The City as a Complex System. Statistical Physics and Agent-Based Simulations on Urban Models

Rémi Lemoy

Université Lumière Lyon 2

The City as a Complex System

Statistical Physics and Agent-Based Simulations on Urban Models

Thèse pour le Doctorat de Sciences Economiques

Université de Lyon
Laboratoire d'Economie des Transports
Institut Rhône-Alpin des Systèmes Complexes

Présentée et soutenue publiquement par

Rémi LEMOY

le 6 octobre 2011

devant la commission d'examen composée de

Directeurs de thèse	Pablo JENSEN, Directeur de recherche au CNRS
	Charles RAUX, Ingénieur de recherche au CNRS
Rapporteurs	Marc BARTHELEMY, Chercheur au CEA
	André DE PALMA, Professeur à l'ENS de Cachan
Examinateurs	Jean CAVAILHES, Directeur de recherche à l'INRA
	Stéphane DOUADY, Directeur de recherche au CNRS
	Florence GOFFETTE-NAGOT, Chargée de recherche au CNRS

Acknowledgements

I thank my supervisors Charles Raux and Pablo Jensen who proposed me to carry out this thesis. Their support and guidance during those three years were an essential help.

I also thank all members of the examining committee, and especially Marc Barthélémy and André de Palma.

I am grateful to Eric Bertin for continuing advice and fruitful collaboration. And indebted to Florence Goffette-Nagot, whose help at different times was precious and selfless. I thank also warmly Sébastian Grauwin for stimulating work and conversations during a common PhD experience. My sincere gratitude goes to John Mc Breen for his advice, help and friendship. I thank Florence Puech for precious advices and a very careful reading of one important chapter. And I am pleased to acknowledge very interesting discussions with Vincent Viguié and also Nicolas Coulombel.

I am grateful to all members of the Transport Economics Laboratory (LET) and of the Rhône-Alpes Complex Systems Institute (IXXI) for making these places pleasant human and scientific environments.

I thank also Serge Latouche, Jacques Testard and Jean-Marc Lévy-Leblond, who allowed me to get in contact with Pablo Jensen during my search for a stimulating thesis project.

I am pleased to thank heartily Jennifer Adam for her love and support.

Finally, I would like to express my gratitude to all the persons who made these three years an enjoyable time, from a scientific and also a human point of view. I hope that they recognize themselves.

Contents

General introduction

The thesis presented here has been carried out at the Transport Economics Laboratory (LET) and at the Rhône-Alpes Complex Systems Institute (IXXI) from 2008 to 2011. This work was co-supervised by Charles Raux, economist, director of LET, and Pablo Jensen, physicist at the Ecole Normale Supérieure (ENS) Lyon. The funding was specifically attributed to a work in urban modeling using agent-based models, so that it is an important part of this work. The project was designed either for an economist interested in social models inspired from statistical physics, or for a physicist interested in social systems. The author corresponds to the second type, as a physicist by training. So that the point of view presented in this work is strongly influenced by this training in physics. In particular, the first part of this work deals with tools coming from statistical physics and used in the context of social models, while the second part is more closely related to economics.

This work at the border between natural and social sciences raises specific questions thanks to the difference of viewpoints. It can lead to a fruitful exchange of objects and methods. Such works are also favored by the growing calculation power of computers and the (relative) abundance of available data about socioeconomic systems, which are trends observed in the last years or decades. However, interdisciplinary works also raise some challenges. In addition to the difficulty of communication and the danger of misunderstandings due to different frameworks and even vocabularies, the funding, publication and evaluation of research is still widely disciplinary. Some infrastructures such as IXXI exist to provide exchange possibilities for interdisciplinary work, but they remain somewhat marginal. And teaching is also very disciplinary, which can also raise some problems discouraging such research.

To cope with the different constraints mentioned, this work focuses on a given object, which is the urban system, but uses different tools in the two main parts of this thesis. The first part relies more on analytical treatment and presents an approach inspired from statistical physics, while the second one uses agent-based simulations on models which are more closely related to the literature on urban economics.

Before presenting more in detail the content of this thesis, and after the description of the scientific inter-disciplinary context, let us give some elements of socioeconomic context. What is at stake in the study of urban systems, linking land use and transport? What are the challenges which cities are facing or will be facing in the next years or decades? We give here a broad overview of the issues which are linked more or less directly to urban systems, and in particular to housing and transport in urban areas.

The context is global, at a world scale, even if very different situations can be encountered in particular cases. Indeed, it is now for instance consensual among scientists that the global climate is changing (America's Climate Choices: Panel on Advancing the Science of Climate Change [2010], Oreskes [2004]). More precisely, the average temperature on earth is increasing, hence this climate change is also named global warming. The average global temperature has increased during the 20th century of approximately 0.7°C (Intergovernmental Panel on Climate Change [2007]), which is very significative. But the effects in particular locations may differ greatly, as the global climate is a perfect example of very complex system.

Nonetheless, there is also a scientific consensus on the fact that greenhouse gases emissions due to human activities, and for instance transport and housing, which will be studied in this work, represent an important if not essential contribution to the global warming (see the same sources).

A second big challenge which humanity is facing or will be in a near future is linked to the decreasing reserves of fossil fuels such as oil and coal (International Energy Agency [2010]) and their subsequent increasing price. Peak oil is especially worrying, as oil is used as a fundamental raw material in many different contexts, not only as fuel. Its increasing price led already to begin the extraction of unconventional oil in spite of its lesser efficiency and greater environmental impact. Biofuels, also less efficient, are becoming a viable source of fuel because of the increasing price of oil. Their growing production is associated with different issues, for instance the "food vs. fuel" debate, as biofuels may be one of the causes of the 2008 food crisis, which caused food riots in a number of countries.

Associated with the decreasing reserves of fossil fuels are the decreasing reserves of many metals (Gordon et al. [2006]), which are important components of many modern devices, for instance those which produce renewable energy like solar panels (Bihouix and de Guillebon [2010]), but also electric cars (or more generally most electric or electronic devices). This complicates the problem even more.

Warnings against these decreasing reserves of fossil fuels and metals reach as far back as Meadows and Club of Rome [1972], Georgescu-Roegen [1986] (these sources providing interesting attempts to promote a science of complex systems, by linking statistical physics and economics for the second one). But almost no political measures have been taken for the moment.

More political consensus can be found on the reduction of greenhouse gases emissions. France for instance pledged to reduce its emissions by a factor of 4 in 2050 compared to 1990. But 2050 seems very far away from a human and political point of view. Internationally, many such commitments have been taken by states and communities.

In France, among other countries, the main response to the increasing price of energy associated to these challenges is the use of nuclear power plants. However, more and more concern is raised by the use of nuclear energy, following the Fukushima Daiichi nuclear disaster, which began in March 2011, and recalled other nuclear accidents, which relatively spared France for the moment. As the question of security is asked anew (several European countries decided recently to abandon progressively the use of nuclear energy), the other problems and constraints of this energy source are also discussed. For instance, the fact that there is no viable solution in sight for nuclear waste, the difficulty to deal with peak hours of electricity consumption and the close link between civil and military use of nuclear power. Nuclear fusion, which produces less toxic waste, will not be a possible energy source in the next years or decades, as it is still at a fundamental research stage for the moment, in spite of massive investments.

In this context of increasing price of energy linked to decreasing reserves of different materials, and of climate change linked to human emissions of greenhouse gases, cities can be seen as focal points. Cities consume over 60% of the global energy production and contribute to 70% of the world's emissions of greenhouse gases. More than half of the world population now lives in cities, and this share is expected to grow to reach 70% by 2050 (United Nations [2008]). Furthermore, inequalities of access to resources are very high between and among countries (Milanovic [2005]). Cities in particular are places where the richest and the poorest coexist. This creates social tensions which are yet another challenge.

The sector of transport produces 20% of total greenhouse gases emissions related to human activities in Europe, and 30% in the USA (OECD [2008]). Housing is also a major sector of emissions. As a consequence, studying the interaction between housing and transport in cities is at the heart of the problems evoked previously. In addition, cities are particularly interesting places to develop environmental friendly transportation, due to the many opportunities available on short distances. Bike use for instance can be very practical in urban areas. However, cities such as Copenhagen or Amsterdam with bicycles accounting for around 30% of all transportation are exceptions in Europe and in the world.

Of course, urban models such as those presented in this thesis are not proposing the solutions to really deal with the challenges we evoked. The main challenge is actually a question of social and political consciousness and will. But all these challenges provide a framework for urban modeling. Hence, the central question regards the "sustainable" city: if the models manage to grasp even roughly the socioeconomic reality of urban systems, or at least to help conceiving their complexity, which urban forms would consume less energy for transport and housing? And what are their consequences on the welfare of different social groups and on social inequalities? More generally, what are the main determinants of city structure, and which factors can explain the differences encountered worldwide?

These questionings constitute the socioeconomic context of this thesis. The scientific context was evoked earlier through the inter-disciplinary collaboration between physicists and economists. One more element should be added regarding the scientific context in economics. Although computer simulations are widely used in physics to make up for analytical treatment when it is limited, the use of simulations is still marginal in economics, and seems to be seen as very unsatisfactory. Hence, an important aim of the part of this thesis which is oriented towards economics is to show that computer simulations, and in particular agent-based models in this work, can provide an interesting complement to analytical resolution.

As stated previously, this work is divided in two main parts. The first part concerns the three first chapters of this thesis, and corresponds to a physicist's point of view on simple socioeconomic models. The second part consists in the three last chapters and deals with agent-based models closely related to the standard urban economics model (Alonso [1964], Muth [1969], Mills [1967], see also Fujita [1989]).

Resolution of a Schelling-like segregation model

The first chapter presents a statistical physics framework which enables us to solve analytically a Schelling-like model. Schelling's model (Schelling [1971]) is a famous toy model dealing with spatial segregation. It can be seen as one of the first agent-based models, as Thomas Schelling studied it even before computer simulations were used in social sciences.

The main ingredient of our resolution is the introduction of a state function, which reflects individual preferences and dynamics of socioeconomic models. In this framework, standard statistical physics tools can be used on simple models coming from social sciences, in a methodological approach which can be linked to potential games (Monderer and Shapley [1996]). The decision to move is taken as a logit rule, which is a standard choice in social models (Anderson et al. [1992]).

Schelling's model gives a counter-intuitive result on the link between "microscopic" behavior of social or economic agents, and the "macroscopic" outcome in the whole system studied. The description of agents' behavior is very simple. Agents make location choices depending on the composition of their neighborhood. These choices are to optimize a welfare function called utility function. For a given agent, this function is in this work a peaked function of the number of agents similar to himself in the neighborhood. Although agents desire a mixed environment, a small asymmetry in the utility function, results in segregated macroscopic patterns. This is a surprising result, as it is quite counter-intuitive to predict such a global equilibrium state of the system when knowing agents' rules of behavior.

We use this model to study the difference between individual and collective dynamics of agents. To this end, we introduce a continuous parameter leading the model to shift continuously from "individualistic" dynamics, where agents care only about their personal welfare, to "collective" dynamics, where the global welfare is the objective. These rules of behavior illustrate an important difference between social sciences models (individual dynamics) and systems usually studied by statistical physics (collective dynamics). The main achievement of this work is to introduce a potential function linking both.

Utility and chemical potential

In the second chapter, this study is brought closer to urban economics modeling. Indeed, we use this statistical physics framework in a different context, where space is not homogeneous, contrary to Schelling's model. The dynamics is still given by a logit rule with a parameter that can be seen as a temperature from a physicist's point of view. In economics, this parameter is interpreted as the width of a distribution of tastes. In this model with non-homogeneous space, we are led to introduce a chemical potential in our social model, as a Lagrange multiplier accounting for the conservation of agents, instead of particles, as is the usual case in statistical physics. When the temperature is not zero, the utility (welfare of agents) is not homogeneous in the whole system at equilibrium, contrary to standard results for socioeconomic models.

However, the chemical potential we introduce, which adds an entropic contribution to the utility term, is homogeneous across the system. To illustrate this framework, we study a model corresponding to a simplified version of the standard urban economics model. The essential simplifying hypothesis consists in supposing that price and density are directly related. As a consequence, the statistical physics framework presented in the first chapter can be used. A numerical resolution of this simple model allows us to link these results to the economic literature on the introduction of a heterogeneity of tastes in the standard urban economics model (De Palma and Papageorgiou [1988], Anas [1990]).

An important condition of the model making analytical resolution possible is that the dynamics respects detailed balance (see Van Kampen [1992], Evans and Hanney [2005]). Some conditions are derived to ensure that it is the case in the model we study. However, we perform agent-based simulations on a model which does not respect these conditions, that is, outside the domain of validity of our statistical physics framework. We observe that our result concerning the homogeneity of the chemical potential can be valid in other cases.

A probabilistic model of housing price formation

The third chapter builds on a hypothesis of the second one, which postulates in the simple model studied a direct relationship between density of inhabitants and price at a given location. We design a simple model of urban housing market which, through the use of an ergodic hypothesis, links the time variations of the price of a representative flat to the spatial variations of price within a virtual city. Price evolves as a function of the occupancy of the flat: it increases when the flat is occupied (highly demanded), and decreases if the flat is empty. Agents moving into or out of this flat will determine the evolution of its price. This is a schematic way to represent urban price dynamics and link it to the density of inhabitants.

However, this work in progress raises a question regarding the relationship between two definitions of the density, one linked to time variations and corresponding to the mean occupancy in time of the flat, and the other being a "spatial" constraint in the virtual city. Indeed, the ergodic hypothesis allows us to link the mean occupancy in time of a given flat to the spatial variation of price in the virtual city. But the simulations show that these two quantities do not coincide, a phenomenon which deserves further investigation.

Still, these simulations show that the model reaches an equilibrium state which is independent of the initial conditions. The evolution of this equilibrium as a function of the different parameters of the model is a perspective of work. The influence on price distributions of the density of agents, which is the most "physical" parameter of the model, is studied in this chapter. An interesting perspective is linked to the simplicity of the dynamics used. It should make an analytical resolution of the model possible in some limiting cases.

An agent-based model of urban economics

The second part of this work relies less on analytical results, in order to study more descriptive models. This research is directly linked to the urban economics literature, but the use of simulations corresponds more to a physicist's approach. This work is inspired for instance by Caruso et al. [2007], Brueckner et al. [1999]. Chapter 4 presents an agent-based model using simple and intuitive dynamics. From a random initial state, a urban system is led by the dynamics to an equilibrium corresponding to a discrete version of Alonso's model (Alonso [1964]). This standard urban economics model presents an analytical, static residential equilibrium of agents in a urban area. The dynamics of the agent-based model are directly inspired by the competition for land of the analytical model, where the highest bidder wins.

We find a good agreement between the results of the analytical and agent-based models. Then, building on this agreement, the agent-based model is used to study phenomena which are less tractable analytically. In particular, this chapter, following Brueckner et al. [1999] or Wu and Plantinga [2003], considers the introduction of a positive amenity in the urban system. The agent-based simulations enable us to study the influence of this amenity on the city. Different variables indicate the behavior of the urban model. The utility of agents is associated to their economic welfare. When two income groups are introduced, the evolution of the gap in welfare between both income groups can be studied when some parameters are changed. The commuting distances of agents are also an important indicator of the urban system.

Different urban social structures can be observed in North American and European cities.

North American cities tend to have rich households in the periphery and poorer ones near the center, while European cities tend to present an inverse configuration, as described by Brueckner et al. [1999]. Attempts are made using the agent-based model to have these social structures emerge from the interactions between agents. But the results seem unsatisfactory. Indeed, the standard urban economics model results usually in a "North American" configuration, with rich agents in the periphery. The "European" city is more difficult to obtain within this model.

The introduction of a travel time cost alone does not induce the "European" configuration, if realistic values of this time cost for rich and poor agents are studied. And the use of a log-linear utility function leads us to define a higher preference of rich agents for the amenity introduced in this chapter, in order to obtain the "European" city configuration. Although this can be sufficient to have this result, the ingredient is exogenous and disturbing from a modeling point of view. The equilibrium obtained is shown to be very sensitive to a difference of location between the work center and the amenity center, in a two-dimensional city.

Exploring the polycentric city

Chapter 5 uses a version of the agent-based model introduced in chapter 4, where the dynamics is simplified, but leads still the model to its equilibrium. The goal of this chapter is to deal with the question of the polycentric city. It is indeed an issue in the economic literature, to determine if a city with several employment centers is more desirable from a point of view of sustainability than a monocentric city. And the polycentric framework seems also closer to the empirical reality of urban systems.

A detailed discussion is given on the existence and uniqueness of the equilibrium of the different models studied. The existence is proved in Fujita and Smith [1987], and uniqueness is proved in Fujita [1985] in a certain monocentric framework. A mapping between mono- and polycentric models allows us to extend this result to some polycentric models we study. And qualitative arguments are used in the other cases.

A comparison is realized between the results of the standard analytical monocentric model with two income groups and those of the agent-based model. The analytical model is solved numerically thanks to a dedicated procedure (see Fujita [1989]). The agreement between analytical results and agent-based simulations is very good. The chapter details the conditions guaranteeing that the agent-based model reaches a discrete version of the analytical equilibrium. We insist on the close link between the analytical results and the agent-based simulations because the latter have still to prove their value in the economic literature, which valuates analytical results more.

The study of the polycentric city in itself is then presented, with different polycentric models where agents have different constraints, and represent single workers in some models and two-workers households in another, with some households composed of two persons working in different employment centers. As in chapter 4, the evolution of certain global variables conveys the economic, environmental and social outcomes of these models.

Polycentrism is globally seen as desirable in these models, as it favors the welfare of agents and tends to decrease commuting distances. But a negative effect in terms of greenhouse gases emissions is the fact that it increases housing surfaces, and thus heating and cooling needs. As a consequence, calibrated urban models are needed to really assess the environmental outcome of such models. A first attempt to calibrate Muth's model, which includes building construction (Muth [1969]) on the Grand Lyon urban area is presented in appendix B.

"European" and "North American" city

In chapter 6, the model presented in chapter 5 is applied to different utility functions, in order to come back to the question of "European" and "North American" cities addressed in chapter 4. Indeed, chapters 4 and 5 use a Cobb-Douglas utility function, which is the most used utility function because of its handy character for analytical resolution. But of course, the form of the utility function determines the results of the model, and other expressions are studied in this chapter 6.

More precisely, taking into account the evolution with income of the value of time of agents and of the budget share of housing yields in the most standard version of Alonso's model (using a Cobb-Douglas utility function) "European" city patterns. A characteristic distance is defined, at which the locational behavior of income groups changes. This allows us to model the emergence of a rich periurban area in "European" cities.

The hypothesis of Brueckner et al. [1999] concerning a central amenity in European cities is then revisited. Indeed, a utility function making rich agents valuate the amenity more than poorer ones can result in a "European" pattern. This result is reproduced with our agent-based model. This work shows in addition that richer social patterns can emerge than just wealthy agents in the center and poorer ones in the periphery (or vice versa), in agreement with the first part of the chapter.

Finally, a logit location choice is introduced in the agent-based model, as a means to lessen the dependence of the results of the model on the choice of the utility function. In fact, including a logit model introduces randomness in the location choice, and blurs the patterns dictated by the utility function. This study bridges the gap with the first part of this thesis, as this last model is also shown to verify the main result obtained in chapter 2. The chemical potential is homogeneous at equilibrium in the agent-based model, which is surprising as this model seems very far from fitting in the domain of validity of the statistical physics framework of chapter 2. More work is needed to link this result with the economic literature, especially De Palma and Papageorgiou [1988] and Anas [1990].

Part I

Statistical physics and urban models

CHAPTER 1

Solution of a Schelling-like segregation model

1.1 Introduction

This chapter presents the contribution of this author to a collective work (Grauwin et al. [2009]). The model we study is a simplified version of the segregation model of Thomas Schelling (Schelling [1971]). The original model is not solvable analytically, so that some simplifications are made here to allow an analytical resolution. The important point is that under these simplifications, the main phenomenon of Schelling's model, a non-intuitive segregation emerging whereas agents are looking for integration, is conserved. We also explore how this "fated" segregation can be broken by introducing a more altruistic behavior of agents, inspired from usual dynamics in physics. This allows us to deal with the question of individual and collective dynamics, which seems to be one important difference between social systems and systems usually studied by statistical physics.

The main idea of this work is detailed in Grauwin et al. [2009]. A function, which is called "link function" in the article, is introduced to describe the dynamics of the model like the dynamics of usual statistical physics models. This can be linked to the literature on potential games (Monderer and Shapley [1996]). From a physicist's point of view, this function can be thought of as an effective Hamiltonian. It allows in the context of this simplified Schelling model to bridge the gap between social sciences and statistical physics. Namely, economic or social agents move in social models to maximize their utility (Mas-Colell et al. [1995]), an individual function which describes their welfare with respect to the variables studied by the model. In statistical physics, particles move to minimize the (free) energy (Goodstein [1985]), which is a global function of the system. This can be interpreted intuitively by observing that physical particles have no individuality, compared with social agents. But this characteristic makes the modeling of their behavior easier. The introduction of this link function allows the use of a statistical physics formalism to solve the problem by finding the equilibrium state.

1.2 Description of the model

This segregation model describes the location choices of agents in a city and how segregation can emerge at a macroscopic level from interactions between agents at a microscopic level. The model is very simple and schematic: two groups of agents are studied, which are distinguished by their "colors", red and green for instance. These agents live on a two-dimensional grid, where each cell can accommodate one agent only, and some cells are left vacant. Agents move to maximize their welfare, which depends only on the number of agents in their neighborhood who have the same color as themselves.

Instead of using interactions with nearest neighbors, for instance Moore neighborhoods, agents are located in quarters and interact only within a quarter, which is an important difference with the standard Schelling model.

Let us first give a brief description of the notations used to characterize this social system.

1.2.1 Notations

The simulation space is a two-dimensional grid, composed of Q blocks labeled by $q = 1, ..., Q$. Each block is in turn composed of H cells. Two groups of agents, described as red and green, live on this space, each agent occupying one cell. There are overall N_r red and N_g green agents, with $N_r + N_g < QH$, the rest of the cells remaining vacant. This is illustrated on figure 1.6.

A microscopic configuration x of the system is given by the knowledge of the state of each cell: red, green or empty. A coarse-grained configuration corresponds to the knowledge of the numbers of red and green agents in each block, described respectively by n_{qr} and n_{qg}. The utility function of a red (respectively green) agent is written $u_R(n_{qr}/H)$ (respectively $u_G(n_{qg}/H)$). It depends only on the number of agents of the same kind in the block where the considered agent lives. Agents move in order to maximize this utility function.

The link function, which allows to describe the individual dynamics by a global quantity, is denoted $L(x)$. This global function must reflect each individual move, which can be written $\Delta u = \Delta L$: for any move, the variation of the (individual) utility of the moving agent is equal to the corresponding variation of the link function L, which is a global function.Then the link function corresponds, in each block, to the sum of agents' utility as they are introduced one by one in the block:

$$L(x) = \sum_q \Big(\sum_{m=0}^{n_{qr}} u(m/H) + \sum_{m=0}^{n_{qg}} u(m/H) \Big)$$

To study the link between individual and collective dynamics, the total utility U is introduced. It is given in a configuration x by the sum of all individual utility functions:

$$U(x) = \sum_q (n_{qr} u(n_{qr}/H) + n_{qg} u(n_{qg}/H))$$

An interesting question can be raised: the difference between collective and individual dynamics. A simple way to study it is to introduce a continuous parameter α, $0 \leq \alpha \leq 1$, which will determine what determines the dynamics of the system: the function driving the system is a cost function denoted by $\mathcal{C}$, and its expression is chosen as $\mathcal{C}(x) = (1-\alpha)L(x) + \alpha U(x)$. For $\alpha = 0$, individual decisions only are taken into account in the dynamics, and for $\alpha = 1$, only the global utility matters.

1.2.2 Dynamics

The evolution of the system is simple. Consider an agent and a vacant cell in the city. Let Δu be the variation of utility the considered agent would experience when moving to the candidate cell and ΔU the corresponding variation of the global utility. Then each agent has a probability $W(x \to y)$ per unit time to move in a given vacant cell, given by a logit rule (Anderson et al. [1992])

$$W(x \to y) = \frac{1}{1 + e^{-\Delta \mathcal{C}/T}} = \frac{1}{1 + e^{-(\mathcal{C}(y) - \mathcal{C}(x))/T}}$$

where x and y are the states of the system before and after the move, and

$$\Delta \mathcal{C} = (1-\alpha)\Delta u + \alpha \Delta U = \Delta u + \alpha(\Delta U - \Delta u)$$

is the variation of the global cost function associated to the considered move. Depending on the value of the "altruism" parameter α between 0 and 1, the agent will move based more on his individual utility gain or rather on the variation of the global utility of the system.

1.2.3 Stationary probability distribution

This system obeys a master equation given by

$$\frac{\partial P(x,t)}{\partial t} = \sum_{y \neq x} \Big(P(y,t)W(y \to x) - P(x,t)W(x \to y) \Big)$$

where $P(x,t)$ is the probability that the system is in state x at time t.

Let us now introduce the probability density $P_s(x)$ given by

$$P_s(x) = \frac{\exp(\mathcal{C}(x)/T)}{\sum_z \exp(\mathcal{C}(z)/T)}$$

It can be verified that for all couples of states x and y this probability density satisfies detailed balance:

$$\begin{aligned} P_s(x)W(x \to y) &= \frac{\exp(\mathcal{C}(x)/T)}{\sum_z \exp(\mathcal{C}(z)/T)} \times \frac{1}{1+\exp[-(\mathcal{C}(y)-\mathcal{C}(x))/T]} \\ &= \frac{\exp(\mathcal{C}(y)/T)}{\sum_z \exp(\mathcal{C}(z)/T)} \times \frac{\exp(\mathcal{C}(x)/T)}{\exp(\mathcal{C}(y)/T)+\exp(\mathcal{C}(x)/T)} \\ &= P_s(y)W(y \to x) \end{aligned}$$

Then $P_s(x)$ is the equilibrium probability density, which describes the probability of each microscopic state x at the equilibrium of the model. Let us now study the coarse-grained description of this system and its equilibrium.

1.2.4 Coarse-grained description

Instead of describing the system by the knowledge of the state of each cell, we describe it now by the knowledge of the numbers of red and green agents in each block, n_{qr} and n_{qg}. We denote this coarse-grained state by $\{(n_{qr}, n_{qg})\}$. There are $\frac{H!}{n_r!n_g!(H-n_r-n_g)!}$ ways of ordering n_r undifferentiated red agents and n_g undifferentiated green agents in H cells. Indeed, there are $\frac{H!}{(n_r+n_g)!(H-n_r-n_g)!}$ ways of placing the vacant cells and $\frac{(n_r+n_g)!}{n_r!n_g!}$ ways of placing the agents' colors. So that the stationary probability of a coarse-grained state $\{(n_{qr}, n_{qg})\}$ can be written:

$$\begin{aligned}
\Pi(\{(n_{qr}, n_{qg})\}) &= \frac{1}{Z}\prod_q \frac{H!}{n_{qr}!n_{qg}!(H-n_{qr}-n_{qg})!} e^{\mathcal{C}(\{(n_{qr},n_{qg})\})/T} \\
&= \frac{1}{Z}\exp\Big(H/T\sum_q f(n_{qr}, n_{qg}, T, H)\Big)
\end{aligned}$$

where Z is a normalization constant, and

$$\begin{aligned}
f(n_{qr}, n_{qg}, T, H) &= -\frac{T}{H}\ln\Big(\frac{n_{qr}!n_{qg}!(H-n_{qr}-n_{qg})!}{H!}\Big) \\
&+ \alpha\frac{n_{qr}}{H}u_r\Big(\frac{n_{qr}}{H}\Big) + \alpha\frac{n_{qg}}{H}u_g\Big(\frac{n_{qg}}{H}\Big) \\
&+ (1-\alpha)\frac{1}{H}\sum_{m=0}^{n_{qr}} u_r\Big(\frac{m}{H}\Big) + (1-\alpha)\frac{1}{H}\sum_{m=0}^{n_{qg}} u_g\Big(\frac{m}{H}\Big)
\end{aligned}$$

with u_r and u_g the utility functions of red and green agents, depending only on the number of agents of the same color in the block. The configurations that maximize the potential $F(x) = \sum_q f(n_{qr}, n_{qg}, T)$ are the more probable to come up. In the limit $H/T \to \infty$, these configurations are even the only ones that will appear in the stationary states (since $\Pi(x)/\Pi(y) = e^{H/T(F(x)-F(y))} \to 0$ for $F(x) - F(y) < 0$ and $H/T \to \infty$).

1.2.5 Continuous limit

In the limit $H \to \infty$, by keeping constant the mean density $\rho_0 = N_0/H$ and the density of each block $\rho_q = n_q/H$ (ρ_q hence becoming a continuous variable), one has thanks to Stirling's formula:

$$\begin{aligned}
\ln\Big(\frac{n_r!n_g!(H-n_r-n_g)!}{H!}\Big) \simeq\; & H\Big[\rho_{qr}\ln\rho_{qr} + \rho_{qg}\ln\rho_{qg} \\
& +(1-\rho_{qr}-\rho_{qg})\ln(1-\rho_{qr}-\rho_{qg})\Big]
\end{aligned}$$

and the stationary distribution can be written as:

$$\Pi(\{(\rho_{qr}, \rho_{qg})\}) = \frac{1}{Z} \prod_q e^{H/T f(\rho_{qr}, \rho_{qg}, T)}$$

where the "block-potential" f is

$$\begin{aligned} f(\rho_r, \rho_g, T) &= -T\rho_r \ln \rho_r - T\rho_g \ln \rho_g - T(1-\rho_r-\rho_g)\ln(1-\rho_r-\rho_g) \\ &+ \alpha\rho_r u_r(\rho_r) + \alpha\rho_g u_g(\rho_g) \\ &+ (1-\alpha)\int_0^{\rho_r} u_r(\rho')d\rho' + (1-\alpha)\int_0^{\rho_g} u_g(\rho')d\rho' \end{aligned}$$

The problem hence results in finding the set $\{(\rho_{qr}, \rho_{qg})\}$ which maximizes the potential $F = \sum_q f(\rho_{qr}, \rho_{qg}, T)$ with the constraints $\sum_q \rho_{qr} = Q\rho_{0r}$ and $\sum_q \rho_{qg} = Q\rho_{0g}$, where ρ_{0r} and ρ_{0g} are the global densities of red and green agents.

When comparing this result with the result of the one color model detailed in Grauwin et al. [2009], it can be remarked that for a zero temperature the two colors model described here is obtained by summing two one color models, one for each color. But at a non-zero temperature the $-T(1-\rho_r-\rho_g)\ln(1-\rho_r-\rho_g)$ term links both groups of agents.

1.3 Phase transitions

The simplest result of the model corresponds to having a homogeneous phase $\rho = (\rho_r, \rho_g)$ in the whole city. But this homogeneous phase may be unstable with respect to phase separation.

Let us split the system into two phases of densities $\rho_1 = (\rho_{1r}, \rho_{1g})$ and $\rho_2 = (\rho_{2r}, \rho_{2g})$. The constraint that the overall densities of particles/agents are $\rho_0 = (\rho_{0r}, \rho_{0g})$ is expressed by the lever rule:

$$\begin{cases} Q_1 + Q_2 &= Q \\ Q_1\rho_1 + Q_2\rho_2 &= Q\rho_0 \end{cases}$$

where Q_1 and Q_2 are respectively the number of blocks of density ρ_1 and ρ_2. The homogeneous phase is stable against phase separation if for all ρ_1 and ρ_2

$$Q_1 f(\rho_1) + Q_2 f(\rho_2) < Q f(\rho_0) \tag{1.1}$$

Geometrically, this inequality corresponds to requiring that $f(\rho)$ is a concave function.

When the concavity requirement is violated, phase separation will occur for certain values of ρ_0. The equilibrium densities ρ_1 and ρ_2 are such that the line that joins the points (ρ_1, $f(\rho_1)$) and (ρ_2,$f(\rho_2)$) is part of the concave hull of the function.

In the two colors model there is also a possibility that the system is split into 3 phases of densities $\rho_1 = (\rho_{1r}, \rho_{1g})$, $\rho_2 = (\rho_{2r}, \rho_{2g})$ and $\rho_3 = (\rho_{3r}, \rho_{3g})$, which is absent in the simpler model presented in Grauwin et al. [2009]. The constraint that the overall densities of particles/agents are $\rho_0 = (\rho_{0r}, \rho_{0g})$ is in the 3 phases case:

$$\begin{cases} Q_1 + Q_2 + Q_3 & = Q \\ Q_1\rho_{1r} + Q_2\rho_{2r} + Q_3\rho_{3r} & = Q\rho_{0r} \\ Q_1\rho_{1g} + Q_2\rho_{2g} + Q_3\rho_{3g} & = Q\rho_{0g} \end{cases}$$

where Q_1, Q_2 and Q_3 are respectively the number of blocks of density ρ_1, ρ_2 and ρ_3.

And the equilibrium densities ρ_1, ρ_2 and ρ_3 are now such that the plane that joins the points (ρ_1, $f(\rho_1)$), (ρ_2,$f(\rho_2)$) and (ρ_3,$f(\rho_3)$) is part of the concave hull of the function. For some values of the parameters there may even be 4 points of the same plane belonging to the f function and its concave hull. In this case there will be a continuum of possible values of Q_1, Q_2, Q_3 and Q_4 verifying the global density constraints.

1.4 With a peaked utility function

1.4.1 Expression of the f function

Let us consider for both color groups the asymmetrically peaked utility function (Pancs and Vriend [2007]) $u_r = u_g = u$ defined for $m < 1$ as:

$$\begin{aligned} u(\rho) &= 2\rho \quad \text{if } \rho \leq 0.5 \\ u(\rho) &= m + 2(1-m)(1-\rho) \quad \text{if } \rho > 0.5 \end{aligned}$$

This function is presented on figure 1.1. Agents have a real taste for integration, as they are most satisfied when only half of their environment consists in agents of the same color as them. But they prefer total segregation, providing them with a utility m, which is usually supposed to be non-negative, to being "alone among strangers" (or vacant cells), which gives a zero utility.

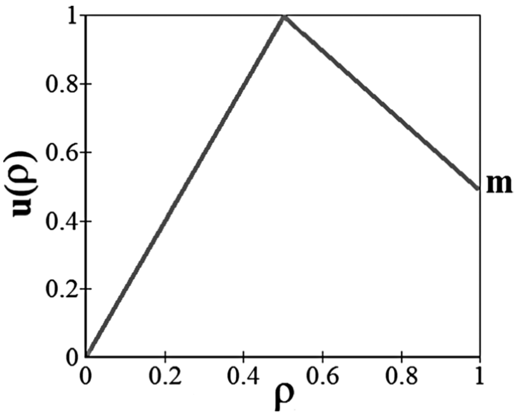


Figure 1.1: Asymmetrically peaked utility function, with $m = 0.5$.

Then for $\rho_r \leq 0.5$ and $\rho_g \leq 0.5$, the f function is given by:

$$\begin{aligned}
f(r,g) &= -T\big(r\ln r - g\ln g - (1-r-g)\ln(1-r-g)\big) + (1+\alpha)(r^2+g^2) \\
\frac{\partial f}{\partial r}(r,g) &= -T\big(\ln r - \ln(1-r-g)\big) + 2(1+\alpha)r \\
\frac{\partial^2 f}{\partial r^2}(r,g) &= -T/r - T/(1-r) + 2(1+\alpha)
\end{aligned}$$

(Partial derivatives relative to g are obtained by replacing $g \leftrightarrow r$)

For $\rho_r > 0.5$ and $\rho_g \leq 0.5$,

$$\begin{aligned}
f(r,g) &= -T\big(r\ln r + g\ln g + (1-r-g)\ln(1-r-g)\big) \\
&\quad -(1+\alpha)(1-m)r^2 + (2-m)r - (1-\alpha)(2-m)/4 + (1+\alpha)g^2 \\
\frac{\partial f}{\partial r}(r,g) &= -T\big(\ln r - \ln(1-r-g)\big) - 2(1+\alpha)(1-m)r - (2-m) \\
\frac{\partial^2 f}{\partial r^2}(r,g) &= -T/r - T/(1-r-g) - 2(1+\alpha)(1-m) \\
\frac{\partial f}{\partial g}(r,g) &= -T\big(\ln g - \ln(1-r-g)\big) + 2(1+\alpha)r \\
\frac{\partial^2 f}{\partial r^2}(r,g) &= -T/g - T/(1-r-g) + 2(1+\alpha)
\end{aligned}$$

The situation $\rho_r \leq 0.5$ and $\rho_g > 0.5$ can be obtained by replacing $g \leftrightarrow r$ in the previous paragraph.

1.4.2 At zero temperature

f is concave in ρ_r and ρ_g for ρ_r and $\rho_g \leq 0.5$. For $\rho_r > 0.5$ and $\rho_g \leq 0.5$, f is concave in ρ_r and convex in ρ_g (and conversely for $\rho_r \leq 0.5$ and $\rho_g > 0.5$, f is concave in ρ_g and convex in ρ_r).

The concave hull of the function has a different form for different values of the parameters α and m:

- for $\alpha \geq \frac{1}{3-2m}$ the points $\left(0, \frac{1}{2}, f(0, \frac{1}{2})\right)$ and $\left(\frac{1}{2}, 0, f(\frac{1}{2}, 0)\right)$ belong to the concave hull whereas for $\alpha \leq \frac{1}{3-2m}$ they are replaced by the points $\left(0, \tilde{\rho}_2, f(0, \tilde{\rho}_2)\right)$ and $\left(\tilde{\rho}_2, 0, f(\tilde{\rho}_2, 0)\right)$ with $\tilde{\rho}_2(\alpha, m) = \frac{1}{2}\sqrt{\frac{1-\alpha}{1+\alpha}\,\frac{2-m}{1-m}}$ (see the resolution of the model with agents of only one color in Grauwin et al. [2009]).

- for $\alpha \geq \frac{m}{4-3m}$ the point $\left(\frac{1}{2}, \frac{1}{2}, f(\frac{1}{2}, \frac{1}{2})\right)$ belongs to the concave hull whereas for $\alpha < \frac{m}{4-3m}$ it does not.

So there are three possible situations, shown on figure 1.2:

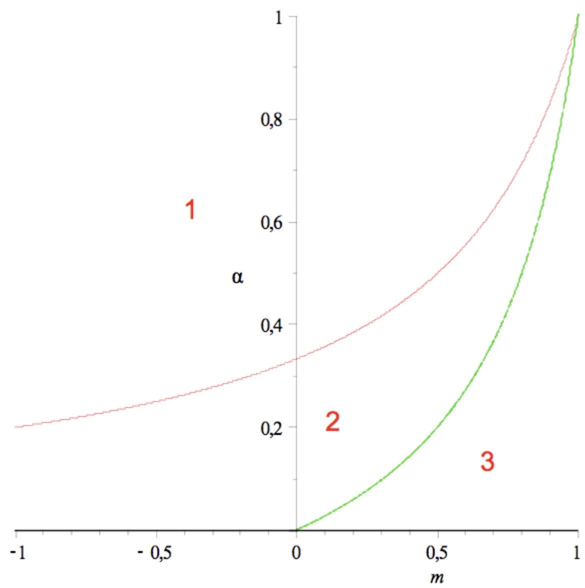


Figure 1.2: The domains of different concave hulls for different values of m and α

- $\alpha \geq \frac{1}{3-2m}$ (which will be case 1)
- $\frac{m}{4-3m} \leq \alpha \leq \frac{1}{3-2m}$ (case 2)
- $\alpha < \frac{m}{4-3m}$ (case 3)

Case 1

The number and composition of the phases depend on the global densities $\rho_0 = (\rho_{0r}, \rho_{0g})$ (see figure 1.3). In part A of figure 1.3, the system separates into 3 or 4 phases of densities $(0, 0)$,

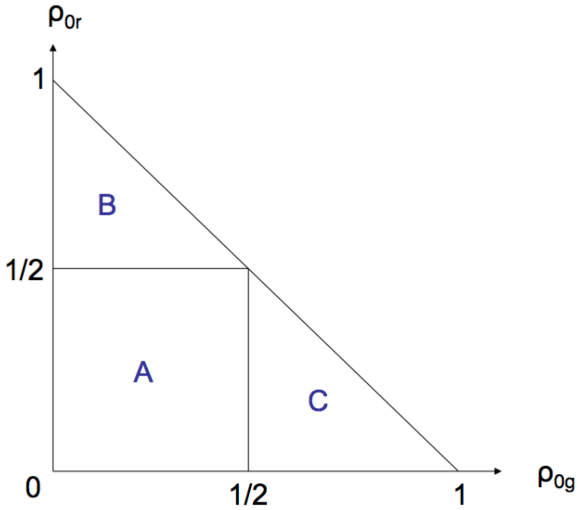


Figure 1.3: Domains of different phases for different global densities $\rho_0 = (\rho_{0r}, \rho_{0g})$ in case 1

$(0, \frac{1}{2})$, $(\frac{1}{2}, 0)$ and $(\frac{1}{2}, \frac{1}{2})$ in respective quantities Q_1, Q_2, Q_3 and Q_4, which must verify

$$\begin{cases} Q_2 + Q_4 & = 2Q\rho_{0g} \\ Q_3 + Q_4 & = 2Q\rho_{0r} \\ Q_1 + Q_2 + Q_3 + Q_4 & = Q \end{cases}$$

The system can build either 3 or 4 phases because red and green agents do not "see" each other: their utility is maximal when half of the block is filled with agents of their color, the other half being either empty or filled with agents of the other color. A simulation of a (discrete) stationary configuration corresponding to this case is shown as illustration on the left panel of figure 1.6.

In part B, the system separates into 2 phases of densities $(\frac{\rho_{0r}-\rho_{0g}}{1-2\rho_{0g}}, 0)$ and $(\frac{1}{2}, \frac{1}{2})$ with respective weights $Q_1 = Q(1-2\rho_{0g})$ and $Q_2 = 2Q\rho_{0g}$. And symmetrically in part C the system separates into 2 phases of densities $(0, \frac{\rho_{0g}-\rho_{0r}}{1-2\rho_{0r}})$ and $(\frac{1}{2}, \frac{1}{2})$ with respective weights $Q_1 = Q(1-2\rho_{0r})$ and $Q_2 = 2Q\rho_{0r}$.

Case 2

The number and composition of the phases depend again on the global densities as shown on figure 1.4. In part A of figure 1.4, the system separates into 3 phases of densities $(0,0)$,

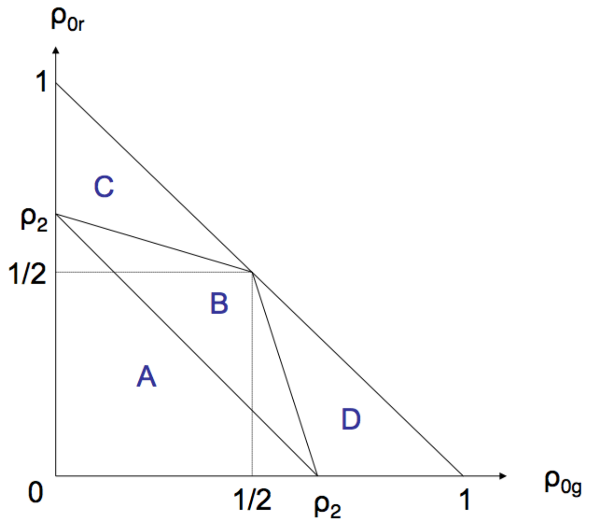


Figure 1.4: Domains of different phases in case 2

$(0, \tilde{\rho}_2(\alpha, m))$ and $(\tilde{\rho}_2(\alpha, m), 0)$ in respective quantities $Q_1 = Q(1 - \frac{\rho_{0r}+\rho_{0g}}{\tilde{\rho}_2})$, $Q_2 = Q\frac{\rho_{0g}}{\tilde{\rho}_2}$ and $Q_3 = Q\frac{\rho_{0r}}{\tilde{\rho}_2}$.

In part B the system is split into 3 phases of densities $(0, \tilde{\rho}_2)$, $(\tilde{\rho}_2, 0)$ and $(\frac{1}{2}, \frac{1}{2})$ in respective quantities $Q_1 = Q\frac{\tilde{\rho}_2(2\rho_{0g}-1)+\rho_{0r}-\rho_{0g}}{2\tilde{\rho}_2(\tilde{\rho}_2-1)}$, $Q_2 = Q\frac{\tilde{\rho}_2(2\rho_{0r}-1)+\rho_{0g}-\rho_{0r}}{2\tilde{\rho}_2(\tilde{\rho}_2-1)}$ and $Q_3 = Q\frac{\tilde{\rho}_2-\rho_{0r}-\rho_{0g}}{\tilde{\rho}_2-1}$.

In part C the phase decomposition is the same as in part B of case 1 and in part D it is the same as in part C of case 1.

Case 3

The different domains of phase decomposition are shown on figure 1.5. In part A of figure 1.5 the system separates into 3 phases like in part A of case 2.

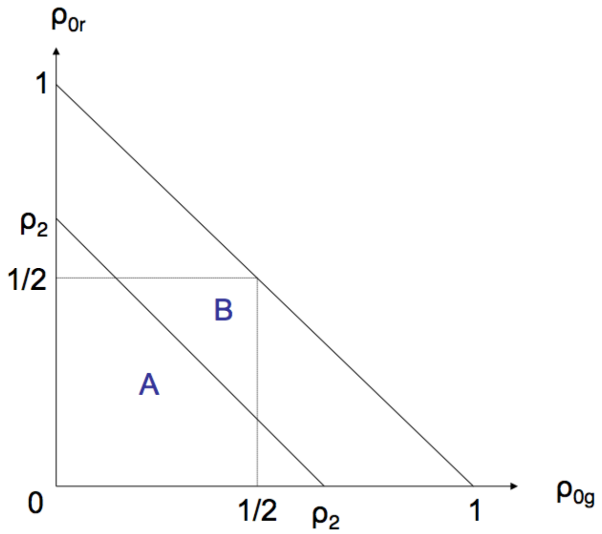


Figure 1.5: Domains of different phases in case 3

In part B there are 2 phases of densities $(\rho_{0r} + \rho_{0g}, 0)$ and $(0, \rho_{0r} + \rho_{0g})$ in respective quantities $Q_1 = Q\frac{\rho_{0r}}{\rho_{0r}+\rho_{0g}}$ and $Q_2 = Q\frac{\rho_{0g}}{\rho_{0r}+\rho_{0g}}$. A (discrete) stationary configuration corresponding to this case is given on the left panel of figure 1.6.

1.4.3 Interpretation

At zero temperature, this two colors model is indeed very similar to the one color model presented in Grauwin et al. [2009]: both groups of agents see each other only as occupied cells (the utility of a red agent does not depend on the number of green agents in the block). The only difference with two superimposed one color models lies in the fact that blocks with density $(\frac{1}{2}, \frac{1}{2})$ are full of agents with maximal utility and cannot take in more agents, so that for certain values of the global densities, in cases 1 and 2, there is a phase of half-red, half-green blocks and another phase containing the excess of the more numerous type of agents with a higher density (and thus an inferior utility). An illustration of equilibrium configurations is given on figure 1.6. Grauwin et al. [2009] provide an illustration of these same equilibrium configurations with continuous neighborhoods, to illustrate the similar behavior of this model with block neighborhoods and of the standard Schelling model.

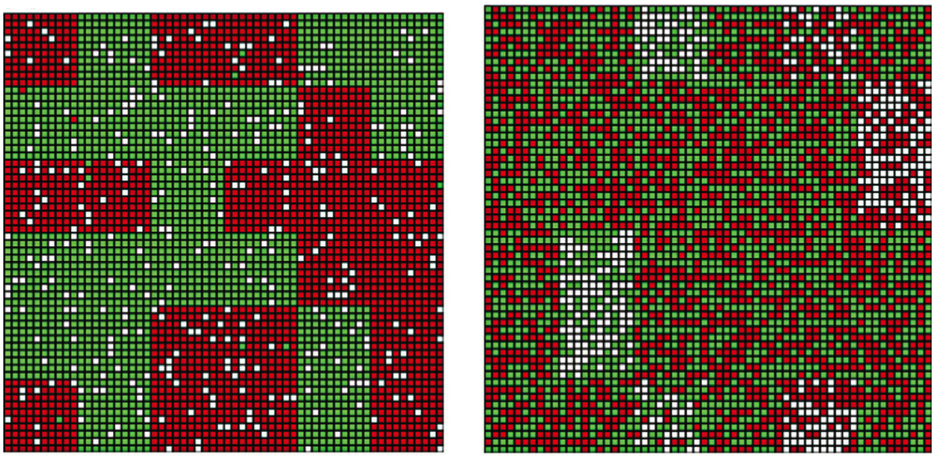


Figure 1.6: Stationary configurations of the city with equal numbers of red and green agents, 10% of the total number of cells being vacant, and $m = 0.5$. On the left panel, $\alpha = 0$, which illustrates case 3, part B: complete segregation with an "individualistic" behavior of agents. On the right panel, $\alpha = 1$, which illustrates case 1, part A: more integrated pattern with an "altruistic" behavior of agents.

1.4.4 For non-zero temperatures

At very high temperatures, the entropic term of the f function is the leading term, and as it is a concave one, the function is concave everywhere : for any density $\rho_0 = (\rho_{0r}, \rho_{0g})$, the system stays in an homogeneous phase because of the strong noise.

For intermediary values of the temperature, the system has a behavior which is intermediary between a noise driven one and the one it has at zero temperature. A numerical resolution such as the one performed in the next chapter on a different model is needed to find the equilibrium configuration of the system, but it is beyond the scope of this work.

1.5 Conclusion

The introduction of a potential function linking individual preferences to the global dynamics of the social system we study allows us to solve analytically a simplified version of Schelling's segregation model. In addition, the simplification we introduce, that is, keeping neighborhoods fixed to have local interactions in a determined environment, conserves the qualitative behavior of the original model. This work can be seen as a first example of simple solvable social models which can be studied with equilibrium statistical physics tools. The method, which consists in introducing a global potential function representing individual dynamics, could be used on other social models. In the next chapter, we show how it can be applied to a model with a non homogeneous space, polarized by a punctual employment center.

Chapter 2

Socio-economic utility and chemical potential

This chapter presents a work which has been carried out in collaboration with Eric Bertin and Pablo Jensen. It is published as Lemoy et al. [2011a].

Socio-economic sciences and statistical physics are both interested in the evolution of systems characterized by a large number of interacting entities. These entities can for instance be economic or social agents in social sciences (Smith [1784], Schelling [1971], Latour [2005]), atoms or molecules in statistical physics (Cotterill [2008], Goodstein [1985], Balescu [1975]). The question of the emergence of macroscopic patterns from the interactions of a large number of microscopic agents is studied by both fields of science. In statistical physics, a quantitative framework has been developed over the last century, allowing the equilibrium behaviour of large assemblies of atoms or molecules to be handled precisely (Balescu [1975]).

In socio-economic models, the preferences of individuals are usually characterized by a utility function, which describes their welfare with respect to their current situation or environment. Each individual or agent wants to maximize his own welfare. Decisions (e.g., moving to a more convenient place) are thus taken in a purely selfish way, while in physics the motion of particles is governed by the variation of the total energy. Recently, a global function linking individual decisions to the variation of a global quantity has been introduced to describe some classes of socio-economic models (see Grauwin et al. [2009], Goffette-Nagot et al. [2009]). This approach then allows such models to be described with statistical physics tools. Importantly, the equilibrium state can then be calculated by maximizing a state function (akin to a free energy) instead of having to solve a complicated Nash equilibrium of strategically interacting agents.

The question we investigate in this chapter is whether this physical description of socio-economic models can be extended to other basic concepts of statistical physics, such as the equalization of thermodynamic parameters like temperature or chemical potential. The equalization of these quantities throughout the system precisely results from the conservation of the conjugated extensive quantities, namely the energy or the number of particles. Although there is no notion of energy in socio-economic models, the dynamics indeed conserves the number of agents. A natural question is thus to know whether a chemical potential can be defined in such models, and what would be its relation to standard socio-economic concepts. This question is further motivated by the following remark. In spatial socio-economic models, the individual dynamics leads to a Nash equilibrium, where no agent has an incentive to move. If all agents are of the same type, the Nash equilibrium results in a spatially uniform utility, even if the environment is spatially inhomogeneous like in cities, where the center plays a specific role. This uniformity property is also expected from the chemical potential (if such a quantity can be defined), suggesting a possible relation between these two notions.

Here, we investigate this issue in the framework of a generic class of exactly solvable models involving a population of locally interacting agents. We define in a precise way a chemical potential for this class of models, and provide a direct link between the chemical potential and the socio-economic utility. Two explicit examples from the field of urban economics are also presented.

2.1 Model and dynamics

This chapter deals with socio-economic models characterized by a large number of interacting agents, residing on a set of sites, labeled by an index $q = 1, \ldots, Q$. Agents are able to move from one site to another in order to increase their utility. In addition, agents belong to m different groups, according for instance to their income, or to their cultural preferences. The variables used to describe the system are the numbers n_{qi} of agents of each group $i = 1, \ldots, m$ at each node q. The configuration of the system is described by the set $x = \{n_{qi}\}$. We assume that agents cannot change group, so that for all i, the total number $N_i = \sum_q n_{qi}$ of agents of group i is fixed. The satisfaction of agents of type i on site q is characterized by a utility

$U_{qi}(n_{q1}, ..., n_{qm})$ that depends only on the numbers of agents of each group on the same site q.

The model is defined with a continuous time dynamics following a logit (or Glauber) rule, which is commonly used in social sciences and in particular economic works (see Anderson et al. [1992]). If transitions between sites q and q' are allowed, agents move from q to q' with a probability per unit time

$$W = \frac{\nu_0}{1 + e^{-\Delta U/T}}, \tag{2.1}$$

where $\Delta U = U'_{q'i} - U_{qi}$ is the variation of the agent's own utility, with

$$\begin{aligned} U'_{q'i} &= U_{q'i}(n_{q'1}, ..., n_{q'i} + 1, ..., n_{q'm}) \quad (2.2) \\ U_{qi} &= U_{qi}(n_{q1}, ..., n_{qi}, ..., n_{qm}). \quad (2.3) \end{aligned}$$

The parameter T plays the role of an effective temperature, introducing noise in the decision process to take into account other factors influencing choices (Anderson et al. [1992]), and ν_0 is a characteristic transition frequency.

In order to obtain analytical results, we assume that the utility function is such that the change of individual utility experienced by an agent during a move can be expressed as the variation of a function of the global configuration $x = \{n_{qi}\}$ (Grauwin et al. [2009]). More precisely, we assume that there exists a function $L(x)$ such that for each agent in group i, moving from node q to node q',

$$U'_{q'i} - U_{qi} = L(y) - L(x) \tag{2.4}$$

where $y = (n_{11}, \ldots, n_{qi}-1, \ldots, n_{q'i}+1, \ldots, n_{Qm})$ and $x = (n_{11}, \ldots, n_{Qm})$ are the configurations of the system after and before the move respectively. Such a function $L(x)$ thus provides a link between the individual behaviour of agents and the evolution of the whole system. In physical terms, it can be thought of as an effective energy. The relevance of this assumption (which bears some similarities with potential games presented in Monderer and Shapley [1996]) for the general class of systems considered above will be discussed at the end of the chapter.

The stationary probability distribution $\mathcal{P}_{\rm s}(\{n_{qi}\}) = \mathcal{P}_{\rm s}(x)$ is obtained by solving the master equation governing the dynamics of the system (Van Kampen [1992]). If Eq. (2.4) holds, detailed balance is satisfied (Van Kampen [1992], Evans and Hanney [2005]), and we obtain the following expression for the distribution $\mathcal{P}_{\mathrm{s}}(x)$:

$$\mathcal{P}_{\mathrm{s}}(x) = \frac{1}{Z_{\mathrm{s}}} \frac{e^{L(x)/T}}{\prod_{q,i} n_{qi}!} \prod_{i} \delta \left(\sum_{q} n_{qi} - N_i \right) \tag{2.5}$$

where Z_{s} is a normalization constant. The product of Kronecker δ functions accounts for the conservation of the total number of agents in each group. The different factors appearing in Eq. (2.5) can be given a simple interpretation. The exponential factor directly comes from the detailed balance associated to the logit rule Eq. (6.4), while the product of factorials appearing at the denominator in Eq. (2.5) results from the coarse-graining of configurations. Namely, given the numbers of agents $\{n_{qi}\}$, there are for each group $N_i!/\prod_q n_{qi}!$ ways to arrange the agents of the group. As the numbers N_i are fixed, $N_i!$ can be reabsorbed into the normalization constant.

Defining a density of agents $\rho_{qi} = n_{qi}/H$, where $H \gg 1$ is a characteristic number (for instance a maximal number of agents on a site), the utility U_{qi} then becomes a function $u_{qi}(\rho_{q1}, ..., \rho_{qm})$. We further assume that the function $L(x)$ can be written in the large deviation form (Touchette [2009])

$$L(x) = H\tilde{L}(\{\rho_{qi}\}). \tag{2.6}$$

To determine $\tilde{L}$, we combine Eqs. (2.4) and (2.6), and expand $\tilde{L}$ to leading order in $1/H$, yielding

$$\frac{\partial \tilde{L}}{\partial \rho_{q'i}} - \frac{\partial \tilde{L}}{\partial \rho_{qi}} = u_{q'i} - u_{qi}. \tag{2.7}$$

By identification, we get for all q

$$\frac{\partial \tilde{L}}{\partial \rho_{qi}} = u_{qi}(\rho_{q1}, ..., \rho_{qm}). \tag{2.8}$$

As the r.h.s. of Eq. (2.8) only depends on densities of agents on node q, $\tilde{L}$ necessarily takes the form

$$\tilde{L}(\{\rho_{qi}\}) = \sum_{q} l_q(\rho_{q1}, ..., \rho_{qm}), \tag{2.9}$$

and one has

$$\frac{\partial l_q}{\partial \rho_{qi}} = u_{qi}. \tag{2.10}$$

If there is a single group ($m = 1$), $l_q(\rho_q)$ is simply obtained by integrating $u_q(\rho_q)$. In contrast, if $m > 1$, l_q (and thus $\tilde{L}$) only exists if the following condition, resulting from the equality of cross-derivatives of l_q, is satisfied:

$$\frac{\partial u_{qi}}{\partial \rho_{qj}} = \frac{\partial u_{qj}}{\partial \rho_{qi}}, \qquad i \neq j. \tag{2.11}$$

If this condition holds, the stationary distribution reads, after an expansion of the factorials using Stirling's formula,

$$\mathcal{P}(\{\rho_{qi}\}) = \frac{1}{Z} \prod_q e^{H f_q(\rho_{q1},\ldots,\rho_{qm})/T} \prod_i \delta\Big(\sum_q \rho_{qi} - Q\overline{\rho}_i \Big) \tag{2.12}$$

where f_q is given by

$$f_q(\rho_{q1}, \ldots, \rho_{qm}) = l_q(\rho_{q1}, \ldots, \rho_{qm}) + Ts(\rho_{q1}, \ldots, \rho_{qm}), \tag{2.13}$$

with

$$s(\rho_{q1}, \ldots, \rho_{qm}) = -\sum_i \rho_{qi} \ln \rho_{qi}. \tag{2.14}$$

In analogy to physical systems, $f_q(\rho_{q1}, \ldots, \rho_{qm})$ can be interpreted as a local free energy (up to a change of sign), and the term $s(\rho_{q1}, \ldots, \rho_{qm})$, which is multiplied by the 'temperature' T, may be seen as an entropic contribution associated to the node q.

2.2 Utility and chemical potential

We now turn to the main result of this chapter. The configurations $\{\rho^*_{qi}\}$ which maximize $F = \sum_q f_q$ under the constraints of fixed global density $\sum_q \rho_{qi} = Q\overline{\rho}_i$ are the most probable (or equilibrium) configurations. Finding the equilibrium densities of agents is then a constrained maximization problem. Let us introduce a Lagrangian

$$\begin{aligned} \mathcal{L}(\{\rho_{qi}\}, \{\lambda_i\}) \;&=\; \sum_q f_q(\rho_{q1}, \ldots, \rho_{qm}) \\ &\qquad -\sum_i \lambda_i \Big(\sum_q \rho_{qi} - Q\overline{\rho}_i \Big), \end{aligned} \tag{2.15}$$

where the parameters λ_i are Lagrange multipliers associated to the conservation of the number of agents in each group. In physical terms, λ_i corresponds to the chemical potential[1] of the agents of group i. The equilibrium densities $\{\rho^*_{qi}\}$ are then determined from the conditions $\partial\mathcal{L}/\partial\rho_{qi} = 0$ for all (q, i), yielding

$$u_{qi}(\rho^*_{q1}, ..., \rho^*_{qm}) + T\frac{\partial s}{\partial \rho_{qi}}(\rho^*_{q1}, ..., \rho^*_{qm}) = \lambda_i, \tag{2.16}$$

which is the main result of this chapter. Equation (2.16) thus provides an answer to the question raised at the beginning of this chapter: there is indeed a direct relationship between the socio-economic utility and the chemical potential defined, in analogy to equilibrium physical systems, from the conservation of the number of particles. At zero temperature, both quantities can be identified. This result might come as a surprise: utility is often thought to be the socio-economic concept most similar to the physical concept of energy (or more precisely, the opposite of the energy), because agents seek to maximize their utility in social systems and physical particles minimize the energy in the zero temperature limit. Hence one might have intuitively expected the homogeneity of utility to be linked to a notion of temperature (the thermodynamic variable conjugated to energy), rather than to a chemical potential.

Interestingly, Eq. (2.16) not only provides a link between two apparently unrelated concepts, but also yields a non-trivial prediction on the variations of utility across the system at non-zero temperature. As the chemical potential remains uniform at any temperature, one sees from Eq. (2.16) that the utility $u_{qi} = \lambda_i - T\partial s/\partial\rho_{qi}$ becomes non-uniform if $T > 0$, and that the corrections to uniformity are given by the derivative of the local entropy.

In a statistical physics language, Eq. (2.12) corresponds to the canonical ensemble, where the number of interacting entities (agents or particles) is fixed. It is sometimes convenient to consider the so-called grand-canonical ensemble, where particles are exchanged with an external reservoir. In the context of agent-based models, the reservoir corresponds to the external world. This means that we implicitly consider a very large set of sites ('the world') and focus only on a small subpart of it ('the system'), still containing a large number of agents. Since the 'world' has a fixed number of agents, it can be described by the stationary distribution Eq. (2.5). Following standard statistical physics methods (see Balescu [1975]),

[1] An equivalent formulation is to define the chemical potential λ_i as the logarithmic derivative of the partition function Z with respect to N_i, a definition that can be extended to some classes of nonequilibrium models (Bertin et al. [2007]). Note also that the standard definition of chemical potential for equilibrium systems differs by a conventional factor $-1/T$ from the one we use here (see Balescu [1975]).

the probability distribution of the considered subpart is given by

$$\mathcal{P}_{\text{ow}}(\{\rho_{qi}\}) = \frac{1}{Z_{\text{ow}}} \prod_q e^{H[f_q(\rho_{q1},\ldots,\rho_{qm}) - \sum_i \lambda_i \rho_{qi}]/T}, \tag{2.17}$$

where λ_i is the chemical potential of group i imposed by the external world. Finding the most probable densities ρ^*_{qi} is now straightforward since the densities on different sites are independent. Maximizing the argument of the exponential in Eq. (2.17), one recovers Eq. (2.16).

In the following, we give two examples of models belonging to the above generic class, in the context of urban economics.

2.3 A simple urban economics model

The model presented here is a simple model of land use and transport interaction in urban economics (Fujita [1989]). In this model, a city is described as a grid composed of Q blocks. In each block, one or several agents (representing households) can live by paying a rent to the landowner. A central business district (CBD) is placed on the grid and all agents commute there for their work (monocentric city model). A transport cost c per unit distance is associated to this commuting. The distance between a block q and the CBD is denoted by r_q. The size of the city is fixed: a constant radius r_f defines the urban fringe, out of which no agent lives[2]. All agents have the same income Y, which is spent on transport, on housing and on a composite good z representing all other consumer goods. This gives a budget constraint for each agent

$$Y = z + cr_q + \sigma p_q \tag{2.18}$$

where σ is the surface of housing, and p_q is the rent per unit surface in block q. We first consider a simple model where all agents have the same surface of housing. Each block of the grid is composed of H cells of surface σ_0. A configuration of the city is then given by the number of agents n_q in each block q. We make the further simplifying assumption that the price p_q of housing in a block q only depends on the density $\rho_q = n_q/H$ of agents in this block, namely $p_q = p(\rho_q)$. Let us emphasize that this hypothesis is an important simplification with

[2]In standard urban economics models, land is used for agriculture outside the city, and the landowners then earn an agricultural rent (Fujita [1989]). These landowners rent to the highest bidder, so that all prices must be greater than the agricultural rent. However, to simplify the presentation, we have dropped the agricultural rent parameter by introducing a fixed city size.

respect to standard urban economics models, in which the price emerges directly from the competition for land between agents, and the density from their utility maximization with respect to the surface of housing (Fujita [1989]). In cases where an explicit expression is required, we will use a logarithmic form

$$p(\rho_q) = p_0 \ln(1 + \rho_q), \tag{2.19}$$

where p_0 is a positive constant.

The utility function has to be specified explicitly. It should be an increasing function $U(z)$ of the quantity of composite good z each agent consumes, that we choose to be simply $U(z) = z$. This means that, in the limit $T \to 0$, each agent wants to maximize the share of his income which is left after transport and housing expenses. Using Eqs. (2.18) and (2.19), the utility U becomes a function $u_q(\rho_q)$ of the local density,

$$u_q(\rho_q) = Y - cr_q - \sigma_0 p(\rho_q). \tag{2.20}$$

Urban economics distinguishes closed city models, where the total number N of agents is fixed, and open city models, where N fluctuates due to exchanges with the external world (Fujita [1989]). We start by considering the closed city model. In the continuous limit where H and $N \to \infty$ with the average density $\overline{\rho} = N/(HQ)$ fixed, the stationary probability distribution takes the form Eq. (2.12), with $f_q(\rho_q)$ given by

$$f_q(\rho_q) = \int_0^{\rho_q} u_q(\rho) d\rho + T s(\rho_q) \tag{2.21}$$

and $s(\rho_q) = -\rho_q \ln \rho_q$.

The most probable density ρ_q^* is then obtained as a function of λ from Eq. (2.16), namely

$$u_q(\rho_q^*) + T \frac{ds}{d\rho_q}(\rho_q^*) = \lambda. \tag{2.22}$$

In the limit $T \to 0$, often considered in socio-economic models, one finds $\rho_q^* = \rho^*(r_q, \lambda)$, with

$$\rho^*(r_q, \lambda) = p^{-1}\left(\frac{Y - cr_q - \lambda}{\sigma_0}\right), \tag{2.23}$$

where p^{-1} is the reciprocal function of p.

The parameter λ is then determined from the density constraint $\sum_q \rho_q^* = Q\overline{\rho}$. Following standard literature (see Fujita [1989]), we focus here on the simplest situation of a one-dimensional city. Using the continuous approximation

$$\frac{1}{Q}\sum_q \rho_q^* \approx \frac{1}{r_f}\int_0^{r_f} \rho^*(r,\lambda)\,dr, \tag{2.24}$$

we compute the average density $\overline{\rho}(\lambda)$, and then determine numerically the reciprocal function $\lambda(\overline{\rho})$.

We now briefly turn to the open city model (similar to the above 'open world' case) where agents can also move to or from a large number of other cities. The stationary distribution is given by Eq. (2.17), which in the present open city model simplifies to

$$\mathcal{P}_{\mathrm{oc}}(\{\rho_{qi}\}) = \frac{1}{Z_{\mathrm{oc}}}\prod_q e^{H[f_q(\rho_q)-\lambda\rho_q]/T}. \tag{2.25}$$

Finding the most probable density is then an unconstrained maximization problem. The relation $df_q/d\rho_q = \lambda$ yields the same equation as (2.22), resulting in the same density profile (2.23) in the limit $T \to 0$. For $T > 0$, the density can be obtained from a numerical resolution of Eq. (2.22). We find that increasing the temperature T progressively blurs the zero temperature profile given by Eq. (2.23), eventually leading to a homogeneous density. The same effect has been observed in urban economics models, see De Palma and Papageorgiou [1988], Anas [1990]. As a consequence, the city is more spread, leading to a utility gain for agents near the city center, and to a loss for agents in the periphery.

2.4 Urban model with two types of agents

In this second model, two income groups are distinguished. Rich agents (group 1) have an income Y_1 and a surface of housing σ_1, while poor agents (group 2) have an income $Y_2 < Y_1$ and a surface of housing $\sigma_2 < \sigma_1$. Each block contains at most H agents, irrespective of their group. A configuration of the city is described by the densities $\rho_{q1} = n_{q1}/H$ and $\rho_{q2} = n_{q2}/H$ in each block q. The price P_{qi} an agent of group i pays for housing in block q depends on his

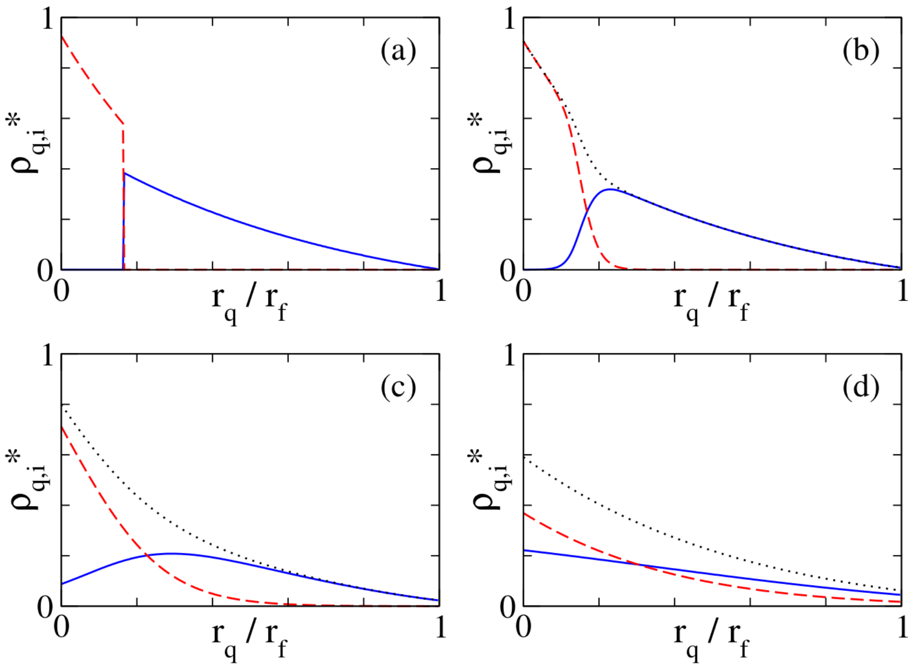


Figure 2.1: Density profile ρ^*_{qi} as a function of r_q for both groups of agents (rich, full line; poor, dashed line) for different temperatures: $T/T'_0 = 0.0018$ (a), 0.018 (b), 0.089 (c) and 0.36 (d), with $T'_0 \equiv p_0\sigma_2$. The dotted lines indicate the total density $\rho^*_{q1} + \rho^*_{q2}$. Parameters: $p_0 = 1.4$, $c = 0.4$, $Y_1 = 452$, $Y_2 = 301$, $\sigma_1 = 6$, $\sigma_2 = 4$, $r_f = 30$, $\overline{\rho}_1 = \overline{\rho}_2 = 0.13$.

surface of housing and on the local density of poor and rich agents:

$$\begin{aligned} P_{q1}(\rho_{q1}, \rho_{q2}) &= \sigma_1\,\tilde{p}(a_1\rho_{q1} + b_1\rho_{q2}) \\ P_{q2}(\rho_{q1}, \rho_{q2}) &= \sigma_2\,\tilde{p}(a_2\rho_{q1} + b_2\rho_{q2}) \end{aligned} \tag{2.26}$$

where a_1, b_1, a_2 and b_2 are given constants, and $\tilde{p}$ a function to be determined. The utility function of an agent of group $i = 1, 2$ in block q has the form

$$u_{qi}(\rho_{q1}, \rho_{q2}) = Y_i - cr_q - P_{qi}(\rho_{q1}, \rho_{q2}). \tag{2.27}$$

The model is analytically solvable if Eq. (2.11) is satisfied. For this condition to hold, one can choose $a_1 = a_2$ and $b_1 = b_2$. Then if $\sigma_1 b_1 = \sigma_2 a_2$, the function $\tilde{p}$ can take any form, for instance the logarithmic form Eq. (2.19) used in the previous model, in which case we get

(choosing $b_1 = \sigma_2$ and $a_2 = \sigma_1$)

$$P_{qi}(\rho_{q1}, \rho_{q2}) = \sigma_i \, p_0 \ln(1 + \sigma_1 \rho_{q1} + \sigma_2 \rho_{q2}). \tag{2.28}$$

The stationary distribution is given by Eq. (2.12), with

$$f_q(\rho_{q1}, \rho_{q2}) = l_q(\rho_{q1}, \rho_{q2}) + T s(\rho_{q1}, \rho_{q2}). \tag{2.29}$$

The expression of $s(\rho_{q1}, \rho_{q2})$ is given by Eq. (2.14), with $m = 2$. Expressing $l_q(\rho_{q1}, \rho_{q2})$ explicitly, we get

$$l_q(\rho_{q1}, \rho_{q2}) = \int_0^{\rho_{q1}} u_{q1}(\rho, 0) \, d\rho + \int_0^{\rho_{q2}} u_{q2}(\rho_{q1}, \rho) \, d\rho. \tag{2.30}$$

The validity of Eq. (2.10), as well as the symmetry of Eq. (2.30) with respect to ρ_{q1} and ρ_{q2}, can be checked using Eq. (2.11). The equilibrium densities $(\rho_{q1}^*, \rho_{q2}^*)$ are determined from Eq. (2.16), yielding a system of two non-linear equations, to be solved numerically. The results of this numerical resolution are shown on Fig. 2.1. One recovers at low temperature the standard separation, typical of north-american cities, between poor agents in the city center, and rich agents in the periphery (Fujita [1989]). The effect of a temperature increase is mainly to blur the zero temperature pattern, hence avoiding total segregation.

Therefore, Eq. (2.16) provides a direct prediction for the utility profile at arbitrary temperature T. It would be interesting to know whether this result remains valid beyond its a priori domain of validity, namely for models satisfying Eq. (2.11) so that a function $\tilde{L}$ can be defined. Considering again the above urban model with two types of agents, we keep the logarithmic form Eq. (2.19) for $\tilde{p}$, and choose as an example $a_1 = a_2 = 1$ and $b_1 = b_2 = 0$. These values imply $\sigma_1 b_1 \neq \sigma_2 a_2$ so that Eq. (2.11) is not satisfied, ruling out the possibility to find a potential function $\tilde{L}$ and to get a simple analytical solution of the model.

Performing numerical simulations of this agent-based model with two income groups, in the case of a one-dimensional closed city, we first validate the simulation thanks to a comparison with the above solvable case. Turning to the non-solvable case, we test the validity of Eq. (2.16), that is, whether the chemical potentials $\lambda_i = u_{qi} + T\partial s/\partial \rho_{qi}$ $(i = 1, 2)$ are uniform over the city for $T > 0$ (when $T \to 0$, the utility should be uniform anyhow). We indeed observe that for a non-zero temperature, the chemical potentials are homogeneous even in this non-solvable model, while the utility is not (see Fig. 2.2). We also note that although the

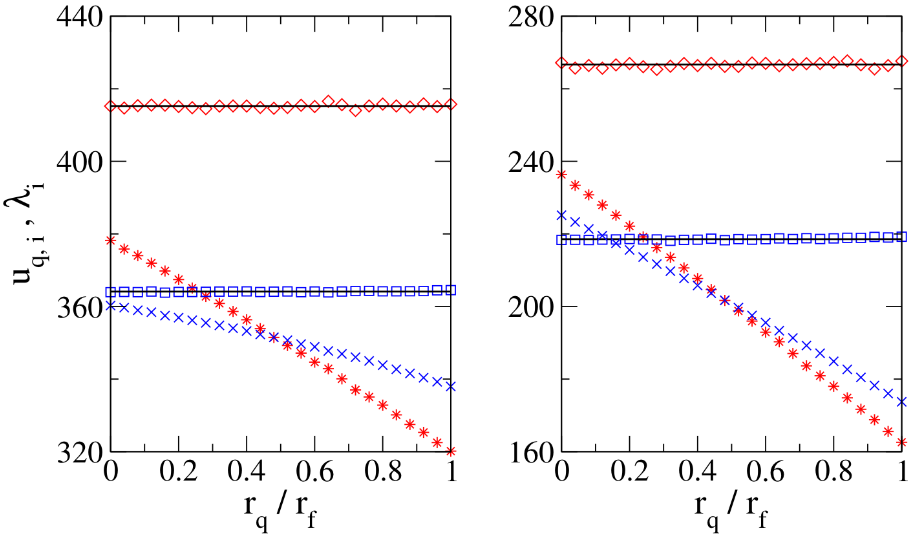


Figure 2.2: Agent-based simulations of the model with two types of agents (left: $i = 1$; right: $i = 2$), when no potential function $\tilde{L}$ exists (one-dimensional closed city). The utility u_{qi} is plotted as a function of r_q for two temperatures ($T = 100$, $*$; $T = 20$, $\times$). Adding to the utility the term $T\partial s/\partial \rho_{qi}$ yields the chemical potential λ_i ($T = 100$, $\diamondsuit$; $T = 20$, $\square$), found to be constant throughout the system although the analytical solution is not known. Full lines are horizontal lines indicating the spatially averaged value of the chemical potential. Parameters: $p_0 = 70$, $c = 2$, $Y_1 = 480$, $Y_2 = 300$, $\sigma_1 = 6.4$, $\sigma_2 = 4$, $r_f = 50$, $N_1 = N_2 = 2000$, $H = 200$, $Q = 51$.

number of agents is relatively small ($H = 200$), no systematic space-dependent correction to the chemical potential is observed.

We further use the obtained values of λ_1 and λ_2 to perform a numerical resolution of the system of non-linear equations (2.16), which we assume to be valid even in the absence of a potential L function. Interestingly, the results of the agent-based model and of the numerical resolution are found to be in very good agreement (see Fig. 2.3), thus yielding a complementary test of the validity of the chemical potential approach.

The validity of Eq. (2.16) in this case can be understood as follows. In this chapter, we focused on cases when the probability distribution has the factorized form Eq. (2.12), which is a consequence of the existence of a potential function $\tilde{L}$. When no function $\tilde{L}$ exists, the stationary distribution is no longer factorized, and we do not know its functional form. However, if the stationary distribution has only short range correlations, a chemical potential can still be introduced, in the same way as a chemical potential can be defined in a physical

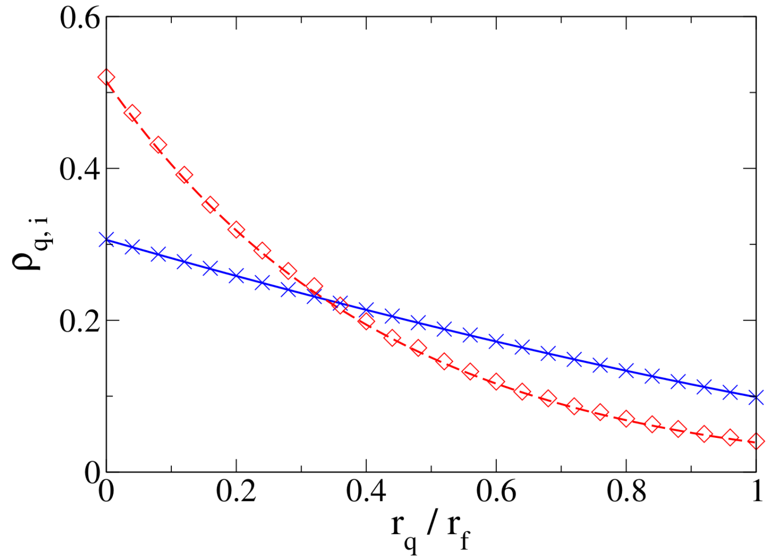


Figure 2.3: Comparison of the density profiles obtained by agent-based simulations ($i = 1$, $\times$; $i = 2$, $\diamond$) and by numerical integration ($i = 1$, full line; $i = 2$, dashed line) using the values of λ_1 and λ_2 measured in the agent-based model (see text). Same parameters as in Fig. 2.2 with $T = 20$.

system with short-range interactions (Balescu [1975]).

This result is also consistent with a phenomenon known in the nonequilibrium statistical physics literature as "restoration of detailed balance" (see Tauber et al. [2001], Bertin et al. [2004]). Namely, starting from a microscopic stochastic dynamics which does not obey detailed balance, a coarse-graining procedure can lead to a detailed balance relation in terms of the effective, coarse-grained, degrees of freedom. This phenomenon can appear in physical systems in cases where no macroscopic flux is present (like fluxes of energy or particles between boundary reservoirs), as macroscopic fluxes cannot be smeared out by the coarse-graining procedure. In the socio-economic models we consider in this chapter, there is no obvious macroscopic flux, which suggest that detailed balance may be restored on large scales.

2.5 Discussion

In this chapter, we have provided a clear relationship between the apparently unrelated notions of socio-economic utility and chemical potential. More specifically, we have shown that the uniformity of utility across the social system can be traced back to the conservation of the number of agents. This result not only provides a conceptually interesting link, but also yields non-trivial and testable predictions on the variations of utility among choices (e.g., sites, blocks) when $T > 0$. We also found numerical evidence that our result extends beyond the class of models in which it was initially derived. It would thus be interesting to explore further its validity through numerical simulations of more realistic models.

The idea of a non-uniform utility at equilibrium (Fig. 2.2) may be counter-intuitive for economists. Indeed, Nash equilibrium for homogeneous agents implies that all have the same utility, which seems not to be the case here when $T > 0$, since agents in the border of the city have a lower utility than those at the center. However, when noise is introduced in the decision process, a static equilibrium picture is no longer valid. Noise allows agents to explore the city, so that the *time average* value of utility is the same for all agents, leading to a macrostate described by Eq. (2.16) through the ergodic hypothesis linking time and ensemble averages. The average utility of agents is then a decreasing function of T. Note that this picture of a "time-averaged agent" is close in spirit to the notion of "representative agent" advocated in discrete choice theory (Anderson et al. [1992]). It would be interesting to investigate further the relation between these two approaches.

Another interpretation of our result is to consider the chemical potential λ_i as an effective utility. We first note that the distribution $\mathcal{P}(\{\rho_{qi}\})$ at $T > 0$ can be obtained from the zero-temperature distribution by replacing l_q by $f_q = l_q + Ts$ [see Eqs. (2.12) and (2.13)], in the same way as the macroscopic energy is replaced by the free energy in a physical system at finite temperature. Then, changing l_q into f_q in Eq. (2.10), we get an effective utility $u_{qi}^{\text{eff}} = \partial f_q / \partial \rho_{qi}$. Hence the Nash equilibrium of an assembly of fictitious agents having this utility would precisely correspond to Eq. (2.16), namely $u_{qi}^{\text{eff}} = \lambda_i$.

The next chapter will present a study dealing with the hypothesis made here as equation 2.19. Chapter 6 will present a study where the framework presented here is used in the context of a more complex model, in an attempt to bridge the gap between this study motivated by statistical physics and research questions coming from economics.

CHAPTER 3

A probabilistic model of housing price formation

3.1 Introduction

The origin of this work is a question raised in the previous chapter (chapter 2, see also Lemoy et al. [2011a]), where a dependency of housing price on density of inhabitants is postulated. This hypothesis is in agreement with the results of the standard urban economics model (also named AMM model), which will be presented in the following chapter (chapter 4) and is described in detail in Fujita [1989]. But the idea here is to have such a dependency emerge directly from simple interactions between agents rather than from an already subtle utility function such as that usually used in the AMM model.

To this end, we build a simple model of urban housing market. An ergodic hypothesis allows us to link the time variations of housing price for a given flat with the spatial distribution of housing prices in the city. That is, the study of only one representative flat is interpreted in terms of the evolution of a whole city. The price of this flat increases or decreases depending on the occupation of the housing lot, with agents moving in and out with given probabilities.

This model provides a very simple tool to study the influence on housing price of different factors associated with schematic, coarse-grained behaviors of landlords and tenants. This model is described in section 3.2. Section 3.3.2 describes the evolution of the system and the equilibrium reached is presented in section 3.3.3. The main parameter we consider is the density of agents in the city. Its influence on our urban system is studied in section 3.3.4.

Though with a different approach, more schematic and closer to analytical treatment, this work can be related to simulations of housing search models such as Mc Breen et al. [2010].

3.2 Simulating a simple housing market

In this section we give a description of the model and its dynamics. The model simulates the evolution of the price of a flat, which is seen as representative of a virtual city consisting of agents and housing lots or flats, thanks to an ergodic hypothesis. The representative flat can be in only two states: either occupied or empty. When it is occupied, there is a probability for the tenants residing in it to move out, and when it is empty, there is (another) probability for some tenants to move in.

In addition, the price tends to rise when the flat is occupied, as the rent is revised periodically by the owner. When the flat is empty, its price tends to decrease until some agents move in. These simple ingredients are related to and consistent with the dynamics of the more complex models presented in the next chapters. However, this work is more closely related to Mc Breen et al. [2010] as the model is not spatial. There is no city center to polarize space and govern housing prices and population densities.

The model is simulated in discrete time and has two time scales, a short one which we will call a "day" for convenience, though it has no real relationship to a 24-hours day. And the long one will be called a year, also for convenience. Let us first describe the dynamics of the model during a "day".

3.2.1 Days are short...

Days correspond to the short-term dynamics of the model. At the beginning of a day, the flat is either occupied or empty. The other variable characterizing the flat is its price p, i.e. the rent that tenants pay to occupy it. The flat is seen as representative of an ensemble of housing lots which are identical except for their price. Let us consider the case where the representative flat is occupied.

Occupied flat

If the flat is already occupied, there is a given probability π_r that the rent is raised by the owner. In case the price is increased, it is multiplied by a given factor $f_r > 1$.

Then, during the same day, there is a probability π_s, independent of the fact that the rent has been raised or not, that the tenants search for another flat in the city with a more interesting price, that is to say, an inferior one. Indeed, rent is the only difference between housing lots in the city.

This probability π_s depends linearly on the price p of the flat: $\pi_s = \mathrm{Min}(p/p_s, 1)$, where p_s is a fixed price giving a scale of prices. Taking the minimum between p/p_s and 1 allows to forbid the probability to be higher than 1. When the price of the flat p is high, that is to say, close to p_s, the probability that tenants look for another flat is high. When p is much smaller than p_s, this probability is low.

Let us now describe how tenants search for another flat in the city, in the case where this event is drawn. As all housing lots are similar except for their price, this search is equivalent to looking for a cheaper rent. Hence, a rent p_d is drawn from a price distribution $\psi_k(p)$ describing the city, where k is an index corresponding to the year. This distribution $\psi_k(p)$ is fixed during a year and will be presented in the next section (3.2.2). If $p_d < p$, tenants move out of the representative flat, which is then empty. Else, if $p_d \geq p$, they remain in the representative flat. The day is finished, the night falls on a representative flat that is empty in the first case or occupied in the second.

Empty flat

If the flat is empty at the beginning of a day, the price of the flat is lowered by the owner with a probability π_l. This probability depends also on the price p of the flat: $\pi_l = \mathrm{Min}(p/p_l, 1)$, where p_l is fixed and gives another scale of prices. Here also, π_l cannot be greater than 1. When the price p of the flat is close to p_l, the probability that the owner lowers the price is high. When p is much smaller than p_l, this probability is low. In case the price is lowered, it is multiplied by a given factor $f_l < 1$.

Then as the day goes on, some agents can visit the flat and possibly move in. The number n of visits is given by a Poisson distribution $P_k(n)$, where k labels the year. This distribution remains fixed during a year:

$$P_k(n) = \frac{(\lambda_k)^n}{n!} e^{-\lambda_k}$$

Its parameter, giving its mean and variance, is a constant throughout a year and is denoted by λ_k. This parameter is presented in the next section.

For each of the n visits successively (n can be zero, in which case there is no visit and the flat remains empty), a rent p_d is drawn in the same distribution $\psi_k(p)$ that determines whether tenants move out of the representative flat or not. This rent represents the price that the visiting agents pay for their current flat, and this rent is compared to the price p of the representative flat. If for (at least) one of the visits, $p_d > p$, then some agents move in the representative flat, which becomes occupied. For the next day, the representative flat is then occupied if some tenants have moved in, or stays empty. A summary of the dynamics of a day is given on figure 3.1.

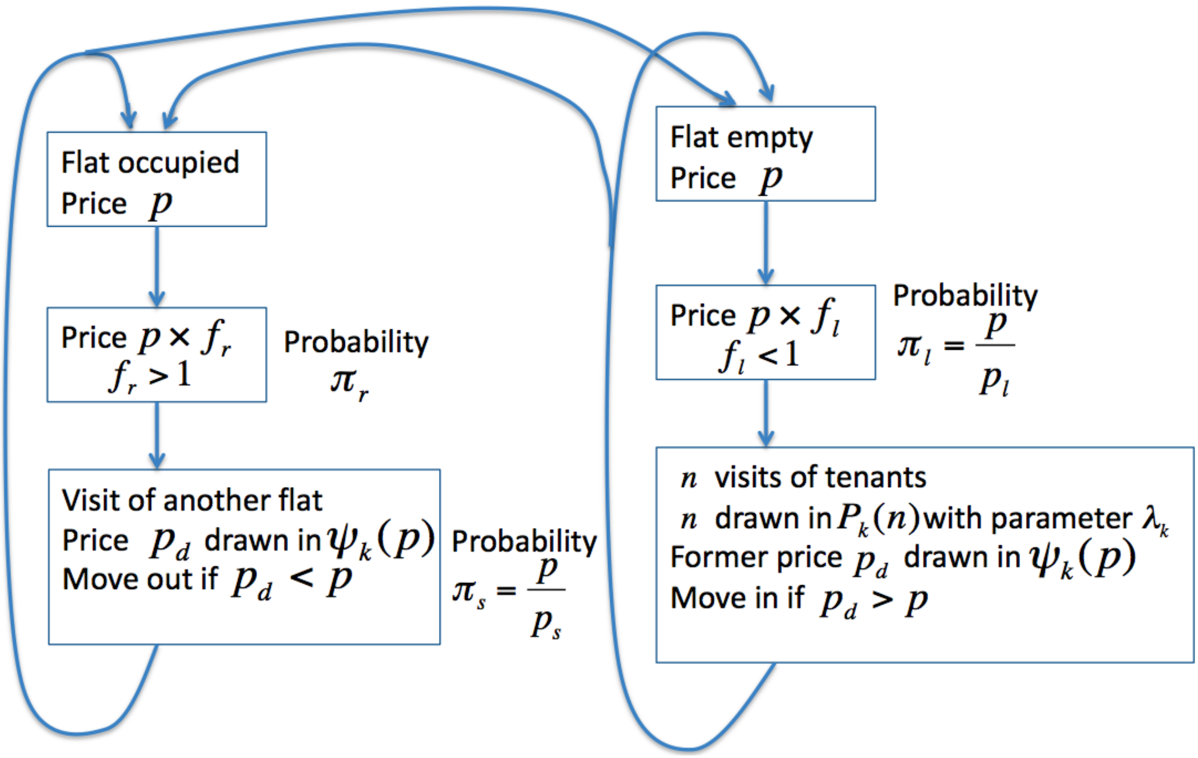


Figure 3.1: Summary of the daily evolution of the model. Note: actually, $\pi_s = 1$ if $p/p_s > 1$, and $\pi_l = 1$ if $p/p_l > 1$.

We describe now what happens during the long time scale of the model, corresponding to a "year".

3.2.2 ...Years are very long

A "year" consists in this model of a high number Y of days, for instance $Y = 10^6$ days.

Long term dynamics

During a year, the distribution in time of the price of the representative flat is observed. It is seen as representative of the space distribution of prices in the city thanks to an ergodic hypothesis. The price distribution $\psi_k(p)$ introduced at the beginning of a year, and remaining constant during a year, represents a urban distribution of prices of housing lots which are identical except for their price. With the dynamics of a "day" described in the previous section (moves in and out of the representative flat, price increases and decreases), repeated a high number of times, the price distribution $\psi_k(p)$ gives rise to a time distribution of the price p of the representative flat.

This time distribution is stored and determines $\psi_{k+1}(p)$, the urban distribution of prices for the next year. In other words, $\psi_k(p)$ is measured as the time distribution of price of the representative flat at year $k-1$.

The long term dynamics determines also how flat visits are distributed through the parameter λ_k. This parameter is also governed by the time evolution at year $k-1$. Indeed, at the end of each year k, the mean rate of flat visits ν_k is computed. It corresponds to the share of days where the representative flat was occupied, and where the tenants living in it visited another flat in the city (as described in section 3.2.1). This rate of visits ν_k determines the Poisson distribution parameter of the next year λ_{k+1}, up to a multiplicative factor:

$$\lambda_{k+1} = \nu_k \frac{N_a}{N_l - N_a} = \nu_k \frac{\rho_s}{1-\rho_s} \tag{3.1}$$

where $\rho_s = N_a/N_l$ is the ratio of the number N_a of agents to the number N_l of housing lots in the virtual city. This formula is motivated by the fact that the mean number $\nu_k N_a$ of agents visiting flats shall be equal to the mean number $\lambda_k(N_l - N_a)$ of flat visits for the housing market to be consistent and representative of a closed city.

To complete the summary of the dynamics provided by figure 3.1, one needs to imagine the whole figure included in a loop corresponding to years. For each year k, λ_k and $\psi_k(p)$ are obtained thanks to statistics on the time evolution of the previous year, $k-1$.

Initialization

We already presented the dynamics of the model, but did not speak about its initialization. At each beginning year k, the state of the representative flat is "occupied" with a probability which is the occupation rate of the flat during the previous year. The price of the flat is drawn in the distribution $\psi_k(p)$.

For the first year $k = 1$, the representative flat is empty and the price distribution $\psi_1(p)$ is a given normal (Gaussian) distribution. The parameter of the Poisson distribution of the first year λ_1 is taken as $\lambda_1 = \rho_s$. However, our results show that this system reaches after several years an equilibrium that is independent of the initial price distribution and of λ_1.

3.3 Results

3.3.1 Discussion on the ergodic hypothesis

As a preliminary result, a study is needed on the ergodic hypothesis of the model. Indeed, if we want to consider one flat which is representative of a closed city composed of numerous flats, a condition linked to equation (3.1) should be verified.

Two densities of agents are present in the model. The first one is the temporal density ρ_{temp}, associated to moves into and out of the representative flat. This flat is occupied during a certain proportion of time, given by the ratio of the number of days during which the flat is occupied to the total number of days, for a given year. The second density is the "spatial" density ρ_s of agents in the virtual city. It corresponds to the ratio of the number of agents to the number of flats. The consistency of the ergodic hypothesis and of the virtual closed city requires that on average $\rho_s = \rho_{\text{temp}}$. However, it is not what happens in the simulations, as figure 3.2 shows. Although ρ_s and ρ_{temp} have close values, the mean value of the temporal density is not equal to the spatial density, once the system is set free of the initial condition. Exploratory simulations show us that the gap between both quantities depends on the values of parameters.

This phenomenon is linked to the fact that a flat visit does not automatically result in a move. Indeed, for each visit, a price is drawn in the distribution $\psi_k(p)$, and depending on the result of this draw and on the price of the representative flat, the move happens or not (see figure 3.1). Condition (3.1) insures that one flat can represent the dynamics of a whole city in terms of visits, but not in terms of moves, which explains the result of figure 3.2.

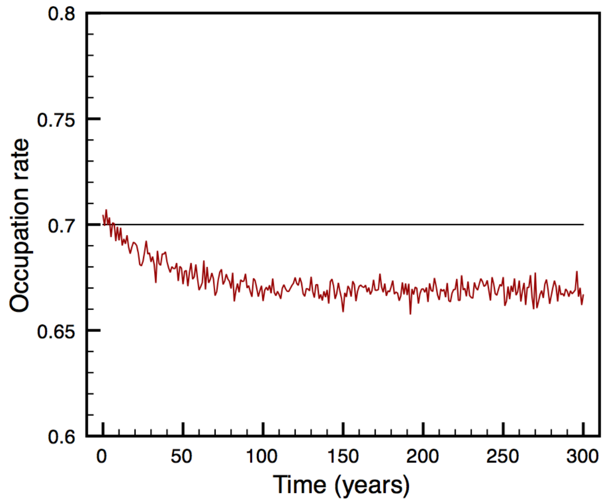


Figure 3.2: Evolution of the temporal density ρ_{temp} over years. The horizontal line indicates the spatial density $\rho_s = 0.7$. Parameters: $p_s = p_l = 400$, $\pi_r = 0.03$, $f_r = 10^{11/1000} \simeq 1.026$, $f_l = 10^{-19/1000} \simeq 0.957$, $Y = 2.5 \times 10^6$.

We perform simulations of the same system where price is absent, that is, a visit results automatically in a move. In this case, we check that the equality $\rho_s = \rho_{\text{temp}}$ is verified on average, as illustrated by figure 3.3, which uses the same parameters as figure 3.2. We perform additional simulations. If visits of agents wishing to move out of the representative flat have a fixed probability to result in a move, and visits of agents wishing to move into this flat have a different fixed probability to result in a move, then the equality is not true. This preliminary result is a problem for our simple interpretation using ergodicity. Further work is needed to deal with this problem in a satisfactory way.

3.3.2 Evolution of the price distribution

The output of main interest of this model is the price distribution, as we want to study a simple housing market. This price distribution is measured as the time distribution of the price of the representative flat, and can also be seen as the distribution of urban price in a simulated city. We present on figure 3.4 its typical evolution during a simulation.

The evolution of this system corresponds to a relaxation to an equilibrium. During the first year, the price variations are driven by the initializing price distribution $\psi_1(p)$. The system is constrained, and the resulting price distribution $\psi_2(p)$ has a shape that is different

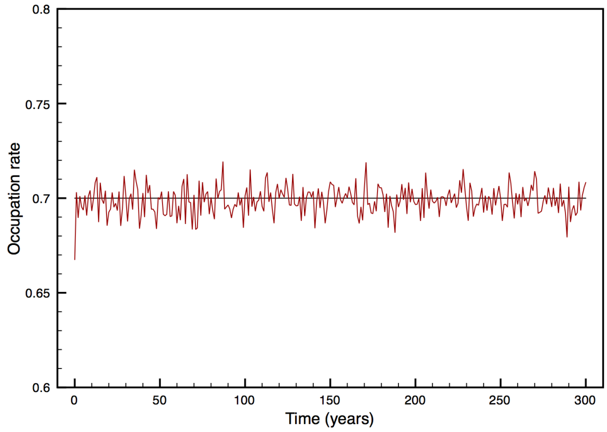


Figure 3.3: Evolution of the temporal density ρ_{temp} over years when the occupation of the flat is not related to price. The horizontal line indicates the spatial density $\rho_s = 0.7$. Same parameters values as figure 3.2

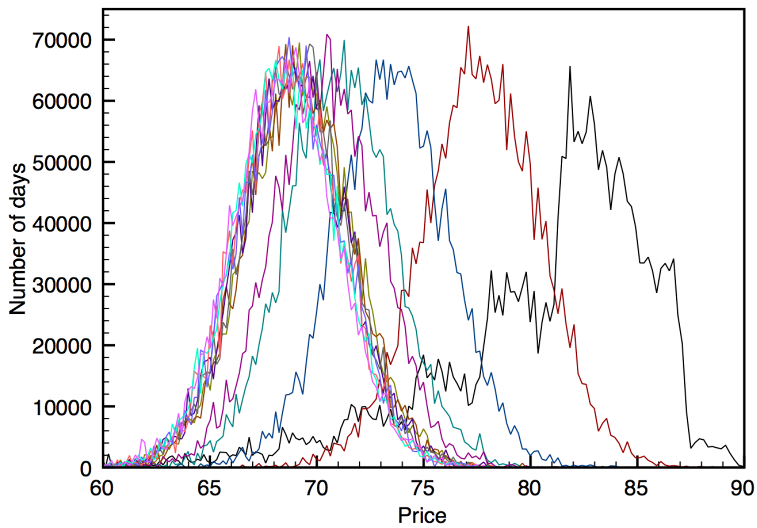


Figure 3.4: Evolution of the price distribution over years. The vertical axis gives the number of days spent at a given price during the year. The successive curves appear from right to left and are separated by 30 years. The first curve on the right corresponds to $\psi_2(p)$. Parameters: $p_s = p_l = 400$, $\pi_r = 0.03$, $f_r = 10^{11/1000} \simeq 1.026$, $f_l = 10^{-19/1000} \simeq 0.957$, $\rho_s = 0.7$, $Y = 2.5 \times 10^6$.

from the distributions of following years. Then the model evolves freely. On figure 3.4, the price decreases progressively, until it reaches an equilibrium distribution, which we will further study in the next section (3.3.3). The corresponding distribution of occupation prices, which presents the distribution of price where only the days during which the flat is occupied are taken into account, will also be studied.

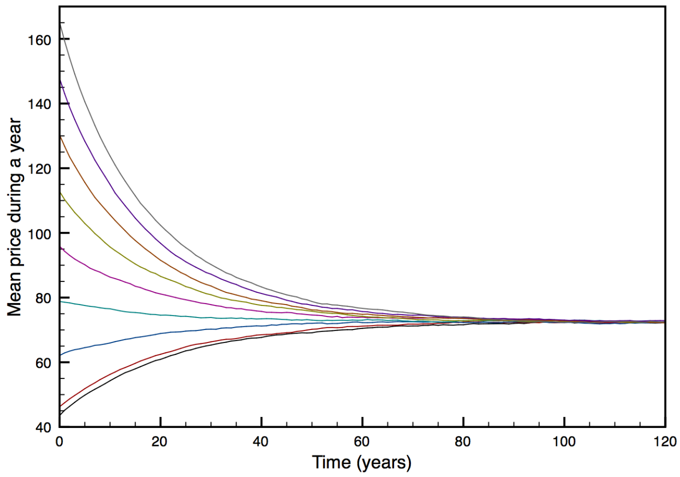


Figure 3.5: Evolution of the mean price (during a year) over years, with different initial price distributions. Parameters: $p_s = p_l = 2000$, $\pi_r = 0.03$, $f_r = 10^{11/1000} \simeq 1.026$, $f_l = 10^{-19/1000} \simeq 0.957$, $\rho_s = 0.7$, $Y = 2.5 \times 10^6$. The initial price distribution is a normal (Gaussian) distribution of variance 10 and of mean $20, 40, ..., 180$.

It can be verified that the equilibrium distribution does not depend on the initial price distribution. As an illustration of this fact, figure 3.5 shows that the system converges to the same equilibrium price up to fluctuations for different initial price distributions.

3.3.3 Equilibrium

As the equilibrium price distribution does not depend on the initial conditions, we study it more closely here. Once the system has reached its equilibrium, we average the price distribution over several years to obtain a smooth curve, corresponding to a statistical representation of the equilibrium (see the left panel of figure 3.6).

The first question asked then is the occupation of flats in the city (or equivalently, of the representative flat): depending on its price, what is the mean share of time during which a flat is occupied, when the market has reached its equilibrium?

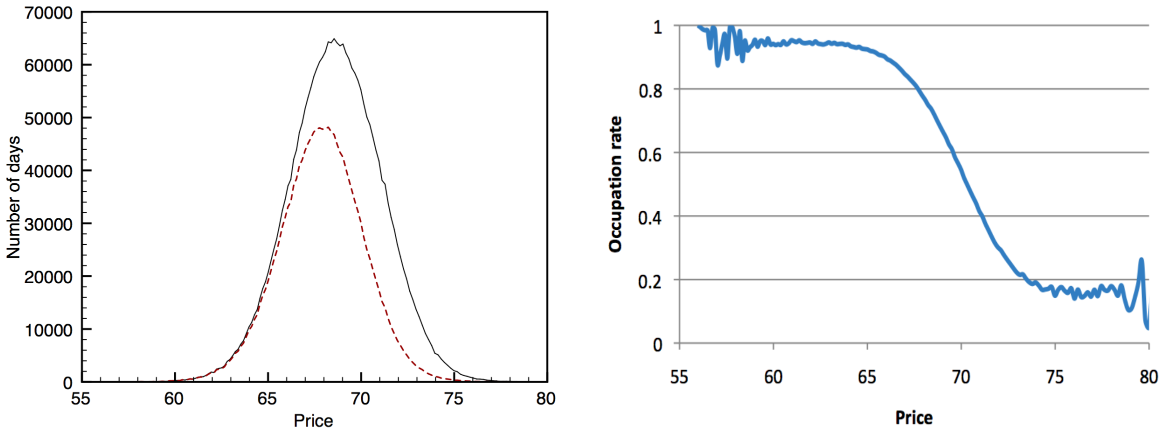


Figure 3.6: Left panel: price distribution (for all flats, black continuous line) at equilibrium, and corresponding price distribution of occupied flats only (red dashed line). The curves are averaged over 50 years, and the vertical axis gives the number of days spent in average at a given price during a year, when the system is at equilibrium. Right panel: corresponding occupation rate as a function of price. Parameters: $p_s = p_l = 400$, $\pi_r = 0.03$, $f_r = 10^{11/1000} \simeq 1.026$, $f_l = 10^{-19/1000} \simeq 0.957$, $\rho_s = 0.7$, $Y = 2.5 \times 10^6$.

The answer to this question is provided by the right panel of figure 3.6. This figure illustrates the dynamics of the model, and corresponds to a simple price formation mechanism. When the price is low, the occupation rate is high, and vice versa. Occupation is governing the price of flats in this model where price is the only variable allowing to distinguish housing lots in the city.

3.3.4 Dependency on the "density" ρ_s

In this section we explore the results of the model depending on the density parameter ρ_s. This parameter is the most physical one and has been shown in the previous chapters to be of a special interest for a physical treatment of simple social models. The variations of the equilibrium price distribution with ρ_s are presented on figure 3.7. The scale of the abscissa is logarithmic. Indeed, it is the most appropriate way to represent the results of this model, as the considered price variations are exponential: the price is multiplied by a factor $f_r > 1$ when it is raised, and by $f_l < 1$ when it is decreased.

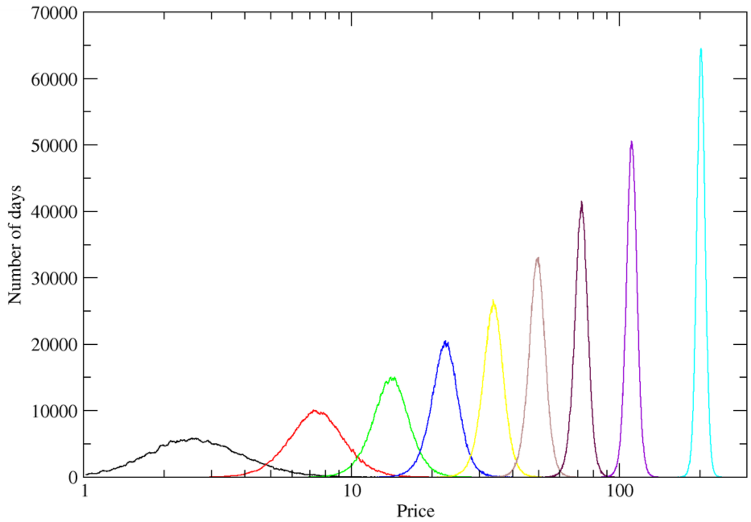


Figure 3.7: Equilibrium price distribution, for different values of the global density ρ_s. Parameters: $p_s = p_l = 2000$, $\pi_r = 0.03$, $f_r = 10^{11/1000} \simeq 1.026$, $f_l = 10^{-19/1000} \simeq 0.957$, $Y = 2.5 \times 10^6$. From left to right, $\rho_s = 0.1, 0.2, ..., 0.9$.

With this logarithmic abscissa, the area under each price distribution curve of figure 3.7 is the same. It can be linked to the number of "days" in a "year" (see also the caption of the left panel of figure 3.6, which has the same vertical axis). These price distributions are well fitted by log-normal distributions. This result can be seen as quite intuitive as the log-normal distribution corresponds under certain conditions to the multiplicative product of many independent random variables.

Two main observations can be drawn from this study on the influence of density. The first is the fact that the price of housing increases when the density ρ_s increases. This is a stylized fact which is well-known in real cities (see Anas et al. [1998]). This stylized fact is also reproduced by the AMM model (which will be presented in the next chapter) and by the model we presented in the previous chapter, though in a different context, where a central business district (CBD) polarizes space and distinguishes as a consequence different urban locations. In the simple model presented here, this phenomenon can be interpreted as the fact that a higher density ρ_s yields a higher competition on the urban housing market, and thus higher prices.

The second observation which can be made on figure 3.7 is the decreasing width of the price distribution when the density ρ_s increases, in logarithmic abscissa. This can be linked to the exponential variations of the price and interpreted as follows. Let us suppose that price p is such that $p < p_s$ and $p < p_l$. For high prices, an increase or a decrease in price p (multiplication by $f_r > 1$ or $f_l < 1$) generates a higher absolute variation of the probability $\pi_s = p/p_s$ that the tenant tries to move out of the representative flat when it is occupied and of the probability $\pi_l = p/p_l$ that the price of the representative flat decreases when it is empty. These higher variations prevent the price from increasing (π_r) or decreasing (π_s) too much. Indeed, the reactions to price variations are quicker for higher prices, and the price is bound to stay within a smaller domain, in logarithmic scale[1].

3.4 Conclusion and perspectives

The work presented in this chapter studies a model of urban housing market. Using an ergodic hypothesis, we identify the variations in time of the price of a given flat with the spatial price variations of a fictitious city. The mechanisms of price evolution are simple and correspond to intuitive dynamics in urban economics: flats which are highly demanded have an increasing price, whereas flats for which agents have a low interest have a decreasing price. This can be seen as a rough representation of the mechanisms which determine the price of flats in the standard urban economics (or AMM) model, which is exposed in the next chapter. It provides us with an illustration or a confirmation of the hypothesis introduced in the previous chapter concerning the increase of housing price with population density (equation 2.19).

As section 3.3.1 showed, further work is needed to have a consistent ergodic hypothesis within this model. One of the main interests of this ongoing work is its simplicity, which should allow to find analytical results in some limiting cases. To this end, a perspective consists in linking this simulation model to analytical treatment of "ratchet" models, which are well-known in out-of-equilibrium statistical physics (see for instance Bena [2006]).

Another perspective is to make this model slightly more complex. It is for instance possible to introduce a cost of moves Δp in the model: agents move into or out of the representative flat if $p_d - p > \Delta p$ (respectively $< \Delta p$). This should have a friction-like effect on the housing market, which could be studied in the frame of this simple model. Moreover, differentiation between flats could allow to link this model with the literature on urban economics. This would produce models which are intermediary between this simple probabilistic model and

[1] In logarithmic scale, the different price values encountered in the model are uniformly distributed on the abscissa of figure 3.7.

the urban economics model presented in the next chapter.

Let us now begin the second part of this thesis, dealing with agent-based models in urban economics.

Part II

Agent-based simulations in urban economics

Chapter 4

An agent-based model of residential patterns and social structure in urban areas

Introduction

The work presented in this chapter has been published as Lemoy et al. [2010a].

Agent-based models have been widely used to simulate traffic at a microscopic level. The goal of the work presented here is to use this tool in Urban Economics to deal with the research question of the location of households with respect to their income.

There are in the literature numerous analytical works on the Urban Economics model of Alonso, Muth, Mills, studying the factors explaining the location choices of households (Brueckner et al. [1999], Gofette-Nagot et al. [2000]). Exogenous amenities have been added to the model (Wu and Plantinga [2003]) to determine their influence.

This work uses a simulation tool which is more and more widely used in social sciences: agent-based programming. This tool allows us to study economic agents living in a predefined simulation space. These agents interact in a simple way, and from their interactions, collective behaviours emerge, which are difficult to predict in an intuitive way or to compute analytically. Agent-based models have three main components: agents, an environment and rules of behaviour. The agent has internal states, some fixed and others that can change, like preferences, and rules of behaviour. The environment is a two-dimensional space supporting resources and can also be a communication network. It is a device that is separated from the agents and that interacts with them. Rules of behaviour determine the interactions between agents, between agents and the environment and within the environment. For this work we use the multi-agent programming language NetLogo and the integrated modeling environment that bears the same name.

The study presented here is based on a model which has been widely studied, the standard Urban Economics model developed by Alonso, Muth, Mills. This model is exposed in section 4.1. Numerous works have been led on this model, theoretical works and also empirical ones, so that this model is suitable to build an agent-based model on it. The first stage of this work is to reproduce the classical results of the Urban Economics model. To build an agent-based model, the analytical model has to be discretized: there is a finite number of agents who interact individually. This is a difference with the analytical model, which is continuous, but it can be argued that social systems are indeed built on individual interactions of a finite number of agents. This even led some authors to study discrete versions of the Urban Economics model, but with different simplifications in order to be able to solve the model analytically. A comparison between the results of the simulations and those of the continuous analytical model is presented in section 4.2.

The use of agent-based models allows us to handle easily agents' states, rules of behaviour and environment. Sets of agents such as neighborhoods can be used, so that it is easy to introduce neighborhood effects. And individual and collective behaviours can be monitored in a simple way. In addition, the simulations and the model are dynamical, which is not the case for the standard urban economics model and for most analytical economic models, which are equilibrium models. This time dependence allows us to see the equilibrium emerge from the interactions between agents, which is described in section 4.3, or to study out of equilibrium dynamics. The emergence of an equilibrium can have no relationship with the historical evolution of a real city, but the interactions between agents can also be used to study specifically features of cities which are linked to their historical evolution. As the results of the agent-based model are validated by the comparison with the analytical model, the model can be made more complex by adding different ingredients, and firstly income groups. This allows us to explore phenomena which are difficult or impossible to treat analytically. This work is presented in section 4.4 with the introduction of transport time costs and amenities. The introduction of transport time cost can change the location pattern of income groups under a certain condition. Amenities for which income groups have the same preference have a complex influence on the city, but do not inverse the location pattern of income groups. Hence differentiated preferences for the amenities are introduced, leading to interesting two-dimensional effects. In section 4.5, several perspectives of this work, for instance the historical evolution of endogenous amenities, are presented.

4.1 Model

4.1.1 Standard Urban Economics model

This model has been developed by Alonso, Mills and Muth to study the location of agents representing households in a city where they compete for land. These agents have daily monetary and time costs to access their workplace, which is located in the central business district (CBD) of the city. Hence this model is called monocentric model. They can also have preferences for amenities in later versions of the model (Wu and Plantinga [2003]). In the simplest version of the model, agents are single workers, but they can represent households in other versions. They rent their housing, and the landowners rent to the highest bidder. Agents compete for land use with one another and with an agricultural land use, represented by an agricultural rent R_a, which is one of the parameters of the model.

Agents have an utility $U = \alpha \ln z + \beta \ln s$, where z is a composite good representing all consumer goods except housing, s is the surface of housing, α and β are parameters describing agents' preferences for composite good and housing surface. These last parameters are chosen so that $\alpha + \beta = 1$, because it makes analytical calculations easier. This "log-linear" form of the utility is often used and we use it in the agent-based model, but analytical calculations are also performed on a Cobb-Douglas function (the exponential of this log-linear function), which gives roughly the same results, or even without specifying the concrete form of the utility function. Agents also have a budget constraint $Y = z + tx + ps$, where Y is their income, t the transport cost per unit distance, x their distance to the CBD and p the price per unit surface of their housing.

The model which is reproduced here with an agent-based system is a closed city model (see appendix A): the population N of the city is chosen exogenously and remains constant. This model can be solved analytically for a two-dimensional space if $R_a = 0$ (see Fujita [1989]). For $R_a > 0$ it can be solved numerically for one income group. With a population divided into several income groups, one needs to build a specific algorithm for the resolution of the model, whose general form is described in Fujita [1989]. We did not write such an algorithm, so that we will only be able to compare quantitatively the results of the agent-based model to the results of the standard model with one income group. But the simulations presented here will also consider the case of two income groups.

The standard urban economics model (or AMM model for Alonso, Muth and Mills) and its results are presented in a more detailed way in numerous works, for instance Fujita [1989] or Fujita and Thisse [2003]. Appendix A presents its analytical resolution.

4.1.2 Agent-based model

In the agent-based system, the space is a two-dimensional grid and each cell can be inhabited by one or several agents, or be used for agriculture. The CBD is represented by a point at the center of the simulation space.

At the beginning, the city population N is fixed and agents are placed at random locations. They can be divided into two income groups or not. Prices are all equal to the agricultural rent $p_0 = R_a$. Agents choose the surface of housing which provides them with the higher possible utility at a given location with price p_0:

$$s = \beta \frac{Y - tx}{p_0}$$

This determines the quantity of composite good they consume and also their utility. This expression for the optimal surface can be obtained by differentiating with respect to surface s the agents' utility at fixed location and price:

$$\begin{aligned} U = \alpha \ln \big(Y - tx - p_0 s\big) + \beta \ln s \\ \frac{\partial U}{\partial s} = \frac{-\alpha p_0}{Y - tx - p_0 s} + \frac{\beta}{s} \\ \frac{\partial U}{\partial s} = 0 \text{ if } \alpha p_0 s = \beta(Y - tx - p_0 s) \end{aligned}$$

Hence the expression of the optimal surface, with $\alpha + \beta = 1$.

Moves

Agents can move with no cost. Iteration n of a simulation, changing the variables from their value at step n to their value at step $n+1$, goes as follows. An agent which will be candidate to a move and a cell are chosen randomly. The price of this cell, located at a distance x of the center, is p_n. The optimal surface that the agent can choose in the candidate cell is $s = \beta\frac{Y-tx}{p_n}$, which allows us to compute the utility that the agent would get thanks to the move and to compare it to its current utility.

If the agent candidate can gain utility $\Delta U > 0$ by moving in the candidate cell, he moves. In this purpose he raises the price of the candidate cell by proposing a bid

$$p_{n+1} = p_n(1 + \epsilon\frac{s_{\text{occ}}}{s_{\text{tot}}}\frac{\Delta U}{U})$$

where ϵ is a parameter that we introduce to control the magnitude of this bid. Prices evolve quicker if this parameter is high. s_{occ} is the surface occupied by other agents in the cell and s_{tot} the total surface of the cell. The factor $\frac{s_{\text{occ}}}{s_{\text{tot}}}$ makes the bid higher if the cell is more occupied, that is to say, attractive. The first agent to move in an empty cell does not raise the price. The price is a price per unit surface and it is a variable linked to a cell. When an agent bids higher, the price is changed immediately for all agents in the cell. Their surface of housing is also changed according to $s = \beta\frac{Y-tx}{p_{n+1}}$ and their utility is computed again.

An agent may want to move in a candidate cell that is already full. This case is treated in the next paragraph.

Displacement and time decreases

In the situation where an agent wants to move into the candidate cell, proposing a higher bid, but this cell is already full, we need a mechanism to still allow prices to evolve, in order to have a fluid housing market and to reach the equilibrium. Then two options are available: either the price of the cell is raised for agents living in it, without the candidate moving in (and he needs to wait until he is offered an attractive cell where there is enough space available), or the agent moves into the cell and then some agents previously in the cell need to be relocated. We choose the second solution.

Thus, one or more agents are randomly chosen within the candidate cell and are sent “to the hotel”, until there is enough space for the agent candidate to move in. These agents which have been displaced are then the next candidates for a move. And they stay priority candidates until they have all found a new location. While they are searching for another cell to live in ("at the hotel"), their utility decreases exponentially following $U_{n+1} = U_n - (U_n - U_{\min})/T_e$, where U_n is the initial utility, U_{n+1} the decreased utility, $U_{\min}$ the minimal utility in the city and T_e a parameter governing the speed of this decrease. Thanks to this mechanism, agents at the hotel find a new cell to live in.

There is another comparable mechanism of decrease in this model: the time decrease of prices of vacant cells. With the bid mechanism presented before (agent and cell chosen at random), the price of a cell where several agents move in successively can increase and reach a value which makes the cell unattractive. In this case, the agents will progressively move to more attractive cells. We choose to decrease exponentially the price of cells which are not completely full[1]. The decrease formula is

$$p_{n+1} = p_n - (p_n - R_a \times 0.9)\frac{\Delta u}{T_p}\frac{s_{\text{free}}}{s_{\text{tot}}}$$

where $\Delta U = (U_{\max} - U_{\min})/U_{\max}$ measures the homogeneity of the utility in the city, T_p is a parameter determining the speed of decrease of prices, $s_{\text{free}} = s_{\text{tot}} - s_{\text{occ}}$ and s_{tot} are the non occupied surface and the total surface of the cell. If no agents moves in, the price decreases according to this formula until it reaches the agricultural rent, where the decrease stops: the cell is then used for agriculture. The factor $\frac{s_{\text{free}}}{s_{\text{tot}}}$ makes this decrease quicker as the cell is emptier and thus less attractive. The ΔU factor makes the decrease slower as the utility becomes more homogeneous, that is, when the system gets closer to the equilibrium.

The different parameters of the model are listed in table 4.1. Their default value has been chosen according to several criteria.

First for the parameters of the model itself: α, β, Y, t, N, R_a, s_{tot}. Their value has been chosen mainly for technical reasons regarding the comparison between the analytical and the agent-based model, but naturally other values could have been chosen, without changing the qualitative behaviour of the model. The calibration of the model with a set of values of parameters that are in agreement with empirical studies is one of the perspectives of this work and will be discussed in section 4.5.

The parameters ϵ, T_e and T_p, which are specific to the agent-based model, have been chosen

[1] With two income groups, it is difficult to determine whether a cell is full or not: we choose to let the price decrease if the mean surface of housing s_{mean} of agents there is smaller than the free surface of the cell s_{free}.

such that the competition between agents on the housing market is efficient and the system reaches the equilibrium in the whole city. The agent-based model has different behaviours and for instance does not reach an equilibrium (the utility of agents does not become completely homogeneous across the city) for certain values of these parameters, but the study of these different behaviours is beyond the scope of the present work.

Additional parameters γ, a_0 and d are introduced to describe the amenities studied in paragraph 4.4.3.

Parameters	Description	Default value
α, β	Preferences for composite good and housing	0.9; 0.1
Y	Income	300
t	Transport cost (unit distance)	5
N	Population	700
R_a	Agricultural rent	10
s_{tot}	Surface of a cell	12
ϵ	Bidding parameter	20
T_e	Time decrease of the utility of displaced agents	7000
T_p	Time decrease of the price of non-full cells	15
γ	Preference for the amenity	0.1
a_0	Attractiveness of the amenity	4
d	Range of attraction of the amenity	2

Table 4.1: Parameters of the model

4.1.3 Socioeconomic outcome

In order to explore the outcomes of the models developed here, we study some variables extracted from these models. By comparison with the values of these variables in a reference simulation (a simulation of a monocentric city with two income groups with the same transport cost, without amenity), we can observe the socioeconomic effects of the modifications introduced in the standard model.

We study mainly some variables of the models: the utility of the individuals, which is associated to their welfare and gives an economic outcome of the models, the cumulated distances of agents' trips to work, which give the environmental outcome of the models, and the social inequalities, which are given by observing the difference in the utility of individuals of different income groups.

4.2 Comparison with the analytical results

As we will see, the simulations allow us to reach the equilibrium of the standard Urban Economics model of Alonso, Muth, Mills. This equilibrium corresponds to a situation where no agent can raise his utility by moving, and therefore no agent has an incentive to do so. In each income group, individuals have an identical utility across the city. The cells which are occupied are those closest to the city center (CBD), the prices at the border of the city are equal to the agricultural rent and prices increase when the distance to the center decreases. The surfaces of housing increase when the distance to the center increases.

The results of the simulations do not match exactly the analytical results because of the effects of the discretization (which leads for instance to a border of the city which is not exactly at the same distance from the center all around the city). The discrete character of the simulation appears in particular on the density curve, which is like a step function in the simulations and a continuous function in the analytical results.

The results of the model can thus be compared to the analytical results when these ones exist: figure 4.1 presents such a comparison for a city with only one income group. It shows the density, rent and surface curves as functions of the distance to the center for the analytical and the agent-based model. Because of the discretization, cells are not completely full in the agent-based model and the density is in general lower than the analytical density. The city is slightly more spread, rents are slightly higher, surfaces slightly lower and the equilibrium utility is slightly lower.

It would be interesting to compare the results of the simulations with the results of discrete models of Urban Economics, which the simulations should reproduce exactly. However there are few or no analytical results concerning discrete models of circular cities. Analytical works deal mainly with continuous models or discrete models of linear (one-dimensional) cities.

4.3 Emergence of a city

We now describe how a city emerges from the interactions between individuals during a simulation. Initially, all agents are located randomly and all rents are equal to the agricultural rent. Then agents move and bid higher, and the rent curve evolves from a flat rent to the equilibrium rent, and the density curve evolves towards the equilibrium density. Figure 4.2 shows how the shape of the city evolves during a simulation. In this simulation two income

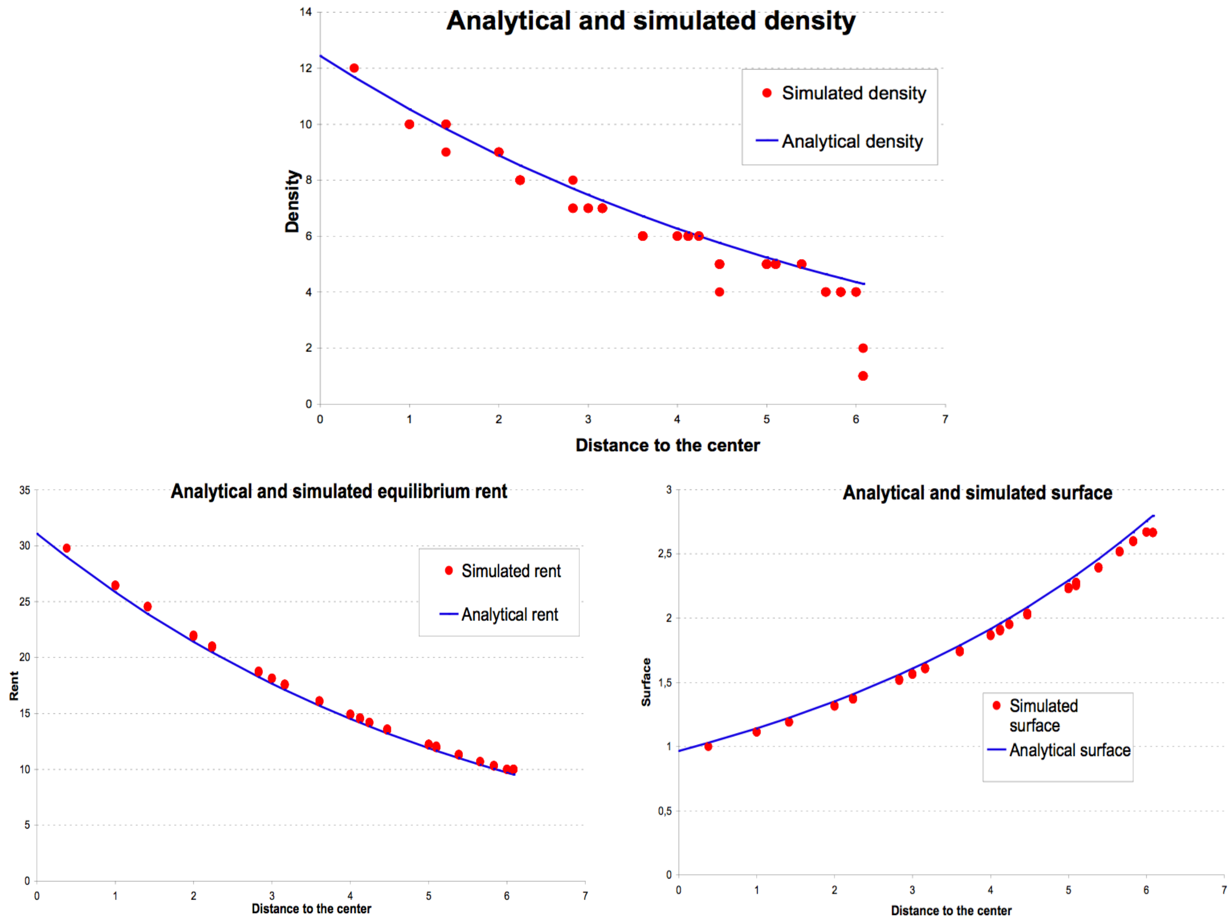
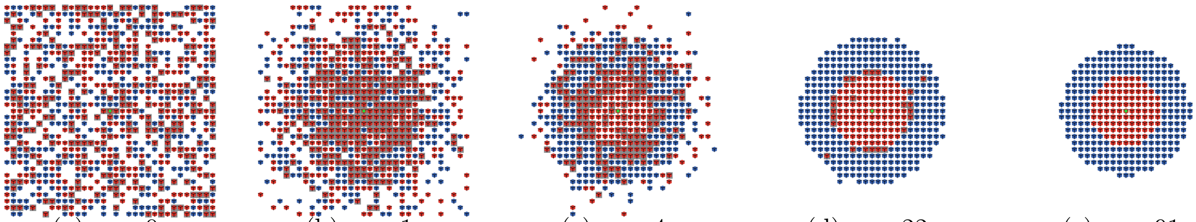


Figure 4.1: Comparison between the results of the agent-based model and the analytical results: density, rents and surface as functions of the distance to the center.

groups are present, "poor" agents in red and "rich" agents in blue. Paragraph 5.2.1 describes the equilibrium configuration in a detailed way.

At the beginning of the simulation (figures 4.2(a) and 4.2(b)), few people are displaced (as described in paragraph 4.1.2), agents gather at the city center without competing much for land, because many cells close to the center are still not full. But when all agents are concentrated around the center (from figure 4.2(d) on), most moves result in one or several agents being displaced, for few cells have a sufficient free surface to allow an agent to move in with an interesting utility. This feature of the model arises because the vacancy rate of the standard Urban Economics model we reproduce is zero.

The variable which shows the proximity to the equilibrium is the homogeneity of the utility



(a) $n = 0$ (b) $n = 1$ (c) $n = 4$ (d) $n = 22$ (e) $n = 91$

Figure 4.2: Evolution of the shape of the city during a simulation. The CBD is a green point. Cells whose background is grey indicate that poor and rich agents live there: at the equilibrium, the city is completely segregated and there are no more such cells. n indicates the mean number of moves per agent since the beginning of the simulation.

of agents. To describe this homogeneity, we use the variable $\Delta U = (U_{\max} - U_{\min})/U_{\max}$, which gives the relative inhomogeneity of the utility. With two income groups, we use the maximum of this variable in both income groups. Initially, this variable has a quite high value as all agents are located randomly, and it decreases during the simulations. We choose to stop the simulations when the relative variations of utility in the income groups are smaller than 10^{-6}, which means that this variable has decreased by approximately five orders of magnitude.

4.4 Results of the simulations

4.4.1 Two income groups: model 1

With two income groups, the utility of "rich" agents is higher than that of "poor" agents at the equilibrium, as they do not have the same income, but utility is homogeneous across the city within each income group. On figure 4.3 we give the shape of the city with two income groups, and the rent as a function of the distance to the center. The values of the parameters are those used in table 4.1. The city population is composed of two groups of 700 individuals each: "poor" agents (in red) with income $Y_p = 300$ and "rich" agents (in blue) with income $Y_r = 300 \times 1.6 = 480$. As in the analytical equilibrium, rich agents are located at the periphery of the city, where they pay lower prices and have higher housing surfaces, but also with higher transport costs.

On figure 4.4 are shown the density and the housing surface as functions of the distance to the center. The agent-based model with two income groups has the same behaviour as the analytical model for the city shape, density curve (discretized), rent and surface curves. But

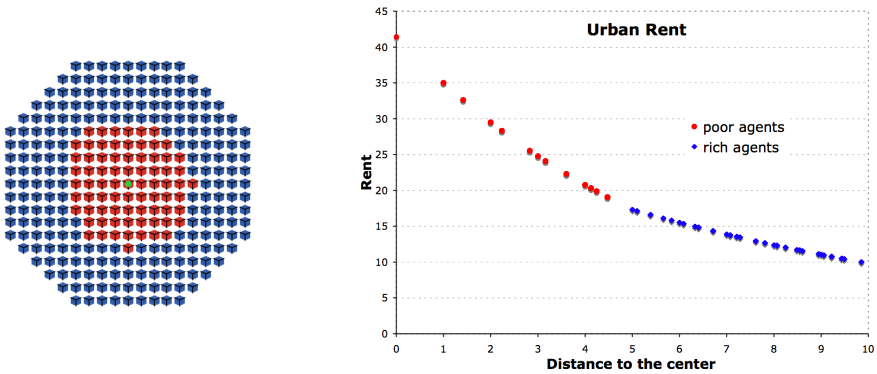


Figure 4.3: City with two income groups: rich in blue and poor in red. The CBD is a green dot. Right panel: equilibrium rent as a function of the distance to the center.

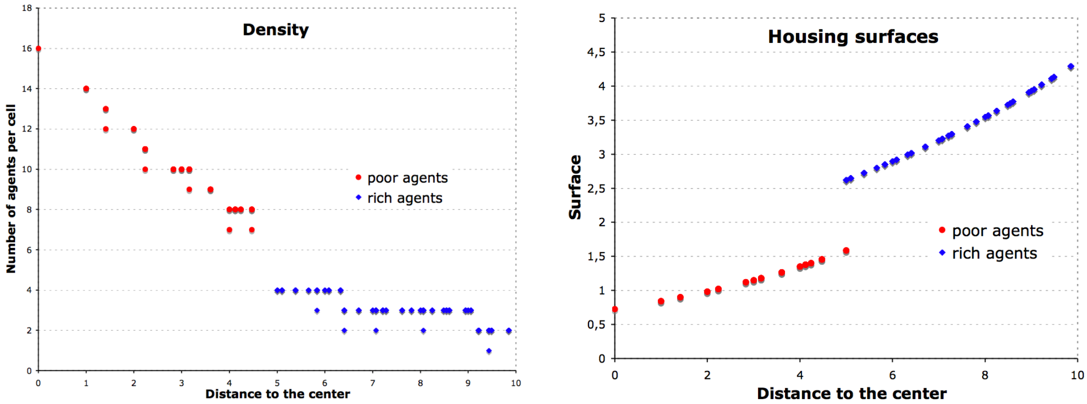


Figure 4.4: Population density (number of agents per cell) and housing surface as functions of the distance to the center.

as said previously we did not build an algorithm for the resolution of the analytical model in this case, so that we can not compare quantitatively the results.

4.4.2 Value of time: model 2

The equilibrium of the standard Urban Economics model we just presented gives a configuration where rich households live in the periphery of the city and poorer households near the center, which is in agreement with empirical results in most North American cities, but many European cities have an inverse configuration (see Brueckner et al. [1999]). One feature of the Urban Economics model that could account for this difference is the introduction of a difference in transport time cost. Rich households have a higher value of time than poorer households, and thus a higher global transport cost per unit distance.

This can be introduced in the model by adding a transport time cost (per unit distance) c_t/v, where c_t is the value of time and v the transport speed. The global transport cost can therefore be written

$$T(x) = tx + \frac{c_t}{v}x$$

where x is the distance to the center and t the monetary cost. Analytical treatment and

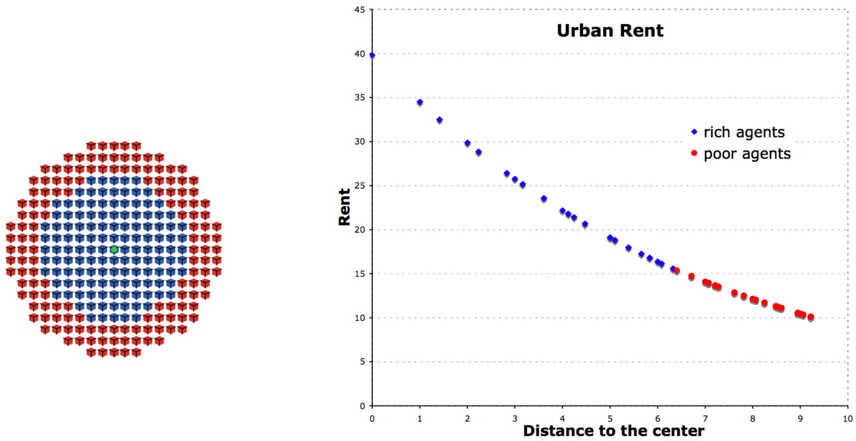


Figure 4.5: Shape of the city and equilibrium rent with a transport time cost much higher for rich agents than for poor agents.

agent-based simulations agree on the results of such a change in the model: to have rich agents located in the center and poor ones at the periphery, the quotient of the global transport costs per unit distance of rich and poor agent must be superior to the quotient of their incomes[2]:

$$T_r/T_p > Y_r/Y_p \tag{4.1}$$

with T_r and T_p the global transport costs per unit distance of rich and poor agents respectively. This situation is represented on figure 4.5, where the monetary cost of transport has a value $t = 2$ for both income groups, and transport time cost has a value $c_t^p/v = 2$ for poor agents and $c_t^r/v = (c_t^p/v) \times 2.4 = 4.8$ for rich agents. Then we have $T_r/T_p = 1,7 > Y_r/Y_p = 1.6$.

The condition (4.1) for the inversion of income groups can also be written as $T_r/Y_r > T_p/Y_p$. This means that the income group which will be located near the center of the city is

[2]This result is presented in appendix A.3. It is valid with the log-linear utility function used here, but also with a Cobb-Douglas function.

the one for which the unit global transport cost represents a higher share of the income.

Empirically, it can be observed that the value of time does rise with income (Palma and Fontan [2001]), but not quickly enough to have the condition 4.1 true if the monetary cost t and speed v are kept identical for both income groups. In this model, the value of time alone can not account for the difference of location of households depending on their income in European and North American cities.

Numerous analytical works deal with the question of the value of time and its influence on the location of households: the importance of this factor is still debated in the literature (see Gofette-Nagot et al. [2000], Brueckner et al. [1999], LeRoy and Sonstelie [1983], Wheaton [1977]).

4.4.3 Introduction of amenities: model 3

We would like now to study the influence of amenities on the location of agents. These amenities can be of different types: environmental amenities (woods or forest for example), historical amenities related to the historical evolution of the city (monuments, parks) or amenities linked to modern entertainment places (for instance cinemas, swimming pools or tennis courts). Several types can naturally be found at the same place.

We introduce in the agents' utility function a term accounting for the presence of amenities:

$$U = \alpha \ln z + \beta \ln s + \gamma \ln (1 + a(r))$$

where $a(r)$ is the amenity function and γ a parameter giving agents' preference for amenities. $a(r)$ is a decreasing function of distance r to the amenity, which we take of the form $a(r) = a_0 \exp(-r/b)$, with the values of parameters given in table 4.1.

Distance to the center

We study the influence of the distance d between the amenity and the CBD. Results are given on figure 4.6.

Depending on the distance d, the influence of the amenity on the shape of the city and on rents is more or less marked. It can be seen that for an amenity at a distance $d = 18$ in this model (figure 4.6(f)), some rich agents form and edge city around the amenity, close to the border of the main city. Such a configuration corresponds to leapfrog development, and

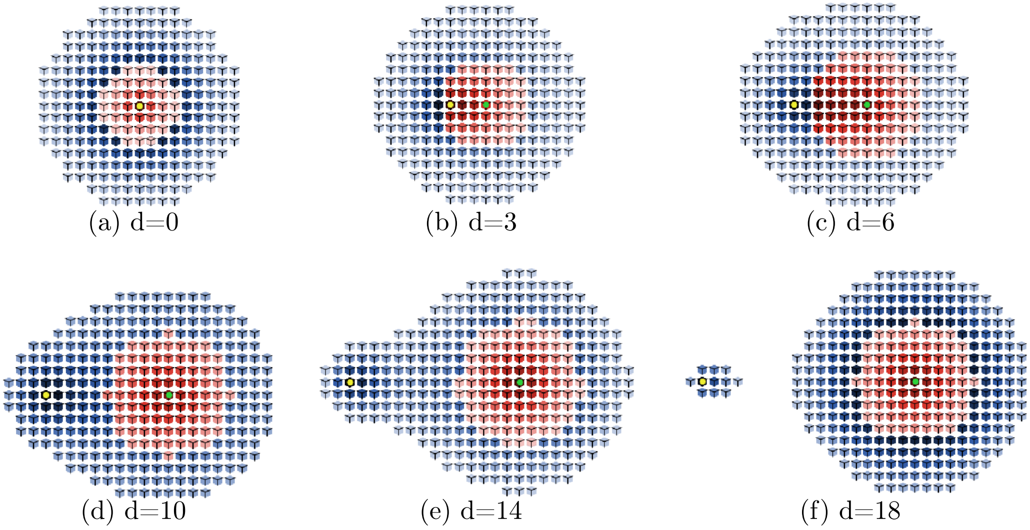


Figure 4.6: Amenity on the west side of the center at a varying distance d. Darker colors indicate higher prices in each income group. Color scales vary from one subfigure to the other.

the presence of an attractive amenity far away from the center can thus be considered as a possible cause of such a phenomenon.

Table 4.2 gives the value of variables allowing us to characterize the equilibrium of each simulation. The variables given in the first columns are the utility of rich agents U_r, of poor agents U_p and their difference. Due to the logarithmic form of the utility function, the variations are very small, so that we present the variations of the exponential of the utility instead. This exponential corresponds to a Cobb-Douglas function , which is also often used in Urban Economics studies and gives the same results as the log-linear utility. The other variables are the mean commuting distance of rich D^r_{mean} and poor D^p_{mean} agents, the total commuting distance D_{tot}, the total rent paid in the city R_{tot}, the mean price p_{mean}, the total surface of the city S_{tot} and the mean density ρ_{moy}.

The economic outcome of the introduction of amenities is naturally positive, because a new source of utility is added. Agents' utility increases when the distance between the amenity and the center decreases (as more and more agents feel the effect of the amenity). From a social point of view, an amenity close to the center reduces the gap in utility between rich and poor agents, as poor agents mostly benefit from it. An amenity far away from the center does just the contrary.

The city is made denser if the amenity is introduced within the city or close to it, and it is

Model	$\exp U_r$	$\exp U_p$	$\exp U_r - U_p$	D^r_{mean}	D^p_{mean}	D_{tot}	R_{tot}	p_{mean}	S_{tot}	ρ_{moy}
2 income groups (§ 5.2.1)	100	100	100	100	100	100	100	100	100	100
Amenity $d = 0$ fig. 4.6(a)	102.1	103.4	98.7	79.4	62.2	74.5	101.8	134.4	75.7	132.1
Amenity $d = 3$ fig. 4.6(b)	101.9	102.8	99.1	83.9	79.5	82.6	101.2	129.7	78.1	128.2
Amenity $d = 6$ fig. 4.6(c)	101.7	102.1	99.6	94.2	98.9	95.6	100.3	117.4	85.5	117.1
Amenity $d = 10$ fig. 4.6(d)	101.3	101.3	100.1	108.9	106.4	108.2	99.4	103.5	96.1	104.1
Amenity $d = 14$ fig. 4.6(e)	100.9	100.9	100.1	116.7	106.8	113.8	99.0	94.7	104.6	95.7
Amenity $d = 18$ fig. 4.6(f)	100.2	100.2	100.0	106.7	102.5	105.5	99.6	97.9	101.7	98.3
$\gamma_p = 0$, $d = 0$ fig. 4.7(a)	102.1	99.2	102.9	44.8	208.0	91.9	100.6	132.1	76.1	131.4
$\gamma_p = 0$, $d = 3$ fig. 4.7(b)	102.1	99.6	102.5	64.7	163.8	93.2	100.5	127.6	78.7	127.1
$\gamma_p = 0$, $d = 6$ fig. 4.7(c)	101.9	100.5	101.4	88.1	125.0	98.8	100.1	116.8	85.7	116.7
$\gamma_p = 0$, $d = 10$ fig. 4.7(d)	101.4	100.8	100.6	108.1	107.8	108.0	99.4	103.7	95.8	104.3
$\gamma_p = 0$, $d = 14$ fig. 4.7(e)	100.9	100.8	100.1	116.8	107.1	114.0	99.0	94.6	104.7	95.7
$\gamma_p = 0$, $d = 18$ fig. 4.7(f)	100.2	100.2	100.0	107.4	103.2	106.2	99.6	91.2	102.2	97.9

Table 4.2: Comparison of the different models with amenities.

spread if the amenity is far away. The environmental outcome is a consequence of this result: commuting distances are reduced in the first case and increased in the second. Amenities close to the center influence more poor agents, and those far away influence mainly rich agents. From looking at the results of an amenity at distance $d = 10$ from the center, it can be understood that the influence of an amenity is not obvious. In this configuration the city is made both denser because the mean density increases (the total surface of the city decreases) and more spread, for commuting distances increase. The reason is that the effect of the amenity has two components: it concentrates agents around the amenity, making the city denser, and it distorts the shape of the city, increasing commuting distances.

According to this model, a city planner wanting to make a city denser and to reduce commuting distances should introduce new amenities as close to the center as possible. It can also be observed that an amenity close to the center results in an increase in rents, which can be a problem in big cities where they are already very high, whereas an amenity far away from the center makes the rents lower.

The parameters a_0 and b which describe the amenity, can naturally change its influence on the city, but changing them does not result in surprising outcomes: increasing them increases simply the influence of the amenity on the city. Leapfrog development occurs only for an attractive amenity which is at a certain distance of the city, weakly attractive amenities do not lead to this phenomenon.

Differentiated preferences for the amenity

We also want to study the influence of individual preferences of agents on their location. Indeed, the model with amenity presents always the same configuration with poor agents close to the center and rich agents at the periphery, following the "North American" configuration of cities. But European cities tend to present the inverse configuration and we would like to see if a difference in agents' preferences for amenities could result in such a configuration.

We first study what happens in the extreme case where poor agents have no preference for amenities $\gamma_p = 0$, and rich agents keep the default value of the preference parameter $\gamma_r = 0.1$. The results of a such model are given on figure 4.7.

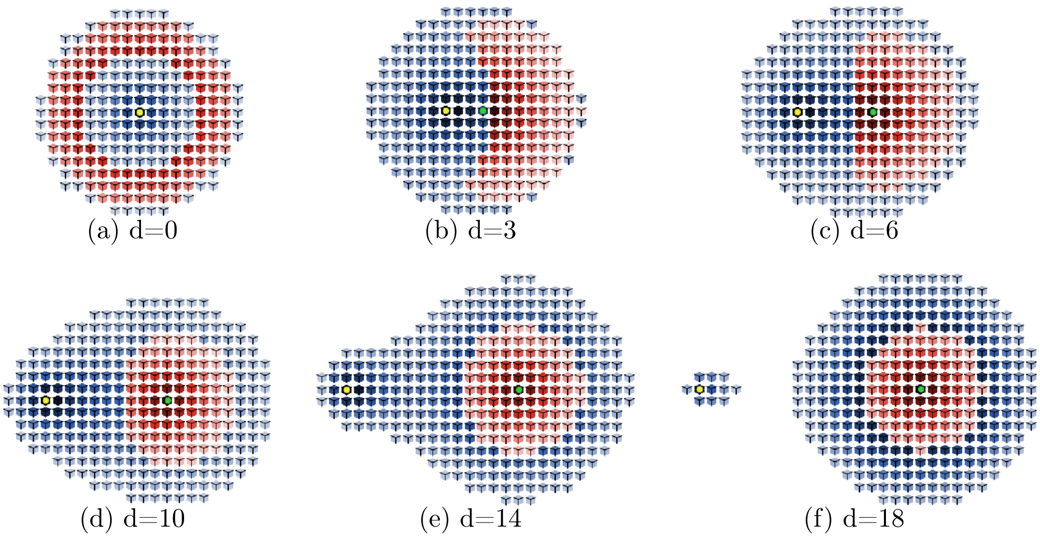


Figure 4.7: Amenity (yellow point) on the west of the center (green point) at a varying distance d, with a zero preference of poor agents for the amenity $\gamma_p = 0$. Darker colors indicate higher prices in each income group; color scales vary from one subfigure to the other.

For a central amenity or an amenity which is very close to the center, there is indeed an inversion of the income groups: rich agents live in the center close to the amenity, then further away live poor agents, and then again rich agents who do not benefit much from the amenity.

Such an amenity has a positive economic outcome for rich agents, and all the more when it is closer to the center. For poor agents, the economic outcome is negative when the amenity is close to the center and it becomes positives when the amenity is further away, because competition with rich agents is reduced. The environmental outcome is roughly the same as for an amenity for which rich and poor agents have the same preference.

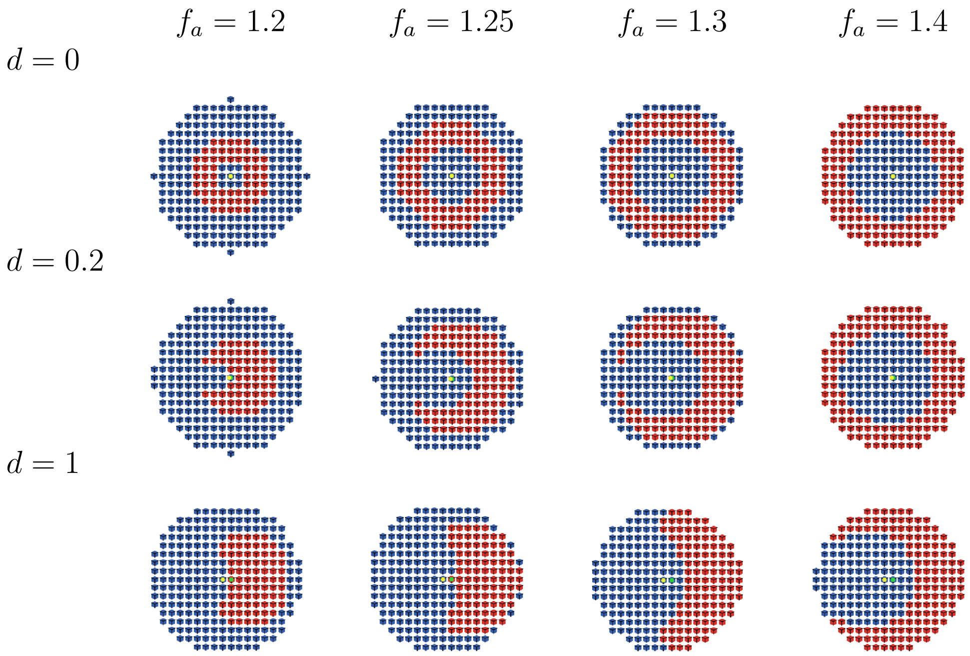


Figure 4.8: Social structure of the city with an amenity (yellow point) for which rich agents have a higher preference than poor agents $\gamma_r = \gamma_p \times f_a$, for different values of the difference factor f_a and of the distance d between the amenity and the center (green point), with $b = 4$ (amenity with a wider range).

Then, as it seems unrealistic to say that poor agents have absolutely no preference for the amenity, we can ask if a differentiated preference between poor and rich agents is sufficient to have rich agents living in the center of the city. We use a strong central amenity for which poor agents have a preference $\gamma_p = 0.1$ (default value) and rich agents a higher preference $\gamma_r = \gamma_p \times f_a$. As shown on the first line of figure 4.8, the location of rich and poor households depends strongly on the factor f_a differentiating the preferences for the amenity, and a value $f_a = 1.4$ is sufficient to have an inversion of the "American" configuration with the values of parameters used here.

The two last lines of figure 4.8 show us that the previous configuration of rich agents in the center and poor agents at the periphery is very unstable with respect to the displacement of the amenity: even a small displacement can change the shape of the city. This feature is specific to a two-dimensional model and shows the importance of two-dimensional simulations rather than linear (one-dimensional) cities.

Comparison of figures 4.6, 4.7 and 4.8 gives a good idea of the influence on the shape of the city of differentiated preferences for the amenity between rich and poor agents.

4.5 Perspectives

Historical evolution with endogenous amenities

This perspective has already been partially carried out, as shown on figure 4.9. The aim is to explore the historical evolution of a "semi-open" city, where population grows in time with the arrival of new inhabitants. We introduce an endogenous amenity which is linked to the presence of rich households: the cells where rich households live become more and more attractive, whereas cells where poor agents live become unattractive. Rich agents have to be more sensitive to this endogenous amenity in order to obtain the "European city" situation presented on figure 4.9. This can be linked to the differentiated preferences for exogenous amenities that we had to introduce previously to account for this structure.

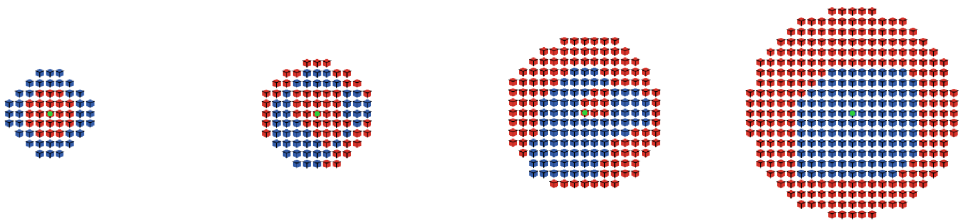


Figure 4.9: Historical evolution of a city with endogenous amenities linked to the presence of rich neighborhoods.

At the beginning of the simulation, the population is ten times lower than at the end, and no endogenous amenity is present, so that the city has a "North American" structure. The "European" structure emerges over time, rich households outbidding progressively poorer ones in the center of the city.

Calibration

Another perspective of this work is the calibration of the agent-based model with values which are coherent with empirical values for the different parameters (income, transport cost, agricultural rent...), and which give coherent results (densities, surfaces, size of the city...). For the moment this calibration is impossible because of the absence of vertical housing in the model: all agents live on the ground, and the densities can not come close to real densities. The city size is also unrealistic as a result.

In order to make such a calibration possible, a mechanism of building construction will have to be added, representing housing industry and determining the housing surface available in each cell. Such models have been introduced analytically by Muth (see Fujita [1989]).

Polycentric city

The monocentric city is a powerful but very rough model and many attempts have been made in the literature to develop analytical models of polycentric cities. This can also be done with the agent-based model presented here. We already studied the influence on the city of several exogenous business districts (but focused here on the influence of amenities), and a perspective of this work is to introduce endogenous centers. To this end, a new type of agents must be created to represent firms, and these new agents should compete with other agents for land. They also make location choices, but trying to maximize a profit function whereas households maximize their utility function. Such analytical models (of one-dimensional cities) are described in Fujita and Thisse [2003].

Conclusion

From a methodological point of view, this work shows the interest of agent-based systems in the study of collective phenomena carried out by social sciences. With the example of the standard Urban Economics model, we use this simulation tool to reproduce the results of an equilibrium model. To this end, we introduce an interaction between agents which allows to lead the system towards the equilibrium. This can be seen as an improvement of the equilibrium model because simple interaction mechanisms can be studied in this way. And as this simulation model is dynamical, it can be an interesting tool to study the dynamics of urban location as a perspective of this work.

In addition, these agent-based simulations allow us to modelize phenomena which are difficult to deal with analytically, like the introduction of amenities in a circular closed city with two income groups. The effects are multiple, on the shape of the city, on rents, surfaces of housing, density and commuting distances. So that the overall effect is difficult to predict. The use of an agent-based system is convenient, because it allows us to handle agents easily and to have an access to individual or global variables characterizing the state of the system.

Thanks to this model we explore research questions such as the influence of a value of time on the location of households depending on their income. We show that the Urban Economics model with a log-linear utility function can not explain the location of households in "European" cities (with richer households living near the center and poorer at the periphery) by introducing a value of time which increases with income, if monetary transport cost and speed are kept identical for both income groups. Indeed, the value of time would have to increase much more rapidly with income than empirical works suggest, to have an inversion of the "North American" city configuration.

Then with the introduction of exogenous amenities, we see that a central amenity for which both income groups have the same preference keeps the "American" configuration. In addition, the influence of the amenity on the city shape and on global variables depends strongly on the distance between the amenity and the center of the city. An attractive amenity far away from the center can cause leapfrog development.

Finally, we show that a central amenity for which richer households have stronger preferences than poorer households can inverse the social structure of the city, with richer households being located in the center. But for a two-dimensional city this situation is unstable with respect to a small displacement of the amenity: moving it slightly away from the CBD breaks the circular symmetry of the city.

The following chapter revisits this agent-based model and uses it to study the question of the polycentric city. Chapter 6 will come back on the question of "North American" and "European" cities.

Chapter 5

Exploring the polycentric city with an agent-based model

Introduction

This chapter is available online as a working paper[1] (Lemoy et al. [2011b]), and an earlier version has been published as a book chapter (Lemoy et al. [2010b]). The model used is very similar to the model presented in the previous chapter, but the eviction mechanism is removed here. Hence repetitions can be found, mainly in the description of this model.

Agent-based models provide an interesting framework to simulate social systems built on the microscopic interactions of social or economic agents. This work uses this simulation tool in urban economics to deal with research questions regarding urban systems, for instance the location of households with respect to their income or the study of the polycentric city.

An important work has been done in the literature on the urban economics model of Alonso, Muth, Mills (AMM model), to study the factors which govern the location choices of households within this model (Brueckner et al. [1999], Gofette-Nagot et al. [2000]), so that it is an interesting benchmark to start an exploration of the polycentric city using agent-based systems.

A first goal of our work is the reproduction of the classical results of the AMM model. To

[1]This work was presented in different conferences and workshops: WCTR 2010 (World Conference on Transport Research), Rencontres de TheoQuant 2011, Journées "Economie et Espace" 2011, MASHS 2011 (Modèles et Apprentissage en Sciences Humaines et Sociales).

this end, we use the agent-based simulation framework to define interactions between economic agents which are inspired from the AMM model. An important difference is the discreteness of the simulated model, whereas the analytical model is continuous. As a result, this work provides an illustration of a discrete model converging to the continuous AMM model for large population sizes, which can be related to a discussion in the literature (Asami et al. [1991], Berliant and Sabarwal [2008]). A comparison between the results of the simulations and those of the AMM model is presented in section 5.2.1.

The simulated model is dynamic: from a simple (random) initial state, interactions between agents on a simulated urban housing market lead progressively the whole system to an equilibrium state. The present work deals only with the study of the equilibrium of the model. However, the dynamics are described in a detailed way in section 5.2.2, and illustrations of the time evolution of the urban system are presented. Studies of dynamic phenomena in the urban housing market can also be performed with this simulation tool (Mc Breen et al. [2010]).

The second aim of this work is to study the urban social structure and the socioeconomic outcomes of non-monocentric models we define. Agent-based models are a convenient tool for a such work, as the monitoring of any local or global variable characterizing the system is easy. It is also an interesting framework to study neighborhood effects, though we do not introduce such mechanisms here.

As the model is shown to lead to an equilibrium corresponding to the analytical one, we separate from the analytical benchmark by relaxing the monocentric hypothesis. Section 5.3 presents the urban forms of the simple polycentric cities we study, and the economic, social and environmental outcomes of these urban forms, compared to the standard monocentric city of the AMM model.

5.1 Description of the model

5.1.1 Urban economics model

The AMM model was developed to study the location choices of economic agents in a urban space, with agents competing for housing (which is identified with land in the simplest version of the model). Agents have a transport cost to commute for work. Their workplace is located in a central business district (CBD), which is represented by a point in the urban space.

Amenities can be introduced in certain versions of the model to study their influence on the location of agents (Wu and Plantinga [2003], Lemoy et al. [2010a]). Agents usually represent single workers, but they can also be used to describe households, which can be made more complex in further versions of the model. Housing is rented by landowners who rent to the highest bidder, which introduces a competition for housing between agents. They also compete with an agricultural use of land, which is represented by an agricultural rent R_a.

Agents have a utility function describing their welfare, which is here a Cobb-Douglas function $U = z^\alpha s^\beta$, where z is a composite good representing all consumer goods except housing and transport, s is the surface of housing, α and β are parameters describing agents' preferences for composite good and housing surface, with $\alpha + \beta = 1$. Agents also have a budget constraint $Y = z + tx + ps$, where Y is their income, t the transport cost per unit distance, x their distance to the CBD and p the price of a unit surface of land at location x. See Fujita [1989] for a more detailed description of this model.

The analytical model reproduced in this work with agent-based simulations is a closed city model, where the population size N is chosen exogenously and remains constant during a simulation. This model can be solved analytically in a two-dimensional space if $R_a = 0$. For $R_a > 0$, a simple numerical resolution can be performed with one income group. With a population divided into several income groups, a specific algorithm is needed for the resolution of the model. The principle of this resolution is described in Fujita [1989].

5.1.2 Agent-based implementation

Let us describe in this section the agent-based implementation of the standard monocentric AMM model. In the agent-based system, the simulation space is a two-dimensional grid where each cell can be inhabited by one or several agents, or used for agriculture. These cells have a fixed land surface s_{tot}. The unit of distance is taken as the side length of a cell. The CBD is represented by a point at the center of the space.

At the initialization, a population of N agents is created. These agents are placed at random locations. The initial land price in each cell p_0 is equal to the agricultural rent R_a. At a given location x, agents occupy a quantity of land which is the optimal consumption of land conditional on price p: $s = \beta \frac{Y-tx}{p}$. This determines the quantity of composite good they consume, and their utility.

Dynamics of moves

The main feature of the model consists in agent-based dynamics of moves and bids in the urban space. The rules defining agents' moves are suggested by the competition for land in the analytical model.

Agents move with no cost, as in the AMM model. Let us describe an iteration n of a simulation, changing the variables from their value at step n to their value at step $n+1$. An agent which will be candidate to a move and a cell are chosen randomly. The price of this cell, located at a distance x of the center, is p_n at step n. The optimal housing surface that the agent can choose in the candidate cell is $s_n = \beta \frac{Y-tx}{p_n}$, which allows us to compute his composite good consumption and the utility that he would get thanks to the move and to compare it to his current utility.

If the agent candidate can have a utility gain $\Delta U > 0$ by moving into the candidate cell, then he moves. In this purpose he raises the price of the candidate cell by proposing a bid $p_{n+1} = p_n(1 + \epsilon \frac{s_{\text{occ}}}{s_{\text{tot}}} \frac{\Delta U}{U})$, where ϵ is a parameter that we introduce to control the magnitude of this bid. Prices evolve quicker if this parameter is high. s_{occ} is the surface of land occupied by other agents in the cell and s_{tot} the total land surface of the cell. The factor $\frac{s_{\text{occ}}}{s_{\text{tot}}}$ makes the bid higher if the cell is more occupied, that is to say, more attractive. Because of this factor, the first agent to move in an empty cell does not raise the price.

The price is a price per unit surface, linked to a cell. When an agent bids higher, the price is changed for all agents in the cell. Their consumption of land is also changed according to $s = \beta \frac{Y-tx}{p_{n+1}}$ and their utility is computed again. This feature of the model defines a competition for land between agents, as in the standard analytical model[2].

Surface constraint, time decrease of price

We described how prices increase in the model. Due to the stochastic choice of agents and cells, prices can rise above their equilibrium level at some locations, making some cells unattractive. Indeed, the price of a cell where several agents move in successively can increase so much that

[2]Specific situations arise which do not appear in an equilibrium (static) model. For instance, an agent may want to move into a candidate cell that is already full, proposing a higher bid on the price of housing there. Then we make the following choice: the price of housing is raised for all agents living in the cell to the level of this new bid, but the agent candidate does not move. Then agents' surfaces of housing and utilities are computed again. As the price is raised, housing surfaces are decreased and there is a chance that enough space is freed for the candidate agent to move in, in which case he does. Else, he has to wait until he is proposed another move.

it reaches a value which makes the cell unattractive. In this case, agents living there will progressively leave the cell for more attractive locations.

Therefore we choose to decrease exponentially the price of cells which are not completely full, according to $p_{n+1} = p_n - (p_n - R_a \times 0.9)/T_p.s_{\text{av}}/s_{\text{tot}}$, where T_p is a parameter determining the speed of decrease of prices, $s_{\text{av}} = s_{\text{tot}} - s_{\text{occ}}$ and s_{tot} are the non occupied surface and the total surface of the cell[3]. If no agent moves in, the price decreases according to this formula until it reaches the agricultural rent, which occurs after a finite time because of the form used. Then the decrease stops, and the cell is used for agriculture. The factor $\frac{s_{\text{av}}}{s_{\text{tot}}}$ makes this time decrease quicker as the cell is emptier and thus less attractive.

Parameters

The different parameters of the model are listed in table 5.1. Their value has been chosen according to several criteria. First for the parameters of the model itself: α, β, Y, t, N, R_a, s_{tot}. Their values have been chosen mainly for technical reasons regarding the comparison between the (continuous) analytical model and the (discrete) agent-based model, but naturally other values could have been chosen, without changing the qualitative behavior of the model. For instance, a higher population N could have improved the agreement between the analytical and the agent-based model, but it would have slowed down the simulations.

Parameters ϵ and T_p are specific to the agent-based model. Their values have been chosen such that the competition between agents on the housing market is efficient and the system reaches the equilibrium in the whole city. This will be described in section 5.2.3. The agent-based model has different behaviors and for instance does not reach an equilibrium within an acceptable simulation time for certain values of these parameters, but the study of these different behaviors is beyond the scope of the present paper. Appendix C.2 provides an illustration of a simulation where the values parameters ϵ and T_p are not chosen so as to have an efficient competition between agents. We refer to section 5.2.3 for a detailed discussion on the equilibrium of the model.

Let us note that the parameters ϵ and T_p are introduced here to provide a minimal framework allowing an agent-based simulation to reach the equilibrium of the AMM model. They do not seem to have an immediate correspondence with relevant or measurable variables

[3]With two income groups, it is difficult to determine whether a cell is full or not: we choose to let the price decrease if the smallest surface of housing s_{min} of agents there is smaller than the available surface of the cell s_{av}.

Parameters	Description	Default value
α, β	Preferences for composite good and housing	0.75; 0.25
Y_p, Y_r	Incomes of poor and rich agents	300, 450
t	Transport cost (unit distance)	10
N	Total population size	10000
R_a	Agricultural rent	10
s_{tot}	Surface of a cell	30
ϵ	Bidding parameter	0.5
T_p	Time decrease of the price of non-full cells	30

Table 5.1: Parameters of the simulations

explaining the dynamics of urban land markets.

5.1.3 Socioeconomic outcome

To study the urban social structure and the socioeconomic outcomes of the models developed here, we focus on some variables of the model, which characterize these outcomes. Our benchmark is a reference simulation of a monocentric city with two income groups. Then we compare the values of the socioeconomic variables in the reference simulation and in more complex models to observe the effects of the modifications which are introduced in the standard model.

To this end, we study variables which we find most relevant to describe the outcomes of the models. The utility of individuals is associated to their welfare and gives an economic outcome of the models. The cumulated distances of agents' trips to work, associated with housing surfaces, give their environmental outcome, which could be conveyed for instance in terms of greenhouse gases emissions associated to transport and land use (heating and cooling). The evolution of social inequalities are given by the difference in the utility of individuals of different income groups.

The use of agent-based systems allows naturally an easy access to any individual or global variable of the model, so that effects of the models on land rents for instance can also be studied.

5.2 Comparison with the analytical model and time evolution

5.2.1 Results with two income groups: model 1

The simulations allow us to reach the equilibrium of the AMM model, as can be seen on figure 5.1. This equilibrium corresponds to a configuration where no agent can raise his utility by

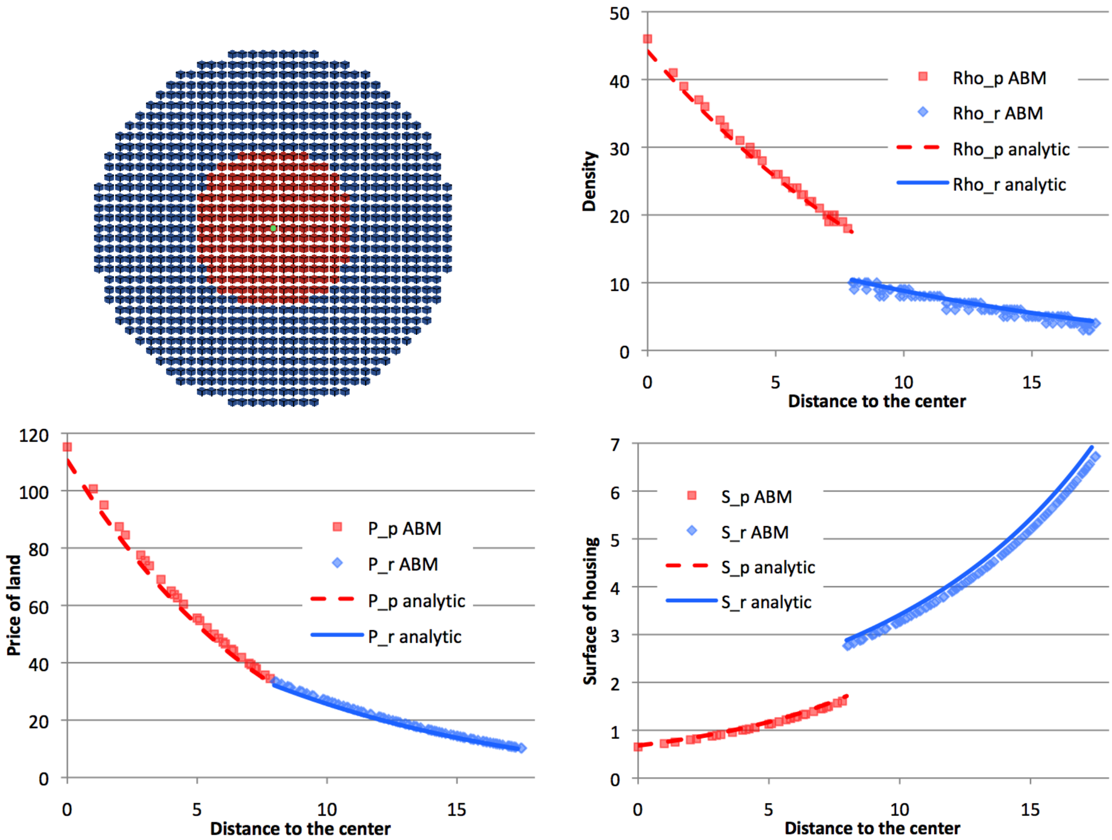


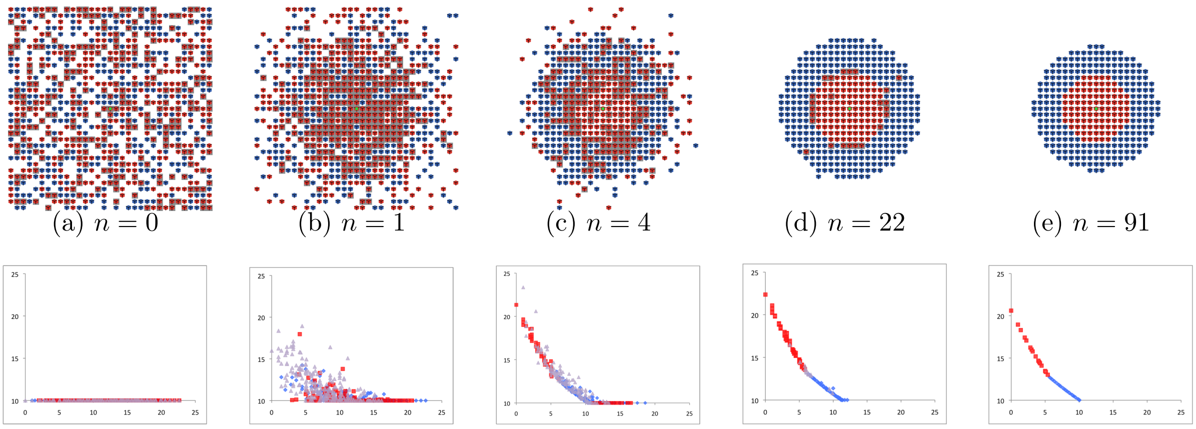
Figure 5.1: Top left panel, shape of the city. Other panels: comparison between the results of the agent-based model and the analytical results. Density, land rents and housing surfaces as functions of the distance to the center. The continuous lines represent the analytical results, whereas the symbols indicate the results of the agent-based model. Poor agents are in red and rich agents in blue, purple symbols on the density graph indicate cells where agents of both income groups live together. Parameters values are given in table 5.1.

moving, and therefore no agent has an incentive to do so. In each income group, individuals have an identical utility across the city. With two income groups, the utility of "rich" agents is still higher than that of "poor" agents, because they do not have the same exogenous parameters (they have different incomes). A gap can be observed on the density and housing surface curves, because of this discrete difference in income.As in the analytical equilibrium, rich agents are located at the periphery of the city, where they pay lower land prices and have higher housing surfaces, but also with higher transport costs. This reproduces the pattern observed in most North-American cities (Fujita [1989]).

The equilibrium of the agent-based model is described in the following sections.

5.2.2 Emergence of a city

The dynamic evolution of a simulation is rapidly presented here, it is very similar to the evolution presented in the previous chapter, in spite of the slightly different dynamics. The initialization is a random choice of locations, all prices are equal to the agricultural rent. Density is also random and quite low as agents are dispersed over the simulation space. Then agents move mainly towards the CBD as shown on figure 5.2 and bid higher, so that the rent curve evolves from a flat rent to the equilibrium rent. At the beginning of the simulation



(a) $n = 0$ (b) $n = 1$ (c) $n = 4$ (d) $n = 22$ (e) $n = 91$

Figure 5.2: Evolution of the shape of the city (first line) and of the price of land as a function of the distance to the center (second line) during a simulation. On the first line, the CBD is indicated by a green point. Cells whose background is grey indicate that poor and rich agents live there; these cells are displayed as purple symbols on the second line. At the equilibrium, the city is completely segregated and there are no more such cells. n indicates the mean number of moves per agent since the beginning of the simulation.

(figures 5.2(a) and 5.2(b)), agents gather at the city center without competing much for land, because many cells close to the center are still not full. But when all agents are concentrated around the center (from figure 5.2(d) on), most bids do not result in an agent moving, for few cells have a sufficient available surface to allow an agent to move in with an interesting utility.

The main variable which indicates the proximity to the equilibrium is the homogeneity of the utility of agents, as described in the previous chapter. To describe this homogeneity, we use the relative inhomogeneity of the utility defined as $\Delta U_{\max} = (U_{\max} - U_{\min})/U_{\max}$. With two income groups, we use the maximum of this variable within income groups. During a simulation, this variable has a decreasing value. We choose to stop the simulations when the relative variations of utility within income groups are smaller than 10^{-6}, which means that $\Delta U_{\max}$ has decreased by approximately five orders of magnitude, as shown on figure 5.3.

The model allows to test explicitly if no agent has an incentive to move: when $\Delta U_{\max} < 10^{-6}$, each agent tests if he can raise his utility by moving into any other cell, regardless of a sufficient or not sufficient available surface in the cell. The relative possible variations of utility are found to be of the same order of magnitude as $\Delta U_{\max}$.

The standard deviation of the utility can also be computed. It gives a more precise idea of the variations of utility in the model. On figure 5.3 are displayed the evolution of the relative inhomogeneity of utility in each income group, and the corresponding evolution of the standard deviation of utility in each income group (also divided by the maximal utility) during a simulation. The latter is always smaller than the former, as should be. This evolution is given as an illustration: because of the stochastic dynamics of the model, it varies across simulations. The equilibrium of the agent-based model is described in more detail in section 5.2.3.

5.2.3 Equilibrium

Existence and uniqueness of the analytical equilibrium

Is is shown in Fujita [1985] that there exists one unique equilibrium for the standard monocentric AMM model with one or more income groups. The proof of this result is based on boundary rent curves between income groups and between agricultural and residential land uses. Although this result seems difficult to extend to any polycentric city, we give arguments to support the fact that there is also one unique equilibrium in the models which are studied in this paper.

The existence of (at least) one equilibrium for all simple polycentric models studied here is proved in Fujita and Smith [1987] using fixed-point methods. Hence, it remains to be argued

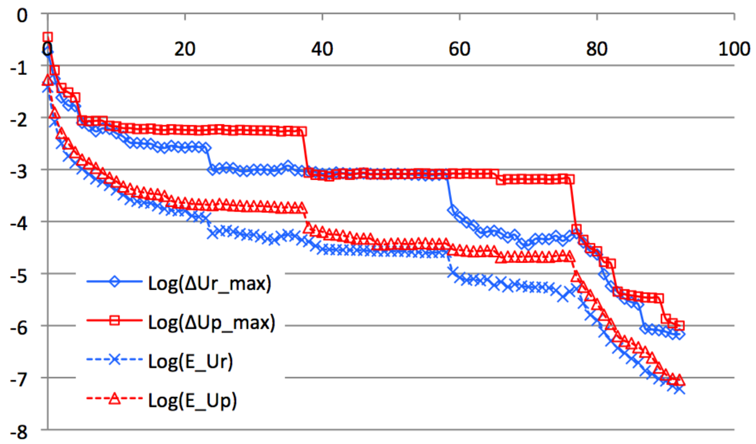


Figure 5.3: Evolution of the common logarithms of the relative inhomogeneity ΔU and of the relative standard deviation E_U of poor (in red) and rich (in blue) agents' utility during a simulation. One unit of the abscissa represents 50000 iterations.

that with the polycentric changes added here to the standard monocentric model, there is no apparition of multiple equilibria, contrary to what can happen with more complicated models, for instance Brueckner et al. [1999] or Fujita and Ogawa [1982]. It can be observed that these works add important changes to the standard model by adding variables to the utility function, while our work only considers a more complex transport. Indeed, introducing several work centers breaks the circular symmetry of the transport cost which is found in the monocentric city.

Let us first consider the case of model 2 with two centers separated by a distance d, the simplest polycentric model we study, where agents work at the employment center which is closest to their housing location (see section 5.3). When centers are sufficiently far apart, two cities are present and do not interact, with an equal share of the whole population residing in each city. In this case, the result of Fujita [1985] ensuring existence and uniqueness of the equilibrium is clearly still valid. When centers are brought closer and cities begin to interact, the situation is a bit more complicated.

Our approach then consists in mapping this simple polycentric model onto a fictitious monocentric model verifying the assumptions required in Fujita [1985] to ensure existence and uniqueness of the equilibrium. This mapping allows us to prove the existence and uniqueness of the equilibrium for the polycentric model. In the urban models studied in Fujita [1985], as well as in our model 2, a given location is completely characterized by the distance of commuting for an agent residing in this place. Equivalently, space is characterized by the amount of land available at each commuting distance x. Let us note $L(x)dx$ the amount of land available between commuting distances x and $x + dx$. A monocentric model with a distribution of land equivalent to that of our simple duocentric model would have $L(x) = 4\pi x$ for $0 \leq x \leq d/2$ and $L(x) = 4x(\pi - \arccos(d/x))$ for $x \geq d/2$. This fictitious monocentric model verifies all assumptions ensuring that it has one unique equilibrium, which is also true for our duocentric model as a consequence. This result could be extended to models with 3 or more centers, as this would only make the function $L(x)$ more complicated. With several income groups, the result still holds.

The case of model 4 with $m = 1$ is almost similar. In this model two work centers (East and West) separated by a distance d are considered, and each agent represents a two-workers household, with each worker of the household working at a different center (see section 5.3). Thus the total commuting distance of the household is the sum $d_E + d_W$ of distances between the household's housing location and both centers. The function $L(x)$ of the corresponding monocentric model is now $L(x) = 0$ for $0 \leq x \leq d$ and $L(x) = 2xE(d/x)$ for $x \geq d$, with $E(e)$ the complete elliptic integral of the second kind. This last formula corresponds to the circumference of an ellipse of major axis x, distance between focal points d, and eccentricity d/x. A similar argument of correspondence proves the uniqueness of the equilibrium in this case[4], and is still valid with several income groups. However, it seems impossible to extend this result to more than two work centers in this case.

In the cases of model 3, model 4 and model 5 with $0 < m < 1$ (see section 5.3), the previous arguments supporting the uniqueness of the analytical equilibrium seem difficult to reproduce. But it remains true that no important change is brought to the utility function when compared with the standard monocentric model. Only the transport cost (seen as a function defined on the two-dimensional space of the model) is changed.

In addition, an important argument in favor of this uniqueness is the fact that for all models presented here, the agent-based model converges to the same equilibrium situation for every run of a simulation, as shown in appendix C.1. This agent-based equilibrium is further described in the next section.

[4]In Fujita [1985], it is assumed that $L(x) > 0$ on $[0, x_1]$ with x_1 a positive number. We assume that the result of existence and uniqueness of the equilibrium is still valid under the condition $L(x) > 0$ on $[x_0, x_1]$ with $0 < x_0 < x_1$, though we do not provide a proof supporting this assertion.

Analytical and agent-based equilibria

Assuming that the previous arguments are convincing of the fact that our simple polycentric models all have one unique analytical equilibrium, it should still be argued that the agent-based model is able to reach a discrete version of this equilibrium. Figure 5.1 shows that it is so for the standard monocentric AMM model with two income groups. Let us describe more precisely the hypotheses ensuring that the agent-based model reaches an equilibrium which is similar to the analytical one. Section 5.2.2 shows that the equilibrium of the agent-based model is defined as a situation in which utility is homogeneous within each group, ensuring that no agent has an incentive to move. But this condition alone does not guarantee that the equilibrium is reached, as shown in appendix C.2. Indeed, a supplementary condition is needed: that every cell is optimally used, either for agriculture or for housing.

From the comparison with the analytical equilibrium, it follows that only two situations should be observed at equilibrium for the cells of the agent-based model. The first is the case of an agricultural cell, whose price should be equal to the agricultural rent, and where no agent should reside. The second case is a residential cell, where no space should be left for another agent to move in. Indeed, if the cell can accommodate (at least) one more agent, it indicates that equilibrium is not reached as the city could be made more compact, providing a higher utility for agents.

We want to monitor the number or share of cells which are not optimally used. Indeed, it follows from the previous discussion that the measure of the distance to the equilibrium in terms of the homogeneity of utility should be completed by another one, linked to residential location. The variable we use is the quotient of a surface we call "empty" S_{empty} to the total housing surface of agents S_{tot}. Let us now describe how this "empty" surface is computed. Each cell of the simulation space is visited. If the cell has inhabitants, the smallest housing surface of the inhabitants s_{min} is stored. If the surface still available in the cell s_{av} is greater than s_{min}, then a part of the cell is considered "empty". To determine how much exactly, it is computed how many agents with housing surface s_{min} could still fit in the cell. The corresponding surface is considered "empty". The values of parameters ϵ and T_p are chosen so as to minimize the quotient $S_{\text{empty}}/S_{\text{tot}}$ (which is checked to be smaller than 0.5% at the equilibrium of the simulations) with an acceptable simulation time. It is also verified that every cell without inhabitants has a price which is equal to the agricultural rent.

With these conditions, the system converges to a unique equilibrium, as described in appendix C.1. Appendix C.2 presents a simulation using values of parameters ϵ and T_p which do not allow the system to reach a state where $S_{\text{empty}}/S_{\text{tot}} < 0.5\%$.

5.3 Additions to the standard model

5.3.1 Polycentric city: model 2

The agent-based mechanism introduced in this work to reproduce the results of the AMM model is robust enough for us to introduce effects which are difficult to treat analytically. For instance, several centers can be introduced. This is an important domain of research in the literature (see Hartwick and Hartwick [1974], White [1976, 1988], Wieand [1987], Yinger [1992]). In the simulations of these models 2 and 3, parameters values are given by table 4.1 of the previous chapter. In the first model studied here, agents work at the center which is the closest to their housing, and as a consequence, can change jobs as they move. This last feature seems unrealistic but allows to prevent market frictions and reach more rapidly the equilibrium. The results of a such model are given on figure 5.4.

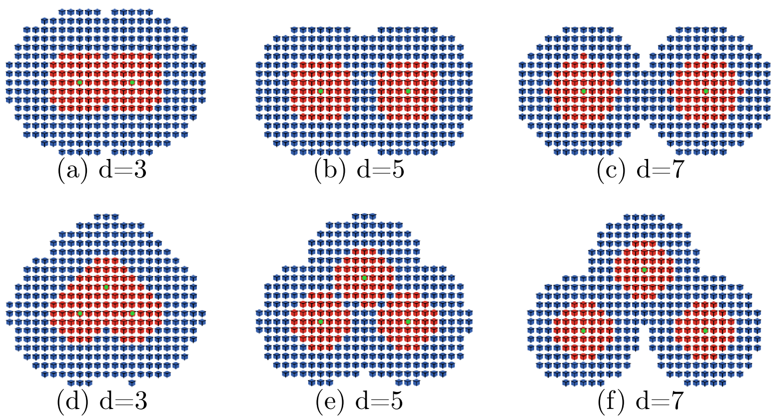


Figure 5.4: Cities with two centers separated by $2d$ cells (first line) and cities with three centers located at $(-d;0)$, $(d;0)$ and $(0;d)$, for different values of d. Centers are indicated by green dots, and agents work in the center closest to their housing location.

Rents, housing surface and density represented as functions of the distance to the nearest center produce curves which are qualitatively similar to those of figure 5.1. Table 5.2 allows to see the evolution of different variables for this polycentric model, such as agents' utility, the mean commuting distance for each income group, the total commuting distance, the total rent and the total surface of the city, compared with the reference configuration with two income groups from paragraph 5.2.1. The mean density is given by the inverse of the total surface, as the population is fixed.

Raising the number of centers amounts to raising the surface available at a given commuting distance in the city, or equivalently to reducing transport costs. As a consequence, this reduces the competition for housing. Agents have greater housing surfaces, smaller commuting distances and a higher utility. The total rent increases, which can seem surprising but can be explained by the fact that housing surfaces are greater. The mean density decreases while housing surfaces increase. These effects are more pronounced when the centers are further away from one another.

Model	U_r	U_p	$U_r - U_p$	D^r_{mean}	D^p_{mean}	D_{tot}	R_{tot}	p_{mean}	S_{tot}
2 income groups (§ 5.2.1)	100	100	100	100	100	100	100	100	100
2 centers $d = 3$ fig. 5.4(a)	101.6	102.6	99.0	83.1	78.1	81.7	101.3	94.9	106.7
2 centers $d = 5$ fig. 5.4(b)	102.2	103.1	99.1	78.2	80.3	78.8	101.5	92.9	109.7
2 centers $d = 7$ fig. 5.4(c)	102.6	103.4	99.3	78.8	81.9	79.7	101.4	88.5	114.5
3 centers $d = 3$ fig. 5.4(d)	102.1	103.4	98.7	77.4	67.0	74.4	101.8	94.3	108.0
3 centers $d = 5$ fig. 5.4(e)	103.1	104.4	98.7	70.6	69.2	70.2	102.1	88.8	114.9
3 centers $d = 7$ fig. 5.4(f)	103.9	105.3	98.7	66.7	71.3	68.0	102.3	83.8	122.1
100% $d = 2$ fig. 5.5(a)	100.0	100.49	99.5	87.2	120.5	96.8	100.2	100	99.5
100% $d = 6$ fig. 5.5(b)	100.8	102.0	99.8	72.0	122.8	86.6	100.9	99.0	102.0
100% $d = 10$ fig. 5.5(c)	101.4	103.1	98.4	70.1	124.4	85.8	101.0	92.5	109.3
80% $d = 2$ fig. 5.5(d)	100.4	100.9	99.6	90.8	97.2	92.7	100.5	100	101.1
80% $d = 6$ fig. 5.5(e)	101.3	102.4	98.9	79.9	91.0	83.2	101.2	96.1	105.2
80% $d = 10$ fig. 5.5(f)	101.9	103.2	98.8	75.5	95.3	81.3	101.3	91.8	110.3

Table 5.2: Comparison between the different polycentric models. Variables are rich and poor agents' utility U_r and U_p, their difference, rich and poor agents' mean commuting distances D^r_{mean} and D^p_{mean}, the total commuting distance D_{tot}, total rent R_{tot}, mean unit surface price p_{mean} and the total surface of the city S_{tot}. Parameters values are given by table 4.1.

Thus economic and environmental outcomes of the introduction of several centers in this model are positive: agents' utility increases and commuting distances decrease. Agents' utility increases when the distance between centers increases, but the effect on commuting distances is more complex (see table 5.2). Commuting distances are always smaller than in the reference simulation, but they can increase again when the centers are moved away from each other, as the decreasing competition for land results in increasing housing surfaces, and thus city size. It should also be noted that bigger housing surfaces result in greater heating (and cooling) needs, which are a major source of greenhouse gases emissions. This puts the environmental outcome of this model in another perspective. The social outcome is not intuitive. Poor agents' utility increases more than rich agents', but the utility gap depends on the number of centers. With two centers, this utility gap is smaller when the centers are closer. With three centers, it remains constant when centers are further away.

The third line of figure 5.8 and appendix C.3 present respectively the shape of the city and the evolution of different variables for this same model 2, with only one income group, as a limiting case of model 5, presented in section 5.3.4.

In model 2, the CBD is splitted in two work centers of equal size. This is the simplest polycentric city which can be imagined. Let us now introduce a version of this simple polycentric model, with two work centers of different sizes.

5.3.2 Constrained polycentric city: model 3

It is also possible to assign agents to a given employment center at the beginning of the simulation and to keep it. The computation of the equilibrium in this configuration on a two-dimensional city has not been done to our knowledge, but the bid mechanism used here allows us to find this equilibrium. For instance, all rich agents work in a given center and all poor ones in another center at another location. The result of a such model is given on the first line of figure 5.5, where the center on the right can be seen as a center with low-skill jobs (or an industrial zone) in the east of the city, and the center of the left, a center with high-skilled workers on the west of the city.

It is also possible, as shown on the second line of figure 5.5, to have only a part of each income group working at each center, that is to say, to suppose that centers are not completely specialized. In this case, as agents in a certain income group have different constraints, their utility is not homogeneous within an income group. Indeed, utility is homogeneous among agents of the same income group working at the same center. On the second line of figure 5.5, the urban system is composed of two work centers, which are not indifferent for agents (contrary to what is done in model 2) and two income groups, that is, four (2×2) utility groups at equilibrium.

As can be seen on table 5.2, the global effect of the introduction of centers with constraints for agents is quite similar to the effect of centers without constraints: the competition for housing decreases. The economic outcome is positive, as agents' utility increases when the distance between centers increases. The housing surfaces increase, and they increase when the distance between centers increases. However, the simulation presented on figure 5.5(a) is an exception: the city surface is reduced and the mean density is higher than in the reference

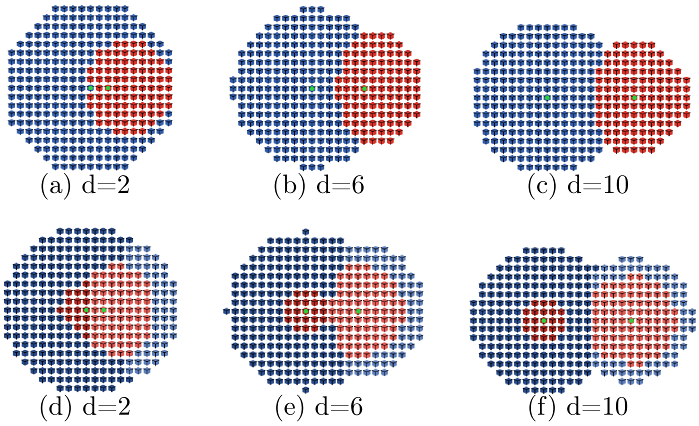


Figure 5.5: First line: cities with two centers where poor agents work in the east center and rich agents in the west center. Second line: 80% of poor agents work in the east center and 20% in the west center, and conversely for rich agents. The distance d between both centers is indicated. On the second line, agents working in the center on the right have paler colors.

configuration.

Partial or total segregation of rich and poor agents in job locations decreases in fact mainly the competition for housing between both income groups: poor agents are less pushed toward the center by their competition with rich agents, and rich agents are less pushed toward the outskirts of the city. Two effects appear on commuting distances. This decrease of the competition between income groups for housing raises poor agent's commuting distances and decreases those of rich agents. And the increase in the surface available at a given commuting distance decreases all commuting distances. So the environmental outcome is positive from the transport point of view, as commuting distances decrease globally when the distance between centers increases. But it has to be balanced with the negative environmental effect of increasing housing surfaces. The effects on commuting distances of each income group are more complex, as can be seen on table 5.2. The social outcome is globally positive. The utility gap between rich and poor agents decreases. When the segregation linked to employment is total, the effect of increasing the distance between centers is not monotonous (see table 5.2). When this segregation is partial, social inequalities decrease when the distance between centers increases. Though it must be remembered that in the last simulations, a new disparity has appeared within each income group.

5.3.3 Spreading the CBD: model 4

The polycentric models presented in the two last sections give interesting indications on the outcomes of polycentrism in the standard urban economics model. However, it is quite unrealistic to imagine a city planner who could for instance have the power to share the CBD of his city in half and put both halves apart, driving as a consequence both halves of the population apart. In this section, we wish to study a more realistic phenomenon: the decentralization of employment locations, which can also be seen as a spread of the CBD.

To this end, we study a model with only one income group, divided into 5 subgroups labeled by $i = 0..4$. Within each of these subgroups, agents work at a distance d_i of the (punctual) center of the city, given by $d_i = d \times i$. We study the influence of the decentralization of employment locations, as a function of this distance d. For $d = 0$, this model corresponds to model 1 with only one income group.

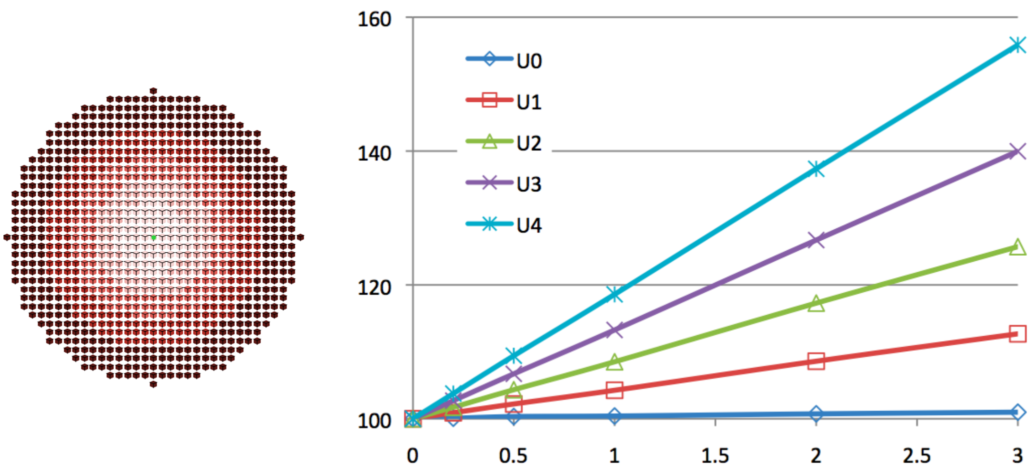


Figure 5.6: Left panel: shape of the city in model 4, where darker colors indicate higher values of i and of the utility. Right panel: evolution of the utility of agents of each group $i = 1..5$ as functions of the difference d of radius of the employment rings of two consecutive groups. Parameters values are given by table 5.1, except $N = 5000$ and $R_a = 5$.

Figure 5.6 shows the location of different employment subgroups at the equilibrium. Agents working further from the center of the city are also located further from the center. This is logical, because they have a comparative advantage there as they benefit of a lower transport cost than agents working closer to the center. They also have a higher utility, as shown on the right panel of figure 5.6, where their utility is compared with their value in the reference simulation $d = 0$, corresponding to model 1 with only one income group. Their utility increases with the distance d, including the utility of agents of group 0, who benefit of a reduction of the competition for housing in the center, even if their transport cost does not change. So that the economic outcome of this model is obviously positive.

Figure 5.7 gives the outcomes of this model on the different variables studied previously. The total commuting distance decreases strongly when d increases, which gives a positive environmental outcome. But it seems unrealistic that agents living and working at the periphery of the city never commute to the center. A simple way to deal with this question is studied in the next section. The total rent increases when d increases, as a result of the decrease of the mean price in the city being overweighed by the increase of the total housing surface. This has a negative environmental effect which mitigates the positive one on commuting distances.

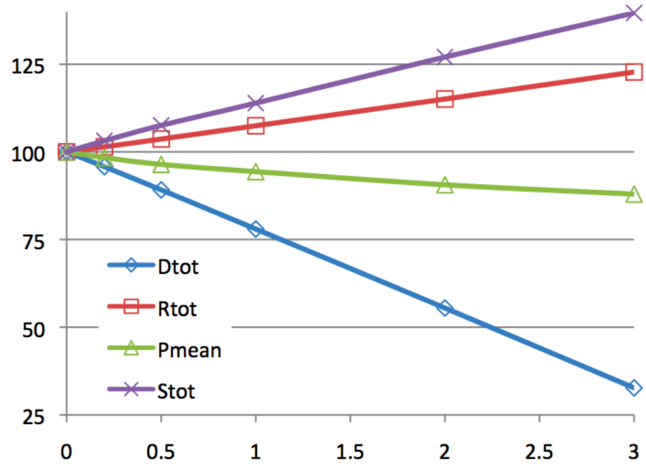


Figure 5.7: Outcomes of model 4 as functions of the difference d of radius of the employment rings of two consecutive groups. Parameters values are given by table 5.1, except $N = 5000$ and $R_a = 5$.

5.3.4 Polycentric city and two-workers households: model 5

One important flaw of the previous models is the fact that as the only transport motive considered is the daily commuting for work, the location choice of each agent is only linked to one work center, so that with several centers, the interaction between centers is poor.

A simple way to add more coherence to the city as a whole while keeping the same framework consists in studying two-workers households. The influence on the results of models with two-workers households is also a research question in the literature (see Madden and White [1980], Kohlhase [1986], Hotchkiss and White [1993]). Each agent described previously represents now a household composed of two workers. For simplicity, we consider only a city with two centers, as the ones of the first line of figure 5.4. In addition, and also for the sake of simplicity, we study a model with only one income group. Households are divided in two groups. In the first group, which we denote by "common", both persons in the household work at the same employment center. In the second group, which we denote by "splitted", they work in different centers. This is imposed exogenously and does not change during a simulation.

We study the outcome of this model depending on two variables: the distance d between centers, and the share $m \in [0, 1]$ of households of the "splitted" group. Note that the case $m = 0$ has already been studied, as it corresponds to model 2 with two centers. Let us label employment centers by "East" and "West", and note d_E and d_W the distances between a given household's location and centers East and West. Then if both persons in this household work at the same employment center ("common" group), the East center for instance, the transport cost associated with the commuting of the household is $2 \times \tilde{t} \times d_E$. If they work at different centers ("splitted" group), their transport cost is $\tilde{t} \times (d_E + d_W)$. The transport cost for a unit distance $\tilde{t}$ is chosen so that in the limit $d = 0$, this model 5 with households corresponds to model 1 (with only one income group): $\tilde{t} = t/2$.

One important consequence of the new ingredient added here is that a minimal commuting distance of d is imposed for all households of the "splitted" group. It is their commuting distance if they are located on the segment linking both employment centers. So that the minimal total commuting distance $D^{\text{tot}}_{\text{min}}$ of agents in the city is $D^{\text{tot}}_{\text{min}} = d \times m \times N$. This minimal distance is exogenously imposed, and is a special feature of this model 5.

To begin with, let us study what happens in the case $m = 1$, where all households belong to the "splitted" group. To minimize their transport cost, agents choose their location by minimizing $d_E + d_W$. As a result, the shape of the city is elliptic with both employment centers as focal points, as can be seen on the first line of figure 5.8. Indeed, the figure defined by the set of points verifying $d_E + d_W = k$, with k a given constant, is an ellipse. The effect of increasing d on the transport cost of agents can be described as follows: the transport cost is increased (everywhere, except at both employment centers themselves, where it does not change when compared when the monocentric case) because of the increasing minimal commuting distance described previously, and the center of the city (seen as the place where transport cost is minimal) is spread on a segment linking both employment centers.

As a consequence, the total commuting distance D_{tot} of agents increases when centers are

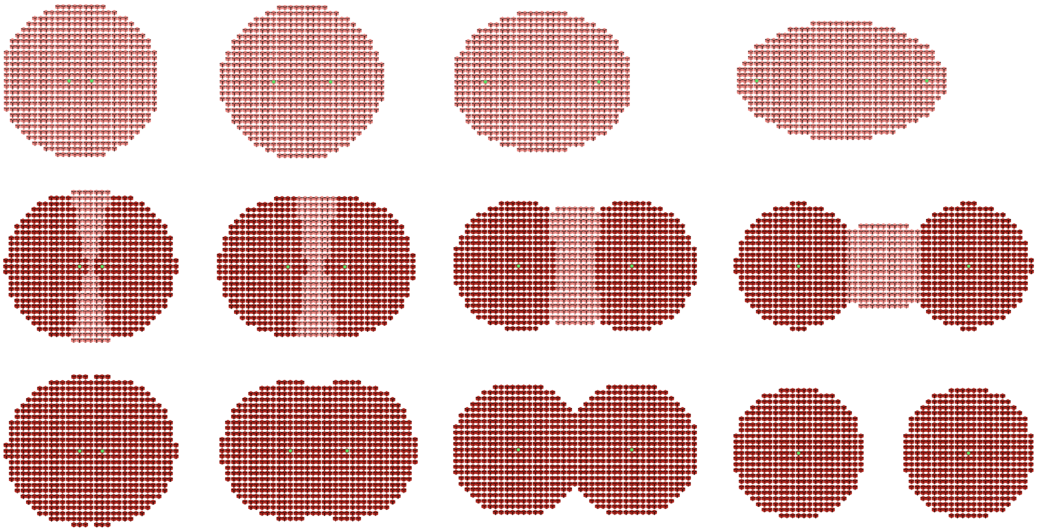


Figure 5.8: Shape of the city with households (model 5), with $m = 1$ (first line), $m = 0.2$ (second line) and $m = 0$ (third line). The different columns correspond to different values of the distance d between centers: $d = 4, 10, 20, 30$ from left to right. Agents of the "common" group have a darker color than agents of the "splitted" group.

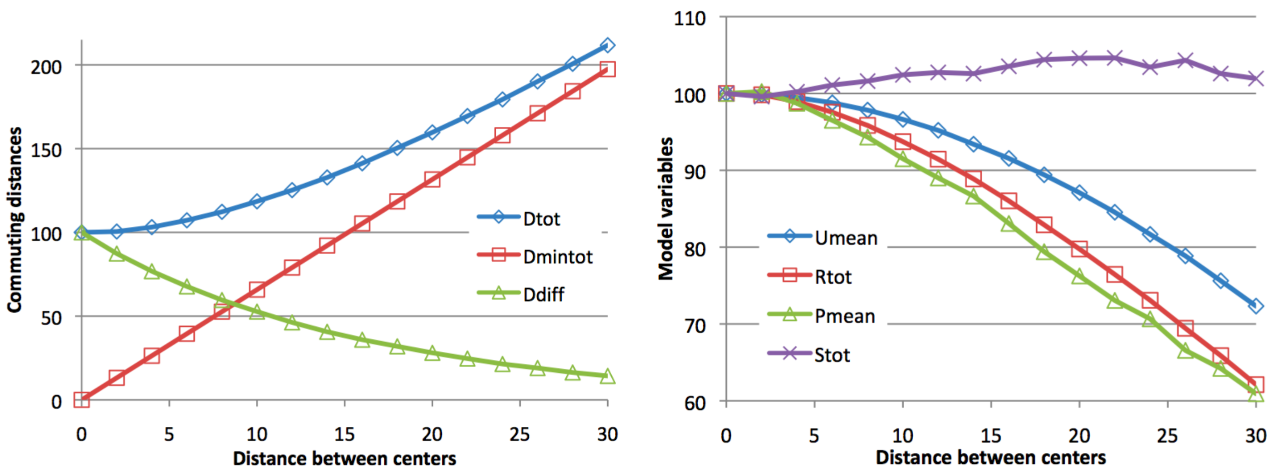


Figure 5.9: Outcomes of model 5 with $m = 1$ as functions of the distance d between centers. Left panel: evolution of the total commuting distance, the minimal distance and their difference. Right panel: evolution of agents' utility, of the total rent, of the mean price and of the total surface of the city. Parameters values are given by table 5.1.

moved apart, mainly because of the contribution of the minimal commuting distance $D_{\min}^{\text{tot}}$, as can be seen on the left panel of figure 5.9. $D_{\text{diff}} = D_{\text{tot}} - D_{\min}^{\text{tot}}$ is also indicated: its decrease when d increases shows that agents are gathering around the segment linking both centers. The variables are given on the basis of their value in a reference simulation with $d = 0$ (corresponding to model 1 with only one income group), to allow an easy comparison. The utility of agents U_{mean} decreases when d increases, very slowly when centers are close to each other and then more rapidly. The total surface of the city is always bigger than in the reference (monocentric) simulation, but it decreases when d is high. The mean price of housing and the total rent decrease when d increases, as the share of income used for transport increases.

Within this model, polycentrism is undesirable. It has both a negative economic outcome with the decreasing utility of agents, and a negative environmental outcome, as housing surfaces increase and commuting distances increase. However, it has to be remembered that commuting distances increase mainly because of the minimal commuting distance shown on the left panel of figure 5.9. This effect could be seen as the worst case scenario of a monocentric city evolving towards a polycentric shape: all households increase their travel distances accordingly. A more realistic scenario is given by the case where only a part of the households increase their travel distances, which we study now.

When $0 < m < 1$, simulations show that the utility of agents of the "common" group is always higher that that of agents of the "splitted" group. This is logical, as agents of the "splitted" group have more constraints, as they want to stay close to two places. The outcomes of this model with $0 < m < 1$ are intermediate between the outcomes of this model with only agents of the "splitted" group, shown on the first line of figure 5.8 and on figure 5.9, and those of model 2 with two centers and only one income group, which are presented on the last line of figure 5.8 and in appendix C.3. The second line of figure 5.8 gives the shape of the city with $m = 0.2$ for different values of d, and figure 5.10 gives the corresponding outcomes of model 5. The city shape is an intermediary between $m = 0$ and $m = 1$, that is to say, between two disks and an ellipse. Agents of the "splitted" group (in a paler shade) are located between both centers, separating agents of the "common" group in two parts.

Figure 5.10 shows that the outcomes of the model in this case $m = 0.2$ are also intermediary between those obtained for $m = 0$ and $m = 1$. The total commuting distance decreases at first when d increases, and then increases again, mainly because of the contribution of the minimal commuting distance imposed on agents of the "splitted" group. So that it becomes higher than its value at $d = 0$ when centers are far away from each other. The utility U_0 of "common" agents increases with d, as their competition for land with "splitted" agents decreases. The utility U_1 of "splitted" agents increases at first when d increases, and then decreases again, below its value at $d = 0$. The total surface of the city increases with d, while the mean price of land decreases. The total rent increases at first when d increases, and then decreases, below its value at $d = 0$.

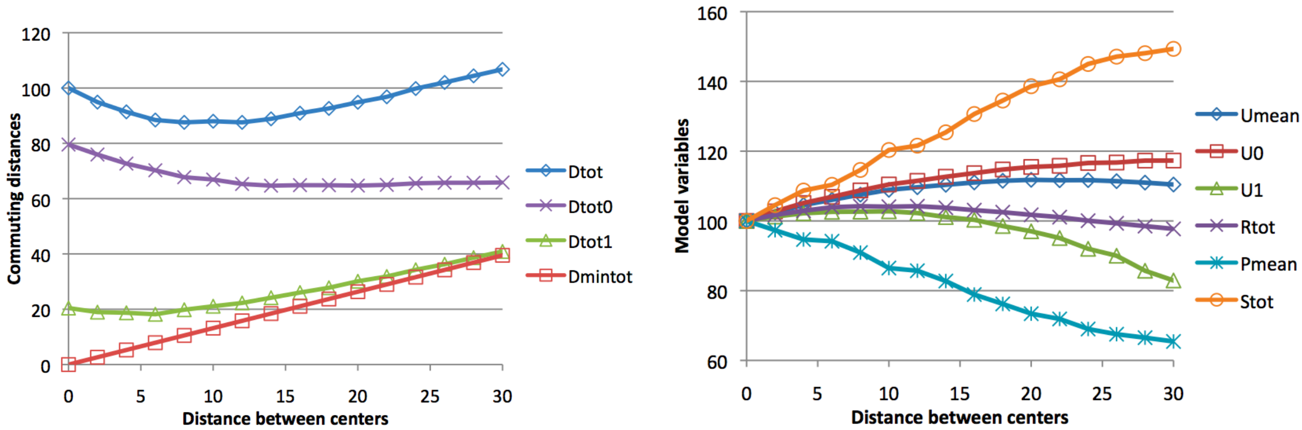


Figure 5.10: Outcomes of model 5 with $m = 0.2$ as functions of the distance d between centers. Same variables as on figure 5.9. On the left pale, the total commuting distances D^0_{tot} and D^1_{tot} of agents of both groups are presented. On the right panel, the mean utility of agents U_{mean}, the utility U_0 of "common" agents and U_1 of "splitted" agents are given. Parameters values are given by table 5.1.

In this case, which seems more realistic than the same model with $m = 1$ presented before, polycentrism is desirable, as long as centers are not moved too far apart from each other. Indeed, the utility of agents of both groups increases when d increases for small values of d, which gives a positive economic outcome of this model. The environmental outcome is also positive, as the total commuting distance decreases when d increases, for small value of d. But this positive effect is mitigated by the fact that housing surfaces increase, which tends to increase emissions of greenhouse gases. Thus this more realistic model tends to confirm the conclusions of model 2, as long as centers are kept not too far away from each other.

5.4 Perspectives and Discussion

Calibration

To have more concrete conclusions on the outcomes of the models we study, an important step is to calibrate the simple urban economics model we use on real data. This would mean that the values of parameters of the agent-based model, such as population size, income or transport cost, should be consistent, at least in a rough way, with real values of a given city (or of a generic city which could be taken as representative of cities of a given size). These parameters would have to give results which are also consistent with real data. The results of the simulated city, like density, housing surface, prices, as functions of the distance to the center, could be compared to empirical datasets.

However, a such work seems impossible using Alonso's model, because this model does not take into account vertical housing: land and housing surface are not distinguished, as all agents live on the ground. A simple solution would be to use an exogenous function of available housing surface depending on the distance to the center, which could be inspired from real-world data. But this solution seems unsatisfactory for the modeler.

Building construction can be introduced in the agent-based model, as it is introduced in Muth's model (see Fujita [1989]): building choices of housing industry are modeled in a simple way to determine the housing surface which is built at a given location. An attempt to calibrate Muth's model on the Lyon urban area is presented in appendix B.

Polycentric city with endogenous centers

This paper presents a study of different simple polycentric models, to explore the outcomes of the AMM model beyond the monocentric framework. But employment centers are still given exogenously, so that the location of jobs can not be studied within this model. It is an interesting perspective of this work to study models with endogenous location of employment centers. The present study can be seen as a first step in this direction. It is indeed important to know what happens with given employment centers, before studying models where these location mechanisms are endogenous.

A model with endogenous location of employment should have a new type of agent which would represent firms, as introduced for instance in Fujita and Ogawa [1982]. In such models, firms compete with a residential use of land and try to maximize their profit. Fujita and Thisse [2003] present such analytical models of one-dimensional cities.

Open and closed city models

As stated in section 5.1, an important choice made in this work is to study a closed city framework, where population size N is fixed. But of course, when looking away from the computer screen to observe economic reality, it is quite obvious that the comparisons made here (corresponding to changes in the models) between mono- and polycentric cities would be in the real world accompanied by changes in population size. This is what is observed in open city models (see Fujita [1989]), where a decrease in transport cost for instance (such as what is done in our model 2, when seen in a schematic way) induces an increase in the population (and city) size, but no increase in agents' welfare.

Indeed, in open city models, the utility of agents at equilibrium is completely determined by the national utility u_{nat} (or utility "of the rest of the world"), which acts as a chemical potential, as seen in chapter 2[5].

From a point a view linked to dynamics, such as that conveyed by the agent-based model, agents arrive in the city following this hypothetical decrease in transport cost while their utility outside (in the rest of the world) is lower than the utility of agents in the city. The increase in population size increases the competition on the housing market and decreases the utility of agents. This evolution stops once the utility in the city is equal to u_{nat}.

The reality of (idealized) urban systems lies certainly in between open and closed city frameworks: supposing that a decrease in transport cost such as that provided by the introduction of polycentrism in our models results indeed in an increasing welfare of inhabitants, it is likely that some agents will migrate into the city, as predicted by open city models, and that the welfare of inhabitants will be globally raised, as predicted by closed city models. A solution to account for both phenomena would be to introduce market frictions in an open city model (for example a cost of moves into the city, preventing some people outside to arrive). Thus, the equilibrium utility of agents in the city would be intermediary between the predictions of open and closed city models. This is beyond the scope of this work.

Broader perspectives and discussion

This work is interesting as a complement of analytical works when analytical results are difficult to obtain. It is not meant to come in competition with a mathematical treatment of urban models. A complete analytical study of the different models presented in this work would surely bring other insights on these simple polycentric models.

[5]It is a drawback of open city models that the parameters determining the population (and city) size is the equilibrium utility, which is difficult to measure to say the least, and is not as intuitive as population size (which is fixed in closed city models). With several income groups, the equilibrium utility levels of the different income groups must be defined.

Hence, two important perspectives can be considered: first, a research perspective is to study dynamic urban models, which are difficult to treat analytically. For instance, once the models presented here have reached an equilibrium, a parameter value is changed (e.g. a raise in transport cost) and the consecutive dynamic changes on the urban systems can be studied, until another equilibrium is reached. Second, a more applied perspective is to design simulation models which could be of an easy use for city planners to help decision-making. Using the robustness of the agent-based dynamics presented here, and applying it to real-world data, for instance a urban road network, simulation models could indeed be designed to study economic, environmental and social consequences of different urban planning policies, within the AMM model.

Conclusion

In this work we present a possible use of agent-based systems in social sciences and in particular in economics. Building on the standard urban economics model (AMM model), we run simulations of a simple model with agents interacting in a urban area. The dynamics of the model consist mainly in agents moving and bidding on housing to represent a competition on the urban housing market. This allows to push our system in the direction of the equilibrium. This equilibrium corresponds to a discrete version of the analytical equilibrium of the AMM model. A comparison shows the very good agreement of the analytical and the agent-based monocentric models with two income groups.

Then we study the evolution of this equilibrium when the monocentric hypothesis is abandoned to explore polycentric cities. Our results present economic, social and environmental outcomes of simple polycentric forms within the agent-based model.

The introduction of several centers, when compared to the monocentric city model, has a positive impact on agents' welfare, as transport expenses and competition on the housing market decrease. Commuting distances are reduced, which gives a positive environmental outcome of the polycentric city in this model. However, the increase of housing surfaces may counterbalance this decrease of greenhouse gases emissions. Although the global effect of a reduction of competition for land between agents is clear, its impact on the different variables of this simple urban model and on different income groups is not obvious, as the results show.

The use of agent-based systems on calibrated urban models could allow to test the effect of different urban policies, and to have a global view of their effect on the urban system. In this goal a calibration of a version of this model where housing construction is endogenous is an interesting perspective of research.

Chapter 6

"European" and "North American" cities

6.1 Introduction

In chapter 4, we presented a discussion on the difficulty to model the social structure of cities using the AMM model. Indeed, as chapter 4 showed, it is difficult to represent in a satisfactory way the "European" type city (like Paris for instance, to take the example used in Brueckner et al. [1999]), with schematically speaking, a rich center and a poor periphery, and the "North American" city (like Detroit), with a poorer center and richer households in the periphery. An important reason for this difficulty is the fact that we only considered a log-linear (or Cobb-Douglas) utility function. This function is very convenient for analytical resolution and has some interesting properties. For instance, it fixes the share of income, net of transport cost, which is used on the different living costs considered by the model: housing and composite good, representing consumptions other than housing. This is interesting in a calibration perspective, if one supposes that the budget shares of composite good and housing do not vary with income. However, this hypothesis can be discussed.

In this chapter, we study the results given by our agent-based model with utility functions which have a form different from the standard log-linear or Cobb-Douglas function used previously. This brings new insights on the question studied of the different social structures between "European" and "North American" cities.

In section 6.2, we study the results of the agent-based model presented in the previous chapter using a Cobb Douglas function, but taking into account the fact that the share of income spent on housing can be shown empirically to vary with income (Accardo and Bugeja [2009]).

The Cobb-Douglas utility function has flaws, and one of them, directly linked to the fact that it fixes budget shares, is that there is no possible substitution between variables. In the case of our study on amenity influence on urban areas (see chapter 4), this is a problem, as Brueckner et al. [1999] showed that functions which allow substitution between variables (coupled with an exogenous amenity) can bring other results about the social structure of cities within the AMM model.

In section 6.3, we illustrate this result thanks to the agent-based model by reproducing a "European" urban social structure with a constant elasticity of substitution (CES, see Chung [1994]) utility function and an exogenous amenity. We also study the behavior of the model when this amenity is not at the location of the central business district (CBD). We show in both sections that a more complex social structure can be found than the usual result of the AMM model. This urban form corresponds to a "European"-type city with a rich suburb, that is, the income as a function of the distance to the city center gives a U-shaped curve.

Finally, in section 6.4, we combine our agent-based model with a logit location model. As a consequence, the choice of utility function is not as crucial anymore, since a part of the locational behavior of agents is determined by the logit rule. This study can be seen as an attempt to bridge the gap between the two main parts of this thesis.

6.2 Evolution with income of housing expenses and value of time

The Cobb-Douglas (or log-linear) utility function has the property to fix the shares of income spent on composite good and housing. These shares are determined by the coefficients of the utility function α and β: $z = \alpha(Y - tx)$ and $ps = \beta(Y - tx)$ (see appendix A). The empirical literature reports that real-world data show an evolution with income of these shares of income spent on different items. On average, when income increases, the share of income spent on housing decreases (see for instance Accardo and Bugeja [2009] or Polacchini [1999]).

Moreover, we already discussed in chapter 4 the fact that the value of time increases with income (see for instance Palma and Fontan [2001]). In this section, we study the influence of these factors together (evolution with income of the value of time and of the share of income spent on housing) on the location choices of households within the AMM model. We still use a Cobb-Douglas utility function $U = z^{\alpha}s^{\beta}$, where z is a composite good, s the surface of housing, α and β are parameters describing agents' preferences for composite good and housing surface, with $\alpha + \beta = 1$. However, we suppose here that the share of income spent on housing and composite good depends on income. With two groups of incomes Y_r and Y_p, $Y_r > Y_p$, this is taken into account by introducing different parameters α_r, β_r and α_p, β_p in the utility functions of rich and poor agents. We still have $\alpha_r + \beta_r = \alpha_p + \beta_p = 1$, but rich agents spend a smaller part of their income on housing: $\beta_r < \beta_p$.

Agents have a budget constraint $Y = z + tx + ps$, where Y is the income, t the transport cost (per unit distance), x the distance to the CBD and p the price of a unit surface of land at location x. We suppose that the value of time is higher for the rich income group. As a consequence, we postulate that the transport cost t_r of rich agents is higher than the transport cost t_p of poor agents, $t_r > t_p$, due to the difference of time costs, supposing that the monetary part of the transport cost is the same for both income groups.

6.2.1 Income elasticities

The condition which determines the social structure of the city is detailed in Fujita [1989] (or Gofette-Nagot et al. [2000]): the income group located near the center is the high income group if the income elasticity of marginal (or unit distance) transport cost is higher than the income elasticity of the demand for housing. If the income elasticity of the demand for housing is higher, then the low income group is located near the center.

The elasticity $e(x, y)$ of variable x with respect to variable y corresponds to the logarithmic derivative of x with respect to y. It is defined as follows:

$$e(x, y) = \frac{\partial x}{\partial y}\frac{y}{x}$$

In our case where only two income groups are considered, we use the discrete version of the elasticity which is called arc elasticity, and is defined as

$$\epsilon(y, x) = \frac{\Delta y}{\Delta x}\frac{\bar{x}}{\bar{y}}$$

where Δy (Δx respectively) is the variation of y (x) between the two points considered, $\bar{y}$ ($\bar{x}$)

is the mean of y (x) between the two points. The arc elasticity is also used in appendix B.

In the case of our model with two income groups, the high income group is located near the center ("European" city) if

$$\epsilon(t, Y) > \epsilon(s, Y) \tag{6.1}$$

that is, if the arc income elasticity of marginal (or unit distance) transport cost is higher than the arc income elasticity of the demand for housing. These two quantities can be written

$$\epsilon(t, Y) = \frac{t_r - t_p}{t_r + t_p} \times \frac{Y_r + Y_p}{Y_r - Y_p}$$

$$\epsilon(s, Y) = \frac{\beta_r Y_r - \beta_p Y_p - (\beta_r t_r - \beta_p t_p)x}{\beta_r Y_r + \beta_p Y_p - (\beta_r t_r + \beta_p t_p)x} \times \frac{Y_r + Y_p}{Y_r - Y_p}$$

Then condition (6.1) can be transformed into

$$\frac{t_r}{t_p} > \frac{\beta_r (Y_r - t_r x)}{\beta_p (Y_p - t_p x)}$$

This inequality can be interpreted as follows: to have rich agents located near the center, the ratio of transport costs of rich and poor agents must be higher than the ratio of housing surfaces. This is a logical translation of the relationship on elasticities in a discrete case with two income groups. An immediate and important consequence of this inequality comes from the fact that the right-hand side depends on the distance x to the center, whereas the left-hand side is constant[1]. This yields richer urban forms in the model than just rich households in the center and poorer ones in the periphery, or vice versa.

Indeed, a distance x_c can be defined, which gives the radius at which arc elasticities $\epsilon(t, Y)$ and $\epsilon(s, Y)$ are equal. At this distance of the center, the order of the inequality (6.1) changes. As a consequence, condition (6.1) can also be written[2]

$$x < x_c = \frac{\beta_p Y_p t_r - \beta_r Y_r t_p}{t_r t_p (\beta_p - \beta_r)}$$

Then different social patterns can be observed in this model, depending on the value of x_c

[1] The left-hand side is constant because we use a linear transport cost, proportional to the distance to the CBD. Studying different transport costs is a perspective of this work.

[2] Dimensional analysis can be used to confirm that x_c is a distance. Incomes are amounts of money, which, divided by a (unit distance) transport cost, give a distance. β is dimensionless.

and on the radius of the city x_b.

If $x_c \leq 0$, the model gives the standard result of the AMM model with two income groups: poor agents are located in the center of the city and rich agents in the periphery. This is usually seen as a representation of a "North American" city. On the contrary, if $x_c \geq x_b$, where x_b is the radius of the city[3], the social pattern is reversed. Rich households are located in the center, and poorer ones in the periphery. This can be seen as a "European"-type city.

A different result can be found if $0 < x_c < x_b$. Indeed in this case, rich households tend to locate closer to the center than poor households when their commuting distance x is such that $0 \leq x < x_c$, and further from the center than poor households when $x_c < x < x_b$. Then a different social pattern can be observed, which would need a complete analytical resolution of the model to be fully characterized, as a perspective of this work. An analytical treatment would for instance compare the bid rent curves (the urban rent as a function of the distance to the center, see Fujita [1989]) of rich and poor agents within this model and show that they can intersect each other two times. In the following section, we use the agent-based model presented in the previous chapter to illustrate the results of the model presented here.

6.2.2 Results of the agent-based model

The dynamics of the agent-based model has already been presented in the previous chapter, so that we only study the equilibrium states of the model here.

The values of parameters we use in the following simulations are presented in table 6.1. The corresponding value of the distance x_c defined in the previous section is $x_c \simeq 9.4$. Let

Parameters	Description	Value
α_p, β_p	Preferences for composite good and housing of poor agents	0.7; 0.3
α_r, β_r	Preferences for composite good and housing of rich agents	0.78; 0.22
Y_r, Y_p	Incomes of rich and poor agents	450, 300
t_r, t_p	Transport cost (unit distance) of rich and poor agents	12, 10
N_r, N_p	Number of rich and poor agents	1000
R_a	Agricultural rent	5
s_{tot}	Surface of a cell	45
ϵ	Bidding parameter	0.5
T_p	Time decrease of the price of non-full cells	100

Table 6.1: Parameters of the model.

us note that in the simulations the side of a cell corresponds to a unit distance. We present

[3]This radius x_b is determined by the boundary conditions of the model, presented in appendix A.

the results of simulations of our model using different population sizes. In this way we model the emergence of a rich periphery in a "European"-type city. The evolution of the urban social structure in this case is shown on figure 6.1. From an initial configuration with a

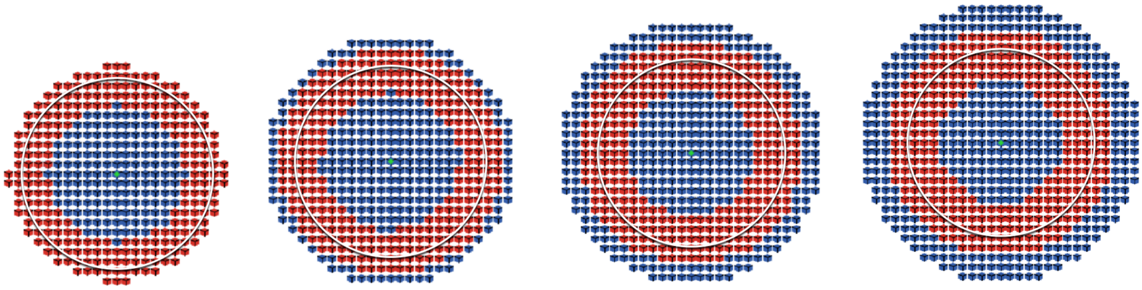


Figure 6.1: Evolution of the equilibrium urban social structure with a growing population size: $N_r = N_p = 1000, 1600, 2200, 2800$ from left to right. The other parameters values are given in table 6.1. Radius $x_c \simeq 9.4$ is indicated by a white circle.

small population size, where rich agents are in the center of the city and poor agents in the periphery, the growth of the population size results in an increase of the city's size and in the formation of a rich periurban area.

It should still be discussed whether the values of parameters we use could be representative of real agents, even in a rough way, as this model is not calibrated on real data. As $Y_r = Y_p \times 1.5$, for a variation of income of 50%, the variation in (unit distance) transport cost is taken as high as 20%, with $t_r = t_p \times 1.2$. We suppose that this variation is mainly due to the variation in transport time cost. And the corresponding variation of the share of income spent on housing is a decrease of roughly 27%, with $\beta_r \simeq \beta_p \times 0.73$.

Supposing that the variation of the global transport cost corresponds only to a variation of the transport time cost, and that monetary and time costs are of the same magnitude for poor agents, the variation of the transport time cost is then an increase of 40% for an increase in income of 50%. This corresponds to an income elasticity of the transport time cost that is smaller than one, which is a standard assumption in the literature (Small [1992], Wardman [2001a,b]).

6.3 Brueckner et al. [1999] revisited

In this section, we first present an illustration of a result which has been shown by Brueckner et al. [1999] to study the difference in urban social structure between "North American" and "European" cities. The AMM model gives easily the "North American" social structure. Indeed, housing is usually considered to be a normal good, whose consumption increases with income. With only one transport mode (associated with a given monetary transport cost) and no time cost, the result of the model corresponds to the "North American" city. The main question is then to have the inverse configuration emerge. In this goal, Brueckner et al. [1999] introduce a central amenity in the AMM model, which can be linked in the reality to the features which make the centers of European cities (with the exception of Brussels for instance) attractive. Brueckner et al. [1999] cite three types of amenities: natural (parks or rivers for instance), historical (e.g. monuments) and modern ones (theaters, swimming pools, etc.). But as chapter 4 showed, a central amenity is not enough to account for a "European"-type city in the AMM model: in a model using a Cobb-Douglas utility function, a central amenity does not inverse the social pattern.

6.3.1 Analytical discussion

To have this inversion, two conditions are given by Brueckner et al. [1999]: the marginal valuation of amenities, after optimal choice of the housing consumption, must rise faster with income than housing consumption, and the gradient of the amenity function must be negative and large in absolute value. An example of utility function satisfying the first condition is given by the constant elasticity of substitution (CES) utility function, which can be written:

$$U_{\mathrm{CES}}(z,s,a) = \left(\alpha z^{-\rho} + \beta s^{-\rho} + (1-\alpha-\beta)a^{-\rho}\right)^{-1/\rho}$$

where the same notations are used as previously for z and s, α and β are now such that $\alpha + \beta \leq 1$, and ρ is a real parameter. The budget constraint is the same as in the previous section, with $t_r = t_p = t$: $Y = z + tx + ps$. $\sigma = 1/(1+\rho)$ is a parameter which is linked to the elasticity of substitution between variables (see Chung [1994]). Brueckner et al. [1999] show that the first condition (the marginal valuation of amenities rising faster with income than housing consumption) is valid if $\sigma < 1$. We use this CES utility function here, with $\rho = 0.3$ ($\sigma \simeq 0.77$), so that this condition is verified.

6.3.2 Agent-based simulations

The amenity function $a(x)$, depending on the distance x to the amenity center, must also be chosen. We take the same function as in chapter 4: a decreasing exponential $a(x) = 1 + a_0 \exp(-x/b)$, with a_0 determining the magnitude of the amenity at its origin and b the characteristic distance of the amenity decrease. The optimal consumption of land (or housing) conditional on price p and location x is

$$s = \frac{Y - tx}{(\alpha p/\beta)^\sigma + p}$$

(see Chung [1994]). This expression is used in the agent-based model instead of the corresponding expression for the Cobb-Douglas function.

Figure 6.2 illustrates the fact that the urban social structure can indeed be inverted under these conditions. The other figures present the urban shape in this model when the work center and the amenity are not at the same location. This model is also sensitive to a small

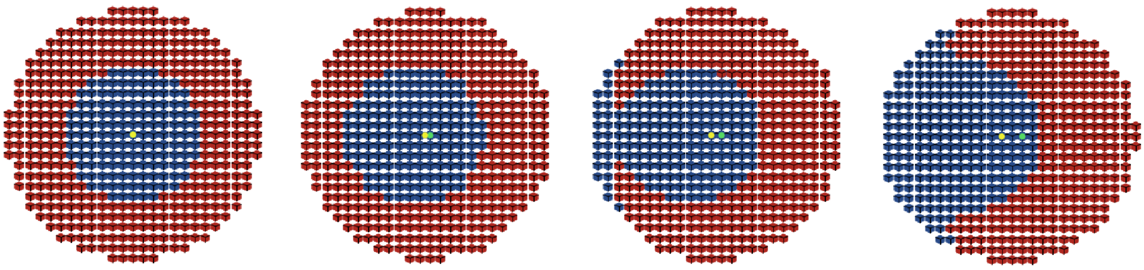


Figure 6.2: Illustration of the result of Brueckner et al. [1999], and evolution of the equilibrium urban social structure when the amenity is pushed to the west of the center. Distance d between the CBD and the amenity is $d = 0; 0.5; 1; 2$ from left to right. Other parameters values are given in table 6.2.

displacement of the amenity, which is a result qualitatively similar to that presented in chapter 4. The parameters used are given in table 6.2.

Then we illustrate a fact which seems to have been overlooked (or simply ignored) by Brueckner et al. [1999] in their study, but appears very easily in our agent-based simulations. This result is in direct correspondence with the one presented in the previous section: the condition which guarantees that the high income group is located near the center depends on the distance to the center. So that it may be verified close to the center and not further away. More precisely, we suppose with Brueckner et al. [1999] that there is a first radius x_1 at which the bid rent curves of rich and poor agents intersect. The condition mentioned before,

Parameters	Description	Default value
ρ	Parameter of the CES utility function	0.3
Y_r, Y_p	Incomes of rich and poor agents	450, 300
t	Transport cost (unit distance)	10
N_r, N_p	Number of rich and poor agents	2000
R_a	Agricultural rent	5
s_{tot}	Surface of a cell	100
ϵ	Bidding parameter	0.5
T_p	Time decrease of the price of non-full cells	100
a_0	Amenity at the origin	5
b	Characteristic distance of decrease of the amenity	4

Table 6.2: Default parameters of the model reproducing the result of Brueckner et al. [1999]

that the gradient of the amenity function is negative and large in absolute value, is verified. So that rich agents are located at $x \leq x_1$ and poorer ones at $x \geq x_1$.

But we suppose now that there is a second radius $x_2 > x_1$ at which these bid rent curves intersect. With the decreasing exponential form of the amenity function we chose, the gradient decreases in absolute value when x increases. This is actually the reason why both curves can intersect again. So that at this second intersection x_2, the gradient of the amenity function can be small in absolute value, and rich agents are now driven to the periphery at $x \geq x_2$, while poor agents are located closer to the center, at $x \leq x_1$.

A urban structure similar to the last configuration presented in the previous section can be observed in this case: a part of the high income group lives in the center, encompassed by a ring where the low income group is located. And further away lives the rest of the high income group, which benefits less of the amenity, but can afford bigger housing surfaces thanks to the lower land prices. This configuration is presented on figure 6.3, which uses the same parameters values as figure 6.2, except that the intensity of the amenity function at the CBD is smaller: $a_0 = 3$. It can be noted that amenity functions more complex than the decreasing exponential form we use could lead to social patterns which are even more complex than the patterns presented on figure 6.3. This should not be really more difficult to handle for the agent-based model used in this work.

Here also, we explore the results of the model when the amenity is not at the same location as the CBD, and we find that the city's social structure is very sensitive to this small displacement of the amenity.

These two first sections give a different viewpoint on the location of income groups within the AMM model, and as a consequence on the question of the "North American" and "Eu-

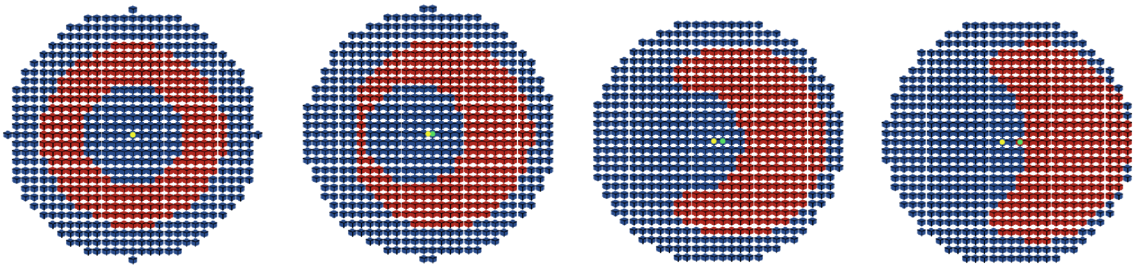


Figure 6.3: Evolution of the equilibrium urban social structure when the amenity is pushed to the west of the center. The distance d between the CBD and the amenity is $d = 0; 0.5; 1; 2$ from left to right. The other parameters values are given by table 6.2, except a_0 which has here a smaller value: $a_0 = 3$.

ropean" urban social structures. The hybrid configuration observed in both models and corresponding to a "European" city structure with a rich suburb, has different origins in both cases.

In the first section, the higher value of time of rich agents, combined with their lower budget share of housing, leads a part of them to have small housing lots in the city center. In the second section, the same outcome is due to a compromise between smaller housing lots and the benefit of a central amenity. However, in each case, the same force drives the other part of rich agents to the periphery: the desire to have bigger housing lots, allowed by smaller price in the periphery. This force, along with the transport cost which attracts each agent to the center, constitutes the fundamental tug-of-war of the AMM model.

The next section gives some preliminary results on a generic way to deal with the fact that utility functions are at the same time the main ingredient and the main unknown when it comes to empirical measurement (and as a consequence the main fragility) of simulated social and economic models: adding a logit choice model to account in a simple way for other factors motivating decisions, i.e. in this case, location choices.

6.4 A logit location choice in the AMM model

The framework used in this section is the same as in chapter 2. Within the agent-based model used until here, the decision to move for a given agent into a certain cell is now probabilistic, where the probability is given by a logit rule

$$P_{\text{move}} = \frac{1}{1 + e^{-\Delta U/T}}$$

with ΔU the variation of the agent's utility associated to this move, and T a parameter which can be seen from a physicist's point of view as a temperature, governing thermal energy in the model. This thermal energy in socioeconomic models represents in a minimal way the phenomena which influence the decision to move in reality, but are not taken into account in the model. From an economist's point of view, an equivalent interpretation is to imagine a representative agent whose utility, in addition to the deterministic part used until now, has a probabilistic term corresponding to a double exponential distribution (see Anderson et al. [1992]). So that the decision to move is not probabilistic, but the utility function is (see also the discussion at the end of chapter 2). We choose the physicist's viewpoint here, the author being better acquainted with this one.

6.4.1 Presentation

Our model using a logit rule bears important similarities with the model presented in chapter 2. However, the present model is a bit more complicated as it does not correspond to a simplified version of the AMM model, but to the full version described in chapter 4. In particular, when agents choose to move into a new cell, they have to propose a bid on the price of housing there. This bid has the same form as in the previous chapters:

$$p_{n+1} = p_n(1 + \epsilon \frac{s_{\text{occ}}}{s_{\text{tot}}} \frac{\Delta U}{U})$$

where p_{n+1} is the increased price, p_n the initial one, $s_{\text{occ}}/s_{\text{tot}}$ the (surface) occupation rate of the candidate cell and $\Delta U/U$ the relative variation of utility during the move (see chapter 4). However, as the logit rule with $T > 0$ may allow some moves which decrease the agent's utility instead of increasing it, the price of land is not changed in this case: if a move yielding $\Delta U < 0$ is chosen, then $p_{n+1} = p_n$.

This model corresponds to an agent-based version of the model studied in De Palma and Papageorgiou [1988], Anas [1990] (see also De Palma et al. [2007]). Its results are qualitatively those predicted by Anas [1990], and are also illustrated on a simpler model in chapter 2. The introduction of a logit choice model blurs the equilibrium pattern of the AMM model by yielding more random decisions. It decreases the urban density of inhabitants close to the center and increases it in the periphery.

In the large T limit (large temperature, or large heterogeneity of tastes), the density is homogeneous, driven by random decisions which can be seen as noise. An important consequence is the fact that when $T > 0$, the deterministic part of the utility is not homogeneous in the system. Agents in the center, where increasing the temperature decreases the competition for housing, have a higher deterministic utility than in the periphery of the city, where the competition for housing is increased. With several income groups, adding noise in the decision process has also the effect to blur the deterministic ($T \to 0$) patterns presented in the previous chapters. In particular, income groups have more and more common locations, that is, segregation is broken, when temperature T increases. Chapter 2 shows that another quantity is homogeneous at the equilibrium in the continuous limit[4] of simpler models: this quantity can be seen as a chemical potential that we denote by λ_{qi}, where q is an index corresponding to the cell and $i = 1, 2$ stands for rich and poor agents (there is one chemical potential for each income group, see chapter 2). Its expression is given by

$$\lambda_{qi} = U_{qi} + T \frac{\partial s}{\partial \rho_{qi}}(\rho_{q1}, \rho_{q2}) \tag{6.2}$$

where U_{qi} is the deterministic part of the utility for income group i in cell q and ρ_{qi}, $i = 1, 2$ are the densities of rich and poor agents in cell q. These densities are defined for each cell q as the quotient $\rho_{qi} = n_{qi}/H$, where $n_{q,i}$ is the number of agents of income group i in the cell, and H a maximal number of agents in the cell, that we introduce in this model. $s(\rho_{q1}, \rho_{q2})$ is the entropic contribution of the cell, which can be expressed as

$$s(\rho_{q1}, \rho_{q2}) = - \sum_{i=1,2} \rho_{qi} \ln \rho_{qi}$$

Hence,

$$\frac{\partial s}{\partial \rho_{qi}}(\rho_{q1}, \rho_{q2}) = - \ln(\rho_{qi}) - 1$$

At equilibrium, in the models studied in chapter 2, the chemical potential is homogeneous: $\lambda_{q,i} = \lambda_i$ for all cells q. But the statistical physics framework we use in chapter 2 does not allow us to predict the outcome of the model studied here, and in particular, the homogeneity of the chemical potential, when the dynamics used does not respect detailed balance, as defined in chapter 2. We can not show that the dynamics of the agent-based model respects detailed balance, as we do not even manage to define a potential L function similar to the one

[4]Chapter 2 presents simulations which confirm this result for a finite number of agents.

presented in chapter 2. So that these results can not be extended to the agent-based model studied in the last chapters.

However, as stated before, this model using a logit location choice can be simulated thanks to the agent-based framework presented in chapter 4. These simulations show that the result of chapter 2 is still true for this more complex agent-based model: the chemical potential λ_i, $i = 1, 2$ defined by equation 6.2 is homogeneous in the city within each income group at equilibrium. The size of the simulation space is fixed exogenously as in chapter 2. It is denoted by r_f. It corresponds to the maximal distance to the center at which housing is available[5]. The result of homogeneity of the chemical potential is illustrated by figure 6.4.

Parameters	Description	Value
α, β	Parameters of the utility function	$0.75; 0.25$
Y_r, Y_p	Incomes of rich and poor agents	450, 300
t	Transport cost (unit distance)	4
N_r, N_p	Number of rich and poor agents	2000
R_a	Agricultural rent	10
s_{tot}	Surface of a cell	2000
ϵ	Bidding parameter	10
T_e	Time decrease of the utility of displaced agents	1000
T_p	Time decrease of the price of non-full cells	500
r_f	Size of the simulation space	50
a, b	Land and capital exponents of the production function	0.5; 0.5
A	Multiplicative factor of the production function	1

Table 6.3: Parameters of the logit models. The three last parameters are specific to Muth's model, presented in appendix B.

In addition, this figure shows that the homogeneity of the chemical potential is also verified for the agent-based version of Muth's model presented in appendix B. This can be more surprising as even the size of cells is not homogeneous in this model. Indeed, the housing industry builds housing using capital and land. More housing is built close to the center than further away, so that the housing surface available in each cell varies with the distance to the center. But the homogeneity of the chemical potential is still verified.

Further work is needed to understand this phenomenon, and link it to the economic literature, for instance De Palma and Papageorgiou [1988], Anas [1990] and Anderson et al. [1992], who present similar results, with a different approach.

[5]There is no such constraint in the standard AMM model. For low values of the temperature T, the city radius may be smaller than r_f because of the agricultural rent. In the high temperature limit, the city is infinitely spread.

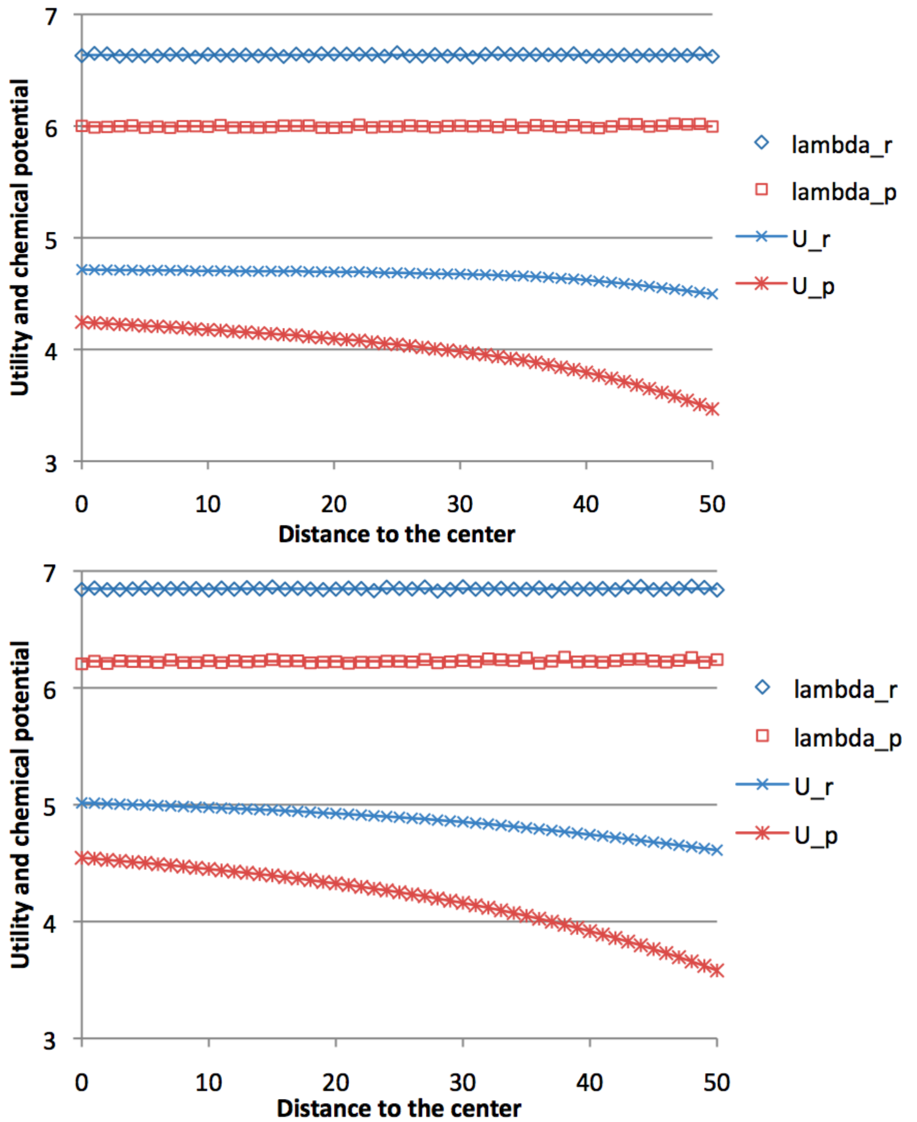


Figure 6.4: Evolution with the distance to the center of the utility and the chemical potential of rich and poor agents. The horizontal lines correspond to the means of the chemical potentials. Parameters values are given by table 6.3. Top panel: standard AMM model, described in chapter 4. Bottom panel: Muth's model, described in appendix B. Additional parameters A, a and b presented in appendix B are used.

6.4.2 Discussion on the equilibrium

It should be noted that the equilibrium of the model using a logit location choice with $T > 0$ is different from the equilibrium studied in detail in the previous chapter. Here, the equilibrium can be seen as "dynamic". Indeed, there are still moves, fluxes, at equilibrium in the system, but these fluxes cancel each other on average. This is expressed for simpler models by the detailed balance (see chapter 1, and Van Kampen [1992], Evans and Hanney [2005]). Even though detailed balance does not hold here, a similar equilibrium is found. This equilibrium has to be studied in a statistical way. Indeed, at each time step, the system is in a configuration which has a certain random character, as illustrated by figure 6.5. This figure can be compared

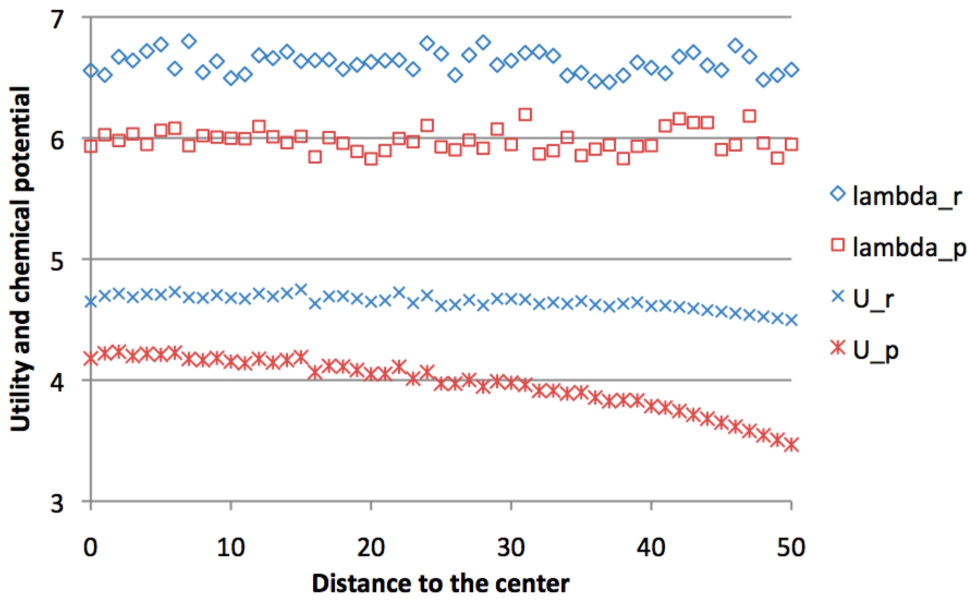


Figure 6.5: Snapshot of the state of the system at a given time: utility and chemical potential of rich and poor agents as functions of the distance to the center. Standard AMM model, described in chapter 4 (to be compared to the top panel of figure 6.4). Parameters values are given by table 6.3.

to the top panel of figure 6.4, which corresponds to a time average of the evolution of the system at equilibrium, corresponding to a high number of configurations similar to the one presented on figure 6.5. Then on average, the behavior of the system at equilibrium can be seen as more predictable. In particular, the chemical potential is found to be homogeneous, as shown by figure 6.4.

The agent-based model presented in the second part of this thesis and used here in a different context has a dynamic evolution which corresponds to a relaxation to an equilibrium. In the two last chapters, this equilibrium is static and is shown to be a discrete version of the static analytical equilibrium of the AMM model. This can be seen as the limit $T \to 0$ of the logit model presented here. This particular situation can be thought of as a situation where no noise is introduced in the system. No agent has an incentive to move and only marginal adjustments on prices and locations take place when the agent-based model approaches the equilibrium[6]. The main problem consists in finding a mechanism of dynamics (and right values of the parameters associated to this mechanism) which brings the system to the equilibrium. Roughly, there is no such difficulty in the model presented in this section (with $T > 0$), where the probabilistic character of moves induces more fluid dynamics and makes the behavior of the model less sensitive to the dynamic mechanism. A more precise study on the behavior of the model as a function of the temperature is a perspective of this preliminary work.

Finally, figure 6.6 presents a preliminary result of a model such as the one studied in section 6.2 (figure 6.3 with $d = 0$), when associated with a logit location choice. The curve

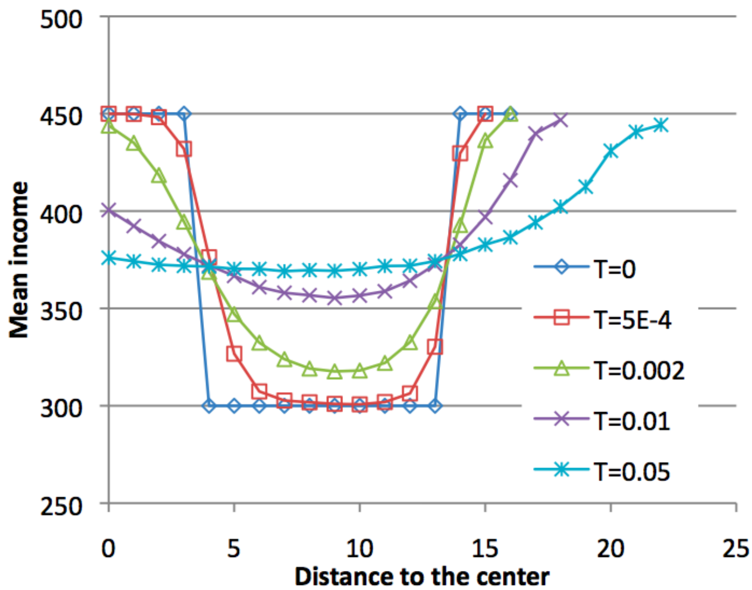


Figure 6.6: Mean income as a function of the distance to the center, for different values of the temperature T. Parameters values are the same as for figure 6.3 with $d = 0$ (see table 6.1), apart from $N_r = N_p = 100$, and with the additional parameter $H = 1000$. The city is linear (one-dimensional). Radius $x_b \simeq 9.4$.

shows the evolution of the mean income with the distance to the center, when the temperature parameter T is changed. For convenience, we use a linear city model, although simulations on a two-dimensional city are also possible. Quite rich periurban areas have developed around some European cities (see Coulombel and Leurent [2011] for Paris for instance), and this U-shaped curve can also be seen in older North American cities, like New York, Chicago or

[6] Indeed, numerically, the equilibrium is never strictly reached. The utility is never completely homogeneous.

