Modeling urban housing market dynamics: can the socio-spatial segregation preserve some social diversity?

Laetitia Gauvin[a,1], Annick Vignes[b], Jean-Pierre Nadal[a,c]

[a]*Laboratoire de Physique Statistique (LPS, UMR 8550 CNRS-ENS-UPMC-Univ. Paris Diderot), Ecole Normale Supérieure, Paris, France*
[b]*Equipe de Recherche sur les Marchés, l'Emploi et la Simulation (ERMES, EA 4441 CNRS-Paris II), University Paris II Panthéon-Assas, Paris, France*
[c]*Centre d'Analyse et de Mathématique Sociales (CAMS, UMR 8557 CNRS-EHESS), Ecole des Hautes Etudes en Sciences Sociales, Paris, France*

Abstract

Addressing issues of social diversity, we introduce a model of housing transactions between agents who are heterogeneous in their willingness to pay. A key assumption is that agents' preferences for a location depend on both an intrinsic attractiveness and on the social characteristics of the neighborhood. The stationary space distribution of income is analytically and numerically characterized. The main results are that socio-spatial segregation occurs if – and only if – the social influence is strong enough, but even so, some social diversity is preserved at most locations. Comparison with data on the Paris housing market shows that the results reproduce general trends of price distribution and spatial income segregation.

Keywords: housing market model, income segregation, social diversity, agent-based model
J.E.L. codes: R31; C61; C62; C63

Don't buy the house, buy the neighborhood (Russian proverb)

1. Introduction

People's choice of residential location and the way they are distributed across cities matter, from both social and economic points of view. This

---

[1]Corresponding author, laetitia.gauvin@lps.ens.fr

article seeks to explain, from the dynamics of price formation in an urban housing market, how individuals who are heterogeneous in their willingness to pay are distributed over a city. It shows that, under certain conditions on social interactions, housing price formation can entail income segregation, even if a space of social diversity remains.

A large literature is concerned with evaluating the extent and impact of housing price discrimination in education, housing and the labor market. While Brueckner et al. (1999) explain that the relative location of different income social groups depends on the spatial pattern of amenities in a city, Gobillon et al. (2007) highlight the adverse labor-market outcomes due to spatial mismatch, with low-skilled inhabitants of the inner-city suffering a greater distance to jobs and consequently a higher level of unemployment. Understanding the formation of prices in the real estate market and the way that prices are distributed over space is clearly an important issue. Following the path opened by Rosen (1974), most studies have focused on explaining prices through hedonic estimations, showing how the price per square meter can be influenced both by variables intrinsic to the apartment or house and by extrinsic variables concerning the surrounding area and its amenities. The role of these extrinsic variables has been particularly explored. Baltagi & Bresson (2010) underline how much the location influences the price of a dwelling. Ioannides & Zabel (2003), Figlio & Lucas (2004), Bono et al. (2007) or Seo & Simons (2009) emphasize the importance of the quality and density of the neighborhood, the reputation of nearby schools and the level of security. Following Tiebout (1956), these authors point out that the decision of where to live is based on families' preferences for the quality of public services and amenities, particularly education. Thus, prices on the real estate market vary with the quality of a bundle of public services which are capitalized into housing prices.

One question underlined by the studies cited above is that of people's preferences. Do they prefer to live with people who are richer than they are, or poorer? In other words, what determines the social component of the attractiveness of a location? The literature on well-being tends to argue that people feel better when those around them are poorer (for a detailed survey see Luttmer (2005)). Clark & Oswald (2002) show empirically that unemployed people are less unhappy when they live with other unemployed people. Goyal & Ghiglino (2010) explore the role of shifting social interactions: they use examples to illustrate how poorer individuals lose while richer ones gain as we move from an economically segregated society towards an

integrated society. In line with the literature on hedonic prices which highlights the importance of the environment, we explore here the consequences of the alternative hypothesis that individuals prefer to live with people who are at least as rich as themselves. Indeed, some preliminary studies suggest that the interplay between social preferences and the preference for a high quality of local amenities has more striking effects when the social preference is to live with richer neighbors.

