Matching dynamics and optima in a multi-agents labor market setting

Christelle Garrouste* and Cyrille Piatecki† and Yvan Stroppa‡§

June 24, 2015

Abstract

One of the greatest difficulties meet on the labor market is that it ensures that all the people who want to work can find a job. But it is not the only one. A second difficulty consists to find a social arrangement which approaches as much as possible the social optimum without forgetting that the optimum can be reached with the rejection of certain candidacies.

It seems impossible to solve a social optimum problem of the size of the french labor market because this one is furthermore constituted of 26 million job-seekers and that no computer is capable of managing the matching problem which, surprisingly is not of an excessive complexity — it is of order n^3 —, but which quickly collides with the computational capacity of machines, however powerful they may be.

It's the reason why it is essential to study restricted size markets, that is markets of less than 10000 job-seekers, to observe how practical solutions which can be organized as, for example, a geographical progressive balkanization either geographic or by skill level, leads to a departure from the Pareto optimum.

In this context, this paper suggests to compare the level of efficiency of a decentralized actually practicable solution with the optimum which would be if only it was possible to collect all necessary information to implement and compute it.

Introduction

Labor markets are notably inefficient. There is always a huge discrepancy between a more or less important reservoir for tenders unfilled jobs and a significant unemployment. For instance, for France the DARES publishes a quarterly indicator of the tension of the labor market — see the figure 1 — which demonstrates that since 1998, and certainly earlier, the 74 professional families[1] have'nt displayed any equilibrium between supply and offer.

*Corresponding author: Laboratory of Economics of Orléans, University of Orléans, CNRS, UMR7322, Orléans, France (e-mail: christelle.garrouste@univ-orleans.fr).

†Laboratory of Economics of Orléans, University of Orléans, CNRS, UMR7322, Orléans, France.

‡Laboratory of Economics of Orléans, University of Orléans, CNRS, UMR7322, Orléans, France.

§Acknowledgements: The authors would like to thank ...

[1] See DARES [2014]. The indicator is calculated as the ratio of the labor offers flow gathered by the french labor agency — *Pôle Emploi* — during the last three month to the demand entries in the same period, corrected for seasonal variations. If in level this indicator is difficult to interpret, on the contrary, in variation it gives the difference between those who find a new job and those who enter in unemployment as measured by *Pôle Emploi*.

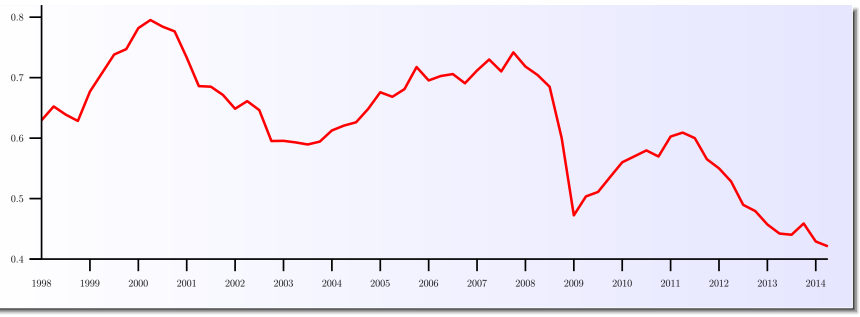

Figure 1: TENSIONS INDICATOR ON THE FRENCH LABOR MARKET — QUARTERLY DATA

As a high level of unemployment has always been considered as dangerous for the social *status quo*, after one century of intervention it is difficult to identify the short term reasons of the discrepancy. If in the long term, one knows that the two main sources come from the weakness of the activity and the impossibility to anticipate which type of formation will be asked by firms, in the short term, it is less clear what are the main sources of the inefficiency.

The reason is elementary: at least in Europa, there is no more countries purely managed through the market or through a *socialist* organization and even if one can see a true tendency to let the market regain the upper hand of the plan, there are good reasons to keep a *status quo*, even if one can not say right now what will be, in the end, the role between centralised allocation of jobs and the market.

Now, since in 1908, Barone — see Barone [1935] — has convincingly argued that a collectivist managed economy achieves, as the market, an efficient allocation of resources and Arrow and Debreu [1954], through the two famous welfare theorems, have shown under which conditions the market is, the Pareto criterion — voir Pareto [1971] — is the core description of what would be an intelligent centralized economy.

Unfortunately, even if, from the theoretical point of view, it is essential, one can doubt, without reopening the Hayekian debate — see von Hayek [1944] —, that a planer managing an organization constituted of agents made of flesh and bones could achieve even an ε-Pareto optimality[2].

Indeed, there is a lot of reasons not to achieve or simply to approach a Pareto optimality. One can begin by the Niskanen [1971] objection that any administration plays for itself — see also Buchanan et al. [1980] —, not necessarily for the collective well-being. So, it can renounce to strictly find a Pareto-optimum for an other objective less dedicated to the collective well-being.

Secondly, there is the *million of equations* debate, which run around the fact that since in a socialist economy there is not a *self-interest* mechanism there is no way to drive the exchange process to an equilibrium. The tentative answer was given as soon as Taylor [1929] that one must not mistaken the property of the goods with the process since, coming back to Barone, the socialist economy and the capitalist economy are

[2]An approximate Pareto optimality.

formally described by the same theoretical system. Two main themes were raised : the first one was about the nature of prices when Lange [1936-1937] made the distinction between prices as exchange ratios and prices as terms in which alternatives are offered. This distinction open the road to the linear programming approach to the optimum, first proposed in the soviet era by Kantorovich in 1939 but which was accessible in occident only after its translation in Kantorovich [1960]. That was to say that if the central planer collects all the information, the problem is *reduced* to solve for prices and quantities by a simple computer program. Unfortunately, in those times, computers were realities only in the brain of a small number of scientists who do not even counted on the fingers of the hand. Therefore, no one could really discuss the time required to complete the calculations required for planning. It's only with Hartmanis and Stearns [1965] that the idea to measure the time and space required to achieve a computation emerged. Later, Klee and Minty [1972] have shown that, in the worst cases, the Dantzig simplex algorithm is of exponential complexity, that is to say that when the dimension of the domain is increased by one, the time necessary to achieve the optimum is lined. The consequence of this is that, in some cases, it's pure utopia to try to achieve an optimum for a huge dimension problem[3], and this is amplified if one search to set aside the hypothesis of absolute divisibility of goods which can be solved only through combinatorial algorithms whose algorithmic complexity is higher[4].

Of course, since the heroic times of the invention of linear programming in two of the most gifted mathematicians of the XXth century — Kantorovitch and Dantzig —, some more efficient interior points algorithms have been conceived which reach solutions in polynomial time.

But, if one follows Shalizi [2012], even if there is now some good commercial linear programming package who could handle a 12 million variables on a desktop computer in a few minutes[5], the problem remains. This 12 million number is not randomly chosen : it appears in Nove [1971] as the number of identifiable distinct products in the world — it is certainly today an under-evaluation[6]. But, as shown brilliantly in 1959 by Debreu — see Debreu [1971] —, argued this doesn't take in account neither the place where goods are produced, neither the place where they are consumed — to forget this simple fact, is to forget the transport services which must also be taken into account — , the time of disposal — one knows that time has a value —, the decided incorporated qualities — think simply of planed obsolescence — and the uncertainties attached to the states

[3]As an example, one can look at the literature on the *Traveling Salesman Problem* which shows that the optimal travel for a limited number of cities can put down the best computers programs — see Applegate et al. [2007].

[4]In 1940, Atanasoff — see Atanasoff [1973] argued that it takes 8 hours to an expert to solve a system of 8 linear equations in 8 unknowns. It seems that the greatest system of equations ever solved by hand emerged on the occasion of the *Principal Triangulation of Great Britain*. Clark [1846] notes a sub-system of 77 equations. Nowadays, computer scientists think that one cannot expect to solve a system of more than 10^5 if the matrix is full and perhaps 10^8 is there are exploitable special structures of the problem, which is not the case here.

[5]For instance the free LP_SOLVE and GLPK or the commercial CPLEX, GUROBI or MOSEK which are freely distributed for academics.

[6]It is not sure that this number includes the services.

of Nature when it will be used — an the value of an umbrella under clear weather is not the same as the value of an umbrella under rain.

How many goods actually exist if one take into account those apparently safe little deviations from reality. Certainly a number so huge that no one could be comfortable with it and certainly no computers present or to come — even in taking in account Moore's law.

One can not hope that markets constitute a solution because they encounter the same difficulty. But, as stated by Axtell [2003], *Arrow-Debreu equilibria are sufficiently difficult to compute that the Walrasian picture of market behavior is simply not plausible*[7].

Of course those are arguments that are easily amenable to static computation, but one can also speak of the necessity to collect information through processes providing strong enough incentives to warrant it's honest transmission. Then one must construct an institution efficiently organized to transform the information in a computable form.

All this take time, and more than that, gathering information is a sequential rather than a synchronic process. So, must one wait to have collected all the information to take a decision and let the system dy in waiting. In short, computing a Pareto-efficient equilibrium is accessible to an omniscient demiurge, not to a human planer. And Pareto-efficiency is by now certainly the criterion by which every economic system is evaluated.

In what follows, we will study highly stylized model of a labor market. Those type of markets are characterized by a two side matching between labor demand and supply which a specific case of *Matching theory*, that is a mathematical account of behavior that is at the basis of many economics and computer science assignments. Stable matching problems are the most prominent models because they capture the aspect of rationality and distributed control that is inherent in most optimization problems (Hoefer [2011]). Examples of two-sided (or pairwise) matching models include the assignment of workers to jobs (Hoefer [2011], Arcaute and Vassilvitskii [2009] and Kelso and Crawford [1982]), organs to patients (Roth et al. [2005]), general buyers to sellers (Lauermann, 2013) and stable marriage and stable roommate problems (Akkaya et al. [2010], Goemans et al. [2006] and Mathieu [2008]).

Initially applied to markets, the classic static Walrasian pairwise matching models (Becker [1973]), assert the existence of a positive assortative matching (PAM) when productively complementary agents are assorted based upon their expected abilities. Recent developments show that the PAM still holds with random matching and search frictions (Atakan [2006] and Shimer and Smith [2000]).

While the literature on static matching has brought many important insights into non-cooperative games for decentralized markets, the need for the development of dynamic matching models is now widely recognized. This need is driven by the existence in the *real world* of recursive spells of games with evolving individual characteristics, as

[7] The problem comes from the *Uzawa equivalence theorem* which states the equivalence between the *general equilibrium* and the *Brouwer theorem* — see Uzawa [1962]. In fact, Hirsch et al. [1989] have shown that the lower bound for worst-case computation of Brouwer fixed points is exponential in the dimension of the problem, that is the size of the commodity space in the Arrow-Debreu version of general economic equilibrium.

demonstrated by Anderson and Smith [2010][8].

In this paper, we consider a dynamic model of stable matching in which players are rational agents seeking to be matched to their optimal partner, chosen within a social network, without any central coordination. The ordinary stable room-mates and marriage problems assert that each player knows the complete player set and can be matched arbitrarily. However, this assumption does not hold for all types of two-sided matchings. For instance, in the case of workers to jobs assignments, any worker can not be assigned arbitrarily to any job. There are restrictions in terms of knowledge and information that will hinder some players to be matched and that will enable certain players to match up faster.

To account for these matching barriers and frictions, our model builds upon the works by Arcaute and Vassilvitskii [2009] and Hoefer [2011]. They developed job-market games in which they considered locally stable matchings. Their models are special cases of stable marriage, where firms strive to hire workers. The main idea is that social links exist only among workers and that each firm can match to k workers, but each worker can only match to one firm.

While we start from similar assumptions Arcaute and Vassilvitskii [2009] and Hoefer [2011], we go one step ahead by generating a computerized optimization algorithm that simulates the spatial dynamics of the convergence process. This unique and innovative computer software enables us to test empirically our theoretical assumptions and to generate descriptive statics about the matching process.

This paper proceeds with a description of the model (section 1). Then we will discuss the interaction space (section 2), before the institutional arrangement preluding to the matches (section 3) and finally one will discuss the simulation results (section 4).

1 Description of the model

One start with two populations $\mathcal{H}$ and $\mathcal{I}$ respectively indexed by h and i — $\sharp\mathcal{H} = n$ and $\sharp\mathcal{I} = m$.

1.1 The matching problem

Within each population, the agents are randomly assigned an individual tag t_h for an $\mathcal{H}$-agent and t_i for an $\mathcal{I}$-agent. The tag, which characterises qualitatively each agent, is randomly determined to ensure that each agent within a population is associated to a unique tag: $t_h \in \{1, \ldots, n\}$ and $t_i \in \{1, \ldots, m\}$.

Then a reservation tag — t_{r_h} for the agents belonging to the $\mathcal{H}$ population and t_{r_h} for the agents belonging to the $\mathcal{I}$ population — are drown in the sets

[8] Anderson and Smith [2010] demonstrate that the PAMs are not validated when dynamics are introduced. They specifically test PAMs in the presence of evolving public Bayesian reputations and stochastic production.

$$\begin{cases} [t_h - \varepsilon_h, t_h + \mu_h] & \text{for the } \mathcal{H}\text{-agents} \\ [t_i - \varepsilon_i, t_i + \mu_i] & \text{for the } \mathcal{I}\text{-agents} \end{cases}$$

and according to which agents accept a match[9] if

$$\begin{cases} t_{r_h} \leq t_i & \text{for the } \mathcal{H}\text{-agent } h \text{ to accept the pairing with the agent } i \\ t_{r_i} \leq t_h & \text{for the } \mathcal{I}\text{-agent } I \text{ to accept the pairing with the agent } h \end{cases}$$

One assumes that the reservation tags are not too far of the own tag of the agents, because it seems natural to reduce the possibility of matching among agents that are too different from one another.

Respectively, if the tag of the other agent fall inside the acceptation set $\mathcal{A}$ of an agent the later one accepts the proposal. It is only when two agents accept a pairing that the match is possible — according to the classical voluntary exchange rule. On the contrary, if at least one tag does'nt enter in its acceptation set the match fails and both agents h, and respectively i, will have to re-initialize their search.

1.2 The demiurge arrangement[10]

Before going further, one must note that, as creator of the electronic world, the authors of the simulation are in a very specific situation : they are omniscient and can calculate the best possible matches.

Whenever a h agent accepts a match he has a surplus $S_h \in [0, \varepsilon_h + \mu_h]$, and whenever an agent i accepts a match he has a surplus $S_i \in [0, \varepsilon_i + \mu_i]$. So, a (h/i)-match generates a surplus $S_{hi} = S_h + S_i$. Define

$$\delta_{hi} = \begin{cases} 1 & \text{if } h \text{ and } i \text{ match} \\ 0 & \text{in the alternate case} \end{cases}$$

The planer then to find the δ_{hi} which satisfy the

$$\mathcal{O}_0 = \max_{\{\delta_{hi}, h \in \mathcal{H}, i \in \mathcal{I}\}} \sum_{h=1}^{H} \sum_{i=1}^{I} S_{hi}\delta_{hi}$$

under the $H + I$ constraints

$$\begin{array}{r} \text{for } \; i = 1, \ldots, I: \; \sum_{h=1}^{H} \delta_{hi} = 1 \\ \text{for } \; h = 1, \ldots, H: \; \sum_{i=1}^{I} \delta_{hi} = 1 \end{array}$$

$$\forall h \in \mathcal{H}, \forall i \in \mathcal{I}, \delta_{hi} \in \mathbb{Z}$$

[9] In this paper, one has retain an uniform drawing such as, for t_h as well as t_i, $t_{\{i\} \vee \{h\}} \rightsquigarrow \mathsf{U}[8, 11]$. For the acceptation intervals one has set $\varepsilon_h = \mu_h = \varepsilon_i = \mu_i = 0.1$.

[10] The *demiurge* is the deity responsible for the creation of the physical universe in diverse cosmogonies.

In a multi-agents simulation, as long as $H+I$ is not too much high[11], one can compute $\mathcal{O}_0$ that is the *Pareto-optimal overall surplus index* as the sum of all total surpluses associated with the matches. This number will be use as benchmark to evaluate the various human organization of the labor market.

As an example, the $\mathcal{T}_{hi}$ table exhibited in the figure 2 gives the aggregated surpluses of all the pair (h, i) up to the restriction that the matches be accepted by the two sides. So, we can evaluate the efficiency of the spontaneous order which emerges from the two-sided search. But, one must also notice that this procedure can be such that some agents stay unmatched. For instance, the $\mathcal{T}_{hi}$ table gives such a *Pareto-optimal matching set* which generates an overall surplus $\mathcal{O}_0$ of 462, but let the agent $7 \in \mathcal{H}$ and the agent $5 \in \mathcal{I}$ unmatched.

	1	2	3	4	5	6	7	8
1	-	97 $_1$	60	34	95	-	12	5
2	84	83	86 $_3$	-	23	74	-	53
3	4	31	64	79	84	-	-	97 $_2$
4	-	-	70	64	59	-	64 $_4$	-
5	-	87	41	44	9	41 $_7$	52	51
6	-	90	-	56 $_6$	2	-	25	52
7	18	2	-	28	-	28	51	-
8	62 $_5$	-	72	64	-	-	32	89

Figure 2: A $\mathcal{T}_{hi}$ TABLE

Let us begin by describing the operating conditions in which the simulations of the demiurge were realized. First, as regards the hardware, the software turned on a HP Z830™ 16 CORE 2 GHZ which is a rather powerful and quick machine. Of course, in an idealized world, the demiurge would have been able to have access to the chinese, right now, the most powerful computer in the world — 3 120 000 core. There are two types of LP-softwares : the non-commercial ones and the commercial ones. For pure doctrinal reasons, one has began by trying to resolve the some random generated situation of matching generated according to the description of the model given in the section 1, with free softwares mainly with LP-SOLVE, because despite it's public domain licence it is a celebrated very versatile soft, running in many platforms and callable nearly from any program. The main observation in running the diverse soft is that they are time consuming. This is not surprising since the matching algorithm is of order $\mathcal{O}(n^3)$ which is harder than a less demanding $\mathcal{O}(n^2)$ algorithm, but stay in the feasible set as

[11] One has not been able to compute the optimal matching allocation with LP_SOLVE for $\mathcal{H}+\mathcal{I} = 1000$ because the program has exhausted the memory in writing the transitory results.

long as n is not too huge[12]. At least, it is not of exponential complexity, but one can easily understand that it may force to kneel down the fastest computers. It is however necessary to note that the analysis of the complexity does not reveal the problem of the size of the memory to realize the calculations.

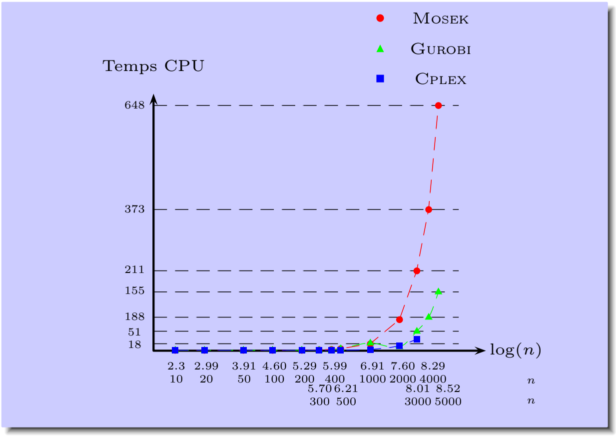

Figure 3: COMPUTATION TIME[13]

All programs, without exception, free or not, met a computational barrier as soon as $n = \sharp\mathcal{H} = \sharp\mathcal{I}$ was set to 7500 and this barrier is in fact the one encountered for the commercial softwares since, for LP-SOLVE, it was 4500. One knows that it is also possible to solve the problem using the *Hungarian algorithm* which is equivalent to the linear programming approach, but it the problem was of the same order. Nevertheless, it seems that CPLEX is the fastest of all the programs one has been able to use[14].

All this means that, without summoning the most powerful available computers in the world, it is likely not possible to calculate the optimal matching of a labor market as the French market. And in all cases, if the the Pareto solution to the French labor market was computable, one necessarily would ask for the European labor market.

Thus it seems necessary to turn towards a second-best solution. As it will be shown in a later section, there are multiple options. The simplest one which emerged was to design an greedy algorithm with began by randomly order both sets then to propose the highest generating surplus match to the first of the $\mathcal{H}$-agents and so on until no possible match remains.

In all the cases, it should be understood that in the actual world the computational dimension is one of the dimensions of the optimal matching problem and perhaps not the more striking. An employment agency has to ask the candidates to attend to a *rendez-vous* with one staff member. This one would have to listen to them, estimate

[12]For $n = 7000$, the order is of 343 $\overline{\overline{M}}$ operations.

[13]This graph may seems weird since expect some nearly linear curves. One must realize that there is no command in those softwares that gives the exact computation time and that one must rely on the human attention to the end of the computation which are heavily soiled by errors.

[14]But CPLEX fails to compute above $n = 3000$.

them and, in one way or an other, he will have to apply a procedure allowing him not only to attribute a tag to the the agent under inspection, but also a reservation tag.

Then, a new *rendez-vous* with the agent turns out to be necessary to ensure that the administrative attributed tags are identical with the those that the agent appropriate to himself[15]. All those operations are time consuming and the human computer, lost in the management of the necessary interaction between the administrative and the private agents, cannot manage it quickly. For all those reasons the Pareto optimal matching could not be really obtained in the actual world.

2 The problem with the interaction space

Actually, the agents doesn't wait in their ivory tour. Specifically, the $\mathcal{H}$-agents — the laborers — move to meet a potential match, when the $\mathcal{I}$-agents — the firms — wait for the first ones. That is not to say that firms have nothing to do to hire a new employee. They must make their offers public. On postulates that it is because of the Nature of the labor market that they do not move : they need not to use a head hunter because on an *a priori basis* there are enough laborers to fulfill all jobs. So this is an imperfect information situation for the both sides : the $\mathcal{H}$-agents do not know where are the offers, the $\mathcal{I}$-agents does not know where are the potential matches.

One must note that, whatever its metaphoric interpretation, the space has always been present in the search literature and more specifically, in the economic literature. Setting aside the land use models "*à la*" Thunen [1966] and the industrial localization models "*à la*" Hoteling [1929], because in those two approaches to economics modeling there was no search of any information — the former was trying to explain the soil rents, and the later, the best place to locate a shop, a plant or any economic activity —, it was necessarily present in the path breaking work of Koopman [1946], since Koopman was mainly interested in hunting submarines. His work has been extended by Guenin [1961] — see also Stone [1975] for a recension of all the work done on the subject.

When Stigler [1961] imported the technology in labor economics, it was also the case because he postulated that the agents were unaware of which are the posted prices but not of where they were located, and Stigler postulated implicitly that any location could be reached from any location for a constant cost. That is all information locations are situated on the same circle centered on the searching agent and this is not acted upon if the agents choose a reservation price strategy "*à la*" Lippman and McCall [1976] rather than a fixed sample strategy discussed by Stigler or even a mixed-one "*à la*" Morgan and Manning [1985] or Benhabib and Bull [1983].

With Sakoda [1949] — see also Sakoda [1971] — and Schelling [1969][16], the agents are searching for the best location possible on a chess or a checker-board according to

[15]This is a well known phenomenon revealed in the *qaly* — *Quality Adjusted Life Years* — assessment in health economics — see, among other Mehrez and Gafni [1991].

[16]Obviously, Sakoda [1949] has been unnoticed before Schelling [1969] and one can conjecture that it was a sentiment of frustration which decided the former to published on the subject, 22 years after his doctoral contribution.

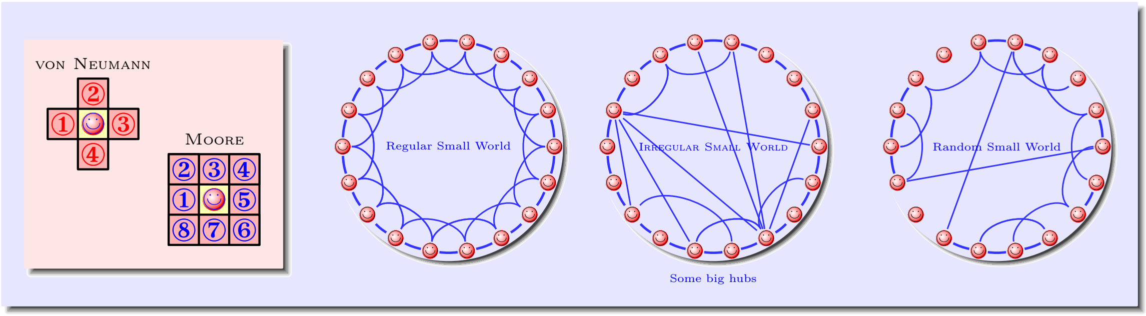

Figure 4: Neighborhood and Small Worlds

the proportion of non desired persons in their neighborhood. The space is then divided in cells and the agents can jump from a cell to an other. This create a discrete space from a continuous one.

In the same spirit, the spatial prisoners' dilemma, initially introduced by Nowak and May [1992], works through a cells organized space but, this time, with fixed agents who play only with their neighbours[17] and adapt their strategy on a local information basis. This simple structure has a natural drawback: to guaranty that some agents are not in a singular situation, which set the number of their neighbours to a lower number that the most part of the other agents — a side effect —, it was necessary to dot the space with a property which make him topologically equivalent to a torus. Later with the introduction of the *small world* network of Watts and Strogatz [1998], the space was structured as a network in which most nodes can be reached from every other by a small number of hops or steps.

In the contrary of all those specific spaces, the space here is a continuous one[18] and the agents can *a priori*[19] explore each atom of the space — *i.e.* each pixel. Even if, in this case, this is not truly necessary, we have imposed a basic torus structure which is no more than a topological transform of a rectangle of whichever dimension[20].

In this space, firms are thrown on fixed locations at random, as well as workers to the restriction that as soon as one simulation begins, workers move toward one objective according to the institutional arrangement which describe the market. In this paper, one will study the institutional arrangement described in the next section, under the voluntary exchange hypothesis.

The nature of the space is but one of the conditions preluding to the matches. A second important property of the space is the number of inhabitant it may welcome.

[17] The neighbourhood may be of the von Neumann style — *i.e.* : in crux to the first order neighbour cells — or more completely through the height direct neighbours.

[18] To be more precise, in the visual version of the simulations, the continuity of the space is limited by the dimension of the pixel. But we have also a batch version for huge simulations in which the space is truly continuous, since there is dimension imposed support.

[19] *A priori*, since one can imposes any type of restriction on the space.

[20] In the representation of this space, one can keep a grid to give an artificial structure to the space which help to localize every agent.

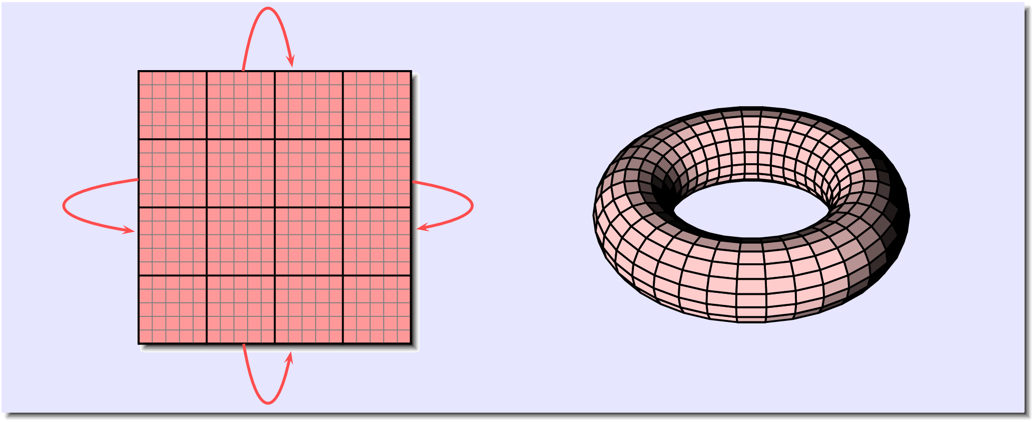

Figure 5: A Torus

Along all simulation, the space has been set to $800 \times 500 = 4000$ pixels. As an electronic agent occupy 50 pixels, one can have at most 8000 electronic agents of one or the other type in the world. In fact, as each firm offers only one job, workers can spatially recover the firms and one reach a maximum number of 16 000 agents. But, let us keep in mind the number of 8000, because this number corresponds to a saturated world where any movement is impossible. This leads to the agents' density described in figure 6. In any cases, this density shouldn't be compared to the population density in the real world.

One should wonder why, after what was described in the former sections, we have chosen a so small space. There are mainly two answers, first because it was the limit of our best disposable screen capacity — for this first exploratory study it was necessary to keep an eye on the agent's behavior —, secondly because due to the overall huge memory requirement needed to simulate each simulation a thousand times, it was impossible to expect to be able to break this limit.

The density here play a fundamental role : as the agents change their displacement direction only when they encounter an other agent — which is obviously an oversimplified way to exchange information, which should be modified in a later version of the model —, if they are too few, an agent who has been assign a bad initial direction, that is a direction which doesn't put a firm on his trajectory, will have a very small probability to ever change direction and, by consequence, to find a job.

In revenge, as the population density rises, the number of direction change increases, offering a better chance to an agent to find a job. But every flow has its ebb. When the population density continue to rise, the agents can eventually change so often direction until to lengthen their search toward the failure to find a job. Theoretically, to every space dimension must an optimal population density be associated.

So, the way agents move and interact in their displacements is an important part of the model which condition the simulation results. This part of the model will be discussed later on this paper.

N^b d'Agents	10	50	100	200	500	600	700	800	900	1000	1200	2000
Densité	0.00125	0.00625	0.0125	0.025	0.0625	0.075	0.0875	0.1	0.1125	0.125	0.15	0.25

Figure 6: Population density for a 500 × 800 grid

3 Institutional arrangement preluding to the matches

No match can be fit outside an institutional arrangement, since institutional arrangements define the rules of the game. The first institutional arrangement preluding to the matches is the *full-centralized* one. The many second institutional arrangements are the *partially-decentralized* ones. The third institutional arrangements correspond to the *markets*.

3.1 The full centralized arrangement

In the full centralized arrangement there is one and only one unemployment agency which collects the labor demands and offers. This agency has a specific location randomly drawn in the interaction space which conditions the arrival of the labor demands, as well as the labor supplies, according to a specific distance — euclidian, Manhattan... — calculated from the initial position of the firms and laborers.

According to the distances, two sets are sequentially constructed : the set of the demands and the set of the supplies. The constitution of those sets takes time, because of the arrival time of each demand as well as each supply but also because each agent's file must be studied by some public agent who, making seriously his job, uses a specific treatment to obtain the various $\mathcal{A}_h$ and $\mathcal{A}_i$ from the agents. For each set, if the information corresponding to two agents arrive at the same time they are ranked according to a random draw.Then two options are possible :

① the records are sequentially treated according to their place in the queue. If they matches, they get out it. If not, one of the two files is down graded of one rank and a new match is tried.

② in the second option, the public agent wait until a certain number of files of the two types are gathered. Then, he tries to find an optimal match among the two constituted files. If there is some agents that cannot be matched, they wait until the correct number of files is anew gathered.

Finally, the center can calculate an index of efficiency $\mathcal{O}_{fc}$ — in fact two indexes can be constructed $\mathcal{O}^s_{fc}$ for the case were the records are sequentially treated, and $\mathcal{O}^{op}_{fc}$ for the case where a partial optimum is attempted — which can be compared with the unaccessible $\mathcal{O}_D$ in constructing a distance[21] $\mathcal{D}_{fc} = |\mathcal{O}_{fc} - \mathcal{O}_D|$.

[21] In that case, there is no matter that the distance be euclidean or L_1.

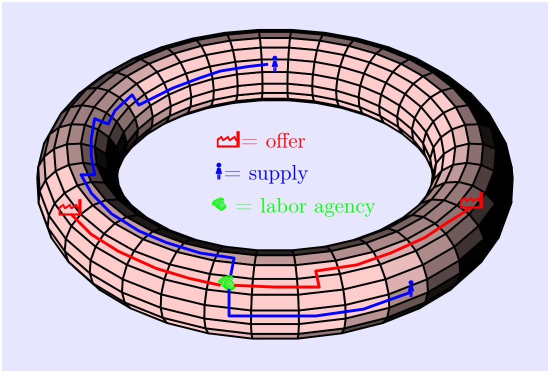

Figure 7: THE FULL CENTRALIZED ARRANGEMENT

3.2 The partial centralized arrangement of degree n

One can observe that in many cases, there are not one labor agency but many, dispersed on the interaction space. This is because the central planer want to diminish the travel cost to the labor agency.

Two arrangements are possible

① in the first arrangement, the agents are attracted by the labor agency which is the nearest of their location.

② in the second arrangement, each labor agency is dedicated to some tag intervals and agents must join a labor agency according to their $\mathcal{A}_h$ and $\mathcal{A}_i$.

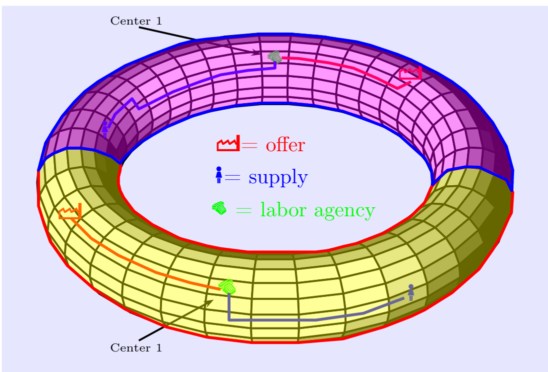

Figure 8: THE PARTIALLY CENTRALIZED ARRANGEMENT — $n = 2$

One more time, one can calculate an index of efficiency an index of efficiency ${}^{n}\mathcal{O}_{pc}$ — more exactly two indexes, the first, ${}^{n}_{l}\mathcal{O}_{pc}$ when the arrangement favors the location, the second, ${}^{n}_{t}\mathcal{O}_{pc}$ when it favors the tags —, then one can construct the distances to the unaccessible $\mathcal{O}_D$ — *i.e.* : ${}^{n}_{l}\mathcal{D}_{fc} = |{}^{n}_{l}\mathcal{O}_{fc} - \mathcal{O}_D|$ and ${}^{n}_{l}\mathcal{D}_{fc} = |{}^{n}_{t}\mathcal{O}_{fc} - \mathcal{O}_D|$.

3.3 The market arrangement

Of course, all this centrally and partially organized arrangements must be compared to the market one under highly imperfect information. In this simulation, the two sides of the market are differentially treated, in the sense that firms are fixed and only laborers move. Whichever be the interpretation of the interaction space, firms use to send messages to recruit not to move. In the contrary, the most part of the labor force uses to move to find a job. Of course, there are exception to those behavior where comportment are reversed but this is rare and more or less reserved to high level employment.

That is not to say that firms adopt a passive mode of acting. They send messages informing of their characteristics the laborers who, in the curse of their moves, enter in their reception field which is raising simulation step after simulation step. This gave to firm an important advantage on laborers : to send a message corresponds to travel in all direction providing that messages diffuse uniformly in the interaction space. But here, the radius of the circle inside which the information of a firm is diffused, is finite and common to all firms. When receiving the information, a searching laborer can decide to change direction to reach the firm which has send the message if it announces a potentially better match than the one he was moving for, if he has yet receive a message — see figure 9.

Obviously, this can be a bad decision since the job can be already given to another laborer when he will reached the firm but it reflect the hazard of the search for a job.

Under the initialization stage, and whichever be the context of the simulation, the directions chosen by the laborers are drown at random in a uniform distribution. This direction is modified, again by a random draw, after any interaction with an other agent.

One must notice that in this paper, laborers can move freely in any direction all at the same displacement speed without to be restrained by any barriers or to be slowed in certain regions of the interaction space, restrictions which could be used to model some sociological weight that must disadvantage a certain part of the laborer population[22].

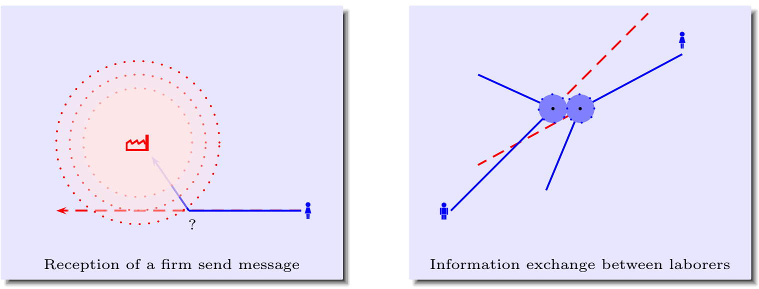

Figure 9: CHANGE OF DIRECTION

That is not to say that the modelling environment is not enough flexible to take in account those subtleties but here one concentrate the focus on the institutional organization of the matches. It's also the reason that their is no breaking of the realized

[22]But we must keep this for a later paper.

matches. One want to abound in the case of a highly stylized market to the debate between centralization and free market taking in account the highly asymmetric level of lack of information from the agents.

An other point, on which it is necessary to insist, come from the space's size. Here, one poses it as fixed as it seems to be in the actual world. In reality, this is not really the case since the space can change over time through public work or many other operations. Here, the space is more metaphoric than real. This has a direct consequence : when the agents' density increases, the number of collision grows accordingly. In the limit, for a very dense population of $\mathcal{H}$-agents some could be prevented to move.

For each simulation, one must also construct an efficiency index — $\mathcal{O}_m$ — , which in turn will be compared with the demiurge index — $\mathcal{D}_D$.

4 Results

To compute the results, one has to use a linear programming solver since one has shown earlier that the two sides matching problems are easily amenable to a linear program. The only problem was to make the link between the market simulation platform in such a way that each configuration can be saved an leads to the desired indicators.

First of all, as the characteristics of the population are randomly drawn — the two tags —, it was necessary to understand the simulation as some Monte-Carlo experiments: for each experiment, the scenario has been played 10000 times which gives a strong credibility to the results. One has successively worked with a couple $(\sharp\mathcal{H}, \sharp\mathcal{I})$ equal to 10×10, $100 \times 100, \ldots$, 1000×1000, 1200×1200 and finally 2000×2000. Of course, those dimensions are very small in comparison with the actual dimension of the true labor markets but, at least for those cases, all information is computable.

In all the cases, the protocol has began by the computation of the Pareto optimal matches, then one has used the greedy algorithm and one has run the moving agents platform.

In the case of all type of pairing, two results are of interest :

① the number of obtained pairs ;

② the level of the index of efficiency.

To evaluate the level of index efficiency, we will use the Hellinger statistic or more exactly the *Hellinger distance* between two discrete distributions p and q is define as :

$$\mathcal{H}(p, q) = \frac{1}{\sqrt{2}} \sqrt{\sum_i (\sqrt{p_i} - \sqrt{q_i})^2}$$

which is no more, to the $(2)^{-1/2}$ normalization constant[23], than the euclidian norm between the square root of the probability of realizations according to both distributions. Because of that it verifies the triangular inequality. One can show that :

[23] It ensures that $0 \leq \mathcal{H}(p, q) \leq 1$.

$$\mathcal{H}(p,q) = \sqrt{1 - \sum_i \sqrt{p_i q_i}}$$

When for all i, $p_i = q_i$, $\mathcal{H}(p,q) = 0$. It is equal to 1, whenever $p_i = 0$ and $q_i \neq 0$ and *vice versa*. So, smaller Hellinger distances give a better proximity between $\boldsymbol{p}$ and $\boldsymbol{q}$. In what follows

4.0.1 The 10×10 market

First of all, before anything else, it is not useless to call back that the optimum of Pareto doesn't look for the maximum number of matchings but the maximal value of the sum of surpluses.

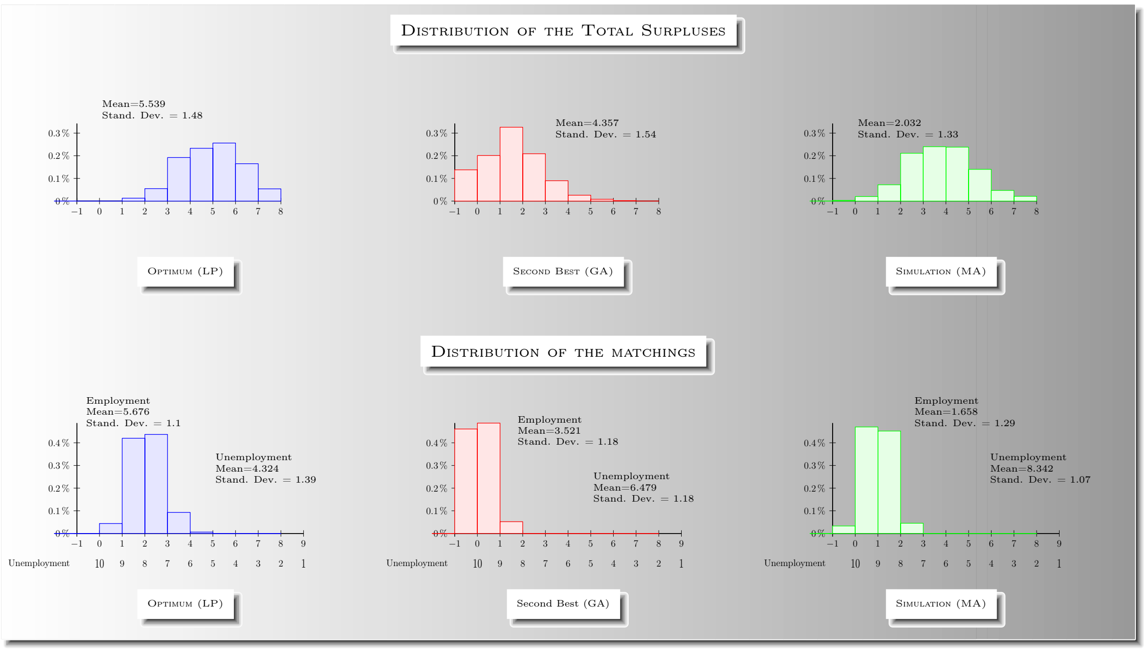

Figure 10: Results for a 10×10 market

In analysing table 10, that is the table obtained for 10×10 markets simulated 1000 times, it is obvious that for a little number of potential matches, the optimal matching linear programming algorithm — LP — largely dominate the two other ways to obtain an arrangement. Not only, in average it finds a greater number of matches, but it can find some perfect arrangements — all 10 jobs are finally allocated to all 10 workers —, when the greedy algorithm — GA — as well as the free search — MA — never find any one. That is to say, that when the two other ways to allocate the jobs are used, it is possible to stop on a non zero level of unemployment which should not be given the size of the market.

From the point of view of the total surpluses, not only it gives a greater mean than the two others, but one can also observed that when $\mathcal{H}$-agents search by themselves the distribution of the surplus is left biased when it should be, according to the Pareto-optimal matchings, right biased.

So, two important points should be noticed that :

① the greedy algorithm and the search are both highly inefficient even if the greedy algorithm — which is less demanding from the computational point of view — seems to be less inefficient.

② One can observe a very unrealistic level of unemployment — more than 40%. It seems to originate in the excessive requirement of the agents of which the mode of determination of the reservation tags creates an over-tension on the pairing possibilities. More specifically, this result seems to ensue from the fact that a lot of $\mathcal{H}$-agents will agree to mate only with $\mathcal{I}$-agents characterized by a greater tag that their own and *vice versa*.

One has two strategies to overcome this unwished effect:

① to change the value of ε_h, μ_h, ε_i and μ_i, mainly to distinguish them from each other and certainly to consider less demanding populations which always agree to mate with agents of tag lower than theirs — *i.e.* : $\varepsilon_h \neq 0$, $\varepsilon_i \neq 0$ and $\mu_h = \mu_i = 0$.

② to change the distribution — until now, uniform — from which those parameters are drawn. One can for instance choose an exponential distribution for the $\mathcal{H}$-agents and a gaussian one for the $\mathcal{H}$-agents.

But one must understand that, that in both cases, the choices are highly arbitrary. That is not to say that it is uninteresting to try to do it. Now one can look at what happen when the dimension of the market rises.

4.0.2 From the 100×100 to the 2000×2000 market

In the annexe, one has collected some histograms showing the evolution of the matching problem according to the population size. What could be said from the histograms collected in the annexe is that from the point of view of the surpluses, even for small dimension the greedy algorithm seems to perform nicely. The most important part of the results can be gathered through the Hellinger distance in the figure 11.

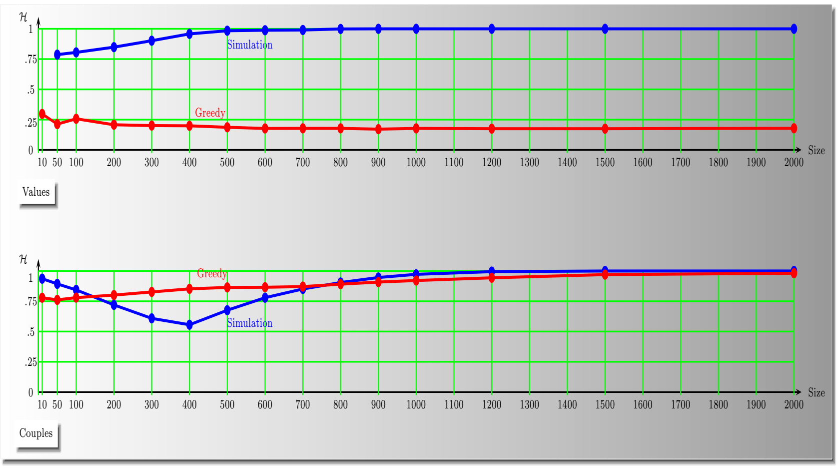

Figure 11: Hellinger distances to Optimum — for a 10 classes histogram

As the optimal matching is calculated on the surplus basis, the main results are linked to the value-Hellinger distance — upper graph in the figure 11. For what concerns the greedy algorithm, obviously, it is never able to implement on a 10 classes histogram — the value for which figure 11 has been implemented — the matching algorithm. But, it performs not so badly since in average the associated Hellinger distance is of 0.11 for a standard deviation of 0.073. That is to say that the greedy algorithm is biased, which is not an unexpected news, but the bias is more or less of constant dimension on the population size, which gives a measure of the performance between the first best and a second best optimum one could expect with it. This results is of interest: according to the complexity and the memory size needed to compute the results, it is clear that reaching the optimum is unattainable with any computer program if the population size is big enough. So, less complex and memory requiring, the greedy algorithm, even if one knows that it will also encounter some limits can be implemented to approach the optimum when it is not possible to compute it. More than that, for a given population size, the Hellinger distance often fluctuates very lightly with the number of classes used to compute the histograms and is nearly invariant for small size. But, when one exceed a size of 500×500, it nearly uniformly fall under 0.1. This shows a nearly adequation on the distribution of the optimal matches and the distribution of the greedy matches[24].

One can also observe that the simulation, that is letting agents searching and finding by themselves their jobs is a very bad solution, since its divergence from the optimum distribution always rises until reaching a complete divergence from it when the population size is 500. According to the fact that, in this model, agents don't really share any information local nether local nor global, it is not so surprising.

Looking at the second aspect of the computations, one can observe that the Hellinger distance for the simulated matches, with is initially falling to a minimum, finish to rise to the maximal divergence. This could be attributed to the invariance size of the space

[24] One should notice that for all the other calculated distributions the Hellinger distances are stable and of high level.

which finish, size by size, by being overpopulated. This leads to two much direction changes which, from a chance to increase the opportunities to meet companies, eventually represent an obstacle to the job search.

5 From one to many labor markets

As underlined earlier, sooner or later one arrives to the limit of the optimum computation. In that case one solution is to break the problem in many subproblems. It is exactly what, in France, Pôle Emploi does when it open many branches. For instance, one can look at what happen when 4 branches of 500×500 instead of one of 2000×2000, are opened.

Formally this should create some problems since in dividing the space, one create artificial boundaries which are restructured as tores but which are not really. In that case, one has been astonished to see that the comparison between the 2000×2000 and $4\ 500 \times 500$ was extremely good with a Hellinger distance falling in a uniform way from 0.089442719 for a 10 classes histogram to 0 for a 40 classes histogram.

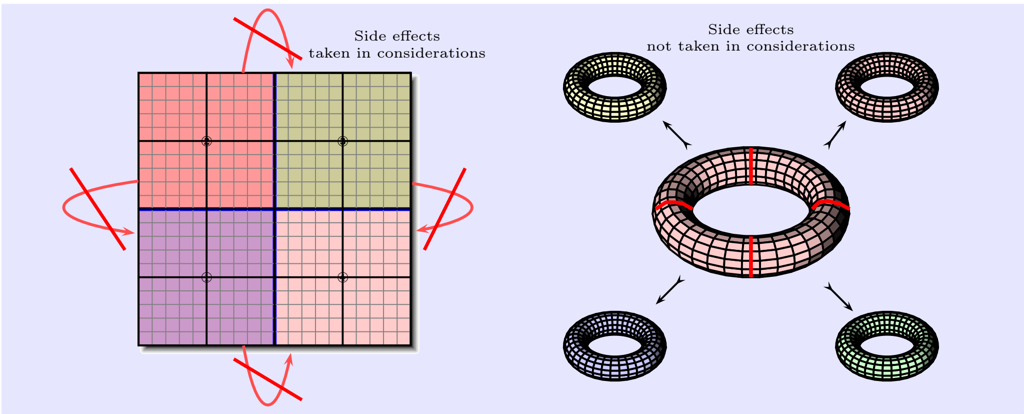

Figure 12: A Division of a Torus

6 Conclusion

The main result of this study is that a population with a so much degree of requirement but managing so badly the information which it could have, should never be left to itself. In that case, the free market could not implement the Pareto optimum, contrary to the fact that lets believe the walrasian dominant approach.

A central planer should intervene and, by procuration, should manage the matching problem. Unfortunately, we have shown from the theoretical view point as well as the empirical view point that the program necessary to accomplish this task is, in all likelihood, non computable because of the size of the problem. Nevertheless, one can conjecture that, in breaking the main problem in subproblem, one can reach a second best optimum which is not so far from the true one. This is the solution to which one resorts in the real world. In this world, it seems not to be effective. In fact, the main difference between the real world and the simulated one is that contrary to the former, the later offer potentially one job to each demand. Obviously the matching of agents to tasks or job is not a simple task whenever their number is huge and one can be. To finish, one are conscious that the agents must be gifted with more capability in the treatment of the information[25]. This will be done in an other paper along with the many sociological barriers to reach employment.

References

K. Akkaya, I. Guneydas, and A. Bicak. Autonomous actor positioning in wireless sensor and actor networks using stable-matching. *International Journal of Parallel, Emergent and Distributional Systems*, 2010.

A. Anderson and L. Smith. Dynamic matching and evolving reputations. *Review of Economic Studies*, 77(1):3–29, 2010.

D. Applegate, R. E. Bixby, V. Chvatal, and W. Cook. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press, 2007.

E. Arcaute and S. Vassilvitskii. Social networks and stable matchings in the job market. In *WINE 2009. LNCS*, 2009.

K. J. Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22:265–90, 1954.

A. Atakan. Assortative matching with explicit search costs. *Econometrica*, 74(3):667–680, 2006.

J. V. Atanasoff. Computing machine for the solution of large systems of linear algebraic equations. In B. Randell, editor, *The Origins of Digital Computers: Selected Papers*, pages 305–325. Springer, 1973.

R. Axtell. The complexity of exchange. Technical report, Center on Social and Economic Dynamics The Brookings Institution, february 2003.

E. Barone. The ministry of production in the collectivist state. In F. A. von Hayek, editor, *Collectivist Economic Planning*, pages 245–290. Routledge & Kegan Paul, 1935.

G. Becker. A theory of marriage: Part i. *The Journal of Political Economy*, 81(4): 813–846, july 1973.

[25] All the simulations presented in this paper have been done through the IMAS — Interacting Moving Agents Simulator — platform, developed in the LEO — the economics laboratory of the University of Orléans —, which can be found in the URL www.execandshare.org/SiteCompanion/Site401.

J. Benhabib and C. Bull. Job search: The choice of intensity. *The Journal of Political Economy*, 91:747–764, 1983.

J. Buchanan, R. Tollison, and G. Tullock, editors. *Toward a Theory of the Rent-Seeking Society*, 1980. A&M University Press.

D. W. Clark. *Elements of Algebra.* Harper & Brothers, 1846.

DARES. Indicateurs. Technical Report 072, Pôle Emploi, septembre 2014.

G. Debreu. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium.* Yale University Press, 1971.

M. Goemans, L. Li, V. Mirrokni, and M. Thottan. Market sharing games applied to content distribution in ad-hoc networks. *IEEE J. Sel. Area Comm.*, 24(5):1020–1033, 2006.

J. Guenin. Optimal allocation of effort: An extension of the Koopman basic theory. *Operation Research*, 9(1):1–7, 1961.

J. Hartmanis and R. E. Stearns. On the computational complexity of algorithms. *Transactions of the American Mathematical Society*, 117:285–306, 1965.

M. D. Hirsch, C. H. Papadimitriou, and A. Vavasis. Exponential lower bounds for finding Brouwer fixed points. *Journal of Complexity*, 5:379–416, 1989.

M. Hoefer. Local matching dynamics in social networks. In *ICALP 2011, Part II, LNCS*, 2011.

H. Hoteling. Stability in competition. 39:41–57, 1929.

L. Kantorovich. The mathematical method of production and planning. *Management Science*, 6(4):363–422, 1960.

A. Kelso and V. Crawford. Job matchings, coalition formation and gross substitute. *Econometrica*, 50:1483–1504, 1982.

V. Klee and G. J. Minty. How good is the simplex algorithm? In *Inequalities III (Proceedings of the Third Symposium on Inequalities held at the University of California, Los Angeles, Calif., September 1-9, 1969, dedicated to the memory of Theodore S. Motzkin)*, pages 159–175. Academic Press, 1972.

B. O. Koopman. Search and screening. Technical report, Operations Evaluation Group Office of the Chief of Naval Operation Navy Department, 1946.

O. Lange. Economic theory of socialism. *The review of Economic Studies*, 1936-1937.

S. A. Lippman and J. J. McCall. The economics of job search: A survey. *Economic Inquiry*, 14(2):155–89, 347–68, 1976.

F. Mathieu. Self-stabilization in preference-based systems. *Peer-to-Peer Netw. Appl.*, (1):104–121, 2008.

A. Mehrez and A. Gafni. Quality-adjusted life years, utility-theory, and healthy-years equivalents. *Medical Decision Making*, 11:140–146, 1991.

P. Morgan and R. Manning. Optimal search. *Econometrica*, 53(4):923–944, july 1985.

W. Niskanen. *Bureaucracy and Representative Government.* Aldine, 1971.

A. Nove. *The Economics of Feasible Socialism.* Routledge, 1971.

M. A. Nowak and R. M. May. Nature. *Evolutionary games and spatial chaos*, 359: 826–829, 1992.

W. Pareto. *Manual of Political Economy.* Augustus M. Kelley, 1971.

A. Roth, T. Sönmezc, and M. U. Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125:151–188, 2005.

J. M. Sakoda. The checkerboard model of social interaction. *Journal of Mathematical Sociology*, 1(1):119–132, 1971.

L. M. Sakoda. *Minidoka: An Analysis of Changing Patterns of Social Interaction.* PhD thesis, University Of California at Berkeley, 1949.

T. C. Schelling. Models of segregation. *The American Economic Review*, 59:488–493, 1969.

C. Shalizi. In soviet union, optimization problem solves you, may 30 2012.

R. Shimer and L. Smith. Assortative matching and search. *Econometrica*, 68(2):343–369, march 2000.

G. J. Stigler. The economics of information. *The Journal of Political Economy*, 69(3): 213–225, june 1961.

L. D. Stone. *Theory of Optimal Search.* Academic Press, 1975.

F. M. Taylor. The guidance of production in the socialist state. *The American Economic Review*, 19(1):1–8, march 1929.

J. von Thunen. *Isolated State: An english edition of Der Isolierte Stete.* Pergamon Press, 1966.

H. Uzawa. Walras existence theorem and Brouwer's fixed point theorem. *The Economic Studies Quarterly*, 8(1):59–62, 1962.

F. A. von Hayek. *The Road to Serfdom.* Routledge, 1944.

D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 6684(393):440–442, june 1998.

Appendix

In this appendix, one first give the data needed to construct the figure 11, then one presents the histograms for the 100×100 market, the 500×500 market and the 1000×1000 market.

Data

Size	$\mathcal{H}_{\text{couples/sim}}$	$\mathcal{H}_{\text{couples/greedy}}$	$\mathcal{H}_{\text{value/sim}}$	$\mathcal{H}_{\text{value/greedy}}$
10	0.872858053	0.558188266	0.744968479	0.294071716
50	0.786861373	0.52312604	0.786812208	0.125884938
100	0.687965749	0.560165283	0.805366463	0.213981183
200	0.439843844	0.602082432	0.848413144	0.115513857
300	0.217776383	0.653623868	0.901711248	0.101630917
400	0.111950019	0.70508775	0.958079384	0.098168908
500	0.351917696	0.728704886	0.984128682	0.073427375
600	0.558647499	0.731492504	0.988196448	0.05433787
700	0.702990725	0.742362313	0.990181025	0.055620966
800	0.809577116	0.781880246	0.998774504	0.055852806
900	0.892693721	0.81720361	1	0.042605932
1000	0.944737569	0.842601753	1	0.055156106
1200	0.989408254	0.88756580	1	0.050084976
1500	1	0.940194407	1	0.202944978
2000	1	0.962359215	1	0.056912437

Histograms

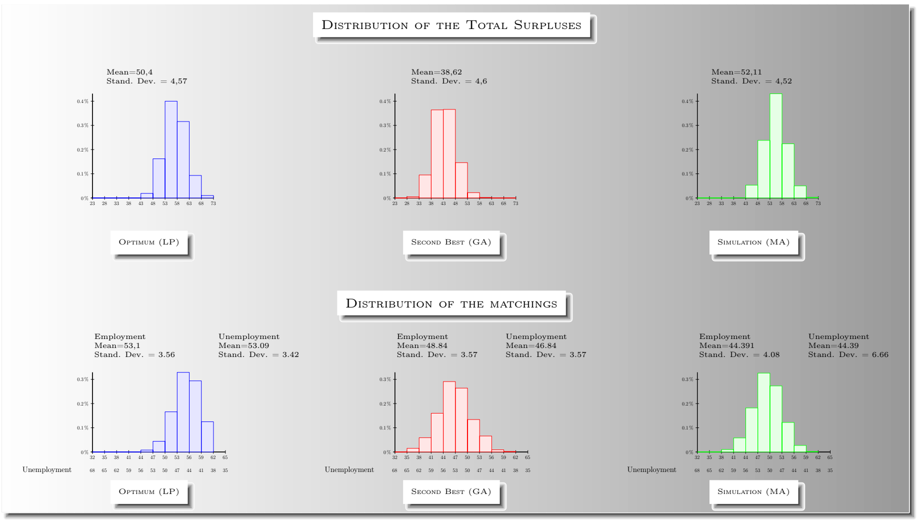

Figure 13: Results for a 100 × 100 market

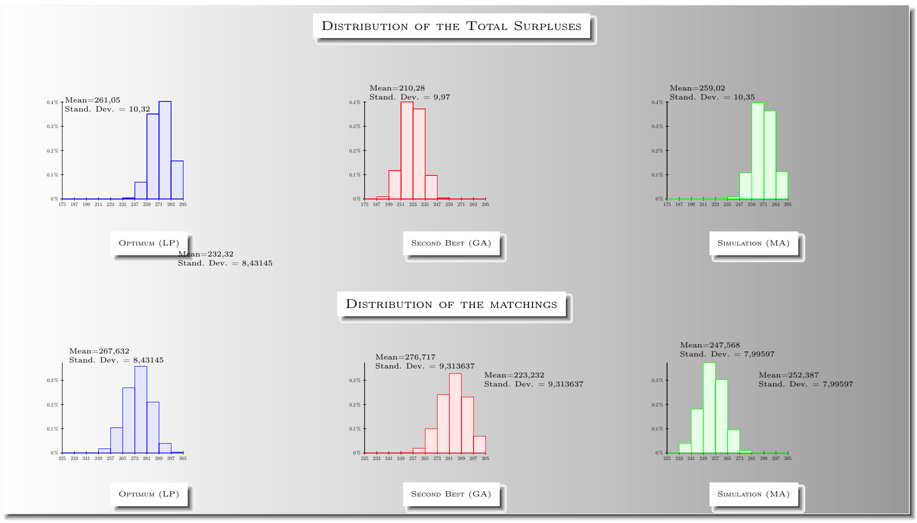

Figure 14: Results for a 500 × 500 market

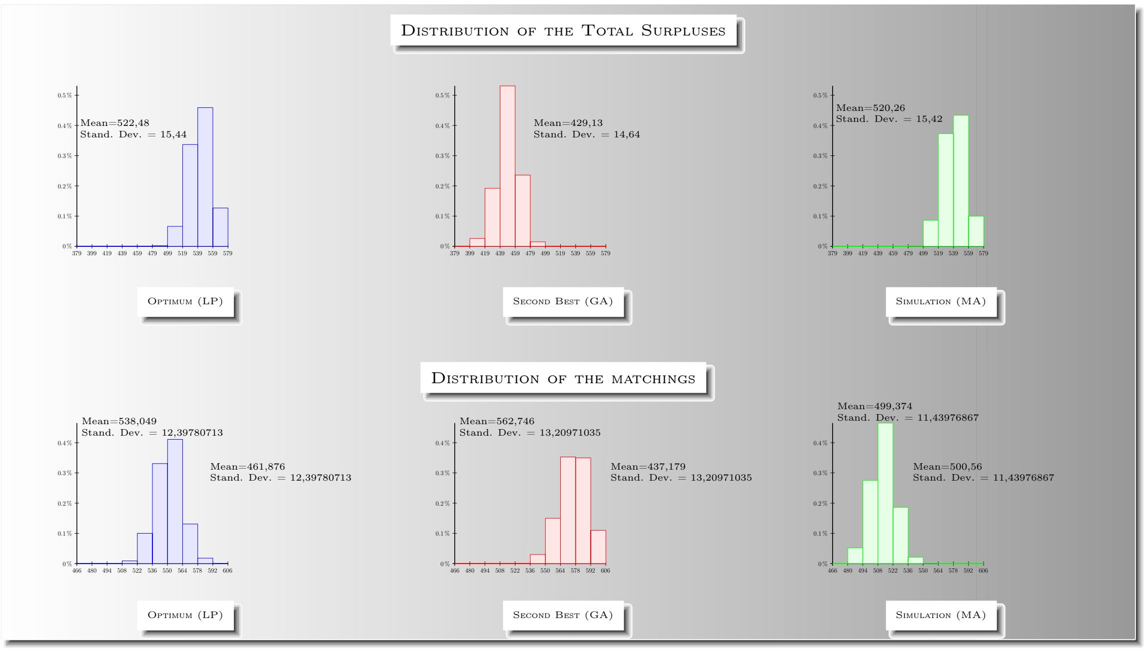

Figure 15: Results for a 1000 × 1000 market