Philadelphia (see Glaeser et al. [2008]), which are closer to the "European" model. So that it would be interesting to confront these curves with empirical curves of mean income as a function of the distance to the center. A question would be for instance to find the level of temperature or taste heterogeneity which gives the best link with empirical data.

With this logit location choice model, we see that two important changes are brought to the standard AMM model. First, the logic of the model changes as the result changes: the equilibrium is not static anymore and must be described in a statistical way. Second, the results of the model depend less on the choice of the utility function, as an important part of the behavior of agents is conveyed by the stochastic part of the model.

6.5 Conclusion

In this chapter, we explore the results given by the AMM model in other contexts than the one used in the previous chapters, where the utility function was a simple Cobb-Douglas function. In particular, we study in the two first parts of this chapter some ways to represent a "European"-type city with two income groups within the AMM model. The two frameworks we use give results which are qualitatively very similar. In the first one, with a Cobb-Douglas utility function where the evolution with income of the budget share of housing is taken into account, analytical calculations on income elasticities lead us to define a characteristic radius at which the locational behavior of income groups changes. In the second part of the chapter, a study of the introduction of a central amenity in the AMM model, combined with a CES utility function, allows us to confirm the results of Brueckner et al. [1999], and to go further. Indeed, like the model of the first part, this second model can have as an outcome a "European"-type urban social structure, with rich agents in the center and poorer ones in the periphery, and in addition an outer ring of rich households, which constitutes a rich suburb. The first framework has the advantage, when compared to the second one, to avoid the exogenous introduction of an amenity, which is a bit frustrating from a modeling point of view.

Finally, the introduction of a logit location choice model in the last part of the chapter leads us to consider a different viewpoint on urban systems, closer to models from statistical physics, where the behavior of agents is probabilistic and the equilibrium of the urban system depends less on the choice of the utility function. The main perspectives of this work, in addition to a more thorough exploration of the results of these different models, are similar to those evoked in the previous chapters. A first one concerns the question of the calibration of urban models, to test more precisely their link with empirical data. In particular, the ideas studied here regarding European and North American city models should be interesting to test on empirical data from both continents, if these can be gathered. The question of the calibration of urban models is also studied in appendix B. Another perspective is linked to the historical evolution of cities. Using as an input the historical evolution of population size and transport cost for instance, a comparison could be carried out between the historical evolution as predicted by the model and that appearing in empirical data. This work could be inspired by LeRoy and Sonstelie [1983]. Moreover, this article proposes an hypothesis on the utility function (corresponding to equation (6.1)) and combines it with two different transport modes. It could be interesting to study a such model with our agent-based model. Including two transport modes in the model presented in section 6.2 is also an interesting perspective of work.

General conclusion

The interdisciplinary work presented in this thesis gives some insights on urban modelling, and more generally, about quantitative social models. Let us remind them now.

Overview of the results

The main result of the first part of this thesis is a physical approach to socio-economic complex systems. The use of a global function reflecting individual preferences links individualistic dynamics in simple social models with the macroscopic state of the system. A statistical physics framework can then be used to solve the models.

Resolution of a Schelling-like segregation model

This approach is developed in the first chapter on a famous toy model dealing with spatial segregation, Schelling's model. We present a first analytical resolution, which is allowed by the potential function evoked previously and by a simplification of agents' neighbourhood in this model.

This simplification keeps the essential idea of Schelling's model: a spatial segregation can occur even if agents do not really desire it. Indeed, with two groups of agents which are distinguished in the model by their colour, a utility or welfare function is defined, which is maximum in a mixed environment. But aside from this high desire for diversity, agents still prefer to be among persons of the same colour as themselves rather than among persons of a different colour. This asymmetry is sufficient to obtain segregated equilibrium patterns in the model, if individualistic dynamics is the rule. A logit rule is added, corresponding to having a non zero temperature in the system in physics. In social models, this can be seen as a minimal way to model statistically phenomena which are not taken into account in the model.

An altruistic disposition is then introduced in the model, leading agents to care not only about their personal welfare, but also about the collective utility. This allows to break the spatial segregation, and to observe a phase transition between individual and collective dynamics.

Utility and chemical potential

This statistical physics framework is developed in the second chapter to include a non-homogeneous space. Indeed, while space is homogeneous in Schelling's model, it is polarized in the AMM model for instance. With the logit rule evoked previously, analytical resolution leads to introducing a chemical potential in the system, as a Lagrange multiplier accounting for the conservation of the number of agents. This quantity is shown to be constant in the system at equilibrium. This chemical potential has a strong link with the utility function, describing the welfare of agents. This can be surprising, as utility in social models seems to correspond to the (free) energy in physical ones. Indeed, these functions determine the dynamics of social and physical models respectively. The utility is usually a constant quantity at the equilibrium in social models, so that it could be expected to have a correspondence with the temperature in physical models, which is the conjugate variable of the energy.

However, utility is directly related to chemical potential in our framework: they even coincide at zero temperature. At non-zero temperature, the chemical potential has in addition an entropic term. This framework is illustrated on simplified versions of the AMM model, where an essential assumption is a direct, positive relationship between price and density. This simple model illustrates results concerning the introduction of a logit location choice in the AMM model (De Palma and Papageorgiou [1988], Anas [1990]). A more surprising result is the fact that the homogeneity of chemical potential, which is predicted by this framework, can be also observed through numerical simulations of models which do not respect detailed balance, an essential ingredient of our analytical resolution. This result will be studied more precisely in further developments of this work.

A probabilistic model of housing price formation

The work presented in the third chapter builds on the assumption of linked price and density in urban areas. The essential ingredient is an ergodic hypothesis, which enables us to relate time variations to spatial ones. The study of a simple urban housing market then concentrates on a given flat, whose price increases or decreases in time following the demand of tenants. This flat is seen through the ergodic hypothesis as representative of a whole urban market. However, more work is needed to solve a problem of consistency of this model linked to ergodicity.

This model presents a minimal description of the combined dynamics of price and density in a urban housing market. The equilibrium is shown not to depend on the initialization. We present a study of the evolution of the price distribution in the city, when the density, which is the most “physical" parameter of the model, varies. The results support our hypothesis on the relationship between price and density. The main perspective of this work, aside from a more thorough study of the equilibrium state, depending on the different parameters, is to perform analytical (or numerical) resolutions in limiting cases, using statistical physics tools. The simplicity of the probabilistic dynamics used should be an advantage in this respect.

An agent-based model of urban economics

The second part of this thesis presents an agent-based model, which can be seen as intermediate in descriptive ability between Schelling's model and heavy urban simulations on land use and transport interaction (LUTI). The approach consists in having a parcimonious model, but still more descriptive than Schelling's model. Two versions of this model are presented respectively in chapters 4 and 5. Both are shown to come in very good agreement with the analytical results of the AMM model. The main difference is of course that the agent-based model is discrete in space (which corresponds to a two-dimensional grid) and because of the finite number of agents, while the analytical model is continuous.

The dynamics of this agent-based model is inspired by the AMM model: a competition for housing between agents is organized by a simple local mechanism of moves and bids in the simulated city. More precisely, agents move to optimize a utility function, which is a log-linear or Cobb-Douglas function in chapters 4 and 5, starting from a random initial state. The consumption of housing is optimal, conditional on location and price. Agents bid on the local price of housing in order to change locations. These bids are taken as proportional to the relative variation of utility during the move.

Two conditions need to be respected to ensure that the agent-based model has reached a discrete version of the analytical equilibrium. These conditions are described in chapter 5: the utility of agents must be homogeneous (that is, present small relative variations numerically), and space must be used optimally (also in a numerical sense).

This agent-based model is implemented in chapter 4 to study the influence of a positive amenity in the AMM model. The research question concerns in particular the different urban social structures in Europe and North America, with, schematically, city centers usually richer than the periphery in the first case, and richer households in the periphery in the second one.

An amenity has two main effects, which are related, on the urban system. First, it concentrates agents around the amenity center, as a higher level of welfare can be achieved thanks to the amenity. This tends to densify the city and reduce the city size as well as commuting distances, if the amenity is close to the center. But a second effect appears if the amenity is not located near the center. In this case, the urban shape is distorted by the amenity, and commuting distances can increase even if the city size decreases. With an attractive amenity located far away from the center, a phenomenon of leapfrog development can be reproduced in the model, with a secondary town appearing on the outskirts of the main city, around the amenity.

On the question of "European" and "North American" city structures, the results are a bit unsatisfactory in this chapter: with two income groups, a transport time cost and a value of time are introduced in the model. This value of time is higher for rich agents than for poor agents, as empirical studies suggest. But realistic values of this parameter do not allow us to have a "European" type city, with the log-linear utility function we use. We conclude that in this model, the value of time alone cannot produce a European city structure. As the standard result of the AMM model is the "North American" city type, the challenge is namely to find a suitable way to obtain "European" city equilibria. With this log-linear utility also, adding a central amenity in the city, for which rich and poor agents have the same preference, does not inverse the social structure. If rich agents have stronger preferences for this amenity than poorer ones, then a "European" type city can appear. However, such different preferences for an amenity would need to be confirmed by empirical studies. As it is, it seems a too exogenous ingredient, and frustrating from a modelling point of view. Nevertheless, in this case, simulations with a non-central amenity show that the urban social structure is very sensitive to the fact that employment center and amenity center are not at the same location, in a two-dimensional space.

Exploring the polycentric city

Then chapter 5 uses a slightly different version of the same agent-based model to focus on the study of the polycentric city, which is an important subject of study in the urban economics literature. Indeed, although the monocentric framework is a powerful one, employment is obviously not concentrated in a single point in reality, even when looking at a city from far away. A polycentric framework can be closer to the reality of urban systems.

Different models are simulated in this chapter. The simplest one consists in defining two work centers and to let agents work at the one closest to their housing location. Different arguments are given to justify the existence and uniqueness of the equilibrium in this model and the other ones studied in the chapter. The evolution of the urban systems is examined thanks to the evolution of different variables characterising agents' welfare, their commuting distances or their housing surfaces for instance. In this simple model, the polycentric city is desirable from an economic point of view, based on the welfare of agents. But the environmental outcome is not as clear. On one side, commuting distances tend to decreases if the centers are not too far away from each other, which has a positive consequence with the reduction of pollution linked to transport. On the other side, housing surfaces increase, which spreads the city and results in higher energy needs for heating and cooling. A calibrated model is necessary to draw clear conclusions on this point.

Other polycentric models are studied where agents have constraints regarding their work place. The outcomes of such models are found to confirm to a certain degree the results of the simplest model evoked previously. In particular, a more realistic model is considered, where agents represent two-workers households. Both persons in the household work at the same employment center, or in different places. Then the city has more consistency and the previous results of positive economic outcome and mitigated environmental one are found again, as long as centers are kept close to each other. If centers are too far away, the polycentric city is undesirable.

"European" and "North American" city

Finally, in chapter 6, our agent-based simulations reproducing the AMM model are used to study again the question of the different urban social structures between "European" and

"North American" cities. But contrary to chapter 4, we do not use only a standard log-linear function. The first part of the chapter considers a Cobb-Douglas function whose parameters vary with income, to take into account the fact that the budget share of housing decreases empirically when income increases. Two income groups are defined, which have in addition different values of time. Hence, the richer income group has a higher global transport cost, supposing that monetary cost is fixed.

The result of such a model, even with a transport cost proportional to the distance to the center, is a bit more complex than standard urban social structures found within the AMM model. We are led to define a distance at which the residential behaviour of agents with respect to income changes. Then depending on the value of this characteristic distance and on the radius of the city, different outcomes can be found. The standard "North American" and "European" city structure can emerge at equilibrium, but also a mixed structure corresponding to a "European" city growing a rich periurban area. This can also be linked to the social structure of old American cities such as New York, Chicago or Philadelphia (see Glaeser et al. [2008]).

In a second part, a similar result is found by building on a model proposed by Brueckner et al. [1999]. A central amenity is considered, and the form of the utility function (a constant elasticity of substitution function in our case) is supposed to let richer agents value it more than poorer ones. This can be linked to our work of chapter 4, but using a utility function which allows substitution between different consumptions. The agent-based model manages to find the equilibrium state, which was not done by the analytical work of Brueckner et al. [1999]. Again, a "European" type city or the mixed structure described previously can result from interactions between agents in this model, with a simple amenity function. The mixed structure seems to have been overlooked by Brueckner et al. [1999]. However, the introduction of an amenity is still an exogenous ingredient, which models such as the one presented in the first part of the chapter can allow to avoid. We also study the behaviour of the urban system when the amenity is not in the exact same location as the work center. The urban social structure is found to depend strongly on the distance between amenity and CBD.

The last section of chapter 6 presents the same agent-based model (the version presented in chapter 4), where a logit location choice is included. Surprisingly, the result of homogeneity of the chemical potential defined in chapter 2 is validated by the simulations, while the statistical physics framework used in chapter 2 cannot be used. This brings together both parts of this thesis, and further work will need to link this last result with existing economic literature, and to explore more precisely the other results of this AMM model with probabilistic moves.

Some elements of discussion

The work presented in the first part of this thesis, using a statistical physics framework, is a good illustration of a certain trend in the domain of social complex systems, a domain still emerging. It is the result of a study with statistical physics tools, which are used outside their original context. A danger of such work is that its results may be interesting neither for physicists, nor for economists. Although it is an important goal that the interdisciplinary exchange may be a benefit to both sides, it is not obvious that the models presented in the first three chapters can bring interesting ideas when returning to objects and phenomena usually studied in physics. The interest of physicists for this kind of work seems mainly due to a non standard use of statistical physics tools, on different objects and phenomena. And these models may be uninteresting to economists, because the point of view and the tools used are too different, or because the level of description is too rough. Nevertheless, making links between disciplinary domains and theories is of course a thrilling intellectual challenge, with a very rich scientific potential, which our work tries to expose. It must be a constant concern in such work to make the interests of both scientific sides meet.

This thesis as a whole can be seen as an attempt to link concretely two disciplines and make them interact. The modelling framework in general is an interesting common feature of both fields of science, and as a consequence, a tempting point to start an exchange. More specifically, linking interactions at a "microscopic" level to an outcome at a "macroscopic" one is a fundamental approach in both domains. An important part of this work relies on an innovative analytical formalism which allows the statistical physics treatment of the relationship between both levels of description to be used on social models. This emphasizes a fundamental difference between physical and social systems, as described by the models: the dynamics of physical systems is collective, while social systems are driven by more individualistic forces.

In the second part of the thesis, the main specificity of the work is the modelling tool used, namely agent-based simulations. Most of the work in social sciences using agent-based modelling relies on a precise description of agents and interactions. It is indeed a constant temptation when using this tool to refine some elements, for instance agents' characteristics, rules of behaviour or the description of the environment. This can lead to models with a twisted link to reality, where for instance the description of agents is very realistic, but the model does not manage to include a transcription of their behaviour which is at the same time inevitably rough, but also descriptive enough.

In this context, the standard urban economics model has the advantage to have benefited of decades of work, which polished it so as to have a coherent level of description on the different entities – that is, a rough one, but still descriptive. In this regard, a good calibration of this model (or attempt to calibrate it) on empirical data regarding a given city is lacking in this work, in spite of the attempt presented in appendix B, where the lack of data is the main obstacle. And even better would be an attempt to calibrate it on different real urban areas. Hence, the link of the AMM model with reality is difficult to assess and provides an interesting research question for further work. In this respect, the study presented in the second part of this thesis is representative of the literature on this model, where no successful attempts to calibrate it can be found to the author's knowledge. This does not stop a rich ongoing literature on this modelling framework.

Another specificity linked to the use of agent-based models, and to computer simulations in general, is the fact that the forms of utility functions for instance need to be chosen. Parameters values are also chosen and because of the tremendous size of the parameters' space, even with simple models, only a very small region of it can be explored. No general analytical work is possible, contrary to what constitutes a large part of this literature. This has at the same time the advantage to encourage a calibration of the model and a close link to reality, and the disadvantage to confine sometimes such work in a very small simulation space. An illustration is given here with the Cobb-Douglas utility function, which is used in the most part of this work (but also in most of the literature on the AMM model).

With the different chapters of this thesis, a panel of urban models is given, of different levels of complexity and descriptive ability, ranging from Schelling's model, which is very close to statistical physics models, to the Alonso, Muth, Mills model, much more descriptive and at the heart of urban economics. The tools used are a combination of analytical treatment, numerical resolution, and computer simulations close to the analytical benchmark. This panel of models and resolution methods is developed as a path to foster dialogue between physics and economics on urban modelling.

Appendix A

Analytical resolution of the standard urban economics model

A.1 Resolution of the AMM model

The utility function of individuals in the models we presented has a log-linear (or equivalently, a Cobb-Douglas) form:

$$u = \alpha \ln(z) + \beta \ln(s)$$

where z is a composite good representing all consumer goods apart from housing, s is the housing surface and $\alpha + \beta = 1$. For fixed s and u, it can be written

$$z(s, u) = s^{-\beta/\alpha} e^{u/\alpha}$$

Each individual has a budget constraint

$$Y = z + tx + ps$$

where Y is the income, t the transport cost for a unit distance, x the commuting distance (distance to the CBD in the monocentric model) and p the price of housing for a unit surface. In the simple versions of the model, agents are all identical.

A.1.1 Resolution of Fujita [1989] using bid rents

In the equilibrium configuration of the city, the utility of agents is homogeneous, no agent has an incentive to move in order to increase his welfare. Agents compete for housing and are led to pay at a given location the maximal price which provides them with a welfare level corresponding to the equilibrium utility. This price is called bid rent. Supposing that this equilibrium utility u is fixed (we show later how it is determined), the bid rent can be written, for a commuting distance x

$$\psi(x,u) = \max_s \frac{Y - tx - z(s,u)}{s}$$

Writing that the derivative with respect to s of the left-hand side, where $z(s,u)$ is replaced by its expression, is equal to zero, we have

$$s(x,u) = \alpha^{-\alpha/\beta}(Y - tx)^{-\alpha/\beta}e^{u/\beta}$$

and then replacing in $z(s,u)$

$$\begin{aligned} z(s,u) &= \alpha(Y - tx) \\ \psi(x,u) &= \alpha^{\alpha/\beta}\beta(Y - tx)^{1/\beta}e^{-u/\beta} \end{aligned}$$

A.1.2 Resolution using the optimal composite good consumption

The utility in the city u is fixed. To determine the optimal composite good consumption at a given commuting distance x and for a given housing surface s and price p, we write s as a function of z thanks to the budget constraint

$$s = \frac{Y - tx - z}{p}$$

The the utility can be written

$$u = \alpha\ln(z) + \beta\ln\big(\frac{Y - tx - z}{p}\big)$$

The optimal composite good consumption is given by

$$\frac{\partial u}{\partial z} = \frac{\alpha(Y - Tx) - z}{z(Y - Tx - z)} = 0$$

which gives $z = \alpha(Y - tx)$. But the form of the utility function gives also $s = z^{-\alpha/\beta} e^{u/\beta}$, which allows to write

$$s = \alpha^{-\alpha/\beta}(Y - tx)^{-\alpha/\beta} e^{u/\beta}$$

and still replacing, the bid rent is

$$\psi = \frac{Y - tx - z}{s} = \alpha^{\alpha/\beta}\beta(Y - tx)^{1/\beta} e^{-u/\beta}$$

A.1.3 Resolution using the optimal housing surface

Using the budget constraint we can write

$$u = \alpha \ln(Y - tx - ps) + \beta \ln(s)$$

We determine which housing surface gives the maximum utility at a given location and price

$$\frac{\partial U}{\partial s} = 0 \Rightarrow ps = \beta(Y - tx)$$

Then the budget constraint gives $z = (Y - tx)(1 - \beta) = \alpha(Y - tx)$, and by replacing just as before we obtain s and ψ.

A.1.4 Resolution using a Lagrangian

The Lagrangian associated to this problem is

$$\mathcal{L} = \alpha \ln(z) + \beta \ln(s) + \lambda\big(Y - tx - ps - z\big)$$

where λ is a Lagrange multiplier. We write the first order conditions

$$\frac{\partial \mathcal{L}}{\partial z} = \frac{\alpha}{z} - \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial s} = \frac{\beta}{s} - \lambda p = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = Y - tx - ps - z = 0$$

Solving this system of equations gives the consumption of households of composite good $z = \alpha(Y - tx)$ and housing $s = \beta\frac{(Y-tx)}{p}$.

These different ways of solving the model give different viewpoints. The most important point of view used in Fujita [1989] is the first one, but the agent-based model of the second part of this thesis chooses rather the third one.

A.1.5 With several income groups

This simple model can easily be generalized to several income groups, as presented in Fujita [1989]: the bid rent curve $\psi(x, u = u_1, Y = Y_1)$ is steeper than the curve $\psi(x, u = u_2, Y = Y_2)$ if $Y_1 > Y_2$ and $u_1 > u_2$[1]. As individuals compete for housing, the highest bidder wins: with this log-linear (or Cobb-Douglas) utility function, poor agents are located in the center of the city, where their bid rent is above that of rich agents, and rich agents are located in the periphery. The border between income groups is located at the commuting distance where their bid rents cross each other.

A.2 Boundary conditions: open and closed city model

Let us go back to the simple model where all agents are identical. This model gives the shape of the bid rent, the housing surface and the composite good, as functions of the commuting distance x. However, this model does not predict the size of the city, nor the population size or the utility of agents. It does not take into account the physical constraint of the problem, which says that between a distance x and a distance $x + dx$ from the center, the housing surface available is not infinite. For a two-dimensional space for instance, it is a disc of area $2\pi x dx$. This constraint is considered in a further version of this same model, which also allows to determine the equilibrium utility.

[1] If two groups are considered, which are only distinguished by their income, the equilibrium utility of rich agents is higher in this model.

To do that, boundary conditions are needed, which determine the equilibrium utility of the city, its size and population size. An agricultural rent R_a is introduced, corresponding to the price of land which is not used for housing. In the model, this land has an agricultural use. At the border of the city, agents compete with an agricultural use of land represented by this agricultural rent, so that their bid rent is equal to R_a. This gives a first boundary condition

$$\psi(x_f, u) = \alpha^{\alpha/\beta}\beta(Y - tx_f)^{1/\beta}e^{-u/\beta} = R_a$$

where x_f is the radius of the city.

For the second condition, two alternatives are used. It can be chosen to fix the population size N of the city, which corresponds to a closed city model. In this case, utility decreases when the population size increases, all equal else (because of a higher competition for housing). The other possible choice corresponds to the open city model, where the utility u_{out} outside the city (or national utility, or utility of the rest of the world) is fixed. This utility level determines the population size of the city: individuals have an incentive to stay in the city while their utility is higher than the national utility u_{out}, and to go if it is lower. So that the utility in the city is equal to the national utility at equilibrium. The population size increases when the national utility decreases, all equal else.

To write mathematically that a given population is located in a given surface, the population density $\rho(x)$ at a given commuting distance x is introduced: it is the inverse of the optimal surface obtained through the resolution of the simple model without boundary conditions

$$\rho(x, u) = \frac{1}{s(x, u)} = \alpha^{\alpha/\beta}(Y - tx)^{\alpha/\beta}e^{-u/\beta}$$

For a circular (two-dimensional) city, the population constraint can be written

$$\int_0^{x_f} 2\pi x\rho(x)dx = N$$

where N is the population size. In the closed city model, this population size is fixed and in this last equation and u is unknown. In the open city model, N is unknown and u is equal to the national utility u_{out}.

Both constraints give a system of two equations with two unknowns, x_f and u for the closed city, x_f and N for the open city. With several income groups, these constraints are easily generalized.

A.3 Elasticity of the demand for housing and of the marginal transport cost

Let us now consider a city with two groups of agents with respective incomes Y_r and Y_p, $Y_r > Y_p$. To find a condition determining which income group is located in the center or in the outskirts of the city, we compare here the income elasticities of the demand for housing and of the marginal transport cost. Indeed, demand for housing and transport cost are the two opposing forces determining the location of agents in urban economics models. Their evolution with income is studied thanks to income elasticities, as is usually done in the literature (see Fujita [1989] or Gofette-Nagot et al. [2000]). Let us consider two income groups with incomes Y_p and Y_r respectively, with $Y_p < Y_r$: we use arc elasticities defined as

$$\epsilon(y,x) = \frac{\Delta y}{\Delta x}\frac{\bar{x}}{\bar{y}}$$

where $\epsilon(y,x)$ is the arc elasticity of variable y with respect to variable x, Δy (Δx respectively) is the variation of y (x) between the two points considered, $\bar{y}$ ($\bar{x}$) is the mean of y (x) between the two points.

The demand for housing in the model of Alonso with a log-linear (or Cobb-Douglas) function is given by equation: $s = \beta(Y - Tx)/p$ (see section 4.1). The income arc elasticity of the demand for housing is then

$$\epsilon(s,Y) = \frac{Y_r - Y_p - (T_r - T_p)x}{Y_r + Y_p - (T_r + T_p)x}.\frac{Y_r + Y_p}{Y_r - Y_p}$$

where p and r subscripts indicate "poor" and "rich" agents.

The income arc elasticity of marginal transport cost T is

$$\epsilon(T,Y) = \frac{T_r - T_p}{T_r + T_p}.\frac{Y_r + Y_p}{Y_r - Y_p}$$

To have rich agents located in the center of the city, the elasticity of marginal transport cost must be greater: $\epsilon(T,Y) > \epsilon(s,Y)$, which can be written

$$\frac{T_r - T_p}{T_r + T_p} > \frac{Y_r - Y_p - (T_r - T_p)x}{Y_r + Y_p - (T_r + T_p)x}$$

and reduced to condition (4.1): $T_r/T_p > Y_r/Y_p$, or in an equivalent way $T_r/Y_r > T_p/Y_p$.

Appendix B

Calibration of a Muth model

B.1 Muth model including housing industry

B.1.1 Analytical model

We consider in this thesis two versions of the AMM model: the first one, which we will call Alonso's version, presented in chapter 4, is simpler because building construction is not studied. Land and housing surfaces are not distinguished, there is no vertical housing. In the second version presented here, following the contribution of Muth [1969], building construction is managed by firms of the housing industry, using land and capital. These firms have a production function F which is assumed to have constant returns to scale. Let q be the quantity of housing service consumed by an agent: the production function of firms has here a Cobb-Douglas form and can be written $q = F(s, k) = As^a k^b$, with s and k the quantity of land and capital used by the housing industry to produce a quantity q of housing service. A, a and b are parameters characterizing this production function, with $a + b = 1$ to ensure the concavity of F in s and k.

Land rent and housing rent are distinguished. Firms compete for land use and seek to maximize their profit $\Pi = p_H q - ps - k$, where p_H is the price of housing and p the price of land (both per unit surface). Land can also be used for agriculture, providing an agricultural rent R_a per unit surface to the landowners. The competition between firms for land implies that land rents p within the city must be higher than the agricultural rent, and that the profit of firms is zero at the equilibrium.

Agents have a utility function which has here a log-linear form $U = \alpha \ln z + \beta \ln q$, where z is a composite good representing all consumer goods except housing and transport, q is the surface of housing, α and β are parameters describing agents' preferences for composite good and housing surface. These last parameters are chosen so that $\alpha + \beta = 1$, without loss of generality. This log-linear form of the utility is often used and gives the same results as a Cobb-Douglas function (the exponential of this log-linear function). Agents also have a budget constraint $Y = z + tx + p_H q = z + tx + ps + k$, where Y is their income, t the transport cost per unit distance, and x their distance to the CBD. The second equality is given by the condition of zero profit of firms of the housing industry, which are supposed to compete.

Alonso's version of the model is somewhat simpler: housing and land surfaces are not distinguished, so we can denote them here by s_A. Agents have a utility function $U = \alpha \ln z + \beta \ln s_A$, and a budget constraint $Y = z + tx + ps_A$ (land and housing prices are not distinguished either). Production of housing is not studied, so that the behavior of housing firms can be ignored (see Fujita [1989] for a more detailed description of both versions of the model).

We consider here a closed city model (see appendix A), where the population size N is fixed.

B.1.2 Agent-based implementation

The agent-based implementation of Muth's model is quite similar to that of Alonso's model, presented in chapter 4.

Following the resolution of Muth's model presented in Fujita [1989], firms of the housing industry are left apart and agents choose land and capital inputs themselves. Instead of competing for housing (bidding on p_H), they compete for land and bid on land price p. At a given location, they choose the quantities of land and capital which provide them with the higher possible utility with price p: $s = a\beta \frac{Y-tx}{p}$ and $k = b\beta(Y - tx)$. This determines the quantity of composite good they consume and also their utility. These expressions for the optimal land and capital quantities can be obtained by differentiating with respect to land surface s and capital k the utility function, at fixed location and price (see appendix B.3). The quantity of housing service they consume is then given by the production function of firms $q = F(s, k) = As^a k^b$, and the price of housing is given by the condition of zero profit of firms $p_H = (ps - k)/q$.

We use the same bid mechanism as for Alonso's model (see chapter 4). However, as stated previously, agents bid on land an not on housing, as land and housing are distinguished here.

B.2 Comparison with the analytical model and calibration

In this paragraph, we present a first attempt of calibration of Muth's version of the standard urban economics model on the Urban Community of Lyon, the city where this work has been carried out. This Urban Community is composed by the city of Lyon and some of its suburbs, with a total of 1.2 million inhabitants. With a mean of 2.3 persons per household in France,

Parameters	Description	Value
α, β	Preferences for composite good and housing	0.7; 0.3
a, b	Land and capital exponents of the production function	0.2; 0.8
A	Multiplicative of the production function	0.025
Y	Income	3×10^4€/year
t	Transport cost	0.3€/m/year
N	Population	5×10^5 inhabitants
R_a	Agricultural rent	8€/m^2/year
s_{tot}	Surface of a cell	200m $\times$ 200m $= 4 \times 10^4$m^2

Table B.1: Parameters of the calibrated Muth model

we perform a simulation with 0.5 million households. This simulation runs for a much longer time than the other simulations presented in this work before it reaches the equilibrium. The simulated city is a grid of square cells which are considered to be 200 m long. Only one income group is considered, to allow the comparison with the analytical results. The income of each households is fixed at 30 k€ per year (2500 € per month). The transport cost is composed of a monetary and of a time cost. The monetary cost is taken as 0.30 €/km (for a one way trip), which is realistic for car travel. And the time cost is taken as 8 € per hour, which gives 0.32 €/km at a speed of 25 km/h. With 250 roundtrips per year, the global transport cost is 300€ per kilometer and per year. The agricultural land rent is taken as 8 € per square meter per year (0.67 €/m^2/month): this is our guessed land rent at the border of the urban community. All values of parameters are given in table B.1. The model only considers the residential use of land. As in reality only a given share of urban land is dedicated to housing, this should be introduced in the agent-based model. Here we use 25% as an estimated value of the mean share of land which is used for housing in the Urban Community of Lyon.

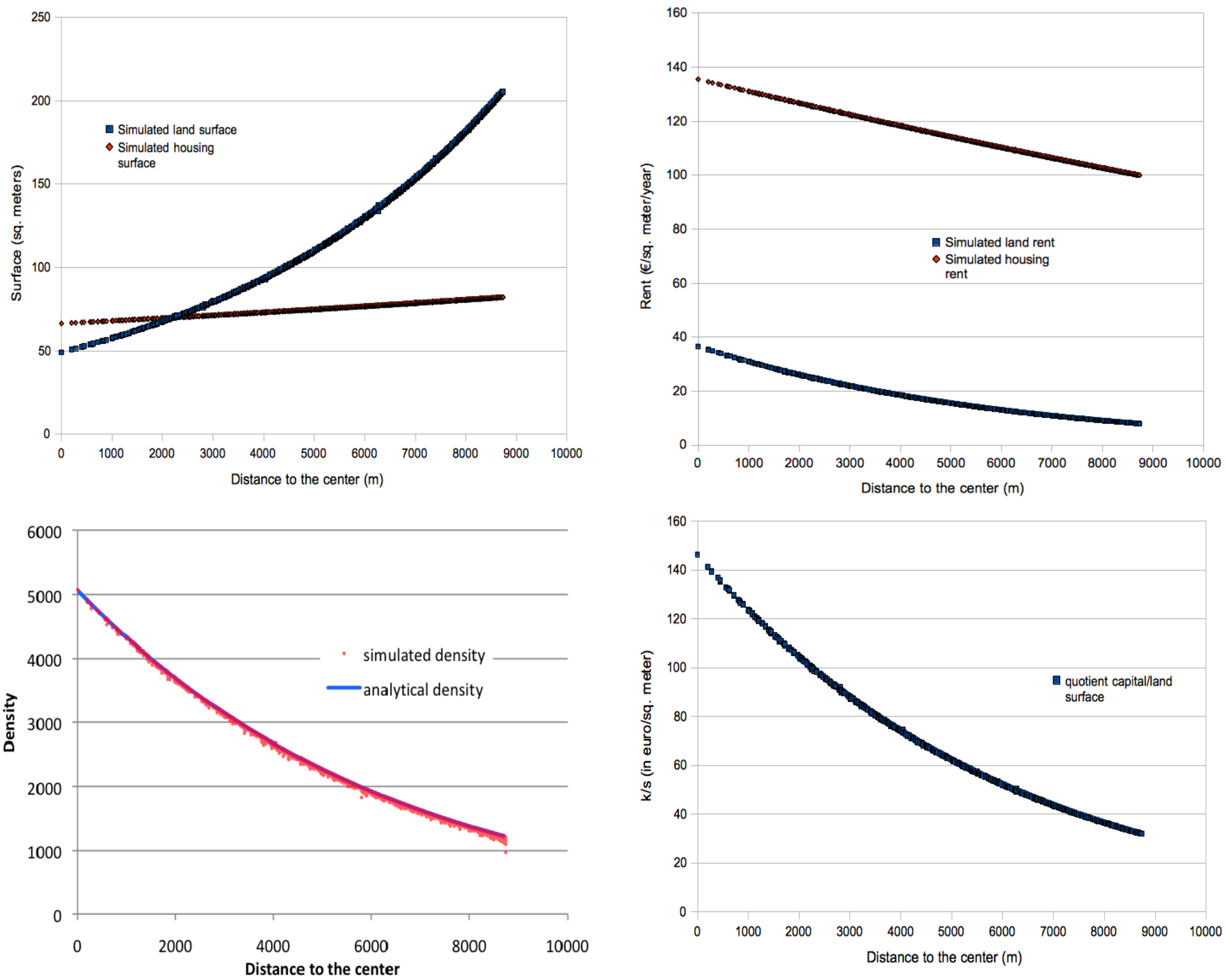


Figure B.1: Results of the calibrated model: surfaces (in m^2), rents (in €/m^2/year), density (in inhabitants/km^2) and capital-land ratio (in €/m^2), as functions of the distance to the center (in m).

On figure B.1 are given the results of this calibrated model. Only the density curve is compared to the analytical curve, the other curves coincide with the corresponding analytical curves. The surface of the Urban Community of Lyon is approximately 500 km^2, so that its radius would be of nearly 13 km if its shape were a disc. The radius of the simulated city is approximately 9 km. This model gives values of variables which are coherent in their evolution with the distance to the CBD, and seem to be of the right order of magnitude. However, land rent is lower than housing rent even in the center of the city, which is certainly unrealistic. In addition, the share of land which is used for housing is taken as constant on the whole territory, which is an unrealistic hypothesis because land is more urbanized near the center of the city than in the suburbs. This contributes to explaining why the density curve is not as steep as it is in reality, because the density is underestimated near the center and overestimated in the periphery.

It is still a perspective of this work to compare these results precisely to real data and understand the differences. What prevented us from doing so until now is the lack of empirical data. Indeed, it is quite difficult to gather enough data to confront all the results of the model, presented on figure B.1, with the reality. In particular, we do not find empirical data concerning the price of land. A confrontation with only some empirical results would have a lot of chances to be successful, as some constraints would be removed.But it would be an underachievement (to say the less) from a modeling point of view, and it could generate misleading interpretations.

B.3 Computation of the optimal land and capital quantities in Muth's model

Price of land p and distance to the center x are fixed. We compute the derivatives of agents' utility U with respect to s and k to find the optimal quantities of land and capital inputs for an agent:

$$\begin{aligned}
U &= \alpha \ln\big(Y - tx - ps - k\big) + a\beta \ln s + b\beta \ln k + \beta \ln A \\
\frac{\partial U}{\partial s} &= \frac{-\alpha p}{Y - tx - ps - k} + \frac{\beta a}{s} \\
\partial U/\partial s &= 0 \text{ if } \alpha ps = \beta a(Y - tx - ps - k) \\
\frac{\partial U}{\partial k} &= \frac{-\alpha}{Y - tx - ps - k} + \frac{\beta b}{k} \\
\partial U/\partial k &= 0 \text{ if } \alpha k = \beta b(Y - tx - ps - k)
\end{aligned}$$

We have a system of two equations in s and k:

$$\begin{cases} \alpha ps + \beta a(ps+k) & = & \beta a(Y-tx) \\ \alpha k + \beta b(ps+k) & = & \beta b(Y-tx) \end{cases}$$

Adding these two equations, knowing that $\alpha+\beta = a+b = 1$, gives $k+ps = \beta(Y-tx)$. Hence the expression of the optimal land and capital quantities:

$$\begin{cases} s & = & \beta a(Y-tx)/p \\ k & = & \beta b(Y-tx) \end{cases}$$

Appendix C

Polycentric city

C.1 Reproducibility of the results

In order to confirm that the equilibrium reached by the agent-based model is unique, we perform the same simulation numerous times. In spite of the stochasticity of the dynamics of the model, each run converges to the same equilibrium, in a sense which is defined more precisely here.

The simulations are stopped only once the two conditions ensuring that the equilibrium is reached, described in section 5.2, are verified : the homogeneity of utility $\Delta U_{\max}$ is smaller than 10^{-6} (section 5.2.2) and the share of "empty" surface $S_{\text{empty}}/S_{\text{tot}}$ is smaller than 0.5% (section 5.2.3).

The results of these simulations are given in table C.1 for two models presented in this work: the first part corresponds to model 1, the reference monocentric model with two income groups. The second corresponds to model 5 with $d = 9$ and $m = 0.2$. The equilibrium values of the variables characterizing the models have only very small variations across different simulations. The maximal variation observed is of approximately 0.1% under the two previous conditions.

These values can be seen as the approximate magnitude of error bars of the results presented in this work.

Model 1	U_r	U_p	$U_r - U_p$	D^r_{mean}	D^p_{mean}	D_{tot}	R_{tot}	p_{mean}	S_{tot}
Variations (in %)	0.009	0.001	0.02	0.09	0.06	0.08	0.02	0.11	0.08
Model 5	U_p	U_0	U_1	D^0_{mean}	D^1_{mean}	D_{tot}	R_{tot}	p_{mean}	S_{tot}
Variations (in %)	0.008	0.01	0.02	0.09	0.05	0.06	0.02	0.09	0.08

Table C.1: Reproducibility of the results: maximal relative variations of the variables characterizing the models across 15 runs of the same simulation.

C.2 Parameters of the agent-based model

In this section, we give an example of stationary configuration[1] of the agent-based model when the parameters specific to the agent-based model, ϵ and T_p, are not chosen so as to minimize the inhomogeneity of utility $\Delta U_{\max}$ and the share of "empty" surface $S_{\text{empty}}/S_{\text{tot}}$. As a consequence, the system does not reach an equilibrium which corresponds to the analytical one.

We keep the values of parameters given in table 5.1, except T_p, which we take as $T_p = 3000$. The results of this simulation are shown on figure C.1. They should be compared with the results of figure 5.1. A much higher value of T_p is used in the present case, so that the price

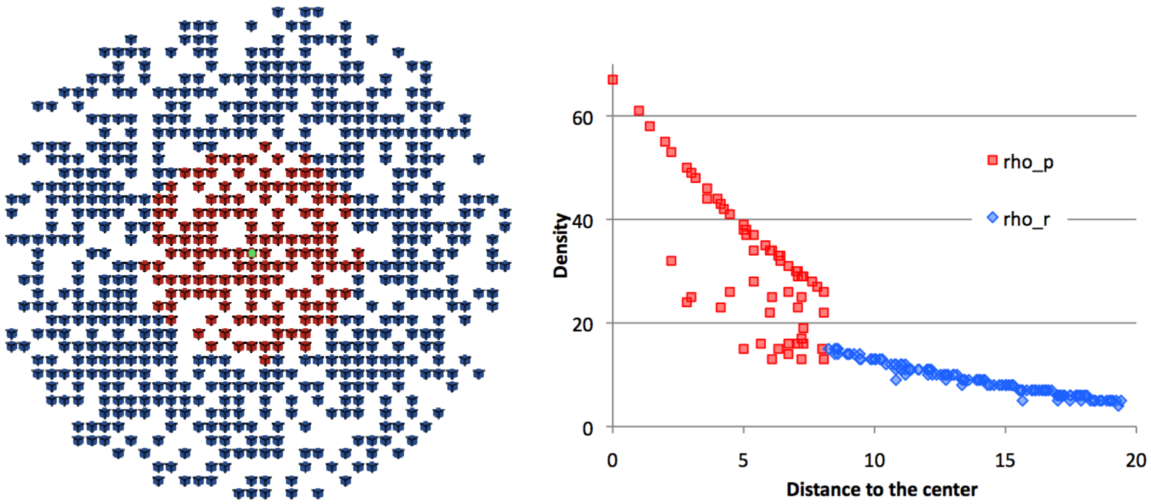


Figure C.1: Monocentric city with two income groups, with $T_p = 3000$. Left panel: shape of the city. Right panel: density as a function of the distance to the center.

of vacant cells decreases very slowly. It even decreases too slowly to manage to compensate

[1]This configuration corresponds actually to a state of the system where the evolution is very slow, so that the configuration seems stationary. We do not study this configuration more precisely here and present it as an illustration of a simulation not converging to an equilibrium corresponding to the analytical one.

price increases due to agents' bids, which prevents the system from reaching an equilibrium corresponding to the analytical one. Indeed, as can be seen on the left panel of figure C.1, as the price of cells decreases too slowly, some cells, even close to the CBD, are left vacant after their price has increased too much. The bid mechanism still manages to bring the system to a state with homogeneous utility, where $\Delta U_{\text{max}} < 10^{-6}$. However, a lot of space is not optimally used, which is indicated by the value of the share of "empty" surface $S_{\text{empty}}/S_{\text{tot}} \simeq 130\%$ (in the other simulations presented in this work, this variable is smaller than 0.5% – see section 5.2.3). Numerous cells where no agents live have a price which is higher than the agricultural rent, a situation which can not be observed if space is optimally used.

C.3 Simple polycentric city with one income group

The model presented in this section corresponds to model 2 with only one income group, or equivalently, to model 5 with $m = 0$. The shape of the city in this framework is given on the last line of figure 5.8. The corresponding outcomes on the different variables characterizing the model are presented on figure C.2. One obvious result can be seen on figure 5.8: when centers

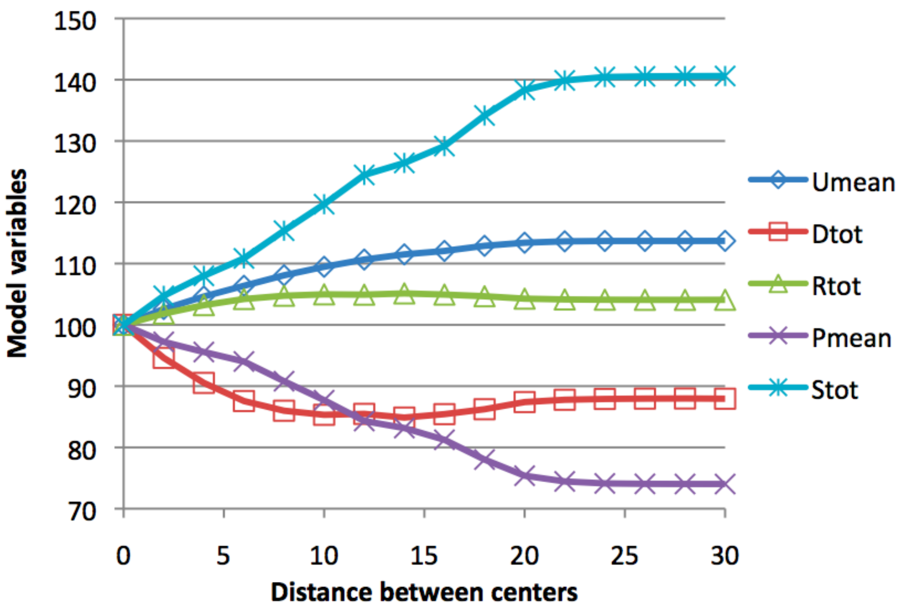


Figure C.2: Evolution of the variables characterizing the model as a function of the distance between centers.

are sufficiently far apart so that both halves of the city do not interact anymore (when the distance between centers is approximately 25 cells), the variables do not evolve when centers are pushed further away from one another.

More generally, these results can be related to those presented in section 5.3.1, on the same model with two income groups.

Résumé

Le travail de doctorat présenté dans ce document a été mené de 2008 à 2011 au Laboratoire d'Economie des Transports (LET) et à l'Institut Rhône-Alpin des Systèmes Complexes (IXXI). Cette thèse a été supervisée par Charles Raux, économiste, directeur du LET, et Pablo Jensen, physicien à l'Ecole Normale Supérieure de Lyon. Son financement concernait spécifiquement la modélisation urbaine utilisant les systèmes multi-agents, qui constitue donc une part importante de ce travail. Le projet pouvait être réalisé soit par un économiste intéressé par la modélisation sociale inspirée de la physique statistique, soit par un physicien voulant étudier les systèmes sociaux. L'auteur correspond au deuxième profil, ce qui influence fortement le point de vue avec lequel ces systèmes sociaux sont abordés.

Ce travail à la frontière entre sciences naturelles et sciences sociales est enrichi par la différence de point de vue entre les deux disciplines. Cela peut mener à un échange fructueux de méthodes et d'objets d'étude. De tels travaux sont de plus fortement encouragés par l'augmentation de la puissance de calcul des ordinateurs, et la relative abondance des données socio-économiques disponibles, qui sont deux tendances importantes des dernières années ou décennies. Cependant, la recherche interdisciplinaire comporte également des défis. En plus des difficultés de communication et des risques de malentendus, dus à des cultures scientifiques, des formalismes et même des vocabulaires différents, le financement, la publication et l'évaluation de la recherche restent très disciplinaires. Des infrastructures telles que l'IXXI fournissent des espaces d'échange et de travail interdisciplinaire, mais elles restent assez marginales. Et l'enseignement également est toujours très disciplinaire, ce qui peut poser des obstacles à la recherche interdisciplinaire.

Face à ces différentes contraintes, et dans le but d'allier intérêt et productivité scientifiques, ce travail est divisé en deux parties principales, l'une directement reliée à la physique statistique, et l'autre plus proche de l'économie urbaine.

Avant de présenter plus en détail son contenu, et après la description du contexte scientifique interdisciplinaire, quelques éléments de contexte socio-économique permettent de donner un cadre de recherche. Quels sont donc les enjeux de l'étude des systèmes urbains, liant transport et utilisation du sol ? Et quels sont les défis auxquels les villes font face ou vont devoir faire face au cours des prochaines années ou décennies ?

Le contexte est global, car beaucoup d'enjeux sont mondialisés, même si les situations particulières sont parfois complètement différentes. En effet, il y a maintenant un consensus scientifique sur le changement climatique (America's Climate Choices : Panel on Advancing the Science of Climate Change [2010], Oreskes [2004]). Ce changement se caractérise d'abord par un réchauffement, puisque la température moyenne sur la planète a augmenté d'environ 0,7°C au cours du 20ème siècle (Intergovernmental Panel on Climate Change [2007]), ce qui a déjà des conséquences significatives. Mais les effets concrets en des lieux donnés peuvent être très différents, car le climat mondial est un parfait exemple de système complexe. Il est également admis par le monde scientifique que les émissions humaines de gaz à effet de serre, et par exemple celles associées au transport et au logement, qui sont étudiés dans ce travail, contribuent de manière importante voire essentielle au réchauffement climatique (voir les mêmes sources).

La baisse rapide des réserves de combustibles fossiles comme le pétrole et le charbon (International Energy Agency [2010]), et conséquemment l'augmentation de leur prix, est un autre défi important auquel l'humanité fait face ou va devoir faire face dans un futur proche. Le pic de pétrole est spécialement inquiétant, car le pétrole est une matière première fondamentale dans de nombreux contextes, et dont l'utilisation ne se limite pas à celle de combustible. L'augmentation de son prix a déjà conduit à l'extraction de pétrole non conventionnel, malgré des rendements plus faibles et des impacts environnementaux plus graves. Les agrocarburants, qui ont également des rendements plus faibles, deviennent rentables à cause de l'augmentation du prix du pétrole. Leur production croissante est associée à différents problèmes, par exemple leur concurrence avec la production alimentaire : les agrocarburants sont une des causes probables de la crise alimentaire de 2008, qui a provoqué des émeutes de la faim dans différents pays.

En parallèle avec la décroissance des réserves de combustibles fossiles, les réserves de nombreux métaux diminuent rapidement (Gordon et al. [2006]), alors qu'ils sont des composants assez incontournables de nombreuses technologies modernes, entre autres de la production d'énergies renouvelables à travers les panneaux photovoltaïques par exemple (Bihouix and de Guillebon [2010]), mais également des voitures électriques pour revenir au transport, ou plus généralement de la plupart des composants électriques ou électroniques. Cela complique encore le problème.

Les avertissements face à de telles menaces remontent par exemple jusqu'à Meadows and Club of Rome [1972] ou Georgescu-Roegen [1986]. Ces références présentent d'ailleurs des tentatives intéressantes de promouvoir une science des systèmes complexes, en reliant physique statistique et économie pour la seconde. Mais quasiment aucune mesure politique n'a été prise pour le moment.

Un consensus politique plus important existe au sujet de la réduction des émissions de gaz à effet de serre. La France par exemple s'est engagée à réduire ses émissions d'un facteur 4 en 2050 par rapport à 1990. Mais 2050 paraît bien lointain d'un point de vue humain et politique. Au niveau international, de nombreux engagements semblables ont été pris par les états et communautés.

En France particulièrement, et dans certains autres pays, la réponse principale face au prix croissant de l'énergie associé à ces défis est le recours massif à l'électricité d'origine nucléaire. Cependant, l'utilisation de l'énergie nucléaire soulève de plus en plus de préoccupations suite à la catastrophe de Fukushima, qui a commencé en mars 2011 et a rappelé d'autres accidents nucléaires, qui ont pour l'instant relativement épargné la France. La question de la sûreté des installations est de nouveau posée, plusieurs pays européens ont d'ailleurs décidé récemment d'abandonner progressivement l'électricité d'origine nucléaire. Et les autres problèmes et contraintes de cette source d'énergie sont mis en question. Par exemple, le fait qu'il n'y a pas de solution viable en vue pour les déchets nucléaires, la difficulté de faire face aux pics de consommation d'énergie même quotidiens et le lien étroit entre nucléaire civil et militaire. La fusion nucléaire, moins dangereuse a priori, et qui produit des déchets moins toxiques, ne devrait pas être une alternative énergétique possible dans les prochaines années ou décennies, car elle se trouve toujours à un stade de recherche fondamentale, malgré des investissements massifs.

Dans ce contexte de prix croissant de l'énergie lié aux réserves décroissantes de différentes matières premières et de changement climatique lié aux émissions humaines de gaz à effet de serre, les villes sont une entité de première importance. En effet, elles consomment 60% de la production mondiale d'énergie et contribuent à 70% des émissions de gaz à effet de serre. Plus de la moitié de la population mondiale vit maintenant dans des zones urbaines, et cette part devrait croître pour atteindre environ 70% en 2050 (United Nations [2008]), alors que la population mondiale elle-même croît. De plus, les inégalités d'accès aux ressources sont extrêmement élevées à l'intérieur et entre les pays (Milanovic [2005]). Les villes en particulier sont des révélateurs de ces inégalités puisque les plus riches et les plus pauvres y coexistent. Cela crée des tensions sociales qui sont un danger supplémentaire.

Le secteur des transports produit 20% des émissions humaines de gaz à effet de serre en Europe, et 30% aux Etats-Unis (OECD [2008]). Le logement est également un secteur majeur d'émissions de gaz à effet de serre. Ainsi, l'étude de l'interaction entre logement et transport touche au cœur des problèmes évoqués précédemment. De plus, les villes sont des espaces particulièrement intéressants pour développer les modes de transport doux, grâce aux nombreuses opportunités présentes sur de courtes distances. L'utilisation du vélo par exemple peut être très pratique dans les espaces urbains. Malgré cela, des villes comme Copenhague ou Amsterdam, où le vélo est utilisé pour environ 30% des déplacements, sont des exceptions en Europe et dans le monde.

Bien entendu, les modèles urbains tels que ceux présentés dans cette thèse sont bien loin de donner des solutions aux défis que nous avons évoqués. La question principale est en fait une question de conscience et de volonté, sociale et politique. Cependant, ces défis constituent un cadre de pensée et une problématique pour le travail de modélisation urbaine. L'interrogation centrale concerne donc la ville "durable" : si les modèles parviennent à saisir même de manière grossière la réalité socio-économique des systèmes urbains, ou au moins aident à concevoir leur complexité, quelles formes urbaines consomment le moins d'énergie pour le transport et le logement ? Et quelles sont leurs conséquences quant au bien être des différents groupes sociaux et aux inégalités sociales ? De manière plus générale, quels sont les déterminants de la structure urbaine, et quels facteurs peuvent expliquer les différences que l'on peut rencontrer au niveau mondial ?

Ces questionnements composent le contexte socio-économique de ce travail. Le contexte scientifique a été évoqué précédemment à travers la collaboration interdisciplinaire entre physiciens et économistes. Un élément supplémentaire devrait être ajouté concernant le contexte scientifique en économie. Les simulations informatiques sont utilisées couramment en physique pour compléter le traitement analytique quand celui-ci est limité. Cependant, l'utilisation de simulations est marginale en économie, et semble être considérée comme très insatisfaisante.

Par conséquent, un objectif important de la partie de cette thèse orientée vers l'économie est de montrer que les simulations informatiques, et en particulier les modèles multi-agents dans ce travail, peuvent fournir un complément intéressant aux résolutions analytiques.

Cette thèse est donc divisée en deux parties principales. La première regroupe les trois premiers chapitres et présente un point de vue de physicien sur des modèles socio-économiques simples. La seconde est constituée des trois derniers chapitres, et étudie des simulations multi-agents étroitement reliées au modèle standard d'économie urbaine (Alonso [1964], Muth [1969], Mills [1967], voir aussi Fujita [1989]).

Résolution d'un modèle de Schelling

Le premier chapitre présente un formalisme de physique statistique qui permet de résoudre analytiquement un modèle de ségrégation proche du modèle de Schelling. Ce modèle (Schelling [1971]) étudie de manière très simple la ségrégation spatiale, et a donné lieu à une littérature abondante. Il peut être vu comme un des premiers modèles multi-agents, car il a été étudié par Thomas Schelling avant même que les simulations informatiques ne soient utilisées en sciences sociales. Il est d'ailleurs souvent utilisé comme exemple canonique de modèle multi-agents.

Sa résolution analytique n'avait pas encore pu être réalisée. L'ingrédient principal qui la rend possible dans ce travail est l'introduction d'une fonction d'état qui reflète les préférences individuelles et la dynamique des modèles socio-économiques. Cela permet d'utiliser des outils standard de physique statistique pour résoudre des modèles venant des sciences sociales, dans une approche qui peut être liée aux jeux potentiels (Monderer and Shapley [1996]). La décision de déménager est donnée par une règle logit, ce qui est un choix standard dans les modèles sociaux (Anderson et al. [1992]), et correspond à avoir une température non nulle dans le système. Dans les modèles sociaux, on peut interpréter cette règle comme une manière minimale d'intégrer tous les phénomènes qui ne sont pas pris en compte mais interviennent en réalité dans les décisions des agents.

La résolution du modèle est rendue possible également par une simplification apportée à la description du voisinage des agents par rapport à la formulation usuelle du modèle de Schelling. Cette simplification préserve cependant l'idée essentielle de ce modèle, à savoir un résultat contre-intuitif sur le lien entre le comportement "microscopique" d'agents sociaux ou économiques, et l'équilibre "macroscopique" résultant à un niveau global dans le système étudié. En effet, deux groupes d'agents sont définis, distincts par leur couleur. Une fonction d'utilité décrit leur bien-être, qui est maximal dans un environnement mixte. Mais à part ce désir important pour la mixité, les agents préfèrent être parmi des personnes de la même couleur qu'eux-mêmes plutôt que parmi des personnes de couleur différente. Cela suffit pour générer un équilibre ségrégué, si les choix de localisation suivent les goûts individuels.

Un penchant pour l'altruisme est ensuite introduit dans le modèle, conduisant les agents à ne pas seulement considérer leur bien-être personnel, mais également l'utilité collective, ce qui ressemble plus à la dynamique des systèmes physiques. Cela permet de casser la ségrégation spatiale, et d'observer une transition de phase entre dynamique individuelle et dynamique collective.

Utilité et potentiel chimique

Ce formalisme de physique statistique est développé dans le second chapitre pour inclure un espace inhomogène, ce qui n'est pas le cas dans le modèle de Schelling. La dynamique est encore donnée par une règle logit, dont le paramètre que nous nommons "température" peut être interprété en économie comme la largeur d'une distribution de goûts individuels. Dans ce modèle où l'espace n'est pas homogène, un potentiel chimique est introduit comme un multiplicateur de Lagrange qui rend compte de la conservation du nombre d'agents (modèle de "ville fermée"), qui est habituellement un nombre de particules en physique statistique.

Quand la température est non nulle, l'utilité ou bien-être des agents n'est pas homogène dans le système à l'équilibre, contrairement au résultat standard dans les modèles sociaux. Cependant, le potentiel chimique, qui a un lien étroit avec l'utilité, est lui homogène. A température nulle, il est confondu avec l'utilité, et son homogénéité est alors un résultat standard. Mais à température non nulle, il comporte en plus un terme entropique. Ce lien est surprenant, car l'utilité dans les modèles sociaux semble correspondre à l'énergie (libre) dans les systèmes physiques. En effet, ces fonctions déterminent essentiellement la dynamique des systèmes sociaux et physiques respectivement. Comme l'utilité est habituellement une quantité constante à l'équilibre dans les modèles sociaux, on aurait donc pu s'attendre à ce qu'elle corresponde à la température dans les modèles physiques, la variable conjuguée de l'énergie.

Ce formalisme est illustré sur une version simplifiée du modèle standard d'économie urbaine (ou modèle AMM). L'hypothèse simplificatrice principale consiste à supposer une relation directe entre prix et densité. Ce modèle simple illustre les résultats concernant l'introduction d'un choix de localisation donné par une règle logit dans le modèle AMM (De Palma and Papageorgiou [1988], Anas [1990]).

La résolution analytique est permise également par le fait que la dynamique respecte un bilan détaillé (voir Van Kampen [1992], Evans and Hanney [2005]). Des conditions qui le garantissent sont dérivées dans le chapitre. Un résultat surprenant est donné par des simulations multi-agents en dehors du domaine de validité de ces conditions et donc de notre formalisme : l'homogénéité du potentiel chimique à l'équilibre est en effet validée par ces simulations.

Un modèle probabiliste de marché du logement urbain

Le troisième chapitre explore l'hypothèse du second, qui suppose une relation directe entre densité d'habitants et prix à une localisation donnée, dans le modèle simple étudié. Nous construisons un modèle minimal de marché du logement urbain qui permet, grâce à une hypothèse ergodique, de relier les variations temporelles du prix d'un appartement représentatif aux variations spatiales de prix dans une ville virtuelle. Le prix varie en fonction de l'occupation de cet appartement. Il augmente quand celui-ci est occupé (c'est-à-dire, fortement demandé), et diminue quand l'appartement est vide. Les agents, qui emménagent dans cet appartement ou en déménagent, déterminent donc l'évolution du prix. Cela définit une représentation schématique de la dynamique urbaine des prix permettant de la lier à la densité d'habitants.

Cependant, ce travail en cours pose une question concernant la relation entre deux définitions de la densité dans ce modèle. L'une est liée aux variations temporelles et correspond à l'occupation moyenne dans le temps de l'appartement, et l'autre est donnée par une contrainte "spatiale" dans la ville. En effet, l'hypothèse ergodique permet de lier l'occupation moyenne dans le temps d'un appartement donné aux variations spatiales de prix dans une ville virtuelle. Mais les simulations montrent que ces deux quantités ne coïncident pas, ce qui doit être étudié plus précisément.

Toutefois le modèle atteint un état d'équilibre indépendant des conditions initiales. L'étude de l'évolution de cet équilibre en fonction des différents paramètres du modèle est une perspective de travail. L'influence sur la distribution des prix de la densité d'agents, qui est le paramètre le plus "physique" du modèle, est étudiée dans ce chapitre et démontre un comportement qui va dans le sens de l'hypothèse introduite dans le chapitre précédent. Une perspective de recherche intéressante est liée à la simplicité de la dynamique utilisée : cela devrait permettre une résolution analytique du modèle dans certains cas limites.

Un modèle multi-agents de l'économie urbaine

La seconde partie de ce travail est directement reliée à la littérature en économie urbaine. Ce travail de recherche est inspiré par exemple par Caruso et al. [2007] ou Brueckner et al. [1999]. Le chapitre 4 présente un modèle multi-agents qui utilise des règles d'évolution dynamique simples et conformes à l'intuition économique. A partir d'un état initial aléatoire, cette dynamique conduit un système urbain vers un état d'équilibre qui correspond à une version discrète du modèle d'Alonso (Alonso [1964]). Ce modèle standard de l'économie urbaine présente un équilibre résidentiel de localisation d'agents dans une zone urbaine. Cet équilibre analytique est statique, obtenu comme résultat d'une optimisation contrainte. La dynamique du modèle multi-agents est directement inspirée par la compétition pour l'utilisation du sol dans le modèle analytique, où le plus offrant l'emporte.

Une comparaison entre les résultats du modèle analytique et des simulations multi-agents permet d'observer un bon accord entre eux. Partant de là, le modèle multi-agents est complexifié pour étudier des phénomènes difficiles à traiter de manière analytique. En particulier, ce chapitre, suivant en cela Brueckner et al. [1999] ou Wu and Plantinga [2003], considère l'introduction d'une aménité positive dans un système urbain. Cette aménité correspond à un élément attractif de l'espace urbain, que ce soit une aménité historique (de beaux bâtiments anciens par exemple), environnementale (un parc) ou plus moderne (cinémas, restaurants, piscine). Les simulations multi-agents nous permettent d'étudier l'influence de cette aménité sur la ville, grâce à différentes variables utilisées comme indicateurs. L'utilité des agents traduit leur bien-être économique. Avec deux catégories de revenu, l'évolution de l'écart de bien-être entre les deux catégories peut être étudiée quand certains paramètres du modèle changent. Les distances domicile-travail des agents sont également des indicateurs importants.

Différentes structures sociales urbaines peuvent être observées dans les villes nord-américaines et européennes. Les villes nord-américaines tendent à avoir des ménages riches en périphérie et des ménages plus pauvres près du centre, alors que les villes européennes ont plutôt une configuration inverse, comme le décrivent Brueckner et al. [1999] ou Gofette-Nagot et al. [2000]. Nous présentons des tentatives de faire émerger ces structures sociales de l'interaction entre agents. Avec deux catégories de revenu, le modèle standard de l'économie urbaine donne en général une configuration "nord-américaine", où les agents riches se localisent en périphérie.

La structure "européenne" est plus difficile à reproduire dans ce modèle.

L'introduction d'un coût temporel de transport seul ne nous permet pas de retrouver la configuration "européenne", si des valeurs réalistes de ce coût temporel pour les agents à haut et bas revenu sont utilisées. L'introduction d'une aménité centrale, avec la fonction utilité log-linéaire que nous utilisons, ne conduit pas non plus à une structure sociale "européenne". Par contre, cela a deux effets principaux sur la ville. D'une part, les agents sont concentrés autour de l'aménité, dont la proximité procure un niveau de bien-être supérieur. Cela tend à densifier la ville et à réduire sa taille ainsi que les distances domicile-travail, si l'aménité est proche du centre. D'autre part, un second effet apparaît quand l'aménité est écartée du centre d'emploi. Dans ce cas, la ville est déformée par la présence de l'aménité, et les distances domicile-travail peuvent augmenter même si la taille de la ville diminue, si on compare les résultats à une situation de référence sans aménité. Si une aménité attractive est localisée suffisamment loin du centre, un phénomène de développement "à saute-mouton" peut être observé dans le modèle, avec une ville secondaire qui apparaît en marge de la ville principale, à proximité de l'aménité.

Pour revenir à la question de la structure urbaine "européenne", nous définissons une préférence plus forte des agents riches pour cette aménité, ce qui permet d'obtenir une structure sociale européenne. Cependant, cette préférence plus forte paraît artificielle et peu satisfaisante du point de vue de la modélisation. Une confirmation empirique serait alors souhaitable. Par ailleurs, les simulations montrent que l'équilibre obtenu est très sensible à une différence de localisation entre le centre d'emploi et le centre d'aménité, pour une ville bidimensionnelle.

Explorer la ville polycentrique

Le chapitre 5 utilise une version du modèle multi-agents introduit dans le chapitre 4, où la dynamique est un peu simplifiée mais permet toujours de converger vers l'équilibre du modèle AMM. Le but de ce chapitre est d'étudier la question de la ville polycentrique. Cela fait l'objet d'un pan de la littérature en économie urbaine, qui cherche par exemple si une ville avec plusieurs centres d'emplois est plus économe en énergie qu'une ville monocentrique. De plus, le formalisme polycentrique paraît plus proche de la réalité empirique des systèmes urbains.

Ce chapitre présente une discussion détaillée concernant l'existence et l'unicité de l'équilibre dans les différents modèles polycentriques étudiés. L'existence de l'équilibre est prouvée par Fujita and Smith [1987], et l'unicité est démontrée par Fujita [1985] dans un contexte monocentrique. Une correspondance entre modèles monocentriques et polycentriques nous permet d'étendre ce dernier résultat à certains des modèles polycentriques que nous étudions. Des arguments plus qualitatifs sont utilisés dans les cas restants.

Une comparaison entre les résultats du modèle analytique monocentrique à deux catégories de revenu et ceux du modèle multi-agents est réalisée. Pour cela, le modèle analytique est résolu numériquement grâce à une procédure particulière décrite par Fujita [1989]. L'accord entre résultats analytiques et simulations multi-agents est très bon. Ce chapitre détaille les conditions qui garantissent que le modèle multi-agents atteint une version discrète du modèle analytique.

Puis l'étude de la ville polycentrique elle-même est présentée à travers différents modèles. Le plus simple consiste à définir deux centres d'emploi et à permettre aux agents de travailler au centre le plus proche de leur domicile. Comme dans le chapitre 4, l'évolution du système urbain est étudiée grâce à différentes variables qui caractérisent les bilans économique, environnemental et social de ces modèles. Dans ce modèle simple, la ville polycentrique est désirable d'un point de vue économique, fondé sur l'utilité des agents. Mais le bilan environnemental n'est pas aussi clair. D'une part, les distances domicile-travail tendent à diminuer si les centres ne sont pas trop éloignés l'un de l'autre, ce qui a des conséquences positives à travers la réduction de la pollution liée au transport. D'autre part, les surfaces de logement tendent à augmenter, ce qui étale la ville et augmente la demande énergétique en termes de chauffage et de climatisation. Un modèle calibré est donc nécessaire pour tirer des conclusions plus tranchées sur ce point. Un premier essai de calage sur l'aire urbaine de Lyon d'un modèle multi-agents intégrant la construction du logement (voir Muth [1969]) est présenté dans l'appendice B.

D'autres modèles polycentriques sont étudiés, dans lesquels les agents ont des contraintes concernant la localisation de leur emploi. Les résultats de ces modèles confirment jusqu'à un certain point ceux du premier modèle étudié. En particulier, un modèle plus réaliste étudie des agents qui représentent des ménages bi-actifs, et non plus des travailleurs célibataires. Les deux travailleurs du ménage ont soit des emplois à la même localisation, soit dans des localisations différentes. La ville est plus cohérente, et on retrouve les résultats de bilan économique positif et de bilan environnemental mitigé, si les centres sont proches l'un de l'autre. Si les centres sont trop éloignés, la ville polycentrique est indésirable dans ce modèle.

Ville "européenne" et ville "nord-américaine"

Dans le chapitre 6, le modèle présenté dans le chapitre 5 est appliqué à différentes fonctions d'utilité, pour reprendre la question des villes "européennes" et "nord-américaines" étudiée dans le chapitre 4. En effet, les chapitres 4 et 5 utilisent une fonction d'utilité de Cobb-Douglas, qui est la plus utilisée en économie urbaine à cause de sa praticité pour la résolution analytique. Mais bien entendu, la forme de la fonction d'utilité est déterminante pour le résultat du modèle, et d'autres expressions sont utilisés dans ce chapitre 5.

La première partie du chapitre considère une fonction Cobb-Douglas dont les paramètres varient avec le revenu, de manière à prendre en compte le fait que la part du revenu utilisée pour le logement décroît quand le revenu augmente, d'après les études empiriques. Deux catégories de revenu sont ainsi définies, et le modèle prend en compte un coût temporel de transport, plus élevé pour les agents de haut revenu. Les agents riches ont donc un coût global de transport plus élevé, en supposant de plus que le coût monétaire de transport est constant.

Le résultat d'un tel modèle, même en considérant simplement un coût de transport proportionnel à la distance au centre, est plus complexe que la structure sociale urbaine obtenue habituellement grâce au modèle AMM. A une certaine distance caractéristique, dépendant des paramètres du modèle, le comportement résidentiel des agents en fonction de leur revenu change. En fonction des valeurs comparées de cette distance et du rayon de la ville, différents cas de figure apparaissent. Les structures "européenne" et "nord-américaine" standard peuvent émerger à l'équilibre, mais également une structure mixte correspondant à une ville "européenne" entourée d'un espace périurbain riche. Cela peut également être vu comme la structure sociale des villes américaines anciennes comme New York, Chicago ou Philadelphie (voir Glaeser et al. [2008]).

L'hypothèse de Brueckner et al. [1999] concernant une aménité centrale dans les villes européennes est ensuite revisitée. En effet, une fonction utilité qui permet aux agents riches de valoriser plus fortement l'aménité que les agents pauvres peut permettre de reproduire une structure "européenne". Notre modèle multi-agents illustre ce résultat, et montre de plus que des structures plus complexes peuvent apparaître, ce qui crée un parallèle avec les résultats de la première partie du chapitre. Cette structure plus complexe correspondant à une ville européenne avec un espace périurbain riche semble avoir été négligée par Brueckner et al. [1999]. Nous étudions également le comportement du système lorsque l'aménité n'est pas confondue avec le centre d'emploi, et là encore, la structure sociale urbaine dépend fortement de la distance entre aménité et centre d'emploi.

Enfin, une règle logit est utilisée dans le choix de localisation des agents, ce qui permet de réduire l'importance du choix de la fonction utilité. Ce choix de déménagement probabiliste introduit de la stochasticité et brouille les structures déterminées par la fonction d'utilité. Cette étude permet également de lier les deux parties de cette thèse, car ce modèle vérifie le résultat principal du chapitre 2. Le potentiel chimique qui y est défini est homogène à l'équilibre dans le modèle multi-agents. Ce résultat est surprenant puisque le modèle paraît loin du domaine de validité du formalisme de physique statistique présenté dans le chapitre 2. Une perspective de travail consiste à relier ce résultat à la littérature économique (De Palma and Papageorgiou [1988], Anas [1990].)

Bibliography

J. Accardo and F. Bugeja. Le poids des dépenses de logement depuis vingt ans. In *Cinquante ans de consommation en France*, pages 33–47. Insee, Paris, 2009.

W. Alonso. *Location and land use.* Harvard University Press, 1964.

America's Climate Choices: Panel on Advancing the Science of Climate Change. *Understanding and Responding to Climate Change.* The National Academies Press, 2010.

A. Anas. Taste heterogeneity and urban spatial structure. *Journal of Urban Economics*, 28: 318–335, 1990.

A. Anas, R. Arnott, and K. A. Small. Urban spatial structure. *Journal of Economic Literature*, 36:1426, 1998.

S. P. Anderson, A. De Palma, and J. F. Thisse. *Discrete Choice Theory of Product Differentiation.* MIT Press, 1992.

Y. Asami, M. Fujita, and T. E. Smith. On the foundations of land use theory : Discrete versus continuous populations. *Regional Science and Urban Economics*, 20(4):473 – 508, 1991.

R. Balescu. *Equilibrium and Nonequilibrium Statistical Mechanics.* Wiley, New York, 1975.

I. Bena. Dichotomous markov noise: Exact results for out-of-equilibrium systems. a review. *Int. J. Mod. Phys. B*, 20:2825, 2006.

M. Berliant and T. Sabarwal. When worlds collide: Different comparative static predictions of continuous and discrete agent models with land. *Regional Science and Urban Economics*, 38(5):438–444, September 2008.

E. Bertin, O. Dauchot, and M. Droz. Temperature in nonequilibrium systems with conserved energy. *Phys. Rev. Lett.*, 93(cond-mat/0406720):230601, 2004.

E. Bertin, K. Martens, O. Dauchot, and M. Droz. Intensive thermodynamic parameters in nonequilibrium systems. *Phys. Rev. E*, 75:031120, 2007.

P. Bihouix and B. de Guillebon. *Quel futur pour les métaux ? Raréfaction des métaux: un nouveau défi pour la société*. EDP Sciences, 2010.

J.K. Brueckner, J.F. Thisse, and Y. Zenou. Why is central Paris rich and downtown Detroit poor? An amenity-based theory. *European Economic Review*, 43:91–107, 1999.

G. Caruso, D. Peeters, J. Cavailhes, and M. Rounsevell. Spatial configurations in a peri-urban city. A cellular automata-based microeconomic model. *Regional Science and Urban Economics*, 37(5):542–567, 2007.

J. W. Chung. *Utility and production functions: theory and applications*. Blackwell Publishers, 1994.

R. Cotterill. *The Cambridge Guide to the Material World*. Cambridge University Press, 2008.

N. Coulombel and F. Leurent. Logement et transport en Île-de-France : une approche par les budgets. Working paper, 2011.

A. De Palma and Y. Y. Papageorgiou. Heterogeneity in tastes and urban structure. *Regional Science and Urban Economics*, 18(1):37 – 56, 1988.

A. De Palma, N. Picard, and P. Waddell. Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region. *Journal of Urban Economics*, 62(2):204 – 230, 2007.

P. R. Evans and T. Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *Journal of Physics A: Mathematical and General*, 38(19):R195, 2005.

M. Fujita. Existence and uniqueness of equilibrium and optimal land use: Boundary rent curve approach. *Regional Science and Urban Economics*, 15(2):295 – 324, 1985.

M. Fujita. *Urban Economic Theory*. Cambridge University Press, 1989.

M. Fujita and H. Ogawa. Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics*, 12:161–196, 1982.

M. Fujita and T. E. Smith. Existence of continuous residential land-use equilibria. *Regional Science and Urban Economics*, 17(4):549 – 594, 1987.

M. Fujita and J.F. Thisse. *Economie des villes et de la localisation*. De Boeck, 2003.

N. Georgescu-Roegen. The entropy law and the economic process in retrospect. *Eastern Economic Journal*, 12(1):3–25, 1986.

E. L. Glaeser, M. E. Kahn, and J. Rappaport. Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, 63(1):1 – 24, 2008.

F. Gofette-Nagot, I. Thomas, and Y. Zenou. *Structure urbaine et revenu des ménages*, chapter 10. Economie Géographique : approches théoriques et empiriques. Economica, Paris, 2000.

F. Goffette-Nagot, P. Jensen, and S. Grauwin. Dynamic models of residential segregation: Brief review, analytical resolution and study of the introduction of coordination, 2009. URL `http://halshs.archives-ouvertes.fr/halshs-00404400/en/`. Working Paper GATE 2009-14.

D. Goodstein. *States of Matter*. Dover, 1985.

R. B. Gordon, M. Bertram, and T. E. Graedel. Metal stocks and sustainability. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5):1209–1214, 2006.

S. Grauwin, E. Bertin, R. Lemoy, and P. Jensen. Competition between collective and individual dynamics. *Proc Natl Acad Sci USA*, 2009.

P. G. Hartwick and J. M. Hartwick. Efficient resource allocation in a multinucleated city with intermediate goods. *The Quarterly Journal of Economics*, 88(2):340–52, 1974.

D. Hotchkiss and M. J. White. A simulation model of a decentralized metropolitan area with two-worker, traditional, and female-headed households. *Journal of Urban Economics*, 34 (2):159 – 185, 1993.

Working Group I Intergovernmental Panel on Climate Change. *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2007.

International Energy Agency. *World Energy Outlook 2010*. 2010.

J. E. Kohlhase. Labor supply and housing demand for one- and two-earner households. *The Review of Economics and Statistics*, 68(1):48–57, 1986.

B. Latour. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press, Oxford, 2005.

R. Lemoy, C. Raux, and P. Jensen. An agent-based model of residential patterns and social structure in urban areas. *Cybergeo: European Journal of Geography*, 2010a. URL `http://cybergeo.revues.org/23381`.

R. Lemoy, C. Raux, and P. Jensen. Un modèle multi-agents de structuration sociale de la ville. In J.-P. Antoni, editor, *Modéliser la ville. Forme urbaine et politiques de transport*, Méthodes et Approches. Economica, Paris, 2010b.

R. Lemoy, E Bertin, and P. Jensen. Socio-economic utility and chemical potential. *EPL (Europhysics Letters)*, 93(3):38002, 2011a.

R. Lemoy, C. Raux, and P. Jensen. Exploring the polycentric city with an agent-based model. 2011b. URL `http://hal.archives-ouvertes.fr/hal-00602087/en/`.

S. F. LeRoy and J. Sonstelie. Paradise lost and regained: Transportation innovation, income, and residential location. *Journal of Urban Economics*, 13(1):67–89, January 1983.

J. F. Madden and M. J. White. Spatial implications of increases in the female labor force: A theoretical and empirical synthesis. *Land Economics*, 56(4):432–446, 1980.

A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic theory*. Oxford University Press New York, 1995.

J. Mc Breen, F. Goffette-Nagot, and P. Jensen. Information and search on the housing market: An agent-based model. In M. Li Calzi, L. Milone, and P. Pellizzari, editors, *Progress in Artificial Economics, Computational and Agent-Based Models*, volume 645 of *Lecture Notes in Economics and Mathematical Systems*, pages 153–164. 2010.

D. H. Meadows and Club of Rome. *The Limits to growth; a report for the Club of Rome's project on the predicament of mankind*. Universe Books New York, 1972.

B. Milanovic. *Worlds Apart: Measuring International and Global Inequality*. Princeton University Press, 2005.

E. S. Mills. An aggregative model of resource allocation in a metropolitan area. *American Economic Review*, 57:197–210, 1967.

D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124 – 143, 1996.

R. F. Muth. *Cities and Housing*. University of Chicago Press, 1969.

OECD. *Highlights of the International Transport Forum 2008: Transport and Energy. The Challenge of Climate Change*. OECD Publishing, 2008.

N. Oreskes. The scientific consensus on climate change. *Science*, 306(5702):1686, 2004.

A. De Palma and C. Fontan. Choix modal et valeur du temps en Ile-de-France. *Recherche Transports Sécurité*, (71):24–48, 2001.

R. Pancs and N. J. Vriend. Schelling's spatial proximity model of segregation revisited. *Journal of Public Economics*, 91(1-2):1–24, February 2007.

A. Polacchini. Les dépenses des ménages franciliens pour le logement et les transports. *Recherche Transports Sécurité*, 63, 1999.

T. Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(1):143–186, 1971.

K. A. Small. Urban transportation economics. In *Regional and Urban Economics II*. Harwood Academic Publishers, 1992.

A. Smith. *An inquiry into the nature and causes of the wealth of nations*. W. Strahan and T. Cadell, London, 1784.

U. C. Tauber, V. K. Akkineni, and J. E. Santos. The effects of violating detailed balance on critical dynamics. *Phys. Rev. Lett.*, 88(cond-mat/0109433):045702. 4 p, Sep 2001.

H. Touchette. The large deviation approach to statistical mechanics. *Physics Reports-review Section of Physics Letters*, 478:1–69, 2009. doi: 10.1016/j.physrep.2009.05.002.

United Nations. *World Population Prospects: The 2007 Revision*. United Nations, Department of Economic and Social Affairs, Population Division, New York, 2008.

N. G. Van Kampen. *Stochastic processes in physics and chemistry*. North Holland, 1992.

M. Wardman. A review of British evidence on time and service quality valuations. *Transportation Research Part E: Logistics and Transportation Review*, 37(2-3):107 – 128, 2001a.

M. Wardman. Inter-temporal variations in the value of time. *ITS Leeds Working Paper*, 566, 2001b.

W. C. Wheaton. Income and urban residence: An analysis of consumer demand for location. *American Economic Review*, 67(4):620–31, September 1977.

M. J. White. Firm suburbanization and urban subcenters. *Journal of Urban Economics*, 3 (4):323 – 343, 1976.

M. J. White. Location choice and commuting behavior in cities with decentralized employment. *Journal of Urban Economics*, 24(2):129 – 152, 1988.

K. F. Wieand. An extension of the monocentric urban spatial equilibrium model to a multicenter setting: The case of the two-center city. *Journal of Urban Economics*, 21(3):259 – 271, 1987.

J.J. Wu and A.J. Plantinga. The influence of public open space on urban spatial structure. *Journal of Environmental Economics and Management*, 46:288–309, 2003.

J. Yinger. City and suburb: Urban models with more than one employment center. *Journal of Urban Economics*, 31(2):181 – 205, 1992.

List of Figures

List of Tables

Abstract

Social and natural sciences share an interest for collective phenomena, which constitute an important part of the domain of complex systems. This thesis focuses on the study of urban systems, using analytical tools inspired by statistical physics, and also simulations, in particular agent-based models.

A first analytical resolution of a Schelling spatial segregation model is presented, using a statistical physics framework linking individual and collective dynamics. This framework shows that utility or welfare in socio-economic models corresponds to a chemical potential in physics, a correspondence which is applied to a urban housing model. The housing market is further studied with a parsimonious price formation model.

The implementation of an agent-based model, which reproduces the results of the standard urban economics (AMM) model, provides a second point of view on urban systems and the interaction between transport and land use. The simulations give results also when analytical resolution is lacking. The model is used to study the economic, environmental and social outcomes of having an amenity in the urban area and of polycentric cities. With two income groups, this work provides insights on the different urban social structures in North American and European cities for instance.

Keywords: statistical physics, urban models, spatial segregation, transport, land use, agent-based.

Résumé

Les sciences "dures" et les sciences humaines ont un intérêt commun pour les phénomènes collectifs, qui sont également un sujet de recherche important dans le domaine des systèmes complexes. Cette thèse se focalise sur l'étude des systèmes urbains, en utilisant des outils analytiques venant de la physique statistique, et des simulations, en particulier la modélisation multi-agents.

Une première résolution analytique d'un modèle de ségrégation spatiale de Schelling est obtenue à l'aide d'un formalisme de physique statistique qui relie la dynamique individuelle et collective. Ce formalisme montre que l'utilité ou le bien être des modèles socio-économiques correspond à un potentiel chimique en physique, ce qu'illustre un modèle de logement urbain. Le marché du logement est étudié plus en détail grâce à un modèle parcimonieux de formation du prix.

La mise en oeuvre d'un modèle multi-agents, qui reproduit les résultats du modèle standard de l'économie urbaine (AMM), donne un deuxième point de vue sur les systèmes urbains et l'interaction entre transport et localisation résidentielle. Les simulations fournissent des résultats là où la résolution analytique fait défaut. Ce modèle est utilisé pour étudier les impacts économique, environnemental et social de l'introduction d'une aménité dans l'espace urbain et de la ville polycentrique. Avec deux catégories de revenu, ce travail fournit des hypothèses quant aux différentes structures sociales urbaines dans les villes nord-américaines et européennes par exemple.

Mots-clés : physique statistique, modèles urbains, ségrégation spatiale, transport, localisation, multi-agents.