When the prices depend not only on the intrinsic characteristics of the goods but also on the level of surrounding amenities, as well as on some social preferences, the space is differentiated and a social mismatch may result. Different measures of segregation have been proposed. Alesina & Zhuravskayay (2011) measure segregation of different ethnic, religious and linguistic groups within the same country from quite an exhaustive data set covering several different countries. Cutler et al. (1999) show the influence of legal barriers on social segregation, while Jenks & Meyer (1990) or Cutler et al. (2008) evaluate the effect of segregation on the socioeconomic performance of minorities. Echenique & Fryer (2007) develop a spectral index of segregation related to the characteristics of social interactions. This index is defined at the individual level and is higher when the considered individual interacts with segregated individuals. Ballester & Vorsatz (2010) propose a random-walk-based segregation measure. For the present work, interesting results are obtained by considering simple measures of multigroup segregation, which compare, for a population in a given local neighborhood, the observed distribution of a given feature (here income) with the uniform distribution. A basic segregation index in the line of Reardon & Firebaugh (2002) and Alesina & Zhuravskayay (2011) is proposed and compared to the information-theoretic measure of Theil (1967), which was introduced into economics for measuring income heterogeneity.

The pioneering modeling work of Schelling (1971) describes residential segregation as emerging from social preferences alone. Since then, extensions have been proposed by several authors in order to integrate a housing market. In Bernard & Willer (2007), the price of a house depends on an intrinsic component of the location, randomly allocated, and on the composition of the neighborhood. In Fossett & Senft (2003), the price only depends on an intrinsic component of the location uniformly distributed over the city. However, the level of quality of locations does not directly impact the choice made by individuals. In Zhang (2004), the price varies according to the density of occupation of the neighborhood, but does not take into account

any measure of house quality. In Bernard & Willer (2007), the choice of a location depends on the mean status of the neighborhood, the status being a wealth-related quantity randomly allocated to the individuals. The present work goes further, dealing with the individual attractiveness of a location and its influence on prices: in our model, the individuals choose a location according to its attractiveness, which is a dynamic quantity depending on both the intrinsic characteristic of the location and the (time-dependent) social composition (measured by the levels of income) of the neighborhood. The prices then depend on the attractiveness through the market dynamics. The framework introduced here can easily be adapted to make the attractiveness reflect different characteristics of the locations and different types of social preferences. It also has the advantage of allowing for detailed mathematical analysis.

The present article focuses on spatial income segregation, leaving aside all other features (such as ethnic characteristics) that may also cause segregation. The model proposed here takes some inspiration from Short et al. (2008) and Berestycki & Nadal (2010), who model the evolution of the spatial distribution of crime in a city, attributing to each location an attractiveness for illegal activity. The originality here is that each agent attributes to each location a specific level of attractiveness. This attractiveness results from a combination of an intrinsic or objective part, and a subjective part. The endogenous (subjective) attractiveness results from the individuals' intrinsic social preferences (for neighbours with similar or higher incomes). The main assumptions are: (*i*) people make decisions according to both their willingness to pay (WTP) and their individual evaluation of the level of attractiveness of the different locations; (*ii*) buyers, who are heterogeneous in their WTP, base their search for housing on the level of attractiveness of the location of the dwelling; (*iii*) agents are both buyers and sellers; (*iv*) the intrinsic attractiveness is spatially heterogeneous.

The results are demonstrated through mathematical analysis and then empirically confirmed in the simulations and in the empirical analysis using data on the Paris housing market. The analysis of the stationary regime reached by the market dynamics is first performed for an arbitrary, spatially heterogeneous, intrinsic attractiveness. For illustrative purposes and to compare with the empirical data, the analysis is then conducted in more detail for specific cases. First, we consider the case of a monocentric city defined by an intrinsic attractiveness which decreases with the distance from the geographical center. This is a simple case motivated by the classical Von

4

Thünen model and Alonso's study of land use (see Alonso (1964)), explaining how, generally, prices are higher in the center of a city or a county. The fact that prices diffuse from the center to the periphery (with higher prices in the center) has more recently been highlighted in a regional context by Clapp et al. (1995), Meen (1996), Berg (2002), Oikarinen (2005) or Holly et al. (2010). The monocentric case also provides a first order approximation to the description of the Paris housing market. Second, we consider a more complex structure of the intrinsic attractiveness, allowing for a better comparison with empirical data on transaction prices in Paris.

This article exhibits three important results. The first concerns the mathematical analysis of the model. This analysis is first presented for an arbitrary intrinsic attractiveness. This provides general results. The results are then specified for the particular case of a monocentric city and this monocentric case is simulated. For the empirical validation, we consider a more complex structure of the intrinsic attractiveness which matches the Paris data quite well. Within this theoretical framework, we provide an analytic description of the spatial distributions of transaction prices and incomes in the city. We prove that, when the market is not saturated, social segregation occurs if and only if the social influence is strong enough.

In this case, and this is the second important result, the general analytical solution underlines the emergence of critical endogenous thresholds in income and intrinsic attractiveness. The income threshold induces a segregation between those with an income above the threshold (whom we shall call the rich agents) and the others. These rich agents can freely choose their location, and they are the only ones to have access to locations above the threshold in intrinsic attractiveness. For agents below the income threshold, the housing possibilities are restricted to a subset of locations. Applied to a monocentric city, this gives that: (i) people with a higher income live closer to the city center while the poorer people live in the periphery; (ii) the rich agents are the only ones to live within a certain endogenous critical distance from the center, but there is no segregation amongst the rich agents within this central domain; (iii) agents below the income threshold have access to locations at distances greater than endogenous thresholds (the poorer the agent the greater the critical distance); (iv) at any location outside the center, except at the very periphery where only the poorest live, there is always some social mixing. And this is the third important result: if the first of these results, (i), is expected and in line with most of the literature, the preservation of social diversity, on the one side amongst the richest agents at the center, and

on the other side in a large domain between the center and the periphery, is particularly original and proved robust for a wide range of parameter values.

The paper is organized as follows. Section 2 presents the assumptions and the dynamics of the agent-based model. Section 3 reviews the key parameters of the model. Section 4 deals with the mathematical analysis of the equilibrium states. Section 5 describes the conditions under which segregation occurs. Section 6 presents numerical simulations for a simple case, that of a monocentric city. Section 7 considers a case allowing for a better comparison between simulated and empirical data. The conclusion follows.

2. The agent-based model: assumptions and dynamics

An agent-based model of residential location is proposed. In this model, agents make their decision according to both their income (characterized by an idiosyncratic willingness to pay) and their individual evaluation of the level of attractiveness of the different locations.

The assumptions of the model are presented below.

2.1. Time, space and goods

- Time is discrete and indexed by t. The time increment is δt (in the numerical simulations in Sections 6 and 7, $\delta t = 1$ is the numerical time step; in the mathematical analysis in Section 4.1, the continuous limit $\delta t \to 0$ will be taken). The horizon is infinite.

- We consider a 'city' defined as a discrete set Ω of locations X uniformly distributed on a bounded open set $\widetilde{\Omega}$ in $\mathbb{R}^2$.

  The origin O is taken as the geographical center of the city, and we denote by $D(X)$ the Euclidean distance from the center to a location X. The total number of locations is $Card(\Omega) = L^2$, and the space is of linear size (diameter) $\mathcal{D} = aL$, where a gives the typical distance between two neighboring locations (we therefore have $D(X) \leq aL/2$).

- In the city, there is a total number $\mathcal{N}$ of goods (housing for sale) with identical intrinsic characteristics. The same number $N = \mathcal{N}/L^2$ of dwellings is available at each location X in Ω.

2.2. Agents

At each period (given time t), there is a finite number of agents in the economy, who can be in one of the three following states: (1) buyer, (2) seller, (3) housed. We assume an infinite "reservoir" of agents outside the city. Agents in the reservoir are heterogeneous in their income - they are indiscernible except for their income category.

From this reservoir, at each time step t a constant number $\Gamma \delta t$ of randomly-chosen agents arrive on the market. They are new buyers who join the buyers who did not buy a dwelling in the previous period. At the same time, housed agents become sellers at a homogeneous rate α. So the goods available for sale at a given location are those put on the market by these agents plus, if any, those that have not yet sold. Then matching occurs between buyers and sellers at each location (see below, Section 2.7, for the detailed rule). Sellers who succeed in selling their good leave the market and return to the external reservoir.

Note that the total number of agents of each type - buyers, sellers and housed - are dynamical variables since they depend on the success rate of the previous period and on the inflow and outflow rates. In what follows, the term "insiders" designates agents who are already housed (former successful buyers), while "outsiders" designates agents looking for a flat.

2.3. Demand prices

Agents are characterized by their willingness to pay (hereafter WTP), which determines the maximum price the agent is ready to pay for an asset. For simplicity, we consider a finite number K of WTP. Agents with the same willingness to pay are denoted by k-agents, $k \in \{0, ..., K-1\}$, and have WTP P_k. These WTP are ordered by increasing values, $P_0 < P_1 < ... < P_{K-1}$. When the agent is acting as a buyer, his demand price is $P_k^d = P_k$. The agent's WTP also determines his behaviour as a seller, as explained below.

Whenever needed, notably in Section 5.6, the theoretical analysis will be done assuming an even distribution of the P_k values between a minimum value, P_0, and a maximum value, $P_0 + \Delta$:

$$P_k = P_0 \ + \ k \, \tfrac{\Delta}{K-1}, \quad k = 0, ..., K-1, \tag{1}$$

so that $\frac{\Delta}{K-1}$ is the constant increment between two consecutive prices. In addition, throughout this paper we assume that these K WTP values are

7

uniformly distributed among the agents in the external reservoir. Hence, among the $\Gamma \delta t$ new incoming agents at any time t, a fraction $1/K$ has WTP equal to P_k, $k \in \{0, ..., K-1\}$. These are not restrictive hypotheses, as shown in the Appendix, Section A.4.

2.4. Attractiveness

The attractiveness $A_k(X, t)$ of a location X at time t, for a k-agent, depends on both (1) an *intrinsic attractiveness*, $A^0(X)$, resulting from the location's intrinsic objective characteristics (e.g. local amenities), independent of the agent category, and (2) subjective characteristics which depend on the agent's social preferences, that is his preferences concerning the social characteristics of the neighborhood. This attractiveness matters both when the agent is a seller (see 2.5 below) and when he is a buyer (see 2.7 below).

2.4.1. The intrinsic attractiveness

The intrinsic attractiveness, $A^0(X)$, idiosyncratic to the position considered, is here assumed to be time-independent. This means that we consider the time scale of transactions to be much shorter than that involved in the transformation of amenities, which we do not take into account.

2.4.2. Attractiveness dynamics

The attractiveness $A_k(X, t)$ is a *subjective attractiveness*, whose value depends on the WTP of the agent looking at the location X. At each time step, new arrivals at the location X modify the social composition at this location, and the attractiveness evolves accordingly. Once the transactions between t and $t + \delta t$ have occurred (as described below), the attractiveness $A_k(X, t)$ of a location X as seen by a k-agent is updated according to:

$$A_k(X, t + \delta t) = A_k(X, t) + \omega \delta t \left(A^0(X) - A_k(X, t) \right) + \delta t \ \Phi_k(X, t) \qquad (2)$$

where $\Phi_k(X, t)$ formalizes how the transactions at time t affect the attractiveness according to the social preferences of the agents of WTP category k. We assume $\Phi_k(X, t) = 0$ whenever no transaction occurred at time t: the time evolution (2) then implies that, when there is no transaction at a given location for a certain amount of time, the attractiveness relaxes towards its intrinsic value $A^0(X)$. In the present work we assume that the attractiveness of a location for an agent increases when more agents with similar or higher

8

incomes are housed at this location. This is done by choosing $\Phi_k(X, t)$ proportional to the number of new buyers with WTP greater or equal to the agent's WTP. Hence we write:

$$A_k(X, t + \delta t) = A_k(X, t) + \omega \delta t \left(A^0(X) - A_k(X, t)\right) + \epsilon \delta t \, v_{k_>}(X, t) \qquad (3)$$

with

$$v_{k_>}(X, t) = \sum_{k' \geq k} v_{k'}(X, t) \qquad (4)$$

where $v_k(X, t)$ is the density of k-buyers in location X who complete a transaction at time t. Here and in all that follows, what we mean by *density* is a number

over the elementary surface area, a^2.

Note that the level of attractiveness for a given WTP category depends on the intensity of the demand from agents of similar or higher category: the higher the demand, the higher the level of attractiveness.

2.5. Offer prices

For each k-agent acting as a seller, his willingness to sell is determined by his WTP, P_k, and by the intensity of the demand, through the level of the mean attractiveness of the location. Specifically, the offer price $P_k^o(X)$ set by a k-agent selling a good at location X is given by:

$$P_k^o(X) = P^0 + (1 - \exp(-\xi \bar{A}(X, t))) \, P_k \qquad (5)$$

where P^0 is the minimum price of an offer, ξ is a parameter, and $\bar{A}(X, t)$ is the average attractiveness of location X at time t, that is:

$$\bar{A}(X, t) \equiv \frac{1}{K} \sum_{k=0}^{K-1} A_k(X, t). \qquad (6)$$

For a good at location X which has not yet been exchanged, the offer price is given by:

$$P^o(X) = P^0 + (1 - \exp(-\xi \bar{A}(X, t))) \, P^1 \qquad (7)$$

Note that for ξ small, the offer prices are all equal to P^0, with, for ξ small enough, $P_k^o(X) = P^0 + \xi \bar{A}(X, t) \, P_k$ and $P^o(X) = P^0 + \xi \bar{A}(X, t) \, P^1$, whereas for ξ large, the offer prices are, respectively, $P_k^o(X) = P^0 + P_k$ and $P_k^o(X) = P^0 + P^1$.

2.6. Transaction prices.

At location X, a transaction between a k-buyer (hence with demand price $P_k^d = P_k$) and a k'-seller with offer price $P_{k'}^o(X)$ given by (5), can be completed if $P_k > P_{k'}^o(X)$. If such a transaction occurs, the transaction price is assumed to be a linear combination of offer and demand prices:

$$P_{tr} = (1 - \beta)P_{k'}^o(X) + \beta P_k \tag{8}$$

where β is a constant coefficient. Similarly, for a good exchanged for the first time, the transaction price is $P_{tr} = (1 - \beta)P^o(X) + \beta P_k$,
    with $P^o(X)$ given by (7).

2.7. Market dynamics: the matching

The market dynamics is then as follows, starting at time $t = 0$ with an empty city and the attractiveness set to its intrinsic value. At each time step (between times t and $t + \delta t$):

1. **On the demand side.** For each $k \in \{0, ..., K - 1\}$, the total number of k-buyers (outsiders) in the city is the sum of a constant number $\frac{\Gamma}{K}\delta t$ of new buyers (coming from the reservoir), and of the k-buyers, if any, who are already present but have not yet bought a home. Each of these agents decides to visit one particular location. The probability $\pi_k(X)$ for a k-buyer to visit a given location X, depends on the attractiveness of the location:

$$\pi_k(X, t) = \frac{1 - \exp(-\lambda A_k(X, t))}{\sum_{X' \in \Omega} 1 - \exp(-\lambda A_k(X', t))}. \tag{9}$$

For small λ (that is for $\lambda \max_{k,X} A_k(X, t) << 1$), this is equivalent to:

$$\pi_k(X, t) = \frac{A_k(X, t)}{\sum_{X' \in \Omega} A_k(X', t)} \tag{10}$$

These decisions, being made in parallel by all the outsiders, determine the demand at each location.

The above rule (9) can be seen as resulting from a search process where the agent requires a minimum acceptable level of quality. This latter is based on the relative attractiveness and on an additional component, corresponding to private information and/or local idiosyncratic characteristics of the good. This can be modeled as a random variable, so that in

the end, the probability of selecting a location is an increasing function of $A_k(X,t)$. The precise rule, Equ. (9), is not meant to assume some specific properties of the unknown random component, but is chosen for its simplicity and for the smoothness properties of the resulting search process.

2. **On the offer side.** At each location, the goods offered are (i) those which have not yet been sold, if any; (ii) those put on sale by the current owner at some previous time step and not yet sold, if any; (iii) goods newly put on the market by a fraction α of the remaining insiders.

3. **The matching.** We presented above the search process on the demand side. On the offer side, the search process is as follows. At each location, one of the sellers is selected by random draw (game against Nature); this seller then picks at random one of the potential buyers among those applying at this location and having a large enough WTP (see 2.6 above), and the transaction occurs. This is repeated until no offer or no buyer remains. Remember that sellers who succeed in selling their good leave the market, returning to the external reservoir.

The matching rule is convenient for the mathematical analysis. In the numerical simulations, we also test an alternative matching rule: sellers are treated in order of increasing offer price, the buyers being randomly chosen as in the previous case.

4. **Updating of the attractiveness** Once the transactions have occurred, the attractiveness $\{A_k(X,t); k = 0, ..., K-1; X \in \Omega\}$ is updated as explained in 2.4 above, Equ. (3).

3. Model parameters and key features

Let us now summarize the set of key parameters that characterize the model.

The model behaviour depends on the following set of control parameters: (1) the incoming rate, that is the number Γ of new agents entering the economy per unit of time; (2) the rate α at which agents leave the city; the parameters in the attractiveness dynamics, that is, (3) the time scale $1/\omega$ of relaxation towards the intrinsic value, and (4) the weight of the social influence, ϵ; (5) the parameter ξ which fixes the relevant scale of the attractiveness values in the offer prices (see Eq. (5)). Finally, there are two

11

structural control parameters: the intrinsic attractiveness, $A^0(X)$, and the social influence term, Φ_k, in the dynamics of the attractiveness, Eq. (2).

As we will see, the five control parameters only matter through three particular (dimensionless) effective parameters:

- The product

$$\widetilde{\xi} \equiv \xi A^0_{max} \tag{11}$$

where A^0_{max} is the maximum intrinsic attractiveness. The regime of interest is the one of not too large values of $\widetilde{\xi}$, for which the offer prices are smoothly modulated by the attractiveness.

- The ratio of incoming to outgoing rates, $\frac{\Gamma}{\alpha}$. We are concerned with a parameter regime for which an equilibrium exists with no saturation (as explained in the next section). This requires an equality between the mean inflow and outflow of agents, with a number of buyers that does not saturate the market. One global necessary condition is that this ratio must be smaller than the number of goods per location, that is:

$$\frac{\Gamma}{\alpha \mathcal{N}} = \frac{\gamma}{\alpha\, n} < 1 \tag{12}$$

where

$$\gamma \equiv \Gamma/\mathcal{D}^2 \tag{13}$$

is the mean incoming rate per surface area, and $n = \mathcal{N}/\mathcal{D}^2 = N/a^2$ is the density of goods (assumed to be uniform here).[2]

- An effective social influence parameter characterizing, in the building of the attractiveness, the strength of the social influence term compared to the intrinsic attractