij} is defined and finite over every arc (i, j).

Our problem is to minimize the total cost of transportation under the feasibility constraint, hence solve the *min-flow problem*

$$\begin{aligned} &\min \sum_{(i,j)\in A} \pi_{ij} k_{ij} \\ &s.t.\ \pi \geq 0 \\ &\mathcal{N}\pi = b \end{aligned} \tag{4.7}$$

We are about to show that the minimum flow problem can be formulated as a matching problem. Together with proposition 4.1, this shows no less than the full equivalence of the matching problem and the network flow problem. The connection between the two problem is that the cost in the optimal assignment problem should be a reduced cost in the network flow problem. But this is not a restrictive requirement, as we reduced costs always exist in the network problem, and whatever the matching cost in the optimal assignment problem, it is easy to set up a network so that the reduced cost coincides with this matching cost.

Proposition 4.6 (Network flows from matchings). *For $i \in S$, let $p_i = -b_i$, and for $j \in T$, let $q_j = b_j$. Then the value of the minimum-flow problem (4.7) coincides with the value of the optimal assignment problem, where the cost function is given by the reduced cost c_{ij}*

$$\begin{aligned} &\min_{\pi_{ij} \geq 0} \sum_{ij} \pi_{ij} c_{ij} \\ &s.t. \\ &\sum_{j \in T} \pi_{ij} = p_i \\ &\sum_{i \in S} \pi_{ij} = q_j \end{aligned} \tag{4.8}$$

Further, for two nodes $i \in S$ and $j \in T$, let $\rho^{ij} \in \mathcal{P}$ be a shortest path from i to j, and c_{ij} be the associated cost. Then the decomposition of an optimal π is given by setting weight π_{ij} on path ρ^{ij}.

Proof. From the Flow Decomposition theorem 4.1, and as cycles we can recast the problem as the minimization of

$$\sum_{a \in A} \pi_a k_a$$

over the set of π_a obtained as

$$\pi_a = \sum_{i \in S, j \in T} \sum_{\rho \in \mathcal{P}_{ij}} h_\rho 1\{a \in \rho\} + \sum_{\mu \in \mathcal{C}} g_\mu 1\{a \in \mu\}$$

with

$$\begin{aligned} \forall i \in S, \ &\sum_{j \in T} \sum_{\rho \in \mathcal{P}_{ij}} h_\rho = -b_i \\ \forall j \in T, \ &\sum_{i \in S} \sum_{\rho \in \mathcal{P}_{ij}} h_\rho = b_j \end{aligned}$$

Clearly, it is optimal to set $g_\mu = 0$ as there are no cycles of negative length. Now given $i \in S$ and $j \in T$, consider $\rho \in \mathcal{P}_{ij}$. One can always weakly improve on ρ by choosing it a shortest path, in which case the associated cost is c_{ij} and one can define $\pi_{ij} = h_\rho$. Thus the value of the program is given by (4.8). ■

The following corollary is immediate given Theorem 2.3.

Corollary 4.1. *Assume that there are as many source points as terminal points, and assume $b_i = -1$ for $i \in S$ and $b_j = 1$ for $j \in T$. Then the optimal flow is integral.*

The minimum flow problem is a linear programming problem, and we now give a useful duality result.

Theorem 4.3 (min-flow, max-cut). *The primal value of the transport problem on the network*

$$\begin{aligned} &\min_{\pi \geq 0} \sum_{(i,j)\in A} \pi_{ij} k_{ij} \\ &s.t.\ \mathcal{N}\pi = b \end{aligned}$$

coincides with its dual value

$$\begin{aligned} &\max_{w} \sum_{i \in V} w_i b_i \\ &w_j - w_i \leq k_{ij} \end{aligned}$$

further, $\pi_{ij} > 0$ implies $w_j - w_i = k_{ij}$.

To give the economic interpretation of this theorem, let us go back to our Silk Road example. This theorem tells us that if w_i is the price charged for silk at a point i in the network, then any new merchant arriving at node i will be indifferent between purchasing or selling the silk at price w_i, or shipping it themselves optimally (assuming they incur transportation costs k_{ij}). Of course, this assumes that there is no externality, either positive (network effect) or negative (congestion effect). We shall investigate what happens in the case of congestion effects in the next paragraph.

Proof of Theorem 4.3. The proof follows from standard Linear Programming duality. The dual of

$$\begin{aligned} &\min_{\pi \geq 0} \pi.k \\ &s.t.\ \mathcal{N}\pi = b \end{aligned}$$

is

$$\begin{aligned} &\max w.b \\ &s.t.\ w'\mathcal{N} \leq k \end{aligned}$$

where the constraint means that for all $a \in A$,

$$\sum_{k \in V} w_k \mathcal{N}_{ka} \leq k_a$$

or equivalently, for each $(ij) \in A$, one has

$$w_j - w_i \leq k_{ij}.$$

■

Stable set. The set of vectors w_i that satisfy the constraint of the dual problem is called the *stable set*, denoted S_k. Remember the interpretation of w_i as the price of silk in city i. If in some city, $w_j - w_i > k_{ij}$, it means that the silk is too expensive in city j: indeed, a entrant can purchase silk in city i at price w_i, ship it to j at cost k_{ij} and offer it in city j at for any price between $w_i + k_{ij}$ and w_j. Thus $w_j - w_i \leq k_{ij}$ holds necessarily. Now if silk is actually shipped from city i to city j, then $w_j \geq k_{ij} + w_i$ otherwise merchants shipping the silk would incur a loss.

This motivates the following definition.

Definition 4.2. *The stable set S_k is defined by*

$$S_k = \left\{ w \in \mathbb{R}^n : w_j - w_i \leq k_{ij} \right\}.$$

As a corollary of Theorem 4.3, we have the following important corollary.

Corollary 4.2. *The following properties are equivalent:*
(i) k has no cycles of negative length, and
(ii) the stable set S_k is nonempty.

Proof. Pick any pair (s, t). Then k has no cycles of negative length if and only if the value c_{st} of the primal of the shortest path problem from s to t is finite, hence if and only if the dual is feasible, which precisely says that S_k is nonempty. ∎

Lattice structure. We are now about to state a fundamental property of the structure of the stable set, namely *lattice structure*. A lattice is a set where for to pairs of elements x and y, one can define a "join" $x \vee y$ and a "meet" $x \wedge y$ operation. Refer to Appendix G for background information on lattices.

Assume that there are two competitive networks, SilkyWay SleekSilk who ship silk from China to the West. They both face the same transportation costs, but for some reason their (stable) prices are different along these two parallel networks – remember, the stable set has not necessarily a unique element.

Assume that the tourists visiting cities of the silk road use two tour operators, the budget Smart Routes and the lavish Sucker Tour. In each city, clients of Sucker Tour are forced to buy the silk at the more expensive price of either network, while clients of Smart Routes buy it at the cheaper price. It turns out that the prices of both Smart Routes and Sucker Tour will be stable. In other words, although clients of Smart Routes sometimes buy from SilkyWay and sometimes from SleekSilk, there is no way they can ship silk between them in order to improve the efficiency of their network. The same is true of the clients of Sucker Tour. Of course, everybody prefers to be a client of Smart Routes. The prices of Smart Routes and Sucker Tour have between them a property that the original prices of SilkyWay and SleekSilk had not: they are ordered. This is formalized into the following property.

Proposition 4.7 (Lattice structure of the stable set). *The stable set is a lattice. In other words, for $w, \tilde{w} \in S_k$, one has*

$$w \vee \tilde{w} \in S_k \text{ and } w \wedge \tilde{w} \in S_k.$$

Proof. Denote $z_i = w_i \vee \tilde{w}_i$. Let $i, j \in \{1, ..., n\}$, and assume w.l.o.g. that $w_j \geq \tilde{w}_j$, hence $z_j = w_j$. Then:

– either $w_i \geq \tilde{w}_i$ thus $z_i = w_i$ and $z_j - z_i = w_j - w_i \leq k_{ij}$,

– or $w_i < \tilde{w}_i$, in which case $z_j - z_i = w_j - \tilde{w}_i < w_j - w_i \leq k_{ij}$.

Either way $z_j - z_i \leq k_{ij}$, hence $z = w \vee \tilde{w} \in S_k$.

Similarly, let $x_i = w_i \wedge \tilde{w}_i$. Let $i, j \in \{1, ..., n\}$, and assume w.l.o.g. that $w_i \leq \tilde{w}_i$, hence $x_i = w_i$. Then:

– either $w_j \leq \tilde{w}_j$ in which case $x_j = w_j$, and thus $x_j - x_i = w_j - w_i \leq k_{ij}$,

– or $w_j > \tilde{w}_j$ in which case $x_j = \tilde{w}_j$, and thus $x_j - x_i = \tilde{w}_j - w_i < w_j - w_i \leq k_{ij}$.

Either way $x_j - x_i \leq k_{ij}$, hence $x = w \wedge \tilde{w} \in S_k$. ∎

Actually, the latter property can be deduced from an even more general property showing that given any set of price vectors indexed by $\omega \in \{1, ..., \Omega\}$, one can rearrange these into an *ordered* set of price vectors of same cardinality. Clearly, one recovers Proposition 4.7 when the cardinality of Ω is two.

Theorem 4.4. *Consider a finite number Ω of stable prices $w(\omega)$*

$$\forall \omega \in \{1, ..., \Omega\}, \ w(\omega) \in S_k.$$

Define $w_i^*(\omega)$ as the ω-th smallest number out of

$$\{w_i^*(1), ..., w_i^*(\Omega)\}.$$

In particular,

$$\begin{aligned} w_i^*(1) &= \min_{\omega \in \{1,...,\Omega\}} (w_i^*(\omega)), \\ w_i^*(\Omega) &= \max_{\omega \in \{1,...,\Omega\}} (w_i^*(\omega)). \end{aligned}$$

Then

$$\forall \omega \in \{1, ..., \Omega\}, \ w^*(\omega) \in S_k.$$

Proof. One has by assumption

$$\forall \omega \in \{1, ..., \Omega\}, \ w_j(\omega) \leq k_{ij} + w_i(\omega)$$

This implies that $k_{ij} + w_i(\omega)$ first-order stochastically dominates $w_j(\omega)$ (see Appendix D), which implies by Proposition D.1

$$\forall \omega \in \{1, ..., \Omega\}, \ w_j^*(\omega) \leq k_{ij} + w_i^*(\omega),$$

hence $w^*(\omega) \in S_k$. ∎

The lattice structure of S_k will have important consequences as it will allow us to apply some results of optimization on lattices. Introduce $T(b)$ the value of transportation problem.

$$\begin{aligned} T(b) &= \max w.b \\ & \quad s.t.\ w \in S_k \end{aligned}$$

One has the following result.

Theorem 4.5. *T is supermodular.*

Proof. The proof of this result follows from Theorem G.2, but we shall prove it by elementary means. Consider b and $\tilde{b}$ two mass distribution which each put equal total mass to source and destinations. Then

$$T(b) + T\left(\tilde{b}\right) = w.b + \tilde{w}.\tilde{b}$$

for some $w, \tilde{w} \in S_k$. Now,

$$\begin{aligned} w.b + \tilde{w}.\tilde{b} &\leq (w \wedge \tilde{w}) . \left(b \wedge \tilde{b}\right) + (w \vee \tilde{w}) . \left(b \vee \tilde{b}\right) \\ &\leq T\left(b \wedge \tilde{b}\right) + T\left(b \vee \tilde{b}\right), \end{aligned}$$

where the first inequality stems from the by supermodularity of the scalar product, and the second one from the fact that $w \wedge \tilde{w}$ and $w \vee \tilde{w}$ are in S_k. QED. ■

4.5 References and notes

Sections 4.2 and 4.3 are mostly taken from Vohra's lecture notes [115]; see a constructive proof of Theorem 4.1 there. See Beckmann (1952) [8] for a continuous version of the model. The min-flow, max-cut theorem (which is more often called max-flow, min-cut theorem) was proved by Ford and Fulkerson (1956) [45].

Part II. Applications of TU models

5 Urban economics

There are many features an economist may want to model in a city.

[...]

In this chapter we shall mostly focus on *congestion.* For a traveler going from A to B, the transportation cost depends on a first approximation of the length of the shortest path from A to B. But it also depends on another parameter: congestion. If many other people take the same itinerary at the same time, there will be a traffic jam, and everyone's transportation cost will increase. In order to study this problem, we need to keep track of the path of each individual, in order to compute how saturated the transportation network is.

5.1 Congestion: social planning problem

In the problem we considered in Section 4.4, we were minimizing a linear transportation cost $\mathcal{W}(\pi)$ under feasibility constraints, i.e.

$$\begin{aligned}&\min \mathcal{W}(\pi)\\&s.t.\ \pi_{ij} \geq 0\\&\mathcal{N}\pi = b\end{aligned}$$

We now would like to relax the assumption that our total total cost function $\mathcal{W}$ should be linear with respect to π. We shall take $\mathcal{W}$ as a separable function

$$\mathcal{W}(\pi) = \sum_{(i,j)\in A} K_{ij}(\pi_{ij})$$

where $K_{ij}(.)$ are real valued functions, one for each arc.

This allows us to model *positive network spillovers*, which is the case where there are positive externalities, captured by the choice of $K_{ij}(x)$ as concave function, which means that path from i to j becomes less and less costly the more people go through it.

Negative externalities, or *congestion effect*, are captured by a choice of convex function for $K_{ij}(x)$. Throughout the sequel, we shall assume that this is the case.

Theorem 5.1. *Assume that $\mathcal{W}$ is a convex function. Then the primal value of the optimal transportation problem on the network*

$$\begin{aligned} & \min \mathcal{W}(\pi) \\ & s.t.\ \pi \geq 0 \\ & \mathcal{N}\pi = b \end{aligned} \tag{5.1}$$

coincides with its dual value, which is

$$\max_{w} \sum_{i} w_i b_i - \mathcal{W}^*(w'\mathcal{N}) \tag{5.2}$$

where

$$(w'\mathcal{N})_{ij} = w_j - w_i$$

and $\mathcal{W}^*$ is the convex conjugate function to $\mathcal{W}$, i.e.

$$\mathcal{W}^*(\kappa) = \sup_{\pi_{ij} \geq 0} \left(\sum_{(i,j)\in A} \pi_{ij}\kappa_{ij} - \mathcal{W}(\pi) \right). \tag{5.3}$$

Proof. This follows from a min-max result, as one has

$$\begin{aligned} & \min_{\pi \geq 0} \max_{w} \mathcal{W}(\pi) + w'(b - \mathcal{N}\pi) \\ = & \max_{w} w'b + \min_{\pi \geq 0} \mathcal{W}(\pi) - w'\mathcal{N}\pi \\ = & \max_{w} w'b - \max_{\pi \geq 0} w'\mathcal{N}\pi - \mathcal{W}(\pi) \\ = & \max_{w} w'b - \mathcal{W}^*(w'\mathcal{N}). \end{aligned}$$

■

Let us give a few examples. First, we show that, unsurprisingly, this theorem is a generalization of the min-flow, max-cut theorem.

Example 5.1. *Consider again the linear case, where*

$$\mathcal{W}(\pi) = \sum_{(i,j)\in A} \pi_{ij} k_{ij}.$$

Then, one has

$$\begin{aligned} \mathcal{W}^*(\kappa) & = 0 \text{ if } \kappa_{ij} \leq k_{ij} \text{ for all } (i,j) \in A \\ & = +\infty \text{ otherwise.} \end{aligned}$$

Hence, Equation (5.2) becomes

$$\max_w w'b$$
$$s.t.\ w'\mathcal{N} \leq k$$

recovering the result from the min-flow, max-cut theorem (Th. 4.3).

We now give a more interesting example, which is fundamental. In the economics of transportation literature, this model is known as Dial's model.

Theorem 5.2. *Consider the case where*

$$\mathcal{W}(\pi) = \sum_{(i,j)\in A} \pi_{ij} k_{ij} + \sigma \sum_{(i,j)\in A} \pi_{ij} \ln \pi_{ij}.$$

In that case, there is a vector $(w_i)_{i\in V}$ *such that for each* $(i,j) \in A$*, the optimal flow* π_{ij} *satisfies the Schrödinger equation*

$$\pi_{ij} = \exp\left(\frac{-k_{ij} + w_j - w_i - 1}{\sigma}\right), \tag{5.4}$$

where the w's exist, are unique up to an additive constant, and are a solution of

$$\max_w \sum_i w_i b_i - \sum_{(i,j)\in A} \sigma \exp\left(\frac{k_{ij} - w_j + w_i - \sigma}{\sigma}\right)$$

and the flow defined by Equation 5.4 is automatically feasible.

The interpretation of this theorem is very interesting. The log-likelihood of a transition from i to j is proportional to minus the direct transportation cost $-k_{ij}$. Hence, all other things equal, all transitions are possible, but less costly transitions will be more likely than others. The potential w_i, on the other hand, adjusts π_{ij} so that it satisfies the feasibility constraint. Hence a terminal node with a high outgoing flow should "pump in" more mass, and therefore transitions to this node should receive higher probability.

It is also, as we shall see in the subsequent chapters, connected with discrete choice models: indeed, it can be interpreted as a random utility model where any path can offer minimal cost with some probability. We stop here as we shall discuss this model, in the case of the marriage market, with much greater detail in Chapter 6.

Proof. Indeed, Equation (5.3) becomes

$$\mathcal{W}^*(\kappa) = \sup_{\pi_{ij} \geq 0} \left(\sum_{(i,j) \in A} \pi_{ij} \left(\kappa_{ij} - k_{ij} - \sigma \ln \pi_{ij} \right) \right),$$

hence by first order conditions,

$$\kappa_{ij} - k_{ij} - \sigma \ln \pi_{ij} - \sigma = 0,$$

hence

$$\pi_{ij} = \exp\left(\frac{\kappa_{ij} - k_{ij} - \sigma}{\sigma} \right),$$

therefore

$$\mathcal{W}^*(\kappa) = \sum_{(i,j) \in A} \sigma \exp\left(\frac{\kappa_{ij} - k_{ij} - \sigma}{\sigma} \right)$$

and when $\kappa = w'\mathcal{N}$, one has $\kappa_{ij} = w_j - w_i$, thus

$$\pi_{ij} = \exp\left(\frac{w_j - w_i - k_{ij} - \sigma}{\sigma} \right),$$

The first order conditions associated to Equation (5.2), one gets

$$b_k = \frac{\partial \mathcal{W}^*(w'\mathcal{N})}{\partial w_k}$$

thus

$$b_k = \sum_{a \in A} \frac{\partial \mathcal{W}^*}{\partial \kappa_a} (w'\mathcal{N}) \mathcal{N}_{ka},$$

hence

$$\begin{aligned} b_k &= \sum_{a \text{ arrives at } k} \exp\left(\frac{\kappa_a - k_a - \sigma}{\sigma} \right) \\ &\quad - \sum_{a \text{ leaves from } k} \exp\left(\frac{\kappa_a - k_a - \sigma}{\sigma} \right) \end{aligned}$$

which is exactly the feasibility equation. ∎

5.2 Congestion: individual decision problem

We now consider the individual decision problem, sometimes called "selfish routing problem". Consider the cost of adding transporting one incremental amount of mass δb in the network from source nodes S to terminal ones T. Let $\delta\pi$ the incremental flow generated.

Assume that the transportation cost of shipping $\delta\pi_{ij}$ through arc (i,j) is a function of the degree of saturation of the network: $k_{ij}(\pi_{ij})\,\delta\pi_{ij}$, where $k_{ij}(.)$ are functions defined over each arcs and assumed to be increasing (in order to model congestion). Clearly, any incremental shipper will face a linear optimization cost with cost $k_{ij} = K'_{ij}(\pi_{ij})$. This rules out cycles, and suboptimal paths in the network flow decomposition and this motivates the notion of a Wardrop equilibrium.

Definition 5.1 (Wardrop equilibrium). *π is a Wardrop equilibrium if given any flow decomposition of π*

$$\pi = \sum_{\rho\in\mathcal{P}} h_\rho 1\{a\in\rho\} + \sum_{\mu\in\mathcal{C}} g_\mu 1\{a\in\mu\},$$

then:

(i) $g_\mu = 0$ for all cycles μ, and

(ii) any path ρ with $h_\rho > 0$ from a source to a terminal node is optimal with respect to cost $k_{ij}(\pi_{ij})$.

One has the following welfare theorem.

Theorem 5.3. *π is a Wardrop equilibrium if and only if it solves problem (5.1)*

$$\begin{aligned} &\min_{\pi\geq 0} \sum_{ij} K_{ij}(\pi_{ij}) \\ &s.t.\ \mathcal{N}\pi = b \end{aligned} \tag{5.5}$$

where K_{ij} is a primitive of k_{ij}, i.e. $K'_{ij}(x) = k_{ij}(x)$.

Proof. By the Kuhn and Tucker conditions, any optimal π is solution of the problem linearized locally around π, hence

$$\begin{aligned} &\min_{\hat\pi\geq 0} \sum_{ij} k_{ij}\hat\pi_{ij} \\ &s.t.\ \mathcal{N}\hat\pi = b \end{aligned}$$

where $k_{ij} = K'_{ij}(\pi_{ij})$. Thus Wardrop equilibria and optimizers of problem (5.1) coincide. ■

Note that π is not optimal. Indeed, the optimal π minimizes instead

$$\begin{aligned}&\min_{\hat{\pi}\geq 0}\sum_{ij} k_{ij}(\hat{\pi}_{ij})\\&s.t.\ \mathcal{N}\hat{\pi} = b\end{aligned}$$

which is a different problem, unless the cost functions k_{ij} are linear.

The function

$$l_{ij}(x) = \frac{k_{ij}(x)}{x} = \frac{K'_{ij}(x)}{x}$$

which captures the cost per unit of traffic is called the *latency function*. With this definition, the optimal π minimizes

$$\begin{aligned}&\min_{\hat{\pi}\geq 0}\sum_{ij} \hat{\pi}_{ij} l_{ij}(\hat{\pi}_{ij})\\&s.t.\ \mathcal{N}\hat{\pi} = b.\end{aligned}\tag{5.6}$$

which is clearly analoguous to (5.5), but l_{ij} is in general different from K_{ij}.

The loss of social welfare due to the difference between the optimal π and the equilibrium π is called in the literature the *price of anarchy*. It can be theoretically bounded. We shall now provide a striking illustration of the departure from social optimum by explaining the celebrated Braess paradox.

Braess paradox. Consider Figure 5.2, where the functions $k_{ij}(x)$ are indicated along the arcs. Thus there is no congestion effect in arcs (a, d) which costs one whatever the traffic is; and there is congestion effect in arcs (a, b) which cost π_{ab} when π_{ab} is the flow through that arc.

One would like to move one unit from node s to node t. In the first picture, the unique Wardrop equilibrium consists in splitting the flow into two halves, one on the path (s, a, b, c, t). Total cost per infinitesimal unit of mass is 3/2 either way, hence total cost is 3/2 and coincides with the optimum.

Let us now consider the second picture, where one has simply added a free arc to the network from b to d. This obviously does not change the optimal flow, and one would anticipate that expanding possibilities has no reverse effect. It turns out that it actually *worsens* the situation. Indeed, irrespective of $x < 1$, the path (s, a, b, d, c, t) is now a shortest path, thus the only Wardrop equilibrium has now all traffic through that path – with a cost of 2.

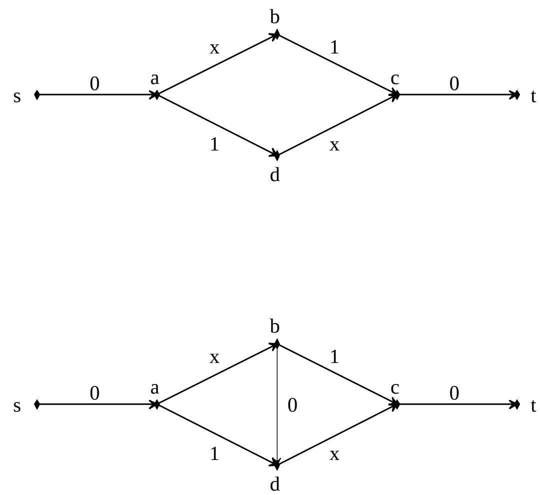


Figure 5.2.

5.3 Urban planning and empirical issues

City planning. The literature on urban economics varies in ambition and scope. The most ambitious and the broader in scope attempts to understand the structure of the cities, by making assumptions on the various positive or negative externalities.

The stylized facts to be captured vary accross periods or countries. Some cities, as most US cities, have a regular street network; in other cities, such as in most European cities, this network is tortuous and irregular. During some periods in time all the businesses were grouped in a single area, in other periods of time there were spread out across the surface of the city. In some cities the downtowns are residential and businesses are the suburbs; in others it is the opposite. Patterns evolve, too: American downtowns have become more residential over the last 20 years. Urban economics models try to capture the most important amongst all sorts of positive and negative externalities between these phonomenons, and make some predictions about the structure of cities.

Network design. In many situations in urban economics, one does not start from stcratch and the problem is not to optimally plan a city, but how to regulate it optimally, e.g. how to regulate traffic optimally. The commuter problem is modelled as a network flow problem with congestion where the set of source nodes S are residential areas and the set of terminal nodes are commercial areas. It is unlikely that the central planner, or the city authorities can direct the traffic in a social optimal way. But the city authorities can design the shape of the network: they can add and remove directions in some arcs (by setting or removing one-way systems) and, in some cases, they can regulate traffic (by setting traffic lights), and, in some rare cases, they can create new vertices by creating new tunnels or highways.

Thus city planners may hope to design the network optimally so as to avoid situations such as the one described in the Braess paradox: sometimes setting a one-way system will improve the efficiency in the network. Hence an important problem in urban economics is the problem of network design: given the transportation network of a city, how to restrict the network so that to minimize the price of anarchy? It turns out that this problem is knwon as a NP-hard problem, which means that there is no efficient algorithm to compute the exact solution in a reasonable amount of time. But there are heuristics that offer reasonable approximations of the solution.

Estimating transportation costs. Discrete choice models in the spirit of Dial's model (Theorem 5.2) are the basis of empirical work on transportation costs.

Estimating congestion. The possibilities of doing empirical work on Urban Economics has fantastically expanded since the arrival of easy-access online tools such as the Google Map API, which gives routes and estimated times depending on traffic conditions. An example of such recent work is Trebbi and Bombardini (2010) who use the Google Map API to estimate congestion costs in United States and Italian cities.

[...]

5.4 References and notes

Theorem 5.2 is due to Dial (1971) [36], see also Sheffi and Powell (1982) [101]. The notion of Wardrop equilibrium was introduced by Wardrop in a 1952 paper [117]. Theorem 5.3 is due to Beckmann, McGuire, and Winsten (1956) [8].

The Braess paradox is due to Braess (1968) [17]. Koutsoupias and Papadimitriou (1999) [73] introduced the notion of "price of anarchy" in the context of Internet traffic. Roughgarden and Tardos (2002) [91] derived the efficiency bounds. For congestion costs estimation, see Trebbi and Bombardini (2010) [107].

6 The economics of marriage

Our goal in this chapter is to carry estimation of preferences based on data on the marriage market. Given oberved marriage patterns, what can be said about the preferences of the participants? what determines marital unions? what is the individual surplus of the participants? these are some of the questions that will be asked and (sometimes partially) answered.

So far we have mostly studied the social planner's problem: assuming that one could define a surplus out of a potential match between two individuals, we have studied the problem of "optimal" matching in the sense that this matching should maximize the total surplus. But in most of today's western societies, marriage is essentially the outcome a decentralized process. (This has not always been always the case, and is still not the case in several societies today, where marriage is often arranged by families subject to strict social rules). Our first concern will be therefore to relate the notions of "optimal matching" and "equilibrium matching".

We'll then move to the identification problem. Given an observed matching, what is the set of surplus functions that rationalize this matching? we'll see that the answer to this question is surprisingly negative, and that we'll need to assume some amount of unobserved heterogeneity if we want to be able to identify anything.

We'll finally move on to estimation, and we'll propose two techniques: one, due to Becker, is Canonical Correlation Analysis, simple and powerful, but not grounded in an equilibrium model. The other one, due to Choo and Siow, makes very specific assumptions about the form of unobserved heterogeneity, allowing for a very powerful identification result.

6.1 The marriage game

We now describe the marriage game of Becker-Shapley-Shubik. Recall the setting described in Section 2.2.4. Consider a population of men M and women W, not necessarily of the same size (remember that in the current setting we allow for singles) to be matched. A *matching* σ maps any individual with his or her partner, if any, and maps singles, by convention, with 0. Thus if man m is matched to woman w, we set $w = \sigma(m)$ and $m = \sigma^{-1}(w)$. If man m is unmatched, we set $\sigma(m) = 0$ and if woman w is unmatched we set $\sigma^{-1}(w) = 0$.

For a given man $m \in M$, let x_m be the observable characteristics of m, and for a woman $w \in W$, let y_w be the observable characteristics of the woman w. Let $\mathcal{X}$ be the set of men's types, and $\mathcal{Y}$ the set of women's types. Let $\mathcal{X}^* = \mathcal{X} \cup \{0\}$ be the set of marital choices of men including singlehood, and $\mathcal{Y}^* = \mathcal{Y} \cup \{0\}$ be the set of marital choices for women. Finally, let $\mathcal{XY}^* = \mathcal{X}^* \times \mathcal{Y}^* \backslash \{(0,0)\}$ be the set of possible households including single households.

Let $\mu(x,y)$ the number of households of type (x,y). In particular, as in Section 2.2.4, we denote by μ_{x0} (resp. μ_{0y}) the number of men of type x (resp. woman of type y) who remain single. One has

$$\begin{aligned} \mu_{xy} &\geq 0,\ \forall x,y \in \mathcal{XY}^* \\ \sum_{y \in \mathcal{Y}^*} \mu_{xy} &= p_x,\ \forall x \in \mathcal{X} \\ \sum_{x \in \mathcal{X}^*} \mu_{xy} &= q_y,\ \forall y \in \mathcal{Y} \\ \mu_{xy} &\in \mathbb{N},\ \forall x,y \in \mathcal{XY}^*. \end{aligned}$$

In fact, we make a "Large Market Assumption": it is assumed that there is a large number of individuals in every observable category of men and women. This will lead us to drop the last constraint imposing the integrality of μ_{xy}. In the rest of the chapter, we shall admit without further justification that it is OK to drop any integrality constraints. Accordingly, we introduce the set of *observable matchings* as

$$\mathcal{M}(p,q) = \left\{ \mu \geq 0 :\ \begin{array}{c} \sum_{y \in \mathcal{Y}^*} \mu_{xy} = p_x,\ \forall x \in \mathcal{X} \\ \sum_{x \in \mathcal{X}^*} \mu_{xy} = q_y,\ \forall y \in \mathcal{Y} \end{array} \right\}$$

If a man of type x and a woman of type y match, they generate joint surplus Φ_{xy}, to be shared between them. If man x (resp. woman y) remains single, it is assumed that his/her surplus is Φ_{x0} (resp. Φ_{0y}). We shall in fact normalize $\Phi_{x0} = \Phi_{0y} = 0$. This is w.l.o.g. as we can always replace Φ_{xy} by $\Phi_{xy}-\ \Phi_{x0} - \Phi_{0y}$, and the resulting new function Φ_{xy}, which is the difference between the surplus of a matched pair of individuals and the sum of the surpluses of the individuals if single, is sometimes called *gains from marriage.*

Note that we have made implicitely an important assumption in this setup: the surplus $\Phi_{x_m y_w}$of a pair (m,w) only depends on observable characteristics of the partners x_m and y_w. We shall relax this assumption in

Section 6.5 when we shall assume that the surplus of a pair also depends on unobservable characteristics.

Let us now give a game-theoretic formulation of the marriage game. An *outcome* (μ, u, v) is the specification of a matching μ, and a vector of payoffs u_x to each man x and v_y to each woman y, which are the utilities that they receive. These utilities have to be feasible: that is, the sum of the utilities across the population has to be equal to the total output under the matching μ. This leads to the following definition of a feasible outcome.

Definition 6.1 (Feasibility). *An outcome* (μ, u, v) *is* feasible *if one has*

$$\sum_{x \in \mathcal{X}} p_x u_x + \sum_{y \in \mathcal{Y}} q_y v_y = \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \mu_{xy} \Phi_{xy}. \tag{6.1}$$

The set of feasible outcomes associated to populations $(\mathcal{X}, p)$ *and* $(\mathcal{Y}, q)$ *is denoted* $\mathcal{F}(p, q)$.

We now turn to the notion of stability. A natural requirement for some outcome (μ, u, v) to be stable is that no one can achieve greater utility either by forming a coalition as a single invididual or as a matched pair than in that outcome. This implies, first, that $u_x \geq 0$ and $v_y \geq 0$ for all x and y. Also, for all pair x and y, there are no two scalars $\bar{u}_x$ and $\bar{v}_y$ such that $\bar{u}_x + \bar{v}_y \leq \Phi_{xy}$ with $\bar{u}_x \geq u_x$ and $\bar{v}_y \geq v_y$, one of these inequalities being strict. A sufficient condition for stability is thus $u_x + v_y \geq \Phi_{xy}$ for all pair x, y. It is also necessary: if $u_x + v_y < \Phi_{xy}$ for some pair x, y, then letting $\epsilon = \Phi_{xy} - u_x - v_y$, and $\bar{u}_x = u_x$, $\bar{v}_y = v_y + \epsilon$, one sees that the pair (u, v) cannot be stable. This motivates the following definition:

Definition 6.2 (Stability). *An outcome* (μ, u, v) *is* stable *if it is feasible, and if for all* $x \in \mathcal{X}$, $y \in \mathcal{Y}$ *(matched or not), one has*
(i) $u_x \geq 0$ *and* $v_y \geq 0$
(ii) $u_x + v_y \geq \Phi_{xy}$.

In the language of mechanism design, condition (i) can be seen as a condition of individual rationality: no one can be forced to participate. Condition (ii) can be seen as an incentive constraint: the utility that man x gets out of the game should be at least as much as the utility that his preferred woman is ready to give him, hence

$$u_x \geq \max_y \left(\Phi_{xy} - v_y\right).$$

One has the following result.

Proposition 6.1. *If (μ, u, v) is a stable outcome, then*
(i) $\mu_{xy} > 0$ implies $u_x + v_y = \Phi_{xy}$,
(ii) $\mu_{x0} > 0$ implies $u_x = 0$
(iii) $\mu_{0y} > 0$ implies $v_y = 0$.

Proof. As μ is feasible, one has

$$\sum_{x\in\mathcal{X}} p_x u_x + \sum_{y\in\mathcal{Y}} q_y v_y = \sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} \mu_{xy}\Phi_{xy}$$

or equivalently,

$$\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}^*} \mu_{xy}u_x + \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}^*} \mu_{xy}v_y = \sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} \mu_{xy}\Phi_{xy},$$

thus

$$\begin{aligned} \sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}}} \mu_{xy}\left(\Phi_{xy} - u_x - v_y\right) \quad & + \sum_{x\in\mathcal{X}} \mu_{x0}\left(-u_x\right) \\ + \quad & \sum_{y\in\mathcal{Y}} \mu_{0y}\left(-v_y\right) = 0. \end{aligned}$$

By stability, all the terms in the sum are nonpositive. Hence as the sum is equal to zero, all these terms are equal to zero. The result follows. ■

This result has a straightforward interpretation. If one observes a positive probability of matches between x and y, then the surpluses of x and y must be generated by this match, hence $\Phi_{xy} = u_x + v_y$. If there is a positive probability that, say, man x remains single, then the surplus of this man must be zero.

Let us now contrast the concept of stability with that of Pareto efficiency.

Definition 6.3. *A feasible outcome (μ, u, v) is Pareto efficient if there is no feasible outcome $(\mu, \bar{u}, \bar{v})$ such that for all x and y, $\bar{u}_x \geq u_x$ and $\bar{v}_y \geq v_y$, one of these inequalities being strict.*

In the present context, stability implies Pareto efficiency, while the converse does not hold true – a fact, as we shall see, which is very general in many matching situations.

Proposition 6.2. *If a feasible outcome is stable, then it is Pareto efficient.*

Proof. Consider a stable outcome (μ, u, v). Assume (μ, u, v) is not Pareto efficient. Then there is a feasible outcome $(\mu, \bar{u}, \bar{v})$ such that $\bar{u}_x \geq u_x$ and $\bar{v}_y \geq v_y$ with one strict inequality. By feasibility of $(\mu, \bar{u}, \bar{v})$, one has

$$\sum_{x \in \mathcal{X}} p_x \bar{u}_x + \sum_{y \in \mathcal{Y}} q_y \bar{v}_y = \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \mu_{xy} \Phi_{xy}.$$

By stability of (μ, u, v), one has for all x, y

$$\begin{array}{rcl} \Phi_{xy} & \leq & u_x + v_y \\ 0 & \leq & u_x \ , \ 0 \leq v_y \end{array}$$

hence

$$\begin{aligned} \sum_{x \in \mathcal{X}} p_x \bar{u}_x + \sum_{y \in \mathcal{Y}} q_y \bar{v}_y &\leq \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \mu_{xy} \left(u_x + v_y\right) \\ &\leq \sum_{x \in \mathcal{X}} p_x u_x + \sum_{y \in \mathcal{Y}} q_y v_y \end{aligned}$$

(where we have used $u_x, v_y \geq 0$ in the second inequality). But this contradicts the fact that $\bar{u}_x \geq u_x$ and $\bar{v}_y \geq v_y$, one of these inequalities being strict. Therefore (μ, u, v) is not stable. ■

It is easy to show by a conterexample that the converse does not hold true.

Example 6.1. *Consider a game with two men and two women, and surplus*

$$\Phi = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then (σ, u, v) where $\sigma(x) = x$, $u = (-1, 0)$, and $v = (1, 0)$ is feasible and Pareto efficient; indeed, any feasible payoff should satisty $u_1+u_2+v_1+v_2 = 0$, and this payoff vector is on the efficient frontier. However it is not stable, as $u_1 < 0$.

The fact that efficient does not imply comes from the fact that stability allows people to run away of situation which hurt them; while in the concept of efficiency, it is possible to sacrifice some individuals for the benefit of others. This example shows that Pareto efficiency is not a very appropriate concept if we want to study equilibrium in this market. Stability is much more appropriate; in fact as we are about to see, stable outcomes coincide with the core of the marriage game.

6.2 The core of the marriage game

We now investigate the core of the marriage game. In cooperative game theory, an outcome is in the core of the underlying game if no coalition (which is merely a subset of the total population) blocks this outcome, which means that no coalition has a feasible way to generate more surplus for all of its members than in the given outcome.

Consider a feasible outcome $(\mu, u, v) \in \mathcal{F}(p, q)$. The notion of stability means that no coalition of a single man or woman is blocking ($u_x \geq 0$ and $v_y \geq 0$), and that no coalition of a pair of man and woman is blocking ($u_x + v_y \geq \Phi_{xy}$). Hence core allocation are stable. But the converse is not obvious. Indeed, even if a stable outcome $(\mu, u, v) \in \mathcal{F}(p, q)$ is (by definition) not blocked by coalition of single individuals and man-woman pairs, it could be blocked by greater coalitions. It turns out that this is not the case, and that if a feasible outcome is not blocked by singles and man-woman pairs, then it cannot be blocked by any coalition. Hence the core of the game is the set of stable outcomes. This is formalized in Theorem 6.1 below.

We need to set up some notation. A coalition is described by a vector $(\bar{p}_x, \bar{q}_y)$ where $0 \leq \bar{p}_x \leq p_x$ and $0 \leq \bar{q}_y \leq q_y$ are the numbers of men of type x and women of type y in the coalition.

This coalition blocks outcome (μ, u, v) if there exists $(\mu, \bar{u}, \bar{v}) \in \mathcal{F}(\bar{p}, \bar{q})$ such that

$$\begin{aligned} \forall x &\in \mathcal{X},\ \bar{p}_x > 0 \Longrightarrow u_x \leq \bar{u}_x \\ \forall y &\in \mathcal{Y},\ \bar{q}_y > 0 \Longrightarrow v_y \leq \bar{v}_y. \end{aligned}$$

The *core of the marriage game* is the set of unblocked outcomes. This set will be related to the set of solutions of the dual optimal assignment problem (2.9), which we just recall:

$$\begin{aligned} T(p,-q) &= \min_{u,v} \sum_{x\in\mathcal{X}} p_x u_x + \sum_{y\in\mathcal{Y}} q_y v_y \\ & s.t. u_x \geq 0\ ,\ \ v_y \geq 0 \\ & u_x + v_y \geq \Phi_{xy} \end{aligned} \tag{6.2}$$

Remember the primal formulation of this problem, given by (2.7)

$$\begin{aligned} T(p,-q) &= \max_{\mu\geq 0} \sum_{x,y} \mu_{xy}\Phi_{xy} \\ & s.t. \\ & \sum_{y} \mu_{xy} \leq p_x \\ & \sum_{x} \mu_{xy} \leq q_y \end{aligned} \tag{6.3}$$

One can then formally define the core of the marriage game.

Definition 6.4. *The core of the marriage game is the set of feasible outcomes* $(\mu, u, v) \in \mathcal{F}(p, q)$ *such that*

$$\sum_{x\in\mathcal{X}} \bar{p}_x u_x + \sum_{y\in\mathcal{Y}} \bar{q}_y v_y \geq T(\bar{p}, -\bar{q})$$

for all $(\bar{p}_x, \bar{q}_y)$ *such that* $0 \leq \bar{p}_x \leq p_x$ *and* $0 \leq \bar{q}_y \leq q_y$.

Clearly, if $(\bar{p}_x, \bar{q}_y)$ is such that $\bar{p}_x = 1\{x = x_0\}$ and $\bar{q}_y = 0$ (one-man coalitions), this implies the restriction

$$u_x \geq 0$$

and similarly, the absence of one-woman blocking coalitions implies

$$v_y \geq 0.$$

Clearly, if $(\bar{p}_x, \bar{q}_y)$ is such that $\bar{p}_x = 1\{x = x_0\}$ and $\bar{q}_y = 1\{y = y_0\}$ (man-woman coalitions), this implies the further restriction that

$$u_x + v_y \geq \Phi_{xy}.$$

We have thus verified that core outcomes are stable. It turns out that the converse holds in the present game–again, this means that there is no other coalition to consider when verifying the belonging of some outcome to the core.

Theorem 6.1 (Shapley and Shubik). *One has:*

(i) The core of the marriage market coincides with the set of stable outcomes.

(ii) The core of the marriage market is the set of outcomes (μ, u, v) such that (u, v) are of solutions of (6.2), and μ is the associated Lagrange multiplier, hence μ is a solution of (6.3).

Proof. Cf. Roth Sotomayor Th. 8.10, p. 208.[...] ■

This theorem has extremely important consequences, both for theoretic purposes and in order to derive our empirical strategies. It states that the outcome of the marriage market modelled as a competitive and decentralized market is the maximizer of a social surplus function.

[...]

6.3 Identification

Our problem is now the following. Assume one observes $\hat{\mu}_{xy}$ the number of matches between men of type x and women of type y. Denote Φ_{xy} the joint surplus from marriage. One would like to identify Φ whenever possible. By the above result, $\hat{\mu}$ is known to maximize

$$\max_{\mu \in \mathcal{M}} \sum_{xy} \mu_{xy} \Phi_{xy} \tag{6.4}$$

if this relation is verified one says that Φ rationalizes matching $\hat{\mu}$.

The minimal requirement for our notion of rationalizability should be that it should have some degree of refutability; otherwise the theory would be devoid of empirical content. For instance, any $\hat{\mu}$ would maximize (6.4) with $\Phi = 0$, therefore a theory based on $\Phi = 0$ has no empirical content: it cannot be rejected. In a similar vein, testing hypothesis about revealed preferences in consumer demand should impose some properties (monotonicity, concavity, etc.) on the consumer utility function to have empirical bite: otherwise, constant utility functions can always rationalize any observations. This leads us to the following (minimalistic) definition:

Definition 6.5. *One says that matching $\hat{\mu}$ is rationalized by Φ if $\hat{\mu}$ is a solution of (6.4), and if the set of solutions of (6.4) is strictly smaller than the set of feasible matchings.*

Let us make a series of observations.

1. First, consider the case where $\Phi_{xy} = a_x + b_y$, i.e. there is no cross interaction term in the joint surplus function. Then the optimal assignment problem rewrites as

$$\begin{aligned}
&\max \sum_{x\in\mathcal{X}} \bar{p}_x a_x + \sum_{y\in\mathcal{Y}} \bar{q}_y b_y \\
&s.t. \\
&0 \leq \bar{p}_x \leq p_x \ \forall x \in \mathcal{X} \\
&0 \leq \bar{q}_y \leq q_y \ \forall y \in \mathcal{Y}
\end{aligned}$$

where $\bar{p}_x$ is the number of married men of type x, and $\bar{q}_y$ is the number of married women of type y. The solution of this problem is therefore $\bar{p}_x = p_x$ if $a_x > 0$ and $\bar{p}_x = 0$ if $a_x < 0$. Therefore, we can set

$$\begin{aligned}
a_x &> 0 \text{ if all men of type } x \text{ are married} \\
a_x &< 0 \text{ if all men of type } x \text{ are single} \\
a_x &= 0 \text{ otherwise.}
\end{aligned}$$

and similarly for women, and our function $\Phi_{xy} = a_x + b_y$ rationalizes $\hat{\mu}$. In particular:

– when everybody is married under $\hat{\mu}$, then any function $\Phi_{xy} = a_x + b_y$ with $a_x > 0$ and $b_y > 0$ rationalizes $\hat{\mu}$;

– when each category of men and women contains both singles and married individuals, then the only function of the form $\Phi_{xy} = a_x + b_y$ that rationalizes $\hat{\mu}$ is the trivial function $\Phi = 0$.

2. Let us assume that Φ rationalizes $\hat{\mu}$, hence $\hat{\mu}$ solves the primal associated to surplus function Φ, with associated payoffs $u_x, v_y \geq 0$ such that $u_x + v_y \geq \Phi_{xy}$. Define

$$\bar{\Phi}_{xy} = u_x + v_y$$

then $\bar{\Phi}$ also rationalizes $\hat{\mu}$.

3. Combining the two previous remarks, assume that under $\hat{\mu}$, each category of men and women contains both singles and married individuals, and assume Φ rationalizes $\hat{\mu}$. Then by the second remark, $\bar{\Phi}$ also rationalizes $\hat{\mu}$, and by the first remark, $u_x = v_y = 0$ for all x and y. Thus $\Phi_{xy} = 0$ for any pair (x, y) such that $\mu_{xy} > 0$.

4. By the preceding remark, if there are singles of every categories, and marriages between every categories (which is the generic case when dealing with empirical matching models), that is, $\mu_{xy} > 0$ for all x, y, then the only Φ that rationalizes $\hat{\mu}$ is $\Phi = 0$. In such case the theory is void of empirical content, as $\Phi = 0$ rationalizes *any* matching. In such case one says that $\hat{\mu}$ cannot be rationalized.

This last point is rather bad news for us. In general, when dealing with empirical data, we shall have observe (if the dataset is rich enough) empirical matchings with singles in every categories of men and women, and with marriages between every categories of men and women. The last remark implies that such matchings cannot be rationalized. We will have to keep this in mind when doing estimation.

6.4 Becker's positive assortative matching

Assume that men and women can be represented by a single-dimensional characteristics index, which we still call x for men and y for women. This index is called *trait* by Becker (1973), and it can be "(...) intelligence, race, religion, education, wage rate, height, aggressiveness, tendency to nurture, or age". And assume that the joint surplus from marriage $\Phi(x, y)$ is strictly submodular, reflecting the *positive assortative matching* property (or, in Becker's words, "positive assortive mating"): agents like to marry their likes. Still quoting Becker (1973) (p. 827):

> "Mating of likes–positive assortive mating–is extremely common, whether measured by intelligence, height, skin color, age, education, family background, or religion (...)."

This theory, which suggests a very strong restriction on Φ, has very strong predictive power: as explained in Section 3.2, the prediction of the theory is

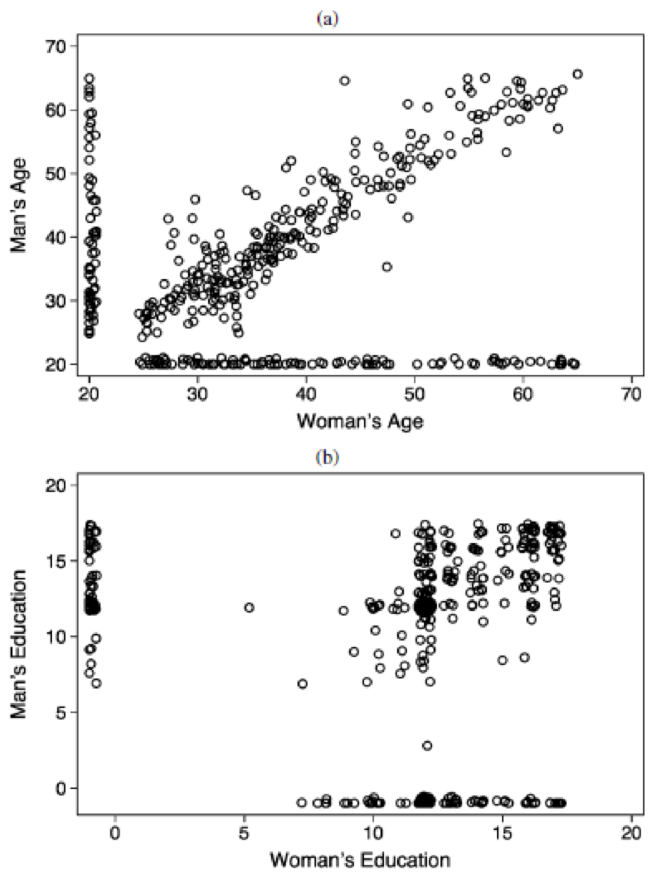


Figure 1. Observed age (a) and education (b) distributions of men and women, NSFH data.

Figure 5: Endogamy in age and education. Source: Logan, Hoff and Newton (2008)

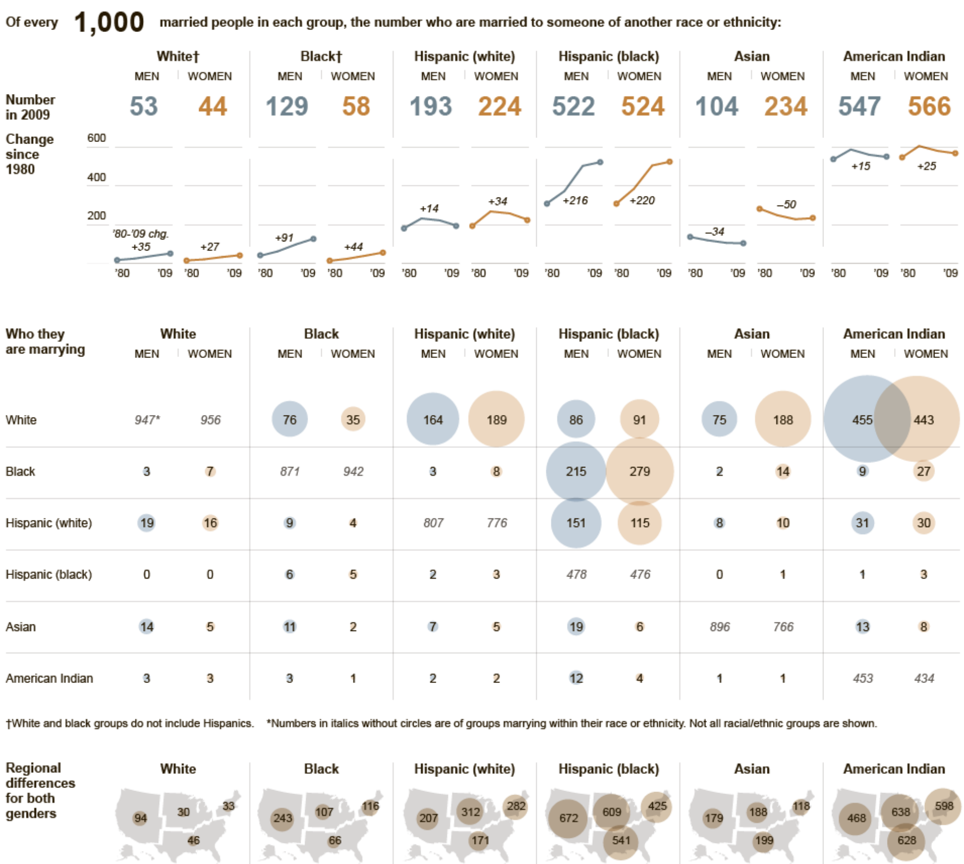


Figure 6: Racial endogamy in the US and its evolution (source: NY Times).

that the traits of the men and the women will be comonotone, namely every man will have the same rank (among the men's ability in the population of men) than his wife (among the women's ability in the population of women).

In other words, if the data exhibit positive assortative matching, then the correlation of man's ability X and woman's ability Y will be highest among all the possible ways to match men and women. This does not mean that this correlation will be equal to one, unless the distribution of men's abilities and the women's abilities is the same. But this does mean that the correlation of the ranks (the *rank correlation*) of the men's abilities and the women's abilities is one.

In terms of identification, if the data exhibit positive assortative matching, then any strictly submodular Φ rationalizes the data.

It is hard to believe that the marriage market clears according to a single-dimensional characteristics. The analysis can in fact be extended to the case where traits aggregate several observable characteristics. Assume the men's (resp. women's) characteristics is a vector X in $\mathbb{R}^n$ (resp. Y in $\mathbb{R}^m$). One may assume, for instance, that the trait on which the market clears is a linear combination of the entries of X (resp. Y). That is,

$$\begin{aligned} x &= \alpha' X \\ y &= \beta' Y \end{aligned}$$

where $\alpha \in \mathbb{R}^n$, $\beta \in \mathbb{R}^m$ are to be determined.

Becker (p. 834) suggests using a canonical correlation technique to determine α and β. This consists in determining α and β that maximize the correlation of $\alpha' X$ and $\beta' Y$. Of course, some normalization is needed. Assuming the vectors X and Y are centered, the problem consists therefore of solving:

$$\begin{aligned} &\max_{\alpha \in \mathbb{R}^n, \beta \in \mathbb{R}^m} \mathbb{E}\left[\alpha' X Y' \beta\right] \\ &s.t.\ var\left(\alpha' X\right) = 1 \text{ and } var\left(\beta' Y\right) = 1 \end{aligned} \tag{6.5}$$

It is well-known that this can be performed by a singular value decomposition. More precisely, let Σ_X be the variance-covariance matrix of X and

Σ_Y that of Y. Let Σ_{XY} be the matrix of general term $cov\left(X_i, Y_j\right)$, so that $\Sigma_{XY} = \mathbb{E}\left[XY'\right]$. Then we can perform the singular value decomposition of matrix $\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1/2}$, that is

$$\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1/2} = U'DV$$

where $D = (r_1, ..., r_d)$ is a diagonal matrix with nonincreasing elements $r_1 \geq ... \geq r_d$, $d = \min(n, m)$, and U and V are orthogonal matrices of dimension n and m respectively.

Then, introducing $\hat{X} = U\Sigma_X^{-1/2}X$ and $\hat{Y} = V\Sigma_Y^{-1/2}Y$, we have $\Sigma_{\hat{X}\hat{Y}} = \mathbb{E}\left[\hat{X}\hat{Y}'\right] = \mathbb{E}\left[U\Sigma_X^{-1/2}XY'\Sigma_Y^{-1/2}V'\right] = U\Sigma_X^{-1/2}\Sigma_{XY}\Sigma_Y^{-1/2}V' = D$. Therefore, α is the vector of the first line of $U\Sigma_X^{-1/2}$ and β is the vector of the first columns of $V\Sigma_Y^{-1/2}$. Thus, x and y are given by the first lines of $\hat{X}$ and $\hat{Y}$, that is

$$\begin{aligned} x &= \hat{X}^1 = \alpha' X \\ y &= \hat{Y}^1 = \beta' Y. \end{aligned}$$

While this is a simple and nice descriptive tool, there are several drawbacks of this approach:

1. This method is purely linear and does not allow for nonlinear transformations of X and Y.

2. The approach it is not grounded in an equilibrium model.

The approach we are about to see alleviates both drawbacks.

6.5 Unobserved heterogeneity

To remedy the identifiability problem, Choo and Siow (2006) assumed that there is unobserved heterogeneity in the individual surpluses.

Assume that each man m draws a vector $\left(\varepsilon_{my}\right)_{y\in\mathcal{Y}^*}$ of unobservable preferences variations for each observable category of women y, such that the utility of a man m of observable characteristics x_m who marries a woman of observable characteristics y_w and who gives her a utility transfer τ can be written:

$$\alpha_{x_m y_w} - \tau + \varepsilon_{m y_w},$$

and if he remains single, then his utility is ε_{m0}. Similarly, each woman w draws a vector $(\eta(x,w))_{x \in \mathcal{X}^*}$ of unobservable preferences variations for each observable category of men x. Hence in a match with the same terms as above, the woman gets

$$\gamma_{x_m y_w} + \tau + \eta_{x_m w},$$

and her utility is η_{0w} if she remains single. One can thus define the joint surplus from matching as the sum of the utilities of a man m and a woman w matched together

$$\tilde{\Phi}_{mw} = \Phi_{x_m y_w} + \varepsilon_{m y_w} + \eta_{x_m w} \tag{6.6}$$

where

$$\begin{aligned}
\Phi_{xy} &= \alpha_{xy} + \gamma_{xy} \text{ for } x \in \mathcal{X}, y \in \mathcal{Y}, \\
\Phi_{x0} &= 0 \text{ for } x \in \mathcal{X}, \\
\Phi_{0y} &= 0 \text{ for } y \in \mathcal{Y}.
\end{aligned}$$

Assuming that the matching surplus is of the form (6.6) requires that conditional on observable types, the surplus exhibit no complementarity across unobservable types.

Finally, assume that conditional on m, the ε_{my}'s are independent and distributed as type-I Extreme Value distribution (Gumbel) with mean zero and scaling parameter σ_M. Similarly, conditional on w, the $\eta(x,w)$'s are independent and distributed as type-I Extreme Value distribution with mean zero and scaling parameter σ_W.

One has the following identification result:

Theorem 6.2. *In equilibrium, if there are very large numbers of men and women within each group, then:*

(i) the marital surplus Φ_{xy} is indentified from $\mu_{xy} > 0$ by

$$\exp\left(\frac{\Phi_{xy}}{2}\right) = \frac{\mu_{xy}}{\sqrt{\mu_{x0}\mu_{0y}}}.$$

hence marriage patterns directly identify the marital surplus Φ in this model.

(ii) Conversely, given the marital surplus Φ_{xy}, then $\mu_{xy} > 0$ is the unique positive solution of the system of $(|\mathcal{X}|+1) \times (|\mathcal{X}|+1) - 1$ equations

$$\begin{aligned}
\ln \frac{\mu_{xy}}{\sqrt{\mu_{x0}\mu_{0y}}} &= \frac{\Phi_{xy}}{2} \\
\sum_{y \in \mathcal{Y}} \mu_{xy} + \mu_{x0} &= p_x \\
\sum_{x \in \mathcal{X}} \mu_{xy} + \mu_{0y} &= q_y.
\end{aligned}$$

In fact this result can be deduced from the more general following result:

Theorem 6.3. *Under the above assumptions, the market equilibrium maximizes the social gain*

$$\mathcal{W}(\mu) = \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \mu_{xy} \Phi_{xy} + \mathcal{E}(p, q, \mu), \tag{6.7}$$

over all feasible matchings $\mu \in \mathcal{M}$, where

$$\mathcal{E}(\mu) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}^*} \mu_{xy} \ln \frac{\mu_{xy}}{p_x} - \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}^*} \mu_{xy} \ln \frac{\mu_{xy}}{q_y}. \tag{6.8}$$

Proof. By the classical dual formulation of the matching problem, the market equilibrium assigns utilities u_{xm} to man $m : x_m = m$ and v_{yw} to woman $y : y_w = w$ so as to solve

$$\mathcal{G} = \min(\sum_{x,m} u_{xm} + \sum_{y,w} v_{yw})$$

where the minimum is taken under the set of constraints

$$\begin{aligned}
u_{xm} + v_{yw} &\geq \Phi_{xy} + \varepsilon_{xym} + \eta_{xyw} \\
u_{xm} &\geq \varepsilon_{x0m} \\
v_{yw} &\geq \eta_{0yw}.
\end{aligned}$$

Denote

$$
\begin{aligned}
U_{xy} &= \min_{m} \left\{ u_{xm} - \varepsilon_{xym} \right\}, \ x \in \mathcal{X}, \ y \in \mathcal{Y}^{*} \\
V_{xy} &= \min_{w} \left\{ v_{yw} - \eta_{xyw} \right\}, \ x \in \mathcal{X}^{*}, \ y \in \mathcal{Y}
\end{aligned}
$$

so that

$$
\begin{aligned}
u_{xm} &= \max_{y \in \mathcal{Y}^{*}} \left\{ U_{xy} + \varepsilon_{xym} \right\}, \ x \in \mathcal{X} \\
v_{yw} &= \max_{x \in \mathcal{X}^{*}} \left\{ V_{xy} + \eta_{xyw} \right\}, \ y \in \mathcal{Y}
\end{aligned}
$$

Then

$$
\mathcal{G} = \min \begin{pmatrix} \sum_{x,m} \max_{y \in \mathcal{Y}^{*}} \left\{ U_{xy} + \varepsilon_{xym} \right\} \\ + \sum_{y,w} \max_{x=0,\ldots,I} \left\{ V_{xy} + \eta_{xyw} \right\} \end{pmatrix}
$$

under the set of constraints

$$
\begin{aligned}
U_{xy} + V_{xy} &\geq \Phi_{xy} \\
U_{x0} &\geq 0 \\
V_{0y} &\geq 0.
\end{aligned}
$$

Assign non-negative multipliers $\mu_{xy}, \mu_{x0}, \mu_{0y}$ to these constraints. By duality, we get

$$
\mathcal{G} = \max_{\mu_{xy} \geq 0} \Big(\sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \mu_{xy} \Phi_{xy} - A(\mu) - B(\mu) \Big)
$$

where

$$
A(\mu) = \max_{U_{xy}} \Big[\sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}^{*}}} \mu_{xy} U_{xy} - \sum_{x,m} \max_{y \in \mathcal{Y}^{*}} \left\{ U_{xy} + \varepsilon_{xym} \right\} \Big]
$$

and a similar expression holds for $B(\mu)$.

Now

$$
\begin{aligned}
& \sum_{x,m} \max_{y \in \mathcal{Y}^{*}} \left\{ U_{xy} + \varepsilon_{xym} \right\} \\
= \ & \sum_{x \in \mathcal{X}} p_x \mathbb{E}_{m \in x} \max_{y \in \mathcal{Y}^{*}} \left\{ U_{xy} + \varepsilon_{xym} \right\}.
\end{aligned}
$$

In this formula $\mathbb{E}_{m\in x}$ denotes the empirical average over the population of men in group x. Now we have assumed that there is a large number of individuals in each group, so for each $x \in \mathcal{X}$,

$$\begin{aligned}\mathbb{E}_{m\in x}\max_{y\in\mathcal{Y}^*}\{U_{xy}+\varepsilon_{xym}\} &\approx \mathbb{E}\left[\max_{y\in\mathcal{Y}^*}\{U_{xy}+\varepsilon_{y}\}\right]\\ &= \log(\sum_{y\in\mathcal{Y}^*}\exp(U_{xy})).\end{aligned}$$

Hence we get

$$\begin{aligned}A(\mu) &= \max_{U_{xy}}\{\sum_{\substack{x\in\mathcal{X}\\ y\in\mathcal{Y}^*}}\mu_{xy}U_{xy}\\ &\quad -\sum_{x\in\mathcal{X}}p_x\log(\sum_{y\in\mathcal{Y}^*}\exp(U_{xy}))\}\end{aligned}$$

and the symmetric expression holds for $B(\mu)$. Finally, one has

$$\begin{aligned}A(\mu) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}^*}\mu_{xy}\ln\frac{\mu_{xy}}{p_x}\\ B(\mu) &= \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}^*}\mu_{xy}\ln\frac{\mu_{xy}}{q_y}.\end{aligned}$$

■

We are now ready to prove Theorem 6.2.

Proof of Theorem 6.2. Let us prove (i). The constraint associated to (6.7) is

$$\begin{aligned}\sum_{y\in\mathcal{Y}^*}\mu_{xy} &= p_x\ \forall x\in\mathcal{X}\\ \sum_{x\in\mathcal{X}^*}\mu_{xy} &= q_y\ \forall y\in\mathcal{Y}.\end{aligned}$$

Letting u_x and v_y the associated Lagrange multipliers, one has

$$\begin{aligned}\Phi_{xy} &= 2+\ln\frac{\mu_{xy}}{p_x}+\ln\frac{\mu_{xy}}{q_y}-u_x-v_y\\ 0 &= 1+\ln\frac{\mu_{x0}}{p_x}-u_x\\ 0 &= 1+\ln\frac{\mu_{0y}}{q_y}-v_y\end{aligned}$$

hence

$$\Phi_{xy} = 2 \ln \frac{\mu_{xy}}{\sqrt{\mu_{x0}\mu_{0y}}},\ QED.$$

To prove (ii), we shoud observe that the surplus function is strictly concave, hence the optimum is unique. ■

6.6 Parametric estimation

The previous section provides a nonparametric identification result. It is useful for both theoretical and empirical purposes, but quite often one may test stylized theories, and a parametric approach is in need.

We have seen that one of the main empirical questions about the marriage market deals with the issue of endogamy, and the broad fact that people tend to "marry their likes". But the marriage market is rarely symmetric, and the socio-economic classes rarely contain equal size of men and women. A classical, and somewhat outdated, example is education: there use to be a sharp dissymetry in the number of educated men and educated women, thus men had on average to "marry down" and women to "marry up". In traditional societies, the existence of doweries also led to dissymetries in the marriage game.

Many dissymetries also mean a trade-off between various sorts of exogamy needed to clear the marriage market. If one feature (say education) is most important for agents, the marriage market will clear so to maximize assortativeness on education, even though this will result in less assortativeness on another characteristics (say income). The marriage patterns are thus quite informative to undestand relative preferences of agents.

Consider the following parameterization of the surplus function

$$\Phi^{\lambda}_{xy} = \sum_{k=1}^{K} \lambda_k \phi^k_{xy}. \tag{6.9}$$

The functions ϕ^k are basis interaction functions, and they describe a certain way in which agents characteristics may interact. For instance, assume that there are two observable dimensions, say education (x^1 and y^1 for the husband and the wife, respectively) and income (x^2 and y^2). One function ϕ^1_{xy} might reflect a taste for endogamy in education, say $\phi^1_{xy} = -\left(x^1 - y^1\right)^2$, or $\phi^1_{xy} = x^1 y^1$. Another function ϕ^2 may similarly reflect a taste for endogamy in income: $\phi^2_{xy} = -\left(x^2 - y^2\right)^2$, or $\phi^2_{xy} = x^2 y^2$. The relative weight of λ_1 and λ_2 reflects agent's relative preferences for endogamy in education or income. Of course, we might also introduce crossed terms $\phi^3_{xy} = x^1 y^2$ and $\phi^4_{xy} = x^2 y^1$ to understand how education and income cross-interact whithin a pair.

Introduce $\mathcal{W}(\lambda)$ as the social surplus function associated to Φ^λ_{xy}. One has

$$\mathcal{W}(\lambda) = \sup_{\mu \in \mathcal{M}(p,q)} \left(\sum_{x,y} \mu_{xy} \Phi^\lambda_{xy} + \mathcal{E}(\mu) \right)$$

where $\mathcal{E}(\mu)$ is given by (6.8), i.e.,

$$\mathcal{E}(\mu) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}^*} \mu_{xy} \ln \frac{\mu_{xy}}{p_x} - \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}^*} \mu_{xy} \ln \frac{\mu_{xy}}{q_y}.$$

Let μ^λ be the solution of the program. One next would like to estimate λ based on the observation of some matching $\hat{\mu}$. Under this matching, the basis functions ϕ^k_{xy} have total values which are the *co-moments*

$$\hat{C}^k = \sum_{x,y} \hat{\mu}_{xy} \phi^k_{xy}.$$

In our education/income example, assuming that the characteristics are centered, then $\hat{C}^1$ is the observed covariance of educations, and $\hat{C}^2$ is the observed covariance of incomes. As we shall see, these will be sufficient statistics for our estimator of λ. The intuition is simple: if $\hat{C}^1$ appears to be maximal among all the covariances of educations produced when hypotetically matching our populations of men and women, then it is likely that people care about endogamy on income infinitely less than they care about endogamy on education, that is $\lambda_2/\lambda_1 = 0$. We shall now show that the position of the observed vector $\hat{C}$ is informative about the parameter vector λm.

Recall the expression of the social surplus. One has

$$\mathcal{W}(\lambda) = \sup_{\mu \in \mathcal{M}(p,q)} \left(\sum_{x,y} \mu_{xy} \Phi^\lambda_{xy} + \mathcal{E}(\mu) \right)$$

this yields an optimal matching μ^λ that solves

$$
\begin{aligned}
\ln \frac{\mu^\lambda_{xy}}{\sqrt{\mu^\lambda_{x0}\mu^\lambda_{0y}}} &= \frac{\Phi^\lambda_{xy}}{2} \\
\sum_y \mu^\lambda_{xy} + \mu^\lambda_{x0} &= p_x \ \forall x \\
\sum_x \mu^\lambda_{xy} + \mu^\lambda_{0y} &= q_y \ \forall y.
\end{aligned}
\tag{6.10}
$$

One would like to do inference on λ.

Maximum likelihood. The most natural approach is maximum likelihood. This consists in finding λ maximizing

$$\sum_{x,y} \hat{\mu}_{xy} \ln \mu^\lambda_{xy} + \sum_x \hat{\mu}_{x0} \ln \mu^\lambda_{x0} + \sum_y \hat{\mu}_{0y} \ln \mu^\lambda_{0y}$$

with respect to λ. However, maximum likelihood estimation is not very easy in this context as the system of equations (6.10) cannot be solved explicitely.

Moment matching inference. It turns out that there is a unique value of λ such that

$$\hat{C}^k = \sum_{x,y} \mu^\lambda_{xy} \phi^k_{xy} \ \forall k = 1, ..., K$$

hence, if the model defined by the functions Φ_λ is well-specified, the vector of C^k's is a sufficient statistics for λ. We call this value $\lambda\left(\hat{C}\right)$ or when there is no possible ambiguity, $\hat{\lambda}$.

Theorem 6.4. *Let $\hat{\lambda} = \lambda\left(\hat{C}\right)$ be the maximizer of*

$$\max_{\lambda \in \mathbb{R}^K} \sum_{k=1}^K \lambda_k \hat{C}^k - \mathcal{W}(\lambda) \tag{6.11}$$

Then:

(i) $\hat{\lambda}$ is the only parameter vector λ such that

$$\sum_{x,y} \mu^{\lambda}_{xy} \phi^{k}_{xy} = \hat{C}^{k}$$

(ii) One has in general

$$\mathcal{E}(\hat{\mu}) \leq \mathcal{E}\left(\mu^{\hat{\lambda}}\right)$$

with equality if and only if $\mu^{\hat{\lambda}} = \hat{\mu}$, that is, if and only if the model is correctly specified by (6.9).

Proof. (i) By the envelope theorem, one has

$$\frac{\partial \mathcal{W}(\lambda)}{\partial \lambda_k} = \sum_{x,y} \mu_{xy} \Phi^{k}_{xy}$$

and by strict convexity of $\mathcal{W}$, $\hat{\lambda}$ is unique.

(ii) One has by definition of $\mu^{\hat{\lambda}}$

$$\sum_{x,y} \hat{\mu}_{xy} \Phi^{\hat{\lambda}}_{xy} + \sigma \mathcal{E}(\hat{\mu}) \leq \sum_{x,y} \mu^{\hat{\lambda}}_{xy} \Phi^{\hat{\lambda}}_{xy} + \sigma \mathcal{E}\left(\mu^{\hat{\lambda}}\right)$$

and, because $\mathcal{E}$ is strictly concave, equality holds if and only if $\mu^{\hat{\lambda}} = \hat{\mu}$. But

$$\sum_{x,y} \hat{\mu}_{xy} \Phi^{\hat{\lambda}}_{xy} = \sum_{x,y} \mu^{\hat{\lambda}}_{xy} \Phi^{\hat{\lambda}}_{xy}$$

by construction, hence

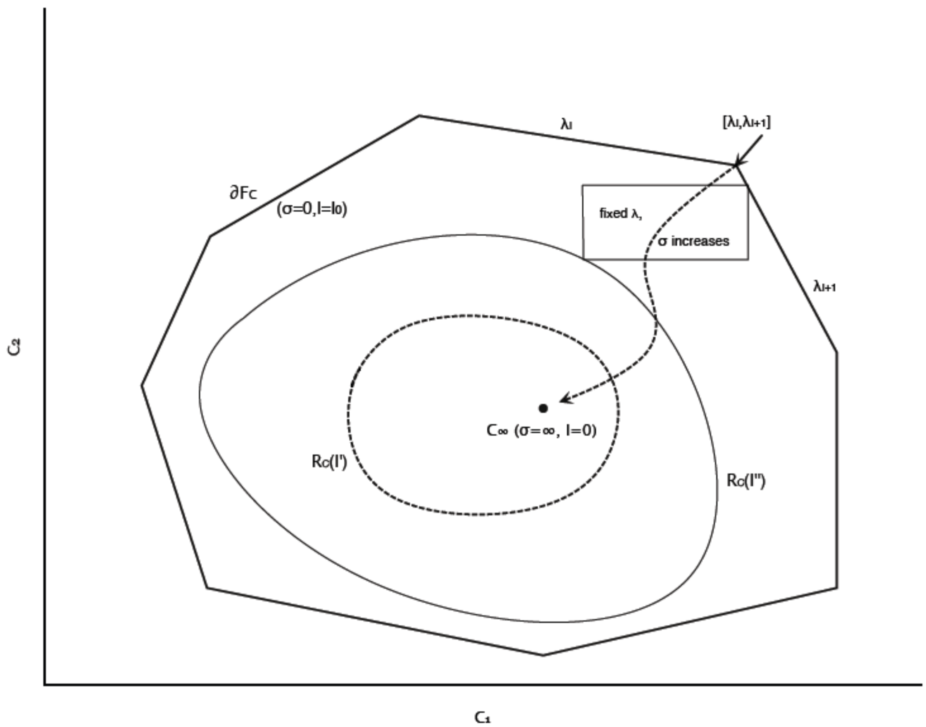
$$\mathcal{E}(\hat{\mu}) \leq \mathcal{E}\left(\mu^{\hat{\lambda}}\right)$$

with equality if and only if $\mu^{\hat{\lambda}} = \hat{\mu}$. ■

The covariogram. We now turn to a geometric interpretation of the identification of the parameter $\hat{\lambda}$. Consider the set of co-moments associated to every feasible matchings

$$\mathcal{F} = \left\{ \left(C^{1}, ..., C^{K}\right) : C^{k} = \sum_{xy} \mu_{xy} \phi^{k}_{xy},\ \mu \in \mathcal{M}(p, q) \right\}$$

This is a convex set, and the observed matching $\hat{\mu}$ belongs to this set. We call this set the *covariogram*.



The covariogram. Source: Galichon and Salanié (2010).

We are going to show that the location of the observed co-moments $\hat{C}$ on the covariogram will let us compute the estimator of parameter vector $\hat{\lambda}$. Define

$$\mathcal{E}_r\left(\hat{C}\right) = \mathcal{E}\left(\mu^{\lambda(\hat{C})}\right).$$

The level sets of $\mathcal{E}_r(.)$ are concentric,

Theorem 6.5. *One has that:*

(i) $\mathcal{E}_r(.)$ is a concave function defined on the covariogram.

(ii) The estimator $\hat{\lambda}$ of the parameter vector is given by the gradient of $-\mathcal{E}_r(.)$ at point $\hat{C}$.

Proof. (i) Let us show that

$$\mathcal{E}_r\left(\hat{C}\right) = \min_{\lambda} \mathcal{W}(\lambda) - \sum_{k=1}^{K} \lambda_k \hat{C}^k \tag{6.12}$$

Indeed, the optimum is reached at $\hat{\lambda} = \lambda\left(\hat{C}\right)$, and there

$$\begin{aligned} \mathcal{E}_r\left(\hat{C}\right) &= \mathcal{W}\left(\hat{\lambda}\right) - \sum_{k=1}^{K} \hat{\lambda}_k \hat{C}^k \\ &= \mathcal{E}\left(\mu^{\hat{\lambda}}\right) \end{aligned}$$

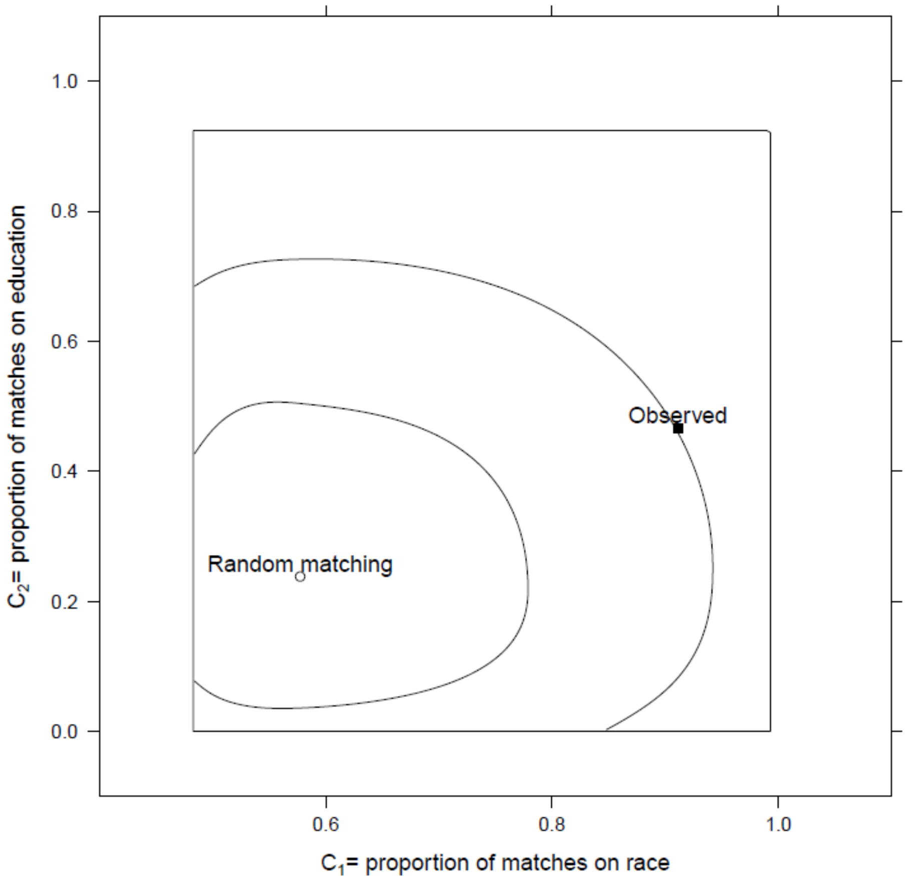
which shows (6.12). This implies that $\mathcal{E}_r\left(\hat{C}\right)$ is a concave function.

(ii) By the enveloppe theorem in (6.12), we get that

$$\frac{\partial \mathcal{E}_r\left(\hat{C}\right)}{\partial C^k} = \hat{\lambda}^k.$$

■

Here is an actual example of a covariogram.

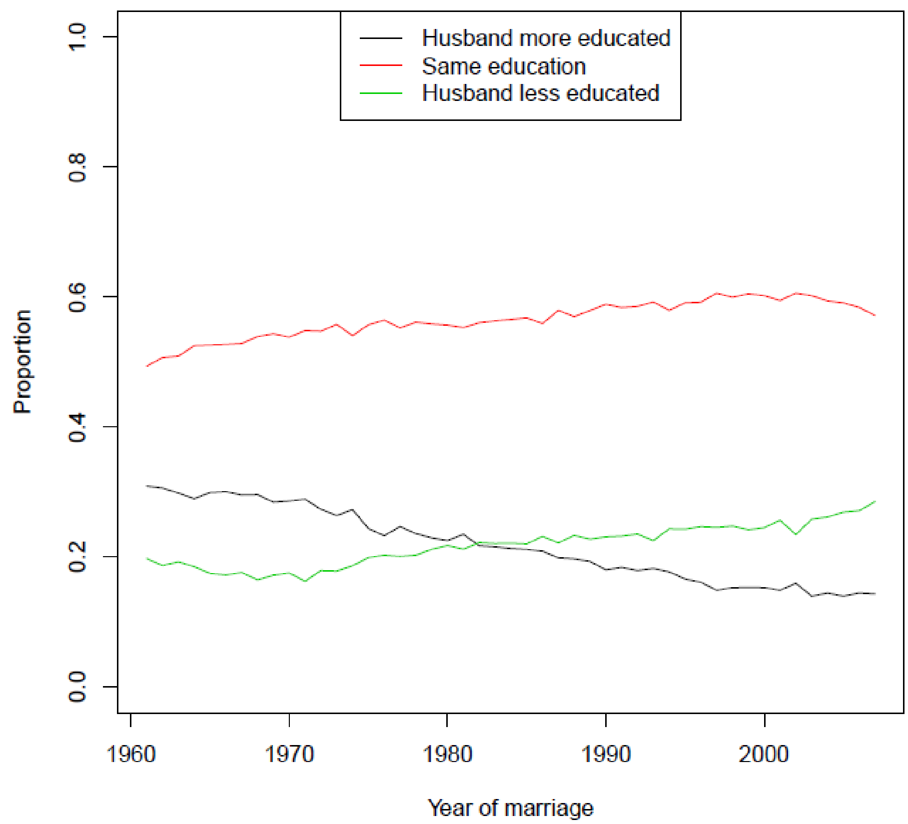


Endogamy on race vs. endogamy on education, covariogram estimated on US census data. Source: Galichon and Salanié (2010).

6.7 Empirical issues

Empirical studies of the marriage market is a topic at the boundary of sociology and economics. Matrimonial arrangement and strategies is also covered in much of the western literature. The traditional picture is evolving towards greater symmetry between men and women, for several reasons:

– access to a higher proportion of female to higher education
– laws and social norms making facilitating divorce
– less social prejudice about of interracial unions
– laws facilitating access to birth control.



Relative education of partners. Source: Chiappori, Salanié and Weiss (2011)

These legal dispositions, along with the accompanying evolution of the social norms, have impacted the marriage surplus in several ways, but often significantly raising the single women's utilities – hence raising their share or surplus at equilibrium. For instance, Chiappori and Weiss (2007) [30] study the impact of welfare laws on the partners' equilibrium shares of utility. Chiappori and Oreffice (2008) [29] study the impact of of birth control on "female empowerment". This has an impact on the marriage patterns, too. Christensen (2010) [31] studies the impact of legal dispositions giving early legal access of the birth-control pill to some young women on the age of marriage (or more broadly speaking, of the first cohabitation).

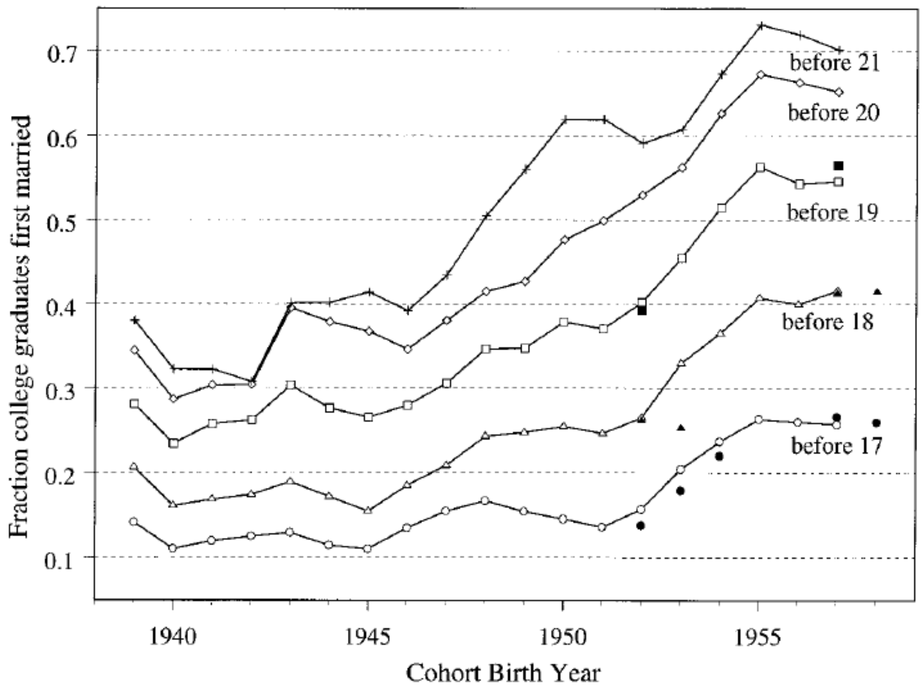
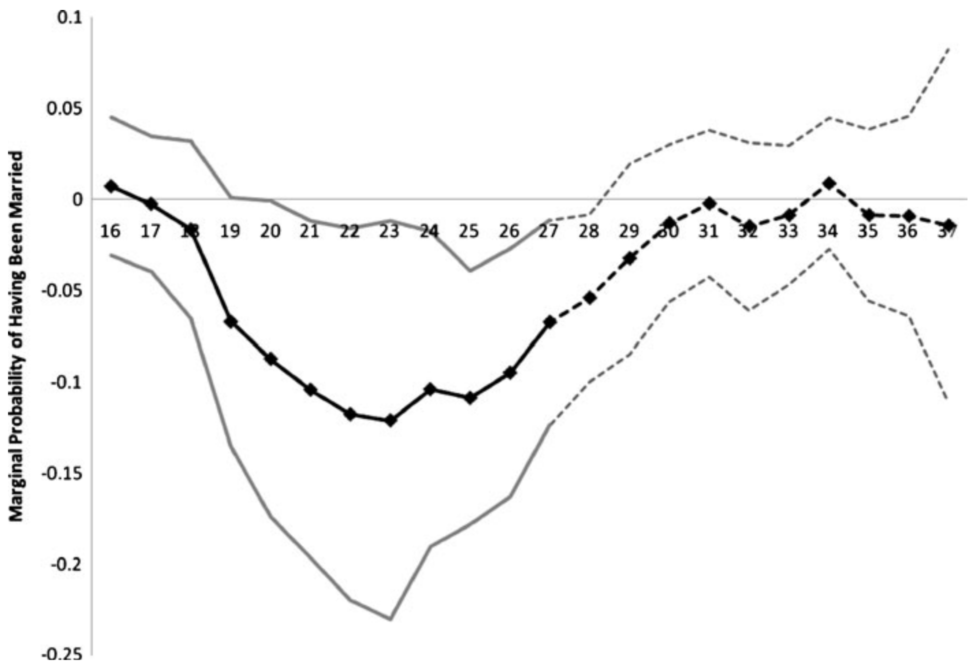


FIG. 6.—Fraction of never-married women having sex before various ages. Source: All but the solid markers: Inter-university Consortium for Political and Social Research (1985). Solid markers for birth cohorts of 1952, 1953, and 1954: Zelnik and Kantner (1989). Solid markers for birth cohorts of 1957 and 1958: Inter-university Consortium for Political and Social Research (1982). Three-year centered moving averages are shown. Solid markers, of the same shape as the open markers, give the values for contemporaneous data.

Source: Goldin and Katz (2002).



The average marginal effect of early legal access to the pill on the probability of having been married by a given age (average and 95% confidence bands). Source: Christensen (2010).

There is also a growing literature on the "Marital Premium College," an effect which was documented by Chiappori, Iyigun and Weiss (2009) [26] who argue that on top on increasing human capital, access to college gives access to a higher-educated set of potential mates, raising the quality of the matches. This is taken to the data by Chiappori, Salanié and Weiss (2011) [28] who find that the marital college premium is decreasing or stagnant for men, but it is increasing for women.

The Marital College Premium. Source: Chiappori, Salanié and Weiss (2011).

6.8 References and notes

Our exposition in sections 6.1 and 6.2 follows from Roth and Sotomayor (1990). Section 6.4 comes from Becker (1973). Theorem 6.2 is from Choo and Siow (2006) [32]. Theorem 6.3 is from Galichon and Salanié (2010) [53], from which 6.6 is also taken.

7 Hedonic models

Until the mid-1950s, neoclassical Economics modelled goods as homogeneous. This raised a number of problems that traditional economic approaches could not address.

– Given the fact that the amenities offered by car constantly evolve over time, how can one construct a price index measuring the evolution of the car prices?

– How can one explain price differentiation in wine?

– What does the correlation between the wage differentials and the level of risk associated to a given job reveal about individuals valuation for their own life?

– How can one account individuals preferences for environmental features?

In the 1950s, Dutch economist Tinbergen and collaborators tackled the issue of modeling heterogeneity and variation in consumer's taste for quality. In the 1970s, Sherwin Rosen offered an equilibrium framework to tackle this issue with greated generality, and provided an empirical strategy.

Applications include:

– Consumer Price Indexes
– Valuation of Environmental goods
– Wage differentiation
– Valuation of Risks to Life
– Economics of highly differentiated consumer goods (Wine, Art, etc.)

7.1 Hedonic equilibrium

In this section as in the next we follow closely Ekeland, Heckman and Nesheim (2004). Denote $\mathcal{X}$ the set of observable types of seller (or workers), and $\mathcal{Y}$ the set of observable types of buyers (or firms) and $\mathcal{Z}$ the set of contract types. Contract types may be the quality of a service that is exchanged between a buyer and a seller; it may be the care with wich it is performed, etc. Assume that the full type of a given worker is (x, ε), where x is observable and ε is not, and (y, η) for the firm. It is assumed that the contract z is fully observed.

The disamenity provided to seller of type (x, ε) by a contract of type z is $U(x, \varepsilon, z) \in \mathbb{R} \cup \{-\infty\}$, and the productivity of the same type of contract to a buyer of type (y, η) is given by $\Gamma(y, \eta, z) \in \mathbb{R} \cup \{-\infty\}$.

Let $P(z)$ price of characteristics z; this price shall be endogenously determined by equilibrium. Assume seller (or worker) of characteristics (x, ε) does

$$\max_z P(z) - U(x, \varepsilon, z)$$

which defines the quality of the service sold by (x, ε), denoted $Z(x, \varepsilon)$.

The buyer (or the firm) of characteristics (y, η) chooses to purchase service of quality z given by

$$\max_z \Gamma(y, \eta, z) - P(z).$$

We shall assume

$$U_{z\varepsilon} < 0 \text{ and } \Gamma_{z\eta} > 0,$$

where we use the notation

$$f_x := \frac{\partial f}{\partial x}, \; f_{xy} := \frac{\partial^2 f}{\partial x \partial y}, \text{ etc.}$$

The econometric problem will be: given the observation of $(P(z), z, x, y)$, recover U and Γ and the distribution of the unobservable characteristics ε and η. It is assumed that x and ε are independent and distributed according to $f_x(.)$ and $f_\varepsilon(.)$; similarly, y and η are independent and distributed according to $f_y(.)$ and $f_\eta(.)$.

The F.O.C. and S.O.C for the seller's program writes

$$\begin{aligned} P_z(z) - U_z(x, \varepsilon, z) &= 0 \\ P_{zz} - U_{zz} &< 0 \end{aligned}$$

The F.O.C. implicitely define a quality supply function

$$z = s(x, \varepsilon),$$

which can be inverted into

$$\varepsilon = \tilde{s}(z, x)$$

Similarly, the F.O.C. and S.O.C. for the buyer's program

$$\begin{aligned} \Gamma_z(y, \eta, z) - P_z(z) &= 0 \\ \Gamma_{zz} - P_{zz} &< 0 \end{aligned}$$

and the F.O.C. implicitely defines a quality demand function

$$z = d(y, \eta),$$

which can be inverted into

$$\eta = \tilde{d}(z, y).$$

Further, one gets

$$\frac{\partial \tilde{s}}{\partial z}(z, x) = \frac{P_{zz} - U_{zz}}{U_{z\varepsilon}} > 0$$

and

$$\frac{\partial \tilde{d}}{\partial z}(z, y) = \frac{P_{zz} - \Gamma_{zz}}{\Gamma_{z\eta}} > 0$$

Let us compute the supply density f_z^s. For an arbitrary function φ, one has

$$\begin{aligned}
\mathbb{E}[\varphi(Z)] &= \mathbb{E}[\varphi(X, \varepsilon)] \\
&= \iint\limits_{x,\varepsilon} \varphi(s(x, \varepsilon)) f_x(x) f_\varepsilon(\varepsilon) dx d\varepsilon \\
&= \int_z \varphi(z) dz \iint\limits_{x,\varepsilon} \delta(z - s(x, \varepsilon)) f_x(x) f_\varepsilon(\varepsilon) dx d\varepsilon
\end{aligned}$$

where we have introduced the Dirac delta function and used one of its property described in Appendix I. Therefore

$$f_z^s(z) = \iint\limits_{x,\varepsilon} \delta(z - s(x, \varepsilon)) f_x(x) f_\varepsilon(\varepsilon) dx d\varepsilon.$$

Using the formula

$$\delta(T(x) - z) = \delta\left(x - T^{-1}(z)\right) \left| \frac{dT^{-1}(z)}{dz} \right|,$$

and the fact that $\frac{\partial \tilde{s}}{\partial z}(z, x) > 0$, we get

$$f_z^s(z) = \iint\limits_{x,\varepsilon} \delta(\varepsilon - \tilde{s}(z, \varepsilon)) \frac{\partial \tilde{s}}{\partial z}(z, x) f_x(x) f_\varepsilon(\varepsilon) dx d\varepsilon,$$

hence

$$\begin{aligned}
f_z^s(z) &= \int_x \frac{\partial \tilde{s}}{\partial z}(z,x)\, f_\varepsilon(\tilde{s}(z,x))\, f_x(x)\, dx, \\
&= \int_x \frac{P_{zz} - U_{zz}}{U_{z\varepsilon}} f_\varepsilon(\tilde{s}(z,x))\, f_x(x)\, dx.
\end{aligned}$$

Similarly, one may compute the demand density f_z^d as

$$\begin{aligned}
f_z^d(z) &= \int_y \frac{\partial \tilde{d}}{\partial z}(z,y)\, f_\eta\left(\tilde{d}(z,y)\right) f_y(y)\, dy, \\
&= \int_y \frac{P_{zz} - \Gamma_{zz}}{\Gamma_{z\eta}} f_\eta\left(\tilde{d}(z,y)\right) f_y(y)\, dy
\end{aligned}$$

At equilibrium, supply and demand for quality z should coincide, hence

$$\int_x \frac{\partial \tilde{s}}{\partial z}(z,x)\, f_\varepsilon(\tilde{s}(z,x))\, f_x(x)\, dx \tag{7.1}$$

$$= \int_y \frac{\partial \tilde{d}}{\partial z}(z,y)\, f_\eta\left(\tilde{d}(z,y)\right) f_y(y)\, dy. \tag{7.2}$$

which can be reexpressed as

$$\int_x \frac{P_{zz} - U_{zz}}{U_{z\varepsilon}} f_\varepsilon f_x dx = \int_y \frac{P_{zz} - \Gamma_{zz}}{\Gamma_{z\eta}} f_\eta f_y dy$$

thus

$$P_{zz} = \frac{\int_y \frac{\Gamma_{zz}}{\Gamma_{z\eta}} f_\eta f_y dy - \int_x \frac{U_{zz}}{U_{z\varepsilon}} f_\varepsilon f_x dx}{\int_y \frac{f_\eta f_y}{\Gamma_{z\eta}} dy - \int_x \frac{f_\varepsilon f_x}{U_{z\varepsilon}} dx}.$$

Thus the curvature of the price curve P_{zz} is a weighted average of the curvature of technology Γ_{zz} and of the preferences u_{zz}.

7.2 The Tinbergen (1956) model

Compared to the model developped above, the Tinbergen model offers greater simplicity, as it assumes a quadratic form for U and Γ. But it also has the advantage of being multivariate. This model will lead to formulas in almost closed form.

7.2.1 Equilibrium

Assume

$$U(x,\varepsilon,z)=\frac{1}{2}z'Az-\theta'z$$

where

$$\theta=\mu_\theta(x)+\varepsilon.$$

By F.O.C.

$$\theta-Az+P_z=0 \tag{7.3}$$

and S.O.C.

$$P_{zz'}-A$$

is negative definite. Assume $\theta\sim N(\mu_\theta,\Sigma_\theta)$.

Assume

$$\Gamma(y,\eta,z)=\nu'z-\frac{1}{2}z'Bz$$

where

$$\nu=\mu_\nu(y)+\eta.$$

By F.O.C.

$$\nu-Bz-P_z=0 \tag{7.4}$$

and S.O.C.

$$-(P_{zz'}+B)$$

is negative definite. Assume $\nu\sim N(\mu_\nu,\Sigma_\nu)$. We shall assume that A, B and π_2 are symmetric.

Guess the following form of equilibrium price $P(z)$

$$P(z)=\pi_1'z+\frac{1}{2}z'\pi_2z$$

(we shall later verify the correctness of this formula). The supply and demand F.O.C. have

$$\begin{aligned}\theta-Az+\pi_1+\pi_2z &= 0\\ \nu-Bz-\pi_1-\pi_2z &= 0\end{aligned}$$

from which we can derive the quality supply and demand

$$\begin{aligned}z^s &= (A-\pi_2)^{-1}(\theta+\pi_1)\\ z^d &= (B+\pi_2)^{-1}(\nu-\pi_1)\end{aligned}$$

thus at equilibrium, we have the equality in distribution

$$(A-\pi_2)^{-1}(\theta+\pi_1) =_D (B+\pi_2)^{-1}(\nu-\pi_1)$$

hence, taking expectations

$$(A-\pi_2)^{-1}(\mu_\theta+\pi_1) = (B+\pi_2)^{-1}(\mu_\nu-\pi_1) \tag{7.5}$$

and taking second moments

$$(A-\pi_2)^{-1}\Sigma_\theta(A-\pi_2)^{-1} = (B+\pi_2)^{-1}\Sigma_\nu(B+\pi_2)^{-1}. \tag{7.6}$$

In the important special case where Σ_θ, Σ_ν A and B are diagonal, then π_2 is diagonal too, and the model boils down to the scalar case. In the scalar case, we get by (7.6) that

$$(A-\pi_2)\Sigma_\nu^{1/2} = (B+\pi_2)\Sigma_\theta^{1/2}$$

thus we can get the following closed-form solution to the model

$$\pi_2 = \frac{A\Sigma_\nu^{1/2} - B\Sigma_\theta^{1/2}}{\Sigma_\theta^{1/2}+\Sigma_\nu^{1/2}} \tag{7.7}$$

and

$$\pi_1 = \frac{-\mu_\theta\Sigma_\nu^{1/2} + \mu_\nu\Sigma_\theta^{1/2}}{\Sigma_\theta^{1/2}+\Sigma_\nu^{1/2}}. \tag{7.8}$$

7.2.2 Estimation

Let us now go back to the identification and estimation problem. Remember, one would like to identify preference and technology parameters A and B, but also worker and firm heterogeneities μ_θ, Σ_θ, μ_ν, and Σ_ν. π_1 and π_2 can be estimated from the data using the observation of function $P(z)$, however from equations (7.7) and (7.8), it is clear that π_1 and π_2 are not sufficient to identify all the parameters of interest. Two "polar" cases arise, discussed in Rosen (1974) [88]:

1. no heterogeneity on the worker side, i.e. $\Sigma_\theta = 0$ and $\Sigma_\nu > 0$, in which case

$$\begin{aligned}\pi_1 &= -\mu_\theta \\ \pi_2 &= A.\end{aligned}$$

2. no heterogeneity on the firm side, i.e. $\Sigma_\nu = 0$ and $\Sigma_\theta > 0$, in which case

$$\begin{aligned} \pi_1 &= -\mu_\nu \\ \pi_2 &= B. \end{aligned}$$

Rosen's two-step estimate procedure: estimate $P(z)$ from the data, then plug in the estimate of P into F.O.C. (7.3) and (7.4) in order to recover preferences u and technology Γ. One gets an estimate of

$$\hat{P}(z) = \hat{\pi}_1' z + \frac{1}{2} z' \hat{\pi}_2 z$$

and one then uses the First Order Conditions of supply resp. demand in order to regress $\hat{P}_z(z)$ on z and x (resp. z and y)

$$\begin{aligned} \hat{\pi}_1' + \hat{\pi}_2 z &= -\mu_\theta(x) + Az - \varepsilon \\ \hat{\pi}_1' + \hat{\pi}_2 z &= \mu_\nu(y) + Bz + \eta \end{aligned}$$

and therefore the parameters can be obtained by regressing $\hat{\pi}_1' + \hat{\pi}_2 z$ on (x, z) and (y, z) respectively. In Rosen's approach, this requires the observation of several markets for identifications of all of the parameters of the model.

7.3 Hedonic models as matching models

7.3.1 First reformulation: assignment problem

In [27], Chiappori, McCann and Nesheim (2010) show that hedonic models are equivalent to matching problems. For the sake of exposition, we shall assume in the sequel that there are no unobservable characteristics ε and η. Let μ be the distribution of the seller's types, which is assumed to have a density $f_x(.)$, and let ν be the distribution of the buyer's types, with density $f_y(.)$.

Consider

$$\Phi(x, y) = \max_z (\Gamma(y, z) - U(x, z)) \tag{7.9}$$

which is the "indirect utility" of the pair (x, y), the maximal joint surplus that they can get together.

Consider the following optimal assignment problem

$$\begin{array}{l} \max \iint \Phi(x,y)\, d\pi(x,y) \\ \pi \in M(\mu,\nu) \end{array} \tag{7.10}$$

which has dual problem

$$\begin{array}{l} \min \int u(x)\, d\mu(x) + \int v(y)\, d\nu(y) \\ s.t.\ u(x) + v(y) \geq \Phi(x,y) \end{array} \tag{7.11}$$

For u and v solution to the dual problem (7.11), define

$$P_{\max}(z) = \inf_x \left(u(x) + U(x,z)\right)$$

and

$$P_{\min}(z) = \sup_y \left(\Gamma(y,z) - v(y)\right).$$

For each x, y and z, one has

$$\begin{aligned} u(x) + v(y) &\geq \Phi(x,y) \\ &\geq \Gamma(y,z) - U(x,z) \end{aligned}$$

thus

$$u(x) + U(x,z) \geq \Gamma(y,z) - v(y)$$

therefore

$$P_{\max}(z) \geq P_{\min}(z).$$

Now, let us consider a measurable function $P(z)$ such that

$$P_{\max}(z) \geq P(z) \geq P_{\min}(z). \tag{7.12}$$

One has:

Theorem 7.1. *There is a hedonic equilibrium supported by $P(z)$ satisfying (7.12) in which buyers and sellers are matched according to $\pi(x,y)$ solution of (7.10).*

Proof. Let $Z(x,y)$ be the optimal z in (7.9).

One has that x chooses $z = Z(x)$ so to maximize

$$u(x) = \max_z \left(P(z) - U(x,z)\right).$$

Similarly, y chooses $z = Z(y)$ so to maximize

$$v(y) = \max_z \left(\Gamma(y,z) - P(z)\right).$$

Hence, for each z,

$$\begin{aligned} u(x) + v(y) &\geq \Gamma(y,z) - U(x,z) \\ &\geq \sup_z \left(\Gamma(y,z) - U(x,z)\right) = \Phi(x,y) \end{aligned}$$

Now, x and y decide to match when equality holds, namely when there exists z such that

$$\begin{aligned} u(x) &= P(z) - U(x,z), \text{ and} \\ v(y) &= \Gamma(y,z) - P(z), \end{aligned}$$

that is, when $Z(x) = Z(y) = Z(x,y)$. Let f_z be the induced distribution of the contract. It is clear that f_z equates supply and demand. ■

7.3.2 Second reformulation: network flow problem

One can actually push the equivalence one step forward and show that hedonic models are formally equivalent to a network flow problem. In order to make exposition as simple as possible, we shall treat the distribution of the sellers and the buyers types as discrete, while they have been treated so far as continuous.

To be consistent with the notations of our exposition of network flows in Section 4.4, we introduce the following notations,

$$k_{xz} = U(x,z) \tag{7.13a}$$
$$k_{zy} = -\Gamma(y,z) \tag{7.13b}$$
$$k = +\infty \text{ otherwise.} \tag{7.13c}$$

The reduced cost function c is given by

$$\begin{aligned} c_{xz} &= k_{xz} \\ c_{zy} &= k_{zy} \\ c_{xy} &= \inf_{z\in\mathcal{Z}} \left(k_{xz} + k_{zy}\right) \end{aligned}$$

so we have

$$\begin{aligned} c_{xy} &= \inf_{z\in\mathcal{Z}} \left(U\left(x,z\right) - \Gamma\left(y,z\right)\right) \\ &= -\Phi\left(x,y\right), \end{aligned}$$

where we recall that

$$\Phi\left(x,y\right) = \sup_{z\in\mathcal{Z}} \left(\Gamma\left(y,z\right) - U\left(x,z\right)\right). \tag{7.14}$$

Let

$$\begin{aligned} b_x &= -f_x\left(x\right) && (7.15a) \\ b_y &= f_y\left(y\right) && (7.15b) \\ b_z &= 0 && (7.15c) \end{aligned}$$

where f_x and f_y are the probability mass function associated to the distributions of x and y (they are the discrete analogs of the pdfs, as we had to discretize the problem).

As before, recall that

$$\begin{aligned} u\left(x\right) &= \max_z \left(P\left(z\right) - U\left(x,z\right)\right), && (7.16) \\ v\left(y\right) &= \max_z \left(\Gamma\left(y,z\right) - P\left(z\right)\right), && (7.17) \end{aligned}$$

and introduce

$$\begin{aligned} w_x &= u\left(x\right) \\ w_z &= P\left(z\right) \\ w_y &= -v\left(y\right). \end{aligned}$$

Note that P is an equilibrium hedonic price if and only if $w \in S_k$, that is

$$w_j - w_i \leq k_{ij} \tag{7.18}$$

Indeed, for $i = x$, $j = z$, this is

$$P(z) - u(x) \leq U(x, z) \tag{7.19}$$

which is (7.16); for $i = z$ and $j = y$, this is

$$-v(y) - P(z) \leq -\Gamma(y, z) \tag{7.20}$$

which is (7.17); and for $i = x$ and $j = z$, this is

$$-v(y) - u(x) \leq c_{xy} = -\Phi(x, y) \tag{7.21}$$

which is (7.14).

Note that equality in (7.19) means that x chooses contract z; equality in (7.20) means that y chooses contract z; and equality in (7.21) means that x and y choose together contract z. This is precisely the complementary slackness conditions associated to the network flow problem!

Hence we have the following result:

Theorem 7.2. *The hedonic equilibrium problem is equivalent to the following network flow problem*

$$\begin{aligned} &\min_{\pi \geq 0} \sum_{(ij) \in A} \pi_{ij} k_{ij} \\ &\mathcal{N}\pi = b \end{aligned} \tag{7.22}$$

and its dual version

$$\begin{aligned} &\max_{w} \sum_{i} w_i b_i \\ &w_j - w_i \leq k_{ij} \end{aligned}$$

where b and k are given by Equations (7.15) and (7.13).

7.4 Empirical literature

Hedonic models and hedonic regressions have been initially designed to build consumer price indexes: see Griliches (1971) [58], Boskin et al. (1996) [16], Lancaster (1966) [65]. This has been especially needed in domains where technology has been evolving fast, such as automotive industry, see Triplett

(1969) [108], or, more recently, personal computers, see Berndt et al. (1995) [10], Berndt and Rappaport (2001) [11], and Pakes (2003) [83].

Since Thaler and Rosen (1976) [103], hedonic models have been used to measure the valuation of risks to life. The job characteristics z usually incorportates many features, such as age, marital status, race, geographical information, sector, part/full-time, and also information of the riskiness of the job. The riskiness of the job is measured as the actuarial risk of death in a given occupation class. A hedonic regression is then conducted to regress the market wage $P(z)$ on these characteristics, hence to provide an estimate $\hat{P}(z)$. However, the differential effect of risk on the market wage cannot be directly interpreted as the differential valuation by individuals of risks to life. Some risky jobs, for instance, may be more productive to the firm, justifying higher market wages. The hedonic approach sets to distinguish variation in markets wages that comes from productivity from variation that comes from individual preferences (valuation of risks to life). The approach set forth by Thaler and Rosen has opened the way to what has been called the literature on Value of Statistical Life (VSL). See surveys in Schelling (1987) [96] and Viscusi (2008) [112].

This approach has been hugely influential in labour economics: Arnould and Nichols (1983) [3], Hersch (1998) [62], Murphy and Topel (2006) [81]. It has also been applied in other area of economics: valuation of safety equipments on the automobile market: Atkinson and Halvorsen (1990) [4], or cigarette smoking: Viscusi and Hersch (2001) [113].

Hedonic pricing models have also been used in wine economics, see e.g. Nerlove (1995) [82], Combris et al. (1997) [34], art economics, see Chanel (1995) [23], in environmental economics, see Greenstone and Gallagher (2008) [57], and to value network externalities, see Gandal (1994).

7.5 References and notes

Sections 7.1 and 7.2 are based on the exposition of hedonic models in Ekeland, Heckman and Nesheim (2004), who build on Tinbergen (1956) [104] and Rosen (1974) [88]. Paragraph 7.3.1 is based on Chiappori, McCann and Nesheim (2010) [27]. The ideas underlying Paragraph 7.3.2 were explained in a talk given by R. T. Rockafellar at the University of Washington in 2005; he credits Maurice Queyranne for them.

Part III. Non-transferable utility models

8 The housing problem

Recall the "housing problem" from Section 1.1, Example 1.1. Given an initial assignment of n houses to n individuals, and assuming individuals form preferences over houses and can trade houses, what is the core of the corresponding game? It is assumed that houses form no preferences over individuals. To frame the question using numbers, let U_{ij} be the utility that house j gives to individual i. Let $j = \sigma(i)$ be the set of assignments of houses to individuals, where $\sigma \in \mathfrak{S}_n$ is a permutation. One wonders what the set of Pareto outcomes is.

For simplicity it will be assumed that people have strict preferences over houses, i.e. $U_{ij} \neq U_{ij'}$ for $j \neq j'$.

8.1 Pareto optimality

Consider n indivisible goods (eg. houses) $j = 1, ..., n$ to be allocated to n individuals. Let U_{ij} be the utility that house j gives to individual i. Assume preferences of individuals are strict, that is $U_{ij} \neq U_{ij}$ for $j \neq j'$.

Consider the assignment where house i is assigned to individual $\sigma_0(i) = i$ (this is without loss of generality, as one can always relabel houses). Question: when is this assignment efficient?

If there are two individuals, say i and j that would both benefit from swapping houses, then assignment is not efficient. Thus if assignment is efficient, then inequalities $U_{ij} \geq U_{ii}$ and $U_{ji} \geq U_{jj}$ cannot hold simultaneously unless they are both equalities. More generally, Pareto rules out trading cycles whose members would benefit from trading (strictly for some).

We thus adopt this as our definition of ex-post Pareto efficiency.

Definition 8.1. *Initial assignment $\sigma_0(i) = i$ is efficient in the housing problem if and only if for every potential "trading cycle" $i_1, ..., i_{p+1} = i_1$,*

$$\forall k,\ U(i_k, i_{k+1}) \geq U(i_k, i_k)$$

implies that all these inequalities are equalities. That is, introducing

$$R_{ij} = U(i, i) - U(i, j), \tag{8.1}$$

we have

$$\forall k,\ R_{i_k i_{k+1}} \leq 0 \textit{ implies } \forall k,\ R_{i_k i_{k+1}} = 0,$$

which is to say that assignment σ_0 is efficient if and only if the matrix R_{ij} is cyclically consistent.

We turn to a characterization of Pareto efficiency, which does not rely on the assumption of strict preferences.

Theorem 8.1 (Afriat). *The following conditions are equivalent:*

*(i) The matrix R_{ij} satisfies **"cyclical consistency"**: for any cycle $i_1, ..., i_{p+1} = i_1$,*

$$\forall k,\ R_{i_k i_{k+1}} \leq 0 \textit{ implies } \forall k,\ R_{i_k i_{k+1}} = 0, \tag{8.2}$$

(ii) There exist numbers (v_i, λ_i), $\lambda_i > 0$, such that

$$v_j - v_i \leq \lambda_i R_{ij},$$

(iii) There exist numbers v_i such that

$$\begin{aligned} R_{ij} &\leq 0 \textit{ implies } v_j - v_i \leq 0, \textit{ and} \\ R_{ij} &< 0 \textit{ implies } v_j - v_i < 0. \end{aligned}$$

Proof of Theorem 8.1. (ii) immediately implies (iii). We now show that (iii) implies (i). Consider a cycle $i_1, ..., i_{p+1} = i_1$, such that

$$\forall k,\ R_{i_k i_{k+1}} \leq 0.$$

Then by (iii) there exist numbers v_i such that

$$\begin{aligned} R_{ij} &\leq 0 \text{ implies } v_j - v_i \leq 0, \text{ and} \\ R_{ij} &< 0 \text{ implies } v_j - v_i < 0. \end{aligned}$$

thus one has $v_{i_{k+1}} - v_{i_k} \leq 0$ for all k, hence all the v_{i_k} are equal. Assume now that there is a k such that $R_{i_k i_{k+1}} < 0$. Then $v_{i_{k+1}} - v_{i_k} < 0$, which is a contradiction. Therefore,

$$\forall k,\ R_{i_k i_{k+1}} = 0,$$

which proves the cyclical consistency of matrix R, that is (i).

For (i) implies (ii), assume R_{ij} is cyclically consistent, and assume by contradiction that (ii) fails to hold. Then the linear program

$$\begin{aligned}
&\min \lambda.0 + v.0 \\
&\text{s.t.} \\
&v_j - v_i \le \lambda_i R_{ij} \\
&\lambda_i \ge 1
\end{aligned}$$

has no feasible solution, thus the value of this program is $+\infty$. The dual of this problem is

$$\begin{aligned}
&\max \sum_i x_i \\
&\text{s.t.} \\
&x_i = -\sum_j \pi_{ij} R_{ij} \\
&\sum_k \pi_{ik} = \sum_k \pi_{ki} \\
&x_i \ge 0,\ \pi_{ij} \ge 0.
\end{aligned}$$

Hence, there is $\pi_{ij} \ge 0$ such that $\sum_k \pi_{ik} = \sum_k \pi_{ki}$ and $\sum_j \pi_{ij} R_{ij} < 0$ for some i. This implies that one can find a cycle $i_1, ..., i_{p+1} = i_1$, such that $R_{i_1 i_2} < 0$ and

$$\forall k,\ R_{i_k i_{k+1}} \le 0,$$

which yields a contradiction. ■

As we have assumed strict preferences, we have $R_{ij} \ne 0$ for $i \ne j$. Then (iii) equivalently reformulates as:

(iii') There exist numbers v_i such that

$$R_{ij} < 0 \text{ implies } v_j - v_i < 0.$$

(i.e., we can drop the first part of (iii): $R_{ij} \le 0$ implies $v_j - v_i \le 0$).

As a consequence, we have a characterization of Pareto efficient outcomes.

Theorem 8.2. *In the housing problem, the following conditions are equivalent:*

(i) Assignment $\sigma_0 = Id$ is Pareto efficient

(ii) $\exists\lambda_i > 0$ *and* $v \in \mathbb{R}^n$ *such that*

$$v_j - v_i \leq \lambda_i R_{ij}.$$

(iii) $\exists\lambda_i > 0$ *such that* $\sigma_0 = Id$ *maximizes*

$$\max_{\sigma \in \mathsf{S}_n} \sum_{i=1}^{n} \lambda_i U(i, \sigma(i))$$

that is

$$\min_{\sigma \in \mathsf{S}_n} \sum_{i=1}^{n} \lambda_i R_{i\sigma(i)} = 0. \tag{8.3}$$

(iv) $\exists v \in \mathbb{R}^n$ *such that*

$$R_{ij} < 0 \text{ implies } v_j - v_i < 0.$$

Proof. Given Theorem 8.1, the only part left to prove is the equivalence between (ii) and (iii). One has:

(iii) $\iff \exists\lambda_i > 0,\ \min_{\sigma \in S} \sum_{i=1}^{n} \lambda_i R_{i\sigma(i)} = 0$

$\iff \exists\lambda_i > 0,\ \min_{\sigma \in S} \sum_{i=1}^{n} \lambda_i R_{i\sigma(i)}$ is reached for $\sigma = Id$

$\iff \exists\lambda_i > 0$, $u, v \in \mathbb{R}^n$

$$\begin{aligned} u_i + v_j &\leq \lambda_i R_{ij} \\ u_i + v_i &= 0 \end{aligned}$$

$\iff \exists\lambda_i > 0$, $v \in \mathbb{R}^n$

$$v_j - v_i \leq \lambda_i R_{ij},$$

which is (ii). ■

Hence, we may state as an important corollary:

Corollary 8.1. *The set of Pareto allocations coincides with the set of maximizers of*

$$\max_{\sigma \in \mathsf{S}_n} \sum_{i=1}^{n} \lambda_i U_{i\sigma(i)}$$

for all possible choices of weights $\lambda_i > 0$.

8.2 Competitive equilibrium

Let us now turn to the study of competitive equilibria. In order to do this, we shall slightly reformulate the problem. We shall assume that there is a market for "shares", which are lottery tickets over houses.

Owning p_j of shares of house j gives a probability p_j of getting house j.

Individual i can purchase quantity

$$\pi_{ij} \geq 0 \tag{8.4}$$

of share of house j. This gives individual i a probability π_{ij} of winning house j.

It is forbidden to detain more than one unit of probability in total: individuals are thus constrained to have

$$\sum_j \pi_{ij} \leq 1. \tag{8.5}$$

The competitive price of one unit of probability of winning house j is q_j. It is assumed that invididual i has an initial monetary amount B_i to spend, so that her budget constraint is

$$\sum_j \pi_{ij} q_j \leq B_i. \tag{8.6}$$

Of course, the market clearing condition is that the sum of "shares" of each house j is equal to one; that is

$$\sum_i \pi_{ij} = 1.$$

Before we study equilibrium, we investigate the notion of Pareto optimality.

Definition 8.2. *An allocation (π_{ij}) is called an Ex-Ante Pareto efficient allocation if no other allocation $(\bar{\pi}_{ij})$ satisfies constraints (8.4), (8.5) and (8.6) with*

$$\sum_j \pi_{ij} U_{ij} \leq \sum_j \bar{\pi}_{ij} U_{ij} \ \forall i$$

one of these inequalities being strict.

It is important to distinguish the notions of Ex-Ante efficiency, and the notion we used in the previous section, commonly called Ex-Post efficiency.

Proposition 8.1. *An allocation* (π_{ij}) *is Ex-Ante Pareto efficient allocation if and only if it maximizes*

$$\begin{aligned} & \max_{\pi \geq 0} \sum_{ij} \lambda_i \pi_{ij} u_{ij} \\ s.t. \qquad & \sum_i \pi_{ij} = 1 \\ & \sum_j \pi_{ij} = 1 \end{aligned}$$

for some any possible weights $\lambda_i > 0$.

Proof. Left as an exercise. ■

Hence, (Ex-Ante) Pareto allocation are the solution of an optimal assingment problem with surplus function $(\lambda_i u_{ij})$. The dual of this problem is

$$\begin{aligned} & \min_{u,v} \sum u_i + \sum v_j \\ s.t. \qquad & u_i + v_j \geq \lambda_i u_{ij}. \end{aligned}$$

Let us now turn to the problem of equilibrium. The invidual maximization problem of individual i is given by

$$\begin{aligned} & \max_{\pi_{ij} \geq 0} \sum_j \pi_{ij} u_{ij} \\ s.t. \qquad & \sum_j \pi_{ij} q_j \leq B_i \\ & \sum_j \pi_{ij} \leq 1 \end{aligned}$$

which is to say

$$\begin{aligned} \max_{\pi_{ij} \geq 0} \min_{\alpha_i \geq 0, \gamma_i \geq 0} \sum_j \pi_{ij} u_{ij} + \qquad & \alpha_i \left(B_i - \sum_j \pi_{ij} q_j \right) \\ & \gamma_i \left(1 - \sum_j \pi_{ij} \right) \end{aligned}$$

so the dual is

$$\begin{aligned} &\min_{\alpha_i \geq 0, \gamma_i \geq 0} \alpha_i B_i + \gamma_i \\ s.t. \quad & u_{ij} \leq \gamma_i + \alpha_i q_j \end{aligned}$$

that is, setting $\lambda_i = \alpha_i^{-1}$ and $u_i = \gamma_i / \alpha_i$,

$$\begin{aligned} &\min_{\alpha_i \geq 0, \gamma_i \geq 0} \lambda_i^{-1} (B_i + u_i) \\ s.t. \quad & \lambda_i u_{ij} \leq u_i + q_j. \end{aligned}$$

8.3 Link with Revealed Preferences

The problem of Revealed Preferences in consumer demand is as follows. One observes a series of choices made by a consumer, i.e. in each of a series of *experiments* one observes the budget set, set of all possible choices, and the choice actually made. The budget set in experiment i is defined by

$$g_i(x) \leq 0$$

and the consumption bundle chosen is $x_i \in \mathbb{R}^d$ where x_i^k is the amount of good k purchased by the consumer in experiment i.

For example, in the case of a linear budget set, one has

$$g_i(x) = x \cdot p_i - x_i \cdot p_i$$

so that x is affordable in experiment i if and only if

$$x \cdot p_i \leq x_i \cdot p_i$$

where $x_i \cdot p_i$ is the amount of money actually spent in experiment i, that is, the wealth of the consumer in experiment i.

The problem of revealed preferences challenges the assumption of rationality of the choices of the consumer. A rational consumer is supposed to have a *preference relation* $\preceq$ over consumption bundles: such a relation satisfies the three following properties:

– *reflexivity*, i.e. for any x, one has $x \preceq x$,
– *completeness*, ie. for any two x, y, $x \preceq y$ or $y \preceq x$ (or both)

– *transitivity*, ie. such that $x \preceq y$ and $y \preceq z$ imply $x \preceq z$.
Denote $x \prec y$ if $x \preceq y$ and not $y \preceq x$.

Assume x_j is such that $g_i(x_j) \le 0$, and denote

$$x_j \preceq_{DRP} x_i \text{ whenever } g_i(x_j) \le 0,$$

in which case x_i is said to be "directly revealed preferred" to x_j. This means that bundle x_j was affordable in experiment i, but consumer preferred x_i.

Let us make several remarks.

1. If $x_j \preceq_{DRP} x_i$, then consumer (weakly) prefers i to j; otherwise consumer would have chosen x_j in experiment i.

2. However, the converse does not necessarily hold: consumer may prefer x_i to x_j while x_j is not affordable under experiment i.

3. $\preceq_{DRP}$ is not a preference relation, as it need not be transitive.

For a consumer to be rational, a preference relation $\preceq_{RP}$ need to exist such that

$$x \preceq_{DRP} y \text{ implies } x \preceq_{RP} y.$$

This will be our first attempt at defining rationalizability of choices: however one sees that this definition is devoid of any meaning, as the trivial preference relation such that $x \preceq_{RP} y$ for any two pair x, y would rationalize any choices. We need to have a sense that if x_j was in the strict interior of the budget set of experiment i, then x_i is strictly preferred to x_j. Hence we define rationalizability of choices as follows.

Definition 8.3. *Choices are rationalizable if there exists a preference relation $\preceq_{RP}$ such that*

$$\begin{aligned} x &\preceq_{DRP} y \text{ implies } x \preceq_{RP} y, \text{ and} \\ x &\prec_{DRP} y \text{ implies } x \prec_{RP} y. \end{aligned}$$

We can represent $\preceq_{DRP}$ by a matrix $R_{ij} = g_i(x_j)$ such that

$$\begin{aligned} R_{ij} &< 0 \text{ whenever } x_j \prec_{DRP} x_i \\ R_{ij} &= 0 \text{ whenever } x_j \preceq_{DRP} x_i \text{ and } x_j \preceq_{DRP} x_i. \end{aligned}$$

A preference relation $\preceq_{RP}$ is representable by a vector v_i such that

$$\begin{aligned} v_j &\leq v_i \text{ whenever } x_j \preceq_{RP} x_i \\ v_j &< v_i \text{ whenever } x_j \prec_{RP} x_i. \end{aligned}$$

Hence, choices are rationalizable if and only if there exist a vector v_i such that

$$\begin{aligned} R_{ij} &\leq 0 \text{ implies } v_j - v_i \leq 0, \text{ and} \\ R_{ij} &< 0 \text{ implies } v_j - v_i < 0. \end{aligned}$$

This is equivalent to the existence of a utility level v_j associated to good j such that consumption x_i results from the maximization of consumer i's utility under budget constraint $g(x) \leq 0$, namely

$$i \in \operatorname{argmax}_j \{v_j : g_i(x_j) \leq 0\}.$$

By Afriat's theorem, this is equivalent to the existence of numbers (v_i, λ_i), $\lambda_i > 0$, such that

$$v_j - v_i \leq \lambda_i R_{ij},$$

which is a set of linear inequalities that can be tested using linear programming. One sees that these results have exactly the same structure as the results in Paragraph 8.1, which means that there is a duality between the housing problem without money, and with the problem of revealed preferences in consumer demand.

To summarize this duality, let us give the following table:

	Revealed prefs.	Pareto indiv. allocs.
setting	consumer demand	allocation problem
budget sets	$\{j : c_{ij} \leq c_{ii}\}$	$\{-v : -v \leq -v_i\}$
cardinal utilities to j	v_j	$-c_{ij}$
# of consumers	one, representative	n, $i \in \{1, ..., n\}$
# of experiments	n	one
goods	divisible	indivisible
unit of c_{ij}	dollars	utils
unit of v_i	utils	dollars

8.4 Application to demand theory

The present paragraph is very informal, but aims at recovering classical notions in Walrasian consumer demand theory, such as the Slutsky matrix. We are in the case where the budget sets are linear, that is

$$g_i(x) = x \cdot p_i - x_i \cdot p_i$$

and we assume that there are a very large number of observations.

First, assume that all the prices p_i have been normalized so that $x_i \cdot p_i = 1$ (one can do this without changing the Direct Revealed Preferences relation).

One has

$$\begin{aligned} R_{ij} &= x_j \cdot p_i - x_i \cdot p_i \\ &= x_j \cdot p_i - 1 \\ &= x_j \cdot p_i - x_j \cdot p_j \\ &= x_j \cdot (p_i - p_j) \end{aligned}$$

thus, by Afriat's theorem, the data is rationalizable if and only if there exist v_i and scalars $\lambda_i > 0$ such that

$$v_j - v_i \leq \lambda_i R_{ij}$$

Now, $p_i = p_j$ implies $v_i = v_j$, thus one may write $v_i = v(p_i)$ for some map $v(.)$. Assume further that $p_i = p_j$ implies that $x_i = x_j$ and $\lambda_i = \lambda_j$. Then one can write

$$\begin{aligned} x_i &= x(p_i) \\ \lambda_i &= \lambda(p_i) \end{aligned}$$

for some map $x(p)$ and $\lambda(p)$, and Afriat's condition rewrites

$$v(p_j) - v(p_i) \leq \lambda(p_i)\, x(p_j) \cdot (p_i - p_j).$$

Taking p_i and p_j can be taken arbitrarily close, and passing to the continuous limit, this yields to

$$\nabla v(p) \cdot dp \leq -\lambda(p)\, x(p) \cdot dp$$

thus

$$\nabla v(p) = -\lambda(p)\, x(p). \tag{8.7}$$

Therefore, differentiating a second time, we get

$$-D^2 v(p) = \nabla \lambda(p) \cdot x(p) + \lambda(p)\, Dx(p)$$

therefore

$$Dx(p) = -\frac{D^2 v(p)}{\lambda(p)} - \nabla \ln \lambda(p) \cdot x(p).$$

We recover a familiar result in demand theory: the differential of the Marshallian demand equals a symmetric matrix plus a rank one matrix.

To make this analogy complete, let us recover the standard results from demand theory. Let $u(x)$ be the utility function of the consumer, one has

$$\begin{aligned} v(p) &= \max_x \{u(x) : x \cdot p \leq 1\} \\ &= \max_x \min_\lambda \{u(x) + \lambda(1 - x \cdot p)\} \end{aligned}$$

Hence, by the envelope theorem,

$$\nabla v(p) = -\lambda(p)\, x(p) \tag{8.8}$$

which is the same as (8.7). Note, however, that

$$-\frac{D^2 v(p)}{\lambda(p)}$$

is not the Slutsky matrix. To compare with the Slutsky equation, we need to recall the notion of *expenditure function*

$$\begin{aligned} E(p, u) &= \min_x \{p.x : u(x) \geq u\} \\ &= \min_x \max_\mu p.x + \mu(u(x) - u) \end{aligned}$$

and *Hickisan demand*

$$H(p,u) = \nabla_p E \tag{8.9}$$

and

$$\begin{aligned} V(p,w) &= \max_x \{u(x) : x \cdot p \leq w\} \\ &= \max_x \min_\lambda (u(x) + \lambda (w - x \cdot p)) \end{aligned}$$

with

$$X(p,w)$$

the associated Marshallian demand, so that

$$\begin{aligned} v(p) &= V(p,1), \\ x(p) &= X(p,1). \end{aligned}$$

Note that for a fixed p, $V(p,.) = E(p,.)^{-1}$, that is

$$E(p, V(p,w)) = w,$$

and we have

$$\begin{aligned} X(p,w) &= H(p, V(p,w)), \\ H(p,u) &= X(p, E(p,u)). \end{aligned}$$

Thus, differentiating the latter, one has

$$D_p H(p,u) = D_p X + \partial_w X. (\nabla_p E)'.$$

Now, we make use of (8.9), and apply the previous inequality where $u = v(p)$ to get

$$D_p^2 E(p, v(p)) = D_p x(p) + \partial_w X(p,1).x'(p)$$

Finally,

$$S(p) = D_p^2 E(p, v(p))$$

is the Slutsky matrix, and we can use homogeneity

$$D_p X.p + \partial_w X.w = X$$

in order to get

$$\partial_w X(p,1) = -D_p x.p$$

thus we can rewrite Slutsky symmetry as

$$S = D_p x (I - p.x')$$

which is the classical way to write Slutsky symmetry.

8.5 References and Notes

The housing problem was first posed by Shapley and Scarf in their 1974 paper [99]. They showed the non-emptiness of the core, as well as an algorithm to arrive to a core allocation: David Gale's method of "top-trading cycles". Afriat's theorem (Theorem 8.1) was initially proved in Afriat (1967) [1], in the case with linear budget sets and with observations strictly ordered in the direct revealed preference relation. The proof we have given here follows the one given by Fostel, Scarf and Todd (2004) [47]. The connection with a matching problem in Theorem 8.2 was made in Ekeland and Galichon (2010) [41].

The study of the assignment problem from a Walrasian equilibrium point of view, as well as the notion of efficient allocation in Section 8.2 is due to Hylland and Zeckhauser (1979) [67].

The Revealed Preference problem discussed in Section 8.3 was initially posed by Samuelson at the end of the 1930's (see [94] and [95]), who left it as an open problem. It was solved in 1950 by Houtakker [64] using a non-constructive argument, and then the solution in terms of inequalities was given by Afriat [1] using the theorem which bears his name. Afriat's solution was then improved by Diewert (1973) [39] who exploited the linear programming nature of the solution, and by Varian [109] who refined the axioms and provided a computationally efficient algorithms. The formal connection between the problem of efficiency in the housing problem and Afriat's theorem in consumer demand discussed in Section 8.3 is from Ekeland and Galichon (2010). A reference for Section 8.4 is Mas-Colell (1977) [76].

9 Stable marriages in a nutshell

Recall the marriage problem without transfers, which we introduced in Chapter 1.1, Example 1.3. In this problem it is assumed that both sides of the market form preferences, and instead of "individuals" and "houses" as in the previous chapter, we shall speak about "men" and "women". There are n men, indexed by $i = 1, ..., n$ and n women, indexed by $j = 1, ..., n$ to be matched. Let U_{ij} be the utility that woman j gives to man i, $U_{i\emptyset}$ utility of man i if single, and similarly let V_{ij} be the utility that man i gives to woman j, and j gets $V_{\emptyset j}$ if she remains single. As in the previous chapter, it will be assumed that preferences of both men and women are strict, that is $U_{ij} \neq U_{ij'}$ for $j \neq j'$, and $V_{ij} \neq V_{i'j}$ for $i \neq i'$.

Given a man i, we denote $\mu(i) = j$ if i is married to i, and $\mu(i) = \emptyset$ if i is single. Similarly, given a woman j, we denote $\mu(j) = i$ the husband of j if j is married to i, and $\mu(j) = \emptyset$ if j is single.

As the problem is purely ordinal, we denote preferences in the following way

$$\begin{aligned} j &\leq {}_i j' \text{ whenever } U_{ij} \leq U_{ij'} \\ i &\leq {}_j i' \text{ whenever } V_{ij} \leq V_{i'j}. \end{aligned}$$

Finally, for two women j and j', we denote

$$\begin{aligned} j \vee_i j' &= j \text{ if } j \geq_i j' \\ &= j' \text{ if } j <_i j' \end{aligned}$$

and

$$\begin{aligned} j \wedge_i j' &= j \text{ if } j \leq_i j' \\ &= j' \text{ if } j >_i j', \end{aligned}$$

and we define $\vee_j$ and $\wedge_j$ for a woman j in a similar fashion.

9.1 Stable marriages

An important concept then is the concept of stable outcomes, i.e. assignments such that (i) no individual would prefer singlehood to his/her current match, and (ii) no pair of man and woman who are not matched would would prefer each other to their current partners. So, assignment $j = \mu(i)$ is stable if and only if for each pair i, j, we get

$$\begin{aligned} &U_{i\emptyset} - U_{i\mu(i)} \leq 0,\ U_{i\emptyset} - U_{i\mu(i)} \leq 0 \\ &\min\left(U_{ij} - U_{i\mu(i)}, V_{ij} - V_{\mu(j)j}\right) \leq 0. \end{aligned}$$

In the latter condition, either this minimum is zero, in which case because of the assumption that preferences are strict, $j = \mu(i)$ and i and j are married; or this minimum is strictly negative in which case one member of the pair i, j strictly prefers his/her current partner to the other member of the pair.

Re-expressed in an ordinal way, stability rewrites as

$$\emptyset \leq_i \mu(i),$$

9.2 Structure of the stable set

Given two stable matchings μ and μ', one can define

$$\begin{aligned} (\mu \vee_M \mu')(i) &= \mu(i) \vee_i \mu'(i) \\ (\mu \vee_M \mu')(j) &= \mu(j) \wedge_j \mu'(j) \end{aligned}$$

and

$$\begin{aligned} (\mu \wedge_M \mu')(i) &= \mu(i) \wedge_i \mu'(i) \\ (\mu \wedge_M \mu')(j) &= \mu(j) \vee_j \mu'(j) \end{aligned}$$

and similarly, one can define $\mu \vee_W \mu' = \mu \wedge_M \mu'$ and $\mu \wedge_W \mu' = \mu \vee_M \mu'$.

One has the following result, which parallels Proposition 4.7 and which we state as a theorem:

Theorem 9.1. *If μ and μ' are stable matchings, then both $\mu \vee_M \mu'$ and $\mu \wedge_M \mu'$ are stable matchings. Hence the set of stable matchings in the marriage problem is a lattice.*

Proof. Let us show that $\mu \vee_M \mu'$ is a matching. Assume it is not: then a woman is chosen by two men under $\mu \vee_M \mu'$, say

$$j = \mu \vee_M \mu'(i_1) = \mu \vee_M \mu'(i_2).$$

Now necessarily and w.l.o.g., one can assume

$$\begin{aligned} j &= \mu(i_1) \geq_{i_2} \mu'(i_1) \\ j &= \mu'(i_2) \geq_{i_2} \mu(i_2) \end{aligned}$$

But j prefers either either i_1 or i_2. If j prefers i_1 then μ' is not stable; if j prefers i_2, then μ is not stable. One sees that μ is stable; if it were not, there would exist a pair (i,j) such that

$$\begin{aligned} j &> {}_i\mu \vee_M \mu'(i) \\ i &> {}_j\mu \vee_M \mu'(j). \end{aligned}$$

If $\mu \vee_M \mu'$ coincide with μ (or μ') on both i and j, we are done. Otherwise, assume w.l.o.g. that

$$\begin{aligned} j &> {}_i\mu \vee_M \mu'(i) = \mu(i) \\ i &> {}_j\mu \vee_M \mu'(j) = \mu'(j) \end{aligned}$$

but then $\mu(i) \geq_i \mu'(i)$, which would imply that μ' is not stable. ■

9.3 The Gale and Shapley deferred acceptance algorithm

0. Start with $v_j = -\infty$.

Loop:

1. Men propose: they announce their *best hopes*, namely the greatest utility they can achieve over the set of women who have not yet rejected them:

$$u_i^{k+1} = \max_j \left\{ U_{ij} : v_j^k \leq V_{ij} \right\} \tag{9.1}$$

2. Women dispose: they announce their *rejection threshold*, namely the utility under which they will reject men:

$$v_j^{k+1} = \max_i \left\{ V_{ij} : u_i^{k+1} = U_{ij} \right\}. \tag{9.2}$$

The algorithm stops when these values remain constant $v^k = v^{k+1}$.

The proof of the algorithm relies on the fact that for each j, v_j^k is nondecreasing with k. Indeed, over the loop, because of (9.1), one has $v_j^{k+1} = V_{ij}$

for some i such that $u_i^{k+1} = U_{ij}$ which means that i proposed j and j did not turn down i. Now, because of (9.1), one has necessarily that

$$v_j^k \leq V_{ij}$$

(this stems from strict preferences), thus $v_j^{k+1} \geq v_j^k$, and v_j^k increases with k. This in turn leads to the fact that u_j^k is nonincreasing with j, as the sets $\left\{i : v_j^k \leq V_{ij}\right\}$ are decreasing with respect to inclusion. One is then led to

$$\begin{aligned} u_i &= U_{i\mu(i)} \\ v_j &= V_{\mu^{-1}(j)j}. \end{aligned}$$

10 Labour economics

Labour Economics has many connections to matching. We shall discuss several directions.

10.1 The economics of superstars

Here are a few papers worth reading:

- Rosen, S. (1981). "The Economics of Superstars", *American Economic Review*, 71 (5), pp. 845–858.

Sherwin Rosen (1981) proposed the first model of "Superstardom". The model makes use of hedonic model in order to give a rationale the market wage or performers (such as artists or athletes), and to understand why their salary has evolved with decreasing marginal cost of broadcasting.

- Gabaix, X., and Landier, A. (2008). "Why Has CEO Pay Increased So Much?," *Quarterly Journal of Economics* 123 (1), pp. 49–100.

Gabaix and Landier apply the same hedonic model as Rosen in order to give a rationale for why the CEO pay has increased so much from 1980 to now.

- Alex Edmans, Xavier Gabaix and Augustin Landier "A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium," *Review of Financial Studies* 22, 2009, p. 4881-4917

Gabaix and Edmans deal with the problem of solving the incentive problem with matching (assuming risk neutrality here).

10.2 Frictions

- Shimer, R. and Smith, L. (2000). "Assortative matching and search". Econometrica 68(2), pp. 343–369.

Shimer and Smith deal with matching in the presence of search costs, with transferable utility. Their approach relies on the Nash bargaining solution. There are also related papers with non-transferable utility.

10.3 Empirical aspects

- Sattinger, M. (1993). "Assignment Models of the Distribution of Earnings," *Journal of Economic Literature* 31(2), pp. 831–880.

Sattinger has a very good review paper of many assignment models used in the labour literature.

10.4 Models of one-to-many matching

The reference here is:

- Kelso, A., and Crawford, V. (1982). "Job matching, coalition formation, and gross substitutes". *Econometrica* 50, pp. 1483–1504.

Assume there are m workers and n firms. Let $u^i(j, s_{ij})$ be the utility of worker i working for firm i with salary s_{ij}. The salaries are assumed to belong to some discrete set. Let C^j be the set of workers working for j. Let $y^j(C^j)$ be the revenue of firm j with set of workers C^j. Its profit is

$$\pi^j\left(C^j, s^j\right) = y^j\left(C^j\right) - \sum_{i \in C^j} s_{ij}$$

The firm's problem is to compute $M^j(s^j)$ solution of $\max_C \pi^j(C, s^j)$.

Gross substitute condition. For some fixed price vectors s and $\tilde{s}$, let $T^j(C^j) = \{i \in C^j : s_{ij} = \tilde{s}_{ij}\}$ so that $T^j(C^j)$, then $C^j \in M^j(s^j)$ and $\tilde{s}^j \geq s^j$ implies $\tilde{C}^j \in M^j(s^j)$ for all $\tilde{C}$ such that $T^j(C^j) \subset \tilde{C}^j$. Economic interpretation: when salaries are higher under tilde, the firm will demand at least more of the workers whose salaries have remained constant – call that set of workers C^j. That is, workers are gross substitute.

An allocation is a function $f : \{1, ..., m\} \to \{1, ..., n\}$. The corresponding salaries are $\left(s_{i,f(i)}\right)_{i=1}^{m}$.

There are two legitimate questions:

1. Are there stable solutions, and if so, how can one attain them?
2. What is the social optimum, and does one have duality results.

The theory to tackle these questions is matroid optimisation. See Vohra (2005) [114], ch. 8.

11 Matching with contracts

We now provide a lattice-theoretic formulation of the Kelso-Crawford model discussed in the previous section. We will follow:

- Hatfield, J. and Milgrom, P. (2005). "Matching with Contracts." *American Economic Review* 95, pp. 913–935.

In this section we shall recall our "labour market" example 1.8. Assume there are $\mathcal{I} = \{1, ..., I\}$ workers and $\mathcal{J} = \{1, ..., J\}$ firms. Let $u_i(j, w)$ be the utility of workers i working for firm j with salary w. Let C^j be the set of workers hired by firm j. Let $v_j(C^j)$ be the valuation by firm j

of set of workers C^j. Its net profit is

$$\pi_j\left(C^j, w_j\right) = v_j\left(C_j\right) - \sum_{i \in C_j} w_j^i.$$

Definition 11.1. *A* contract *is the specification of a worker, a firm, and a wage* $(i, j, w) \in \mathcal{I} \times \mathcal{J} \times \mathbb{R}$*.*

We are going to assume that there is a finite number of possible contracts available, which we shall denote $X \subset \mathcal{I} \times \mathcal{J} \times \mathbb{R}$, and that each workers can sign at most one contract with at most one firm, but can also sign no contract at all.

On the other hand, firms can only sign at most one contract with each worker, and they can also sign no contract at all.

We shall assume that both firms and workers have strict preferences $\geq_i$ and $\geq_j$ over the sets of contracts that include them, and that they are indifferent over contracts that do not include them.

For an arbitrary set of contracts Z, and given a worker i and a firm j,

- let $C_i(Z)$ be the contract that worker i prefers among Z, possibly the empty set (the empty set will be chosen if no contract in Z contains i);

- let $C_{\mathcal{I}}(Z) = \bigcup_{i \in \mathcal{I}} C_i(Z)$ be the set of contracts that, among Z, are preferred by at least one worker;

and similarly,

- let $R_{\mathcal{I}}(Z) = Z - C_{\mathcal{I}}(Z)$ the set of contracts that, among Z, have been rejected by all the doctors.
- let $C_j(Z)$ be the contract that firm j prefers among Z, possibly the empty set (the empty set will be chosen if no contract in Z contains j);
- let $C_{\mathcal{J}}(Z) = \bigcup_{j \in \mathcal{J}} C_j(Z)$
- let $R_{\mathcal{J}}(Z) = Z - C_{\mathcal{J}}(Z)$.

Definition 11.2. *An allocation is a set of contracts $Z \subset X$ where no worker appears in more than one contract.*

Definition 11.3. *An allocation Z is stable when the following two conditions are met:*

(i) No contract is unilaterally rejected, i.e.

$$Z = C_{\mathcal{I}}(Z) = C_{\mathcal{J}}(Z)$$

(ii) No alternative set of contracts Z' is preferred by a firm j and the set of workers it hires under this alternative, ie. there is no j and $Z' \neq C_j(Z)$ such that

$$Z' = C_j(Z \cup Z') \subset C_{\mathcal{I}}(Z \cup Z').$$

Let us comment briefly on this. One has $C_{\mathcal{I}}(Z) \subset Z$; if this holds strictly, it means that there is a contract in Z that is rejected by every worker, in which case we may drop it from Z: even if it is chosen by some firm, participation is voluntary, so the allocation would not be stable. Hence one should have $Z = C_{\mathcal{I}}(Z)$, and, for similar reasons, $Z = C_{\mathcal{J}}(Z)$.

As for understanding the second condition, assume there are contracts strictly preferred by a firm j when expanding the set of contracts by these new ones, that are such that these contracts are also preferred by workers involved.

Introduce the following operators over the set of contracts:

$$\begin{aligned} F_1(Z) &= X - R_{\mathcal{J}}(Z) \\ F_2(Z) &= X - R_{\mathcal{I}}(Z) \\ F(Z_{\mathcal{I}}, Z_{\mathcal{J}}) &= (F_1(Z_{\mathcal{J}}), F_2(F_1(Z_{\mathcal{J}}))) \end{aligned}$$

One has:

Theorem 11.1. *$(Z_{\mathcal{I}}, Z_{\mathcal{J}})$ is a a fixed point of F if and only if $Z_{\mathcal{I}} \cap Z_{\mathcal{J}}$ is a stable allocation.*

We now make an important assumption: the substitution condition.

Definition 11.4. *Contracts are substitutes for firm j if*

$$Z \subset Z' \implies R_j(Z) \subset R_j(Z'),$$

or in other words, when R_j is monotone with respect to set inclusion.

That is, if a contract is rejected by a firm j among a given set of proposed contracts, then it is still rejected when the set of contracts proposed is expanded.

11.1 A generalized Gale and Shapley algorithm.

Algorithm.

1. Initialize $Z_{\mathcal{I}}(0) = X$
   and $Z_{\mathcal{J}}(0) = \emptyset$
2. At step t, workers choose the contracts they prefer among those that have not been rejected already.
3. Firms then hold the best offers and reject others.
4. $Z_{\mathcal{I}}$ is updated by removing rejected offers, and $Z_{\mathcal{J}}$ is updated by adding newly made offers.

5. Stop when $Z_{\mathcal{I}} \cap Z_{\mathcal{J}} = C_{\mathcal{I}}(Z_{\mathcal{I}}) = C_{\mathcal{J}}(Z_{\mathcal{J}})$, and $Z_{\mathcal{I}} \cap Z_{\mathcal{J}}$ is a stable allocation.

One can show that at step t,

$$\begin{aligned} Z_{\mathcal{I}}(t) &= X - R_{\mathcal{J}}(Z_{\mathcal{J}}(t-1)) \\ Z_{\mathcal{J}}(t) &= X - R_{\mathcal{I}}(Z_{\mathcal{I}}(t)) \end{aligned}$$

that is,

$$\begin{aligned} Z_{\mathcal{I}}(t) &= F_1(Z_{\mathcal{J}}(t-1)) \\ Z_{\mathcal{J}}(t) &= F_2(F_1(Z_{\mathcal{J}}(t-1))) \end{aligned}$$

thus

$$(Z_{\mathcal{I}}(t), Z_{\mathcal{J}}(t)) = F(Z_{\mathcal{I}}(t-1), Z_{\mathcal{J}}(t-1)).$$

11.2 References and notes

The main reference for this section is Hatfield and Milgrom (2005).

Appendix

A Notations and conventions

Throughout the course, we adopt the following conventions.

For X a compact convex subset of $\mathbb{R}^d$, $\mathcal{U}(X)$ denotes the uniform probability distribution on X. For μ a probability distribution on $\mathbb{R}^d$, F_μ denotes the cumulative distribution function (cdf), and, if $d = 1$, $Q_\mu(t)$ denotes the associated quantile function. A function φ is *lower semi-continuous* (l.s.c.) whenever $\varphi(\lim x_n) \leq \lim \varphi(x_n)$ for each converging sequence x_n. "pdf" means *probability density function*. "iid" means *independent and identically distributed*. $|z|^2$ denotes the squared Euclidian norm $\sum_{i=1}^d z_i^2$, while $x.y$ denotes the scalar product of x and y $\sum_{i=1}^d x_i y_i$. The operator ∇ denotes the gradient; for a function $u : \mathbb{R}^d \to \mathbb{R}$, $\nabla u(x)$ is the vector $\left(\frac{\partial u(x)}{\partial x_i}\right)_i$, and u^* denotes its convex conjugate (Legendre-Fenchel) function. The notation $X \sim \mu$ means that random variable (or random vector) X has probability distribution μ. For $x_i \geq 0$, not all zero, the notation $p_i \propto x_i$ means

$$p_i = \frac{x_i}{\sum_j x_j}.$$

B Facts from Linear programming

Consider the generic linear programming (LP) problem called the *primal program*

$$\begin{aligned} &V_P = \max c.x \\ &s.t.\ Ax = b \\ &x \geq 0 \end{aligned} \tag{P}$$

This form is pretty general. Indeed, if x_j is unrestricted we can introduce two variables x_j^+ and x_j^-, $x_j^+, x_j^- \geq 0$, and set $x_j = x_j^+ - x_j^-$. We can transform constraints of the form $d_i.x \leq b_i$ by adding a slack variable $s_i \geq 0$ such that $d_i.x + s_i = b_i$; for constraints of the form $d_i.x \geq b_i$ we subtract the slack variable $s_i \geq 0$: $d_i.x - s_i = b_i$; $\min c.x$ is replaced by $-\max c.x$, etc.

The primal program is *infeasible* if the set of x that verify the constraints is empty; in that case, we take the convention of setting its value equal to $-\infty$.

Otherwise it is *feasible*; its value is either finite if the set $\{c.x : Ax = b,\ x \geq 0\}$ is bounded from above, or equal to $+\infty$ if not.

To each linear programming problem (P) one can associate another one (D) called its dual program

$$\begin{aligned} &V_D = \min b.y \\ &s.t.\ A'y \geq c \\ &y \text{ unrestricted} \end{aligned} \tag{D}$$

where A' is the transposed matrix of A.

Consistently with the convention taken above for the primal program, when the dual program is infeasible its value is set to $+\infty$.

One has in general the inequality

$$V_D \geq V_P,$$

which is called the *weak duality inequality*. In the case it holds strictly, one is said to have a *duality gap*. The following result asserts that there cannot be a duality gap unless both the primal and the dual programs are unfeasible.

Theorem B.1. *Exactly one of the following statements is true*
(i) Both the primal and dual programs are infeasible, or
(ii) Strong duality holds

$$V_D = V_P.$$

Proof. We only sketch a proof. Consider the primal problem

$$\begin{aligned} &V_P = \max c.x \\ &s.t.\ Ax = b \\ &x \geq 0 \end{aligned} \tag{P}$$

This problem can be written

$$V_P = \max_{x \geq 0} \min_y c.x + y'(b - Ax)$$

But by the minmax theorem,

$$\max_{x \geq 0} \min_{y} \leq \min_{y} \max_{x \geq 0}$$

thus

$$\begin{aligned} V_P &\leq \min_y \max_{x \geq 0} c.x + y' (b - Ax) \\ &\leq \min_y y'b + \max_{x \geq 0} (c - y'A) x = V_D. \end{aligned}$$

It turns out that when either the primal or the dual problem is feasible, then the weak duality inequality is actually an equality (a result that shall be admitted without a proof). ■

Finally, we should also mention *complementary slackness* relations: if x and y are solutions to the primal and dual problem, respectively, and if one constraint of the dual program is not saturated, say $\sum_j A_{ij} y_j > c_i$ for some i, then we have $x_i = 0$.

C Facts from Convex Analysis

C.1 Convex sets

The space here is $\mathbb{R}^d$.

Definition C.1 (Convex set). *A set C is convex if for all $x, y \in C$ and $t \in [0,1]$, $tx + (1-t) y \in C$.*

Theorem C.1 (Separating hyperplane theorem). *Let C be a closed convex set and $z \notin C$. Then there is a hyperplane strictly separating C and z.*

In what follows X is a set of $\mathbb{R}^d$.

Definition C.2 (Convex hull). *The convex hull of X, denoted conv (X) is the set of $x = \sum_{i=1}^{p} \alpha_i x_i$ for all $x_1, ..., x_p$ in X and $\alpha_i \geq 0$ such that $\sum_{i=1}^{p} \alpha_i = 1$.*

The convex hull of X is the smallest convex set containing X. When X is finite, $conv\,(X)$ is always closed. When it is infinite, it may be open (for instance, the open ball).

Hence we define the closed convex hull of X as the closure of $conv\,(X)$:

Definition C.3 (Closed convex hull). *The closed convex hull of X, denoted $cch\,(X)$, is the closure of the convex hull of X.*

In what follows C is a closed convex set of $\mathbb{R}^d$.

Definition C.4 (Extreme point). *An extreme point of C is a point $z \in C$ such that $z = tx + (1-t)\,y \in C$, $t \in (0,1)$ implies $x = y = z$.*

Definition C.5 (Boundary point). *A boundary point of C is a point x such that there exists a vector b such that*

$$x.b = \max_{y \in C} y.b.$$

The set of boundary points of X is also the topological boundary of C.

Extreme points are boundary points. The converse is false in general. However, the set of extreme points of C is the set of points $x \in C$ such that there exists a vector b such that x is the unique maximizer of $\max_{y \in C} y.b$.

Theorem C.2 (Krein-Milman). *A convex set is the convex hull of its extreme points.*

C.2 Convex functions

Cf. Hiriart-Urrut and Lemaréchal (2001) for a thorough exposition of the topic. See also Villani (2003), pp. 52–50, from which we have borrowed much of the subsequent exposition.

Definition C.6 (Convex function). *A function $\varphi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex if it is not identically $+\infty$, and if for all $x, y \in \mathbb{R}^d$ and $t \in [0,1]$, $\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y)$.*

The domain of φ is set set of x such that $\varphi(x) < +\infty$. It is a convex set.

Proposition C.1. *If $(y_1, c_1, ..., y_p, c_p)$ are vectors of $\mathbb{R}^d$, then*

$$\varphi(x) = \max_k (x.y_k + c_k)$$

is convex.

To a given set X, one can associate a function ι_X such that

$$\iota_X(x) = \begin{cases} 0 \text{ if } x \in X \\ +\infty \text{ otherwise} \end{cases} .$$

Note that ι_X is a convex function if and only if X is a convex set.

Differentiability. A convex function is continuous and locally Lipschitz on the interior of its domain, hence the set of points where it is not differentiable is of Lebesgue measure zero. Let $\nabla\varphi(x)$ be the gradient of φ at x; it is the row vector of derivatives $\left(\frac{\partial\varphi(x)}{\partial x_1}, ..., \frac{\partial\varphi(x)}{\partial x_d}\right)$.

For all points x where φ is differentiable, one has

$$\varphi(y) \geq \varphi(x) + \nabla\varphi(x).(y-x).$$

Subdifferentials. We define the subdifferential $\partial\varphi(x)$ as the set of $z \in \mathbb{R}^d$ such that $\forall y \in \mathbb{R}^d$, $\varphi(y) \geq \varphi(x) + z.(y-x)$. If φ is differentiable at x, then the subdifferential of φ at x is restricted to a single element which is the gradient at x: $\partial\varphi(x) = \{\nabla\varphi(x)\}$.

Conjugate functions.

Definition C.7 (Conjugate function). *For a function φ that is not identically $+\infty$, one defines its convex conjugate (or Legendre-Fenchel transform) function φ^* as*

$$\varphi^*(y) = \sup_{x \in \mathbb{R}^d} (x.y - \varphi(x)).$$

It is easy to see that φ^* is a convex function.

Let us take some examples. For $\varphi(x) = |x|^2/2$, one gets $\varphi^*(y) = |y|^2/2$.
For $\varphi(x) = \sum_i \lambda_i x_i^2/2$, $\lambda_i > 0$, one gets $\varphi^*(y) = \sum_i \lambda_i^{-1} y_i^2/2$.
For $\varphi(x) = \ln\left(\sum_i e^{x_i}\right)$, one gets

$$\varphi^*(y) = \begin{cases} \sum_i y_i \ln y_i \text{ for } y \geq 0,\ \sum_i y_i = 1 \\ +\infty \text{ otherwise} \end{cases}$$

Duality. We have in general

$$x.y \leq \varphi(x) + \varphi^*(y)$$

and equality holds if and only if $y \in \partial\varphi(x)$, or equivalently $x \in \partial\varphi(y)$. In particular, if φ is differentiable at x, if $y = \nabla\varphi(x)$ and if φ^* is differentiable at y, one gets

$$x = \nabla\varphi^*(y).$$

If φ is convex and lower semi-continuous (that is if it is such that $\varphi(\lim x_n) \leq \lim \varphi(x_n)$), then one gets that

$$\varphi^{**} = \varphi.$$

Support functions. The *support function* S_Y of Y is defined as

$$S_Y(x) = \sup_{y \in Y} x \cdot y$$

for any x in Y. It is a convex function, and it is homogeneous of degree one. Note that

$$S_Y = \iota_Y^*.$$

Moreover, $S_Y = S_{cch(Y)}$ where $cch(Y)$ is the closed convex hull of Y, and $\partial S_Y(0) = cch(Y)$.

D Stochastic orders

Let μ and ν be two probability measures on the real line. One says that μ is first-order stochastically dominated by ν if and only if

$$\exists \pi \in \mathcal{M}(\mu, \nu) : \mathbb{E}[1\{X > Y\}] = 0.$$

Proposition D.1. *μ is first-order stochastically dominated by ν if and only if*

$$\forall t \in [0,1]\,,\ Q_\mu(t) \leq Q_\nu(t)\,.$$

Proof. Consider the optimal transportation problem

$$\min_{\pi \in \mathcal{M}(\mu,\nu)} \mathbb{E}\left[c\left(Y - X\right)\right]$$

with cost

$$\begin{aligned} c(u) &= +\infty \text{ if } u < 0 \\ &= 0 \text{ otherwise.} \end{aligned}$$

By assumption, there is a zero-cost solution. But cost function c being submodular, it is known from Theorem 3.3 that the solution can be taken as

$$X = Q_\mu(U)\,,\ Y = Q_\nu(U)$$

for $U \sim \mathcal{U}([0,1])$. Therefore $Q_\mu(U) \leq Q_\nu(U)$, QED. ■

References for stochastic orders are Shaked and Shanthikumar (2007) [98], and Müller and Stoyan (2002) [80].

E Generalized convexity

Our exposition follows the three papers by Carlier [20], [21] and [22], to which we refer for proofs. We make the same assumptions as in Section 3.4, which we briefly recall. We set $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$. We assume:

Assumption A0. μ and ν are two Borel probability measures, and μ is absolutely continuous with respect to the Lebesgue measure, with a pdf denoted f_μ.

Assumption A1. Φ is twice continuously differentiable and for every compact set $\Omega \subset \mathcal{Y}$, there is $c_\Omega > 0$ such that

$$\sup_{y \in \mathcal{Y}} \left|\Phi(x_1, y) - \Phi(x_2, y)\right| \leq c_\Omega \left|x_1 - x_2\right|.$$

Assumption A2. Φ is such that:

$$\frac{\partial \Phi(x, y_1)}{\partial x} = \frac{\partial \Phi(x, y_2)}{\partial x} \Longrightarrow y_1 = y_2. \tag{E.1}$$

Throughout this section, we shall maintain Assumptions A0, A1 and A2 above.

We are ready to develop a framework to generalize convexity which boils down to classical convexity when $\Phi(x, y)$ is the scalar product $\Phi(x, y) = x \cdot y$ (which satisfies assumptions A1 and A2).

Definition E.1 (Φ-convexity). *Let $u : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ and $v : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be two functions. Then $u(x)$ is Φ-convex if and only if there exists a nonempty subset A of $\mathcal{Y} \times \mathbb{R}$ such that for all $x \in \mathcal{X}$,*

$$u(x) = \sup_{(y,t) \in A} (\Phi(x, y) + t).$$

Similarly, $v(y)$ is Φ-convex if and only if there exists a nonempty subset B of $\mathcal{X} \times \mathbb{R}$ such that for all $y \in \mathcal{Y}$,

$$v(y) = \sup_{(x,t) \in B} \Phi(x, y) + t.$$

Note that this indeed boils down to the standard definition of convexity when $\Phi(x, y) = x \cdot y$.

Definition E.2 (Φ-conjugate function). *Define the Φ-conjugate functions of u and v, respectively, by*

$$\begin{aligned} u^{\Phi}(y) &:= \sup_{x \in \mathcal{X}} \Phi(x, y) - u(x) \\ v_{\Phi}(x) &:= \sup_{y \in \mathcal{Y}} \Phi(x, y) - v(y) \end{aligned}$$

Again, this coincides with the standard definition of convex conjugate functions when $\Phi(x, y) = x \cdot y$.

Proposition E.1. *One has for all $x \in \mathcal{X}$,*

$$\left(u^{\Phi}\right)_{\Phi}(x) = \sup\{f(x) : f \leq u \text{ and } f \text{ is } \Phi\text{-convex}\}$$

and for all $y \in \mathcal{Y}$,

$$\left(v_{\Phi}\right)^{\Phi}(y) = \sup\{g(y) : g \leq v \text{ and } g \text{ is } \Phi\text{-convex}\}$$

thus u is Φ-convex if and only $u = \left(u^{\Phi}\right)_{\Phi}$, and similarly for v.

Hence, $\left(u^{\Phi}\right)_{\Phi}$ and $\left(v_{\Phi}\right)^{\Phi}$ are straightforward generalizations of the standard notion of convex enveloppe of a function.

Definition E.3 (Φ-subdifferential). *Define for $x \in \mathcal{X}$ the Φ-subdifferential of u at x*

$$\partial^{\Phi} u\left(x\right) := \left\{ \begin{array}{c} y \in \mathcal{Y} : \forall x' \in \mathcal{X}, \\ u\left(x'\right) - u\left(x\right) \geq \Phi\left(x', y\right) - \Phi\left(x, y\right) \end{array} \right\}.$$

One says that u is Φ-subdifferentiable at x if and only if $\partial^{\Phi} u\left(x\right) \neq \emptyset$. Alternatively, $\partial^{\Phi} u\left(x\right)$ is characterized by

$$\partial^{\Phi} u\left(x\right) = \left\{ y \in Y : u\left(x\right) + u^{\Phi}\left(y\right) = \Phi\left(x, y\right) \right\}.$$

Similarly, define for $y \in \mathcal{Y}$

$$\partial_{\Phi} v\left(y\right) := \left\{ \begin{array}{c} x \in \mathcal{X} : \forall y' \in \mathcal{Y}, \\ v\left(y'\right) - v\left(y\right) \geq \Phi\left(x, y'\right) - \Phi\left(x, y\right) \end{array} \right\}.$$

This definition generalizes the classical notion of subdifferential of a function. As in the standard convex case, we have:

Proposition E.2. *If u is Φ-convex and $x \in \mathcal{X}$, then $y \in \partial^{\Phi} u\left(x\right)$ if and only if $x \in \partial_{\Phi} u_{\Phi}\left(y\right)$.*

The link with the classical notion of gradient is given by the following proposition. This result is important, as it sheds some light on the concept of Φ-subdifferential.

Proposition E.3. *Let $u : \mathcal{X} \to \mathbb{R}$ be a function differentiable at $x \in \mathcal{X}$. Let $y \in \partial^{\Phi} u\left(x\right)$. Then*

$$\nabla u\left(x\right) = \frac{\partial \Phi\left(x, y\right)}{\partial x}. \tag{E.2}$$

It turns out that every locally bounded Φ-convex function u is almost-everywhere differentiable, and everywhere Φ-subdifferentiable, so (E.2) holds for almost every x and for every $y \in \partial^{\Phi} u\left(x\right)$.

Hence, we can characterize the solutions $u(x)$ and $v(y)$ of the dual of the optimal transportation problem (3.13) in terms of the apparatus developed above: u and v are Φ-convex functions, and verify

$$\begin{aligned} v &= u^{\Phi}, \\ u &= v_{\Phi}. \end{aligned}$$

Further, the optimal transportation plan $T(x)$ defined by (3.14) satisfies the inclusion

$$T(x) \in \partial^{\Phi} u(x)$$

for all x.

Applying this to multidimensional contracts, one has the following characterization of incentive-compatible constraints, with the notations of Section 3.5:

- A contract curve (T, w) is incentive-compatible if and only if $U_{T,w}$ is Φ-convex, and $T(x) \in \partial^{\Phi} U_{T,w}(x)$ for all $x \in \mathcal{X}$.
- A map $T : \mathcal{X} \to \mathcal{Y}$ is implementable of and only if there exists some Φ-convex function u such that $T(x) \in \partial^{\Phi} u(x)$ for all $x \in \mathcal{X}$.

F Computational methods

There are dozens, possibly hundreds of algorithms for the assignment/optimal transportation problem. Here we have chosen to present three of them, which seemed to us to bear the most interesting connections with Economics. The first one (discussed in Section F.1) can be interpreted in terms of a Walrasian tâtonnement process, and is specific to the optimal transportation problem seen in Section 2.1. The second one (Section F.2) is related to an auction mechanism and is specific to the discrete optimal assignment problem. As we shall see, there are very strong similarities between these two first methods. Finally, the last one (Section F.3) is connected to the optimal matching problem with unobservable heterogeneities discussed in Chapter 6, and very powerful and general, as it applies both to optimal assignment and optimal transportation problems.

F.1 The Walras algorithm

In this setting, P is a measure on $\mathcal{X}$ which is absolutely continuous with respect to the Lebesgue measure, while Q is a measure on $\mathcal{X}$ with a finite number of support points $\{y_1, ..., y_N\}$. Let q_j be the mass of y_j. For $x \in \mathcal{X}$, let $\Phi(x, y_j)$ the utility provided to individual situated at x by fountain y_j (this utility might depend negatively on the transportation cost).

Because of Theorem 3.1, one may look for the solution under the form of $v_1, ..., v_N$ which are the potentials associated to points $y_1, ..., y_N$, and

$$u(x; v) = \max_j (\Phi(x, y_j) - v_j)$$

and the right v's should satisfy

$$\inf_v \int_{\mathcal{X}} u(x; v)\, dP(x) + \sum_{j=1}^{N} v_j q_j.$$

As it was seen in Section 2.1 when discussing the fountain example, function

$$W(v) = \int_{\mathcal{X}} u(x; v)\, dP(x) + \sum_{j=1}^{N} v_j q_j$$

is concave, and its gradient is given by

$$\frac{\partial W(v)}{\partial v_j} = q_j - P(\{j = \operatorname{argmax}_k (\Phi(x, y_k) - v_k)\})$$

Hence Walras algorithm works as follows. Pick any constant $\varepsilon > 0$.
Start from any initial vector v^0.

1. At step t, update v by by
$$v^{t+1} = v^t - \varepsilon \frac{\partial W(v^t)}{\partial v_j}.$$
2. Repeat, until $|v^{t+1} - v^t|$ is close enough.

The algorithm converges to the Kantorovich potentials $\varphi(x) = \max_j (\Phi(x, y_j) - v_j)$ and v_j.

F.2 The auction algorithm

In the auction algorithm, we are going to adopt the optimal assignment of Section 2.2, and let the agents bid for machines in a generalized second-price Vickrey auction. We will need the following result.

Hall's marriage theorem. Let $E = \{1, ..., N\}$ be the set of labels of machines and agents, and for each agent $i \in E$ denote D_i the demand set of i, which is the set of machines that are acceptable by i (possibly empty). D is called the *demand correspondence*. The question is whether there exists an assignment $\sigma \in \mathfrak{S}_N$ *compatible* with this demand correspondence, i.e. such that under this assignment, each agent should be matched with a machine that is acceptable to him, that is, such that

$$\sigma(i) \in D_i \text{ for all } i.$$

To do this, we shall introduce the following notation. For a set of agents $S \subset E$, introduce

$$D_S = \bigcup_{i \in S} D_i$$

which is the set of machines that are acceptable by members of S. We are ready to state Hall's theorem.

Theorem F.1 (Hall). *There is an assignment compatible with demand correspondence if and only if for any set of agents, the set of machines acceptable by members of this set has no smaller cardinality than the set of agents. Equivalently, there exists $\sigma \in \mathfrak{S}_N$ such that*

$$\sigma(i) \in D_i \text{ for all } i$$

if and only if

$$|S| \leq |D_S|$$

for every $S \subset E$.

Proof. The direct implication is obvious. Indeed,

$$\begin{aligned} |S| = \sum_{i=1}^{N} 1\{i \in S\} &\leq \sum_{i=1}^{N} 1\{\sigma(i) \in D_S\} \\ &\leq |D_S|. \end{aligned}$$

The proof of the reverse implication is more involved, and we do not give it here. ■

We are now ready to explain the auction mechanism. Let Φ_{ij} be the production function of agent i and machine j. It will be assumed that these numbers are integers and nonnegative. At each round, given the current price v_j of machine j, agents state their list of acceptable machines at this price, i.e. for agent i,

$$D_i(v) = \text{argmax}_{j=1,\dots,N} (\Phi_{ij} - v_j)$$

The auction mechanism works as follows.

Start from zero initial price $v^0 = 0$.

1. At step t, given price t, agents state their list of acceptable machine $D_i(v^t)$. If there is an assignment σ compatible with demand correspondence $D(v^t)$, then stop, and return σ and v^t.

2. Otherwise, by Hall's theorem, there exist some *overdemanded set* of machines T which is such that the number of agents demanding only objects in this set is greater than the number of machines in this set, that is

$$|T| < \{i : D_i \subset T\}.$$

   Take a minimal such set S, i.e. such that no strict subset S' of S satisfies the same property.

3. For each j, define

$$v_j^t = v_j^t + 1\{j \in T\}$$

   which raises prices of machines by one in the overdemanded set, keeping them constant outside.

This algorithm returns an optimal assignment σ and a supporting price vector v_j. Set

$$u_i = \max_{j=1,\dots,N} (\Phi_{ij} - v_j),$$

so that (u, v) are solution to the dual optimal assignment problem.

Finally, let us give an example to illustrate how the algorithm works.

Example F.1. *Consider*

$$\Phi = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 2 & 3 \\ 2 & 3 & 2 \end{pmatrix}$$

In the first step, the price of all the machines is zero. Hence, the machines demanded by agents are the following:

$$\Phi = \begin{pmatrix} 2 & \underline{4} & 1 \\ \underline{3} & 2 & \underline{3} \\ 2 & \underline{3} & 2 \end{pmatrix}$$

and clearly, no assignment is compatible with this demand, as agents 1 *and* 3 *are in conflict for machine 2, so* $T = \{2\}$ *is an overdemanded set. Hence one shall raise the price of machine 2 by one, in which case the agents utilities and new demands look like*

$$\Phi_{ij} - p_i = \begin{pmatrix} 2 & \underline{3} & 1 \\ \underline{3} & 1 & \underline{3} \\ \underline{2} & \underline{2} & \underline{2} \end{pmatrix}$$

and therefore $p = (0, 1, 0)$*, and there are two matchings compatible with the demand at this price, which are both optimal assignments:*

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Thus $u_1 = 3$*,* $u_2 = 3$*, and* $u_3 = 2$*, and the value of the optimal assignment problem is* 9*.*

F.3 The IPFP

The IPFP (Iterated Projection Fitting Procedure), also sometimes called RAS algorithm, is based on the idea of alternating projections, discovered by von Neumann. The difference with the original von Neumann method is that instead of using Euclidian projections one uses the Kullback-Leibler divergence. Here is the basic idea of it.

This algorithm is a *simulated annealing* method for the optimal transport problem, meaning that instead of solving the actual transport problem, we are going to solve that problem with an additional entropic perturbation.

The advantage of this perturbation is that it will provide a lot of smoothness, allowing for more efficient computations. Of course, this comes at a cost, and the cost is that because of the perturbation, the solution one is led to is not quite the solution of the transportation problem; but, as the perturbation can be made as small as one wants, this shortcoming will be manageable. In fact, simulated annealing methods are known to this day to be among the most efficient methods for dealing with convex programming problems. Throughout this section it will be assumed that p and q are discrete probability measures with a finite number of support points. (The method actually extends to more general measures, but one then needs to worry about integrability conditions, which we do not want to do here.)

Simulated annealing. Let $\sigma > 0$ be a parameter which shall be called the *temperature*, and consider the variational problem

$$\max_{\pi \in \mathcal{M}(p,q)} \mathbb{E}_\pi \left[\Phi(X,Y) - \sigma \ln \frac{\pi(X,Y)}{p(X)\,q(Y)} \right] \tag{F.1}$$

which is the classical optimal transport problem if one sets $\sigma = 0$. It is shown in Section 6.5 that there there exist $u(x)$ and $v(y)$, which are unique up to a constant, such that the optimal π can be written as

$$\pi(x,y) = p(x)\,q(y) \exp\left(\frac{\Phi(x,y) - u(x) - v(y)}{\sigma} \right).$$

Our goal will be to compute $u(x)$ and $v(y)$, which amounts to determining $\pi(x,y)$. When σ is very close to zero, Problem (F.1) is very close to the optimal transport problem, and $\pi(x,y)$ will turn out to be close to the optimal transport plan.

Reformulation as a projection problem. It turns out that one can rewrite Problem (F.1) in an insightful way. Introduce

$$\pi_0(x,y) = K p(x)\,q(y) \exp\left(\frac{\Phi(x,y)}{\sigma} \right)$$

where K is an irrelevant normalizing constant which will be omitted. Hence,

$$\Phi(x,y) = \sigma \left(\ln \pi_0(x,y) - \ln\left(p(x)\,q(y)\right) \right)$$

and Problem (F.1) can be equivalently formulated as

$$\min_{\pi\in\mathcal{M}(p,q)} D_{KL}\left(\pi||\pi_0\right)$$

where

$$D_{KL}\left(\pi||\pi_0\right) = E_\pi\left[\ln\frac{\pi\left(X,Y\right)}{\pi_0\left(X,Y\right)}\right] \tag{F.2}$$

is the *Kullback-Leibler (KL) divergence* of π and π_0. Hence, we have reformulated our problem as the problem of projecting π_0 with respect to the KL divergence on the set $\mathcal{M}\left(p,q\right)$.

Alternating Projections. Let E be a finite dimensional vector space.

For any closed set K of E, let $\mathcal{P}_K$ be a projection operator over K; in the original method $\mathcal{P}_K$ was assumed to be the projection associated to the Euclidian distance, but the method extends to more general object than distances, namely *Bregman divergences*, which have similar properties as distances but do not satisfy the triangle inequality nor symmetry. This extension will be important for us as will need to work with the KL divergence which is not a Euclidian distance but is a Bregman divergence. Hence, $\mathcal{P}_K$ is defined by

$$\mathcal{P}_K\left(x\right) = \mathrm{argmin}_{z\in K}\,\delta\left(z,x\right).$$

Let C_1 and C_2 be two closed convex sets with a nonempty intersection $C = C_1\cap C_2$. One considers the problem of projecting a point x on $C_1\cap C_2$, that is

$$x^* = \mathcal{P}_C\left(x\right) = \mathrm{argmin}_{z\in C_1\cap C_2}\,\delta\left(z,x\right) \tag{F.3}$$

It will be assumed that it is "hard" to determine the projection of points on $C_1\cap C_2$, but that it is easy to project points either on C_1 or on C_2.

Then the Alternating Projections algorithm consists in:
Start from $x_0 = x$.

1. At step t, set

$$\begin{aligned} x_{2t+1} &= \mathcal{P}_{C_1}\left(x_{2t}\right) \\ x_{2t+2} &= \mathcal{P}_{C_2}\left(x_{2t+1}\right) \end{aligned}$$

2. Repeat until $|x_{2t+2} - x_{2t+1}|$ is small enough.

It turns out that the sequence (x_t) obtained from this algorithm converges to x^*.

The IPFP algorithm. The IPFP (Iterated Proportional Fitting Procedure) algorithm, a.k.a. RAS algorithm, is nothing else than the Alternating Projections algorithm applied to projection problem (F.2). Indeed, consider

$$\begin{aligned}\mathcal{M}_p &= \left\{\pi \geq 0 : \sum_y \pi(x,y) = p(x)\right\},\\ \mathcal{M}_q &= \left\{\pi \geq 0 : \sum_x \pi(x,y) = q(y)\right\}.\end{aligned}$$

Then one has

$$\mathcal{M}(p,q) = \mathcal{M}_p \cap \mathcal{M}_q$$

and the problem becomes

$$\min_{\pi \in \mathcal{M}_p \cap \mathcal{M}_q} D_{KL}(\pi || \pi_0)$$

which is exactly of the form (F.3). It remains to be verified that projecting either on $\mathcal{M}_p$ and $\mathcal{M}_q$ is indeed an easy task.

One has

$$\pi_{2t+1} = \arg\min \{KL(\pi || \pi_{2t}) : \pi \in \mathcal{M}_p\}.$$

thus by F.O.C., one gets

$$\log \pi_{2t+1}(x,y) = \log \pi_{2t}(x,y) + \phi(x)$$

thus

$$\pi_{2t+1}(x,y) = \pi_{2t}(x,y)\, e^{\phi(x)}$$

where ϕ should be adjusted so that $\pi_{2t+1} \in \mathcal{M}_p$, that is

$$\sum_y \pi_{2t+1}(x,y) = p(x),$$

thus

$$e^{\phi(x)} \sum_y \pi_{2t}(x,y) = p(x)$$

hence the first updating formula

$$\pi_{2t+1}(x,y) = \pi_{2t}(x,y) \frac{p(x)}{\sum_{y'} \pi_{2t}(x,y')} \tag{F.4}$$

which is indeed very easy! similarly, one projects π_{2t+1} on $\mathcal{M}_q$ by the second updating formula

$$\pi_{2t+2}(x,y) = \pi_{2t+1}(x,y) \frac{q(y)}{\sum_{x'} \pi_{2t+1}(x',y)}. \tag{F.5}$$

F.4 References and notes

The most complete reference for the descent algorithm is Aurenhammer (1991) [5]. See also references in Section 2.3.

The auction algorithm we described in Section F.2 is a variant of the celebrated Hungarian algorithm, or Kuhn and Tucker algorithm, see Dantzig (1963) [35]. Despite its name, Hall's theorem was actually first proved by Frobenius in 1917 [48]. Hall's paper appeared in the 1935 paper [60]. Our exposition follows Demange, Gale and Sotomayor (1986) [37]. We chose to present this algorithm as its economic interepretation is the simplest, yet this algorithm is not computationally efficient, due to the assumption that the production function should be integer. In fact, one of the most efficient assignment algorithms to this day is due to Bertsekas (1979) [12], which already relied on the idea of solving the assignment algorithm by an auction mechanism.

The IPFP algorithm was introduced by Deming and Stephan (1940) [38] in the context of contingency tables, and later generalized by Ireland and Kullback (1968) [68], Rüschendorf (1995) [93]. See also Bauschke and Combettes (2003) [6]. A similar entropic relaxation problem is studied independently, and with no apparent connection made, by Yuille and Kosowsky (1994) [118]. See the very complete literature survey in Pukelsheim, and Simeone (2009) [84]. The IPFP algorithm has been applied in Econometrics by Galichon and Salanié (2010) [53]. For a comparison of various computational method, see Bosc (2010) [15].

G Facts from Lattice theory

The results are stated in this section without proofs; see Section G.5 for references with proofs.

G.1 Definitions

Posets. A partially ordered set (*poset*) is a set endowed with a partial order $\leq$. $\leq$ is assumed to have the following three properties:

- Reflexive: $x \leq x$ for all x
- Antisymmetric: $x \leq y$ and $y \leq x$ imply $x = y$
- Transitive: $x \leq y$ and $y \leq z$ imply $x \leq z$.

We shall sometimes use the notation $x < y$ to mean $x \leq y$ and $x \neq y$.

Lattices. A *lattice* is a poset with the property that for any to elements x and y, we may define:

- the *join* of x and y, denoted $x \vee y$ which is the least upper bound, or supremum of $\{x, y\}$, and
- the *meet* of x and y, denoted $x \wedge y$ which is the greatest lower bound, or infimum of $\{x, y\}$.

Example G.1. *If $(X, \leq)$ and $(Y, \leq)$ are lattices, then one defines the product lattice $(X \times Y, \leq)$ as*

$$\begin{aligned}(x, y) \vee (x', y') &= (x \vee x', y \vee y') \\ (x, y) \wedge (x', y') &= (x \wedge x', y \wedge y').\end{aligned}$$

Example G.2. *As a result, $\mathbb{R}^d$ with the componentwise order is a lattice, where*

$$\begin{aligned}x \vee y &= (\max(x_1, y_1), ..., \max(x_d, y_d)), \\ x \wedge y &= (\min(x_1, y_1), ..., \min(x_d, y_d)).\end{aligned}$$

Finite lattices. A *finite lattice* $(X, \leq)$ is a lattice such that X is finite. Any subset $Z \subset X$ has a supremum $\bigvee_{x \in Z} x$, and an infimum $\bigwedge_{x \in Z} x$. Also, X has a maximum point $\bigvee_{x \in X} x$, and a minimum point $\bigwedge_{x \in X} x$.

Example G.3. *Let Ω be a finite set. Then 2^{Ω} the set of subsets of Ω is a finite lattice when endowed with partial order of the inclusion. The join and meet operations are the set union and intersection, respectively. Its maximum point is Ω and its minimum point is $\emptyset$.*

Complete lattices. A *complete lattice* is a lattice such that any subset $Z \subset X$ has both a supremum $\bigvee_{x \in Z} x$, and an infimum $\bigwedge_{x \in Z} x$ in X. In particular, a complete lattice has a maximum point and a minimum point.

Example G.4. $(\mathbb{R} \cup \{-\infty, +\infty\})^d$ *with the componentwise order is a complete lattice, where*

$$
\begin{aligned}
\bigvee_{x \in Z} x &= \left(\sup_{x \in Z} x_1, ..., \sup_{x \in Z} x_d \right) \\
\bigwedge_{x \in Z} x &= \left(\inf_{x \in Z} x_1, ..., \inf_{x \in Z} x_d \right).
\end{aligned}
$$

In particular, we have the following result:

Proposition G.1. *Any finite lattice is complete.*

G.2 Monotonicity

Definition G.1. *Given two posets X and Y, a function $f : X \to Y$ is called* monotone*, or* isotone*, if*

$$x \leq y \text{ implies } f(x) \leq f(y).$$

Tarski's fixed point theorem. One has the following result:

Theorem G.1 (Tarski). *Let $(X, \leq)$ be a complete lattice, and suppose $f : X \to X$ is monotone. Then:*

(i) The set of fixed points Z of f is a complete lattice with respect to $\leq$. As a result, f has a lowest fixed point $\underline{z}$ and a greatest fixed point $\overline{z}$, which are the minimum and maximum elements of Z respectively.

(ii) For all $x \in X$, $x \leq f(x)$ implies $x \leq \overline{z}$, while $f(x) \leq x$ implies $\underline{z} \leq x$.

(iii) If $\underline{x}$ and $\overline{x}$ are the minimum and maximum elements of X, respectively, then the sequence $f^k(\underline{x})$ is increasing and converges to $\underline{z}$, while the sequence $f^k(\overline{x})$ is decreasing and converges to $\overline{z}$.

G.3 Supermodularity

Definition G.2. *Let $(X, \leq)$ be a lattice, and $f : X \to \mathbb{R}$. Then f is called* supermodular *whenever for any x and y in X,*

$$f(x \vee y) + f(x \wedge y) \geq f(x) + f(y)$$

If $X = \mathbb{R}^d$ with the componentwise order, and if f is twice continously differentiable, then f is supermodular if and only if

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0 \text{ for all } i \neq j.$$

Example G.5. *Let $\alpha_i \geq 0$. Then $f(x) = \prod_{i=1}^{n} x_i^{\alpha_i}$ is supermodular on $\mathbb{R}_+^n$.*

Example G.6. *Let $a_i \geq 0$. Then $f(x) = \min_i (a_i x_i + b_i)$ is is supermodular on $\mathbb{R}_+^n$.*

Supermodularity is somehow an extension of the notion of convexity, as explained in the following proposition.

Proposition G.2. *Let Ω be a finite set, and endow 2^Ω with the inclusion order. Consider $f : 2^\Omega \to \mathbb{R}$. Then f is supermodular if and only if it satisties the* increasing marginal returns property*: for all $A \subset B \subset \Omega$ and $a \in A$,*

$$f(B) - f(B \backslash \{a\}) \geq f(A) - f(A \backslash \{a\})$$

Somehow similar to convexity, supermodularity is preserved by maximization; however it does not suffice that both f and g be supermodular so that $\max(f, g)$ to be supermodular. The following result holds.

Theorem G.2. *Let X and Y be two lattices, and $f : X \times Y \to \mathbb{R}$ supermodular. Then assuming*

$$h(y) = \max_{x \in X} f(x, y)$$

is well defined for all y, then h is supermodular on Y.

Proof. Let $y_1, y_2 \in Y$ and $x_1, x_2 \in X$ such that $h(y_i) = f(x_i, y_i)$ for $i = 1, 2$. One has

$$\begin{aligned} h(y_1) + h(y_2) &= f(x_1, y_1) + f(x_2, y_2) \backslash \\ &\leq f(x_1 \wedge x_2, y_1 \wedge y_2) + f(x_1 \vee x_2, y_1 \vee y_2) \\ &\leq h(y_1 \wedge y_2) + h(y_1 \vee y_2). \end{aligned}$$

■

G.4 Monotone Comparative Statics

We now turn to Monotone Comparative Statics. First we give Topkis's theorem. Before we do that, we should introduce the notion of strong set order.

Definition G.3. *Let X be a lattice, and A and B two subsets of X. One says that $A \leq B$ in the strong set order if and only if for every $a \in A$ and $b \in B$,*

$$b < a \implies b \in A \text{ and } a \in B.$$

Example G.7. *When $A = [a_1, a_2]$ and $B = [b_1, b_2]$, $A \leq B$ in the strong set order if and only if $a_1 \leq a_2$ and $b_1 \leq b_2$.*

Theorem G.3 (Topkis). *Let $f : X \times Y \to \mathbb{R}$, where $(X, \leq)$ is a lattice, and Y a poset. Suppose that f satisfies the increasing difference property: for all $x' > x$ and $y' > y$*

$$f(x', y') + f(x, y) \geq f(x, y') + f(x', y).$$

Let

$$x^*(y) = \text{argmax}_x f(x, y).$$

If $x^*(y)$ is nonempty for each y, then we have

$$y' > y \implies x^*(y') \geq x^*(y)$$

in the strong set order.

Proof. Take $x \in x^*(y)$ and $x' \in x^*(y')$. Then $f(x,y) \geq f(x',y)$ and $f(x',y') \geq f(x,y')$, hence by summation

$$f(x,y) + f(x',y') \geq f(x',y) + f(x,y').$$

Assume that $x' < x$. Then by increasing differences, we have the reverse inequality. Thus we have an equality, thus $f(x,y) = f(x',y)$ and $f(x',y') = f(x,y')$, hence $x \in x^*(y')$ and $x' \in x^*(y)$. ■

It turns out that this result simplifies with some further assumptions. When on top of the previous assumptions, $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, Y is a lattice and $x \to f(x,y)$ is supermodular for each fixed y, then the following result holds.

Theorem G.4 (Milgrom and Shannon). *Let $f : X \times Y \to \mathbb{R}$, where $(X, \leq)$ is a compact lattice, and Y a lattice with $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$. Suppose that for each $y \in Y$, $f(.,y)$ is continuous and supermodular, and that f satisfies the increasing difference property: for all $x' > x$ and $y' > y$,*

$$f(x',y') + f(x,y) \geq f(x,y') + f(x',y).$$

Let

$$x^*(y) = \operatorname{argmax}_x f(x,y).$$

Then:
(i) $x^*(y)$ is a non-empty compact lattice and has a greatest element $x(y)$.
(ii) $x(y) \geq x(y')$ for $y > y'$.
(iii) Assume $y \geq y'$. Then for any $x \in x(y)$ and $x' \in x(y')$, then $x \geq x'$.

Proof. See the proof in Vohra (2005) [114], Theorem 7.9. ■

G.5 References and notes

A good general reference for lattice theory is Birkhoff (1967) [14], and more specifically for supermodularity, Topkis (1998) [106]. For the proof of Tarski's theorem (Theorem G.1), see Tarski (1955) [102]. For the proof Proposition G.2, see Moulin (1988) [79]. For the proof of Topkis theorem (Theorem G.3) see [106], or the original paper Topkis (1978) [105]. The proof of the Milgrom and Shannon theorem (Theorem G.4 is in Vohra (2005) [114], Chapter 7.

H Facts from Extreme Value Theory

The Gumbel (or type-I Extreme Value) distribution has two parameters (μ, β): μ is the location parameter and β is the scale parameter. The c.d.f. of the Gumbel of parameters (μ, β) is $F(x) = \exp\left(-\exp\left(-\frac{x-\mu}{\beta}\right)\right)$. Its mean is $\mu + \beta\gamma$, where $\gamma = 0.577215665...$ is Euler's constant, and its variance is $\frac{\pi^2}{6}\beta^2$.

The $(0,1)$-Gumbel distribution has location and scale parameters $\mu = 0$ and $\beta = 1$. It has c.d.f. $F(x) = \exp(-\exp(-x))$. In that case, if $(\varepsilon_1, ..., \varepsilon_n)$ are independent draws from the $(0,1)$-Gumbel distribution, then the c.d.f. F of this vector is

$$F(x_1, ..., x_n) = \exp\left(-\sum_{i=1}^{n} \exp(-x_i)\right).$$

The fundamental property of the Gumbel distribution is that for $(\varepsilon_1, ..., \varepsilon_n)$ independent draws from the $(0,1)$-Gumbel distribution, and given a vector $(u_1, ..., u_n)$, one has

$$\mathbb{E}\left[\max_i \{u_i + \varepsilon_i\}\right] = \log\left(\sum_{i=1}^{n} e^{u_i}\right) + \gamma.$$

Therefore, when ε is a Gumbel with zero-mean and scale parameter 1 (that is $\mu = -\gamma$), one has

$$\mathbb{E}\left[\max_i \{u_i + \varepsilon_i\}\right] = \log\left(\sum_{i=1}^{n} e^{u_i}\right).$$

I Properties of the Dirac Delta Function

It is well known that only distributions that are absolutely continuous with respect to the Lebesgue measure have probability distribution functions (pdfs); in particular the constants are very degenerate distributions that have no pdf.

However the theory of *distributions*, a.k.a. *hyperfunctions*, sets to extend the notion of classical functions in order to associate objects that behave like pdfs to distributions that are not necessarily absolutely continuous. The "pdf" of the constant equal to zero is called the "Dirac delta function" and is denoted δ.

The following three basic facts, informally stated, are useful to keep in mind about the Dirac function. For details and formal statements, we refer to the original text [97] (in French) or [92].

1. For $k(x) \geq 0$ such that $\int k(x)\,dx = 1$, one has
$$\frac{1}{h} k\left(\frac{x}{h}\right) \to \delta(x).$$
2. One has
$$\int \delta(x-a) f(x)\,dx = f(a).$$
3. Assume that $\phi(x) = a$ implies $\phi'(x) \neq 0$. Then one has
$$\delta(\phi(x) - a) = \sum_{k=1}^{n} \frac{\delta(x - x_k)}{|\phi'(x_k)|},$$
where $(x_k)_{k=1..n}$ are the [single] roots of equation $\phi(X) = a$. In particular when ϕ is increasing, one gets
$$\delta(\phi(x) - a) = \frac{\delta\left(x - \phi^{-1}(a)\right)}{\phi'\left(\phi^{-1}(a)\right)}.$$

References

[1] Afriat, S. (1967). "The construction of a Utility Function from Demand Data". *International Economic Review* 8, pp. 67–77.

[2] Anderson, S., de Palma, A., and Thisse, J.-F. (1992). *Discrete Choice Theory of Product Differentiation*. MIT Press.

[3] Arnould, R., and Nichols, L. (1983) "Wage-Risk Premiums and Workers' Compensation: A Refinement of Estimates of Compensating Wage Differential". Journal of Political Economy 91, pp. 332–340.

[4] Atkinson, S., and Halvorsen, R. (1990). "The Valuation of Risks to Life: Evidence from the Market for Automobiles". *Review of Economics and Statistics* 72, pp. 133–136.

[5] Aurenhammer, F. (1991). “Voronoi diagrams – a survey of a fundamental geometric data structure”. *ACM Computing Surveys* 23 pp. 345–405.

[6] Bauschke, H. and Combettes, P.L. (2003). “Iterating Bregman retractions”. *SIAM Journal on Optimization* 13(4), pp. 1159–1173.

[7] Becker, G. (1973). “A theory of marriage, part I”. *Journal of Political Economy*, 81, pp. 813–846.

[8] Beckmann, M. (1952). “A continuous model of transportation”. *Econometrica* 20, pp. 643–660.

[9] Beckmann, M. , McGuire, C., Winsten, C. (1956). *Studies in Economics of Transportation*. Yale University Press.

[10] Berndt, E., Griliches, Z. and Rappaport, N., (1995). “Econometric Estimates of Price Indexes for Personal Computers in the 1990’s”. *Journal of Econometrics* 68, pp. 243–268.

[11] Berndt, E., and Rappaport, N. (2001). “Price and Quality of Desktop and Mobile Personal Computers: A Quarter-Century Historical Overview”. *American Economic Review* 91 (2), pp. 268–273.

[12] Bertsekas, D. (1979). “A distributed algorithm for the assignment problem”. Technical report, MIT.

[13] Birkhoff, G. (1946). “Three observations on linear algebra”. *Univ. Nac. Tucumán. Rev.* Ser. A, 5, pp. 147–151.

[14] Birkhoff, G. (1967). *Lattice Theory*, 3rd ed. Amer. Math. Soc.

[15] Bosc, D. (2010). “Numerical Computation of the Brenier Map”. Preprint.

[16] Boskin, M., Dulberger, E., Gordon, R., Griliches, Z. and Jorgensen, D. (1996). “Toward a More Accurate Measure of the Cost of Living”. U.S. Government Printing Office, Washington.

[17] Braess, D. (1968). “Uber ein paradoxon der verkehrsplanung”. *Unternehmensforschung* 12, pp. 258–268.

[18] Brenier, Y. (1987). "Déomposition polaire et réarrangement monotone des champs de vecteurs". *C.R. Acad. Sci. Paris*, Série I, 305, pp. 805–808.

[19] Brenier, Y. (1991). "Polar factorization and monotone rearrangement of vector-valued functions". *Communications on Pure and Applied Mathematics* 44, pp. 375–417.

[20] Carlier, G. (2001). "A general existence result for the principal-agent problem with adverse selection". *Journal of Mathematical Economics* 35, pp. 129–150.

[21] Carlier, G. (2002). "Nonparametric adverse selection problems". *Annals of Operations Research* 114, pp. 71–82.

[22] Carlier, G. (2003). "Duality and existence for a class of mass transportation problems and economic applications". *Advances In Mathematical Economics* 5, pp. 1–21.

[23] Chanel, O. (1995). "Is art market behaviour predictable?" European Economic Review Volume 39, Issues 3-4, April 1995, Pages 519-527.

[24] Chernozhukov, V., Fernandez-Val, I., and Galichon, A. (2010). "Improving Estimates of Monotone Functions By Rearrangement". *Biometrika* 96, pp. 559–575.

[25] Chernozhukov, V., Fernandez-Val, I., and Galichon, A. (2010). "Quantile and Probability Curves without Crossing". *Econometrica*, 78(3), pp. 1093–1125.

[26] Chiappori, P.-A., Iyigun, M., and Weiss, Y. (2009). "Investment in Schooling and the Marriage Market". *American Economic Review* 99, pp. 1689–1713.

[27] Chiappori, P.-A., McCann, R., and Nesheim, L. (2010). "Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness". *Economic Theory* 42(2), pp. 317–354.

[28] Chiappori, P.-A., Salanié, B., Weiss, Y. (2011). "Partner Choice and the Marital College Premium". Working paper.

[29] Chiappori, P.-A., and Oreffice, S. (2008). “Birth control and female empowerment. An equilibrium analysis”. *Journal of Political Economy*, 116, pp. 113–140, 2008.

[30] Chiappori, P.-A. and Weiss, Y. (2007). “Divorce, Remarriage and Child Support”. *Journal of Labor Economics* 25, pp. 37–74.

[31] Christensen, F. (2010). “The pill and partnerships: the impact of the birth control pill on cohabitation”. Journal of Population Economics 117(3): 997–1038.

[32] Choo, E., and Siow, A. (2006). “Who Marries Whom and Why”. *Journal of Political Economy* 114, pp. 175–201.

[33] Collinsa, A., Scorcub, A., and Zanolac, R. (2009). “Reconsidering hedonic art price indexes”. *Economics Letters* 104 (2), pp. 57–60.

[34] Combris, P., Lecocq, S., and Visser, M. (1997). “Estimation of a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter?” *The Economic Journal* 107, No. 441, pp. 390–402.

[35] Dantzing (1963). *Linear Programming and Extensions*. Princeton University Press.

[36] Dial R.B. (1971). “A probabilistic multipath traffic assignment model which obviates path enumeration”. *Transportation Research* 5, pp. 83–111.

[37] Demange, G., Gale, D., and Sotomayor, M., 1986. “Multi-Item Auctions”. *Journal of Political Economy* 94(4), pp. 863–872.

[38] Deming, W.E. and Stephan, F.F. (1940). “On a least squares adjustment of a sampled frequency table when the expected marginal totals are known”. *Annals of Mathematical Statistics* 11, pp. 427–444.

[39] Diewert, W. (1973), “Afriat and Revealed Preference Theory”. *Review of Economic Studies* 40, 419-425.

[40] Echenique, F. Lee, S.M., and Shum, M. (2010) “Aggregate Matchings”. Preprint.

[41] Ekeland, I., and Galichon, A. (2010). "Pareto indivisible allocations, revealed preferences and duality". Preprint.

[42] Ekeland, I. Galichon, A., and Henry, M. (2011). "Comonotonic Measures of Multivariate Risks". *Mathematical Finance*, forthcoming.

[43] Ekeland, I., Heckman, J., and Nesheim, L. (2004). "Identification and estimation of Hedonic Models". *Journal of Political Economy* 112 (S1), pp. S60–S109.

[44] Feenstra, R., and Levinsohn, J. (1995). "Estimating Markups and Market Conduct with Multidimensional Product Attributes". *Review of Economic Studies* 62, pp. 19–52.

[45] Ford, L. R.,. and Fulkerson, D. R. (1956). "Maximal flow through a network". *Canadian Journal of Mathematics* 8 (3), pp. 1778-1780.

[46] Forges, F. and Minelli, E. (2009). "Afriat's Theorem for General Budget Sets". *Journal of Economic Theory* 144, pp. 135–145.

[47] Fostel, A., H. Scarf and M. Todd (2004), "Two New Proofs of Afriat's Theorem". *Economic Theory* 24, pp. 211–219.

[48] Frobenius G. (1917). "Über zerlegbare Determinanten". *Sitzungsbericht der Königlich Preussischen Akademie der Wissenschaften* XVIII, pp. 274–277.

[49] Gale, D. (1960). *The Theory of Linear Economic Models*. McGraw-Hill.

[50] Gale, D. (1984). "Equilibrium in a discrete exchange economy with Money". *International Journal of Game Theory* 13(1), pp. 61–64.

[51] Gale, D., and L. Shapley (1962). "College Admissions and the Stability of Marriage". *American Mathematical Monthly* 69, pp. 9–14.

[52] Galichon, A., and Henry, M. (2011). "Inference in models with multiple equilibria" *Review of Economic Studies*, forthcoming.

[53] Galichon, A., and Salanié, B. (2010). "Matchings with Trade-Offs: Revealed preferences over Competing Characteristics". Preprint.

[54] Gandal, N. (1994). "Hedonic Price Indexes for Spreadsheets and an Empirical Test for Network Externalities". *RAND Journal of Economics* 25 (1), pp. 160–170.

[55] Goldin, C. and Katz, L. (2002). "The power of the pill: oral contraceptives and women's career and marriage decisions". *Journal of Political Economy* 110, pp. 730–770.

[56] Granot, D. (1984). "A note on the roommates problem and related revenue allocation problem". *Management Science* 30 (5), 633–643.

[57] Greenstone, M., and Gallagher, J. (2008). "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program". *Quarterly Journal of Economics* 123(3), pp. 951–1003.

[58] Griliches, Zvi, (ed.) (1971). *Price Indexes and Quality Change.* Cambridge, Mass.: Harvard Univ. Press.

[59] Gusfield, D. (1988). "The Structure of the Stable Roommate Problem: Efficient Representation and Enumeration of All Stable Assignments". *SIAM J. Comput.* 17, pp. 742–769.

[60] Hall, Philip (1935). "On Representatives of Subsets". *J. London Math. Soc.* 10 (1), pp. 26–30.

[61] Hardy, G., Littlewood, J. and Polya, G. (1952). *Inequalities.* Cambridge: Cambridge University Press.

[62] Hersch, J. 1998. "Compensating differentials for gender-specific job injury risks". *American Economic Review* 88, 598–627.

[63] Hiriart-Urrut, J.-B., and C. Lemaréchal (2001). *Fundamental of Convex Analysis.* Springer.

[64] Houtakker, H., (1950). "Revealed preferences and the utility function". *Economica* 17, 159-174.

[65] Lancaster, K. (1966). "A New Approach to Consumer Theory". *Journal of Political Economy* 74, pp. 132–157.

[66] Logan, J., Hoff, P., and Newton, M. (2008). “Two-Sided Estimation of Mate Preferences for Similarities in Age, Education, and Religion”. *Journal of the American Statistical Association* 103 (482), pp. 559–569.

[67] Hylland, A., and Zeckhauser, R. (1979). “The Efficient Allocation of Individuals to Positions”. *Journal of Political Economy* 87(2), pp. 293–314.

[68] Ireland, C.T. and Kullback, S. (1968). “Contingency tables with given marginals”. *Biometrika* 55, pp. 179–188.

[69] Irving, Robert W. (1985). “An efficient algorithm for the "stable roommates" problem”. *Journal of Algorithms* 6 (4), pp. 577–595.

[70] Kelso, A. S. and V. P. Crawford (1982). “Job matching, coalition formation, and gross substitutes.” *Econometrica* 50, pp. 1483–1504.

[71] Knott, M. and Smith, C.S. (19840. “On the optimal mapping of distributions”. *Journal of Optimization Theory and its Applications* 43, pp. 39–49.

[72] Koopmans, T. and Beckmann, M. (1957). “Assignment Problems and the Location of Economic Activities”. *Econometrica* 25, No. 1, pp. 53–76.

[73] Koutsoupias, E. and Papadimitriou, C. (1999). “Worst-case equilibria”. In *16th Annual Symposium on Theoretical Aspects of Computer Science*, pp. 404–413.

[74] Lambert, N. (2009). “Eliciting Information on the Distribution of Future Outcomes”. Preprint.

[75] Lorentz, G. (1953). “An Inequality for Rearrangements”. *American Mathematical Monthly* 60, pp. 176–179.

[76] Mas-Colell, A. (1977). “On revealed preference analysis”. *Review of Economic Studies* 45, pp. 121–131.

[77] McAfee, R.P., and McMillan, J., (1988). “Multidimensional Incentive Compatibility and Mechanism Design”. *Journal of Economic Theory* 46 (2), pp. 335–354.

[78] Monge, G. (1781). *Mémoire sur la théorie des déblais et des remblais.* In *Histoire de l'Académie Royale des Sciences de Paris*, pp. 666–704.

[79] Moulin, H., (1988). *Axioms of Cooperative Decision Making.* Econometric Society Monographs, Cambridge University Press.

[80] Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks.* Wiley.

[81] Murphy, K., and Topel, R. (2006). "The Value of Health and Longevity". *Journal of Political Economy*, 114, pp. 871–904.

[82] Nerlove, M. (1995). "Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers". *European Economic Review* 39 (9), pp. 1697–1716.

[83] Pakes, A. (2003). "A Reconsideration of Hedonic Price Indexes with an Application to PC's". *American Economic Review* 93 (5), pp. 1578–1596.

[84] Pukelsheim, F., Simeone, B. (2009). "On the Iterative Proportional Fitting Procedure: Structure of Accumulation Points and L1-Error Analysis". Universität Augsburg Preprint Nr. 05/2009.

[85] Quinzii, M (1984). "Core and Competitive Equilibria with Indivisibilities". *International Journal of Game Theory,* 13(1), pp. 41–60.

[86] Rachev, S. T., and Ruschendorf, L. (1990). "A characterization of random variables with minimum L2-distance". *Journal of Multivariate Analysis* 32 (1), pp. 48–54. [Corrigendum in 34 (1), p. 156, 1990].

[87] Rachev, S.T., and Ruschendorf, L. (1998). *Mass transportation problems* Vol. I: *Theory* and II: *Applications* Springer.

[88] Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition". *Journal of Political Economy* 82, No. 1, pp. 34–55.

[89] Roth, A.E., T. Sönmez, and Ünver, U. (2004). "Kidney exchange". *Quarterly Journal of Economics* 119, 457–488.

[90] Roth, A., and Sotomayor, M. (1990). *Two-Sided Matching. A study in Game-Theoretic Modeling and Analysis.* Cambridge University Press.

[91] Roughgarden, T., Tardos, É. (2002). "How bad is selfish routing?". *Journal of the ACM* 49 (2), pp. 236–259.

[92] Rudin, W. (1991). *Functional Analysis* (2nd ed.), McGraw-Hill.

[93] Rüschendorf, L. (1995). "Convergence of the iterative proportional fitting procedure". *Annals of Statistics* 23, pp. 1160–1174.

[94] Samuelson, P. (1938), "A Note on the Pure Theory of Consumer's Behavior". *Economica* NS 5, 61-71.

[95] Samuelson, P. (1948). "Consumption theory in terms of revealed preference". *Economica* 15, pp. 243–253.

[96] Schelling, T. (1987). "Value of life". In *The New Palgrave Dictionary of Economics* v. 4, pp. 793–796.

[97] Schwartz, L. (1950–1951). *Théorie des distributions* 1–2, Hermann.

[98] Shaked, M., and Shanthikumar, J. (2007). *Stochastic Orders.* Springer.

[99] Shapley, L. and H. Scarf (1974). "On cores and indivisibility". *Journal of Mathematical Economics* 1, pp. 23–37.

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

[101] Sheffi, Y. and Powell, W. (1982). "An Algorithm for the Traffic Assignment Problem with Random Link Costs". *Networks* 12, pp. 191–207.

[102] Tarski, A. (1955). "A Lattice-Theoretical Fixpoint Theorem and Its Applications". *Pacific J. Math.* 5, pp. 285–309.

[103] Thaler, R., and Rosen, S. (1976). "The Value of Saving a Life: Evidence from the Labor Market." In *Household Production and Consumption*, Nestor E. Terleckyj (ed.). New York: NBER pp. 265–298.

[104] Tinbergen, J. (1956). "On the Theory of Income Distribution". *Weltwirtschaftliches Archiv* 77, pp. 155–173.

[105] Topkis, D. (1978). "Minimizing a Submodular Function on a Lattice". *Operations Research* 26, No. 2, pp. 305–321.

[106] Donald M. Topkis (1998), *Supermodularity and Complementarity*, Princeton University Press.

[107] Trebbi, F., and Bombardini, M. (2010). "City Structure and Congestion Costs". Working paper.

[108] Triplett, J. (1969). "Automobiles and Hedonic Quality Measurement". *Journal of Political Economy* 77, No. 3, pp. 408–417.

[109] Varian, H. (1982), "The Nonparametric Approach to Demand Analysis". *Econometrica* 50, pp. 945–973.

[110] Villani, C. (2003). *Topics in Optimal Transportation.* Lecture Notes in Mathematics, AMS.

[111] Villani, C. (2009). *Optimal transport, old and new.* Grundlehren der Mathematischen Wissenschaften, 338, Springer-Verlag, Berlin.

[112] Viscusi, W.K. (2008). "The Value of Life". In *The New Palgrave Dictionary of Economics and the Law*, 2nd edition, Steven N. Durlauf & Lawrence E. Blume (eds.), 2008.

[113] Viscusi, W. and Hersch, J. (2001). "Cigarette smokers as job risk takers". *Review of Economics and Statistics* 83, pp. 269–280.

[114] Vohra, R. (2005). *Advanced Mathematical Economics.* Routledge.

[115] Vohra, R. (2008). *Paths, cycles and mechanism design.* Technical report, Northwestern University.

[116] von Neumann, J. (1953). "A Certain Zero-Sum Two-Person Game Equivalent to the Optimal Assignment Problem". In H.W. Kunh and A.W. Tucker (Eds.), *Contributions to the Theory of Games*, Vol. II, Princeton University Press, pp. 5–12.

[117] Wardrop, J.G. (1952). Some theoretical aspects of road traffic research". *Proc. Inst. Civ. Eng.* 2, pp. 325–378.

[118] Yuille, A., and Kosowsky, J. (1994). "The invisible hand algorithm: Solving the assignment problem with statistical physics". *Neural Networks* 7, Issue 3, pp. 477–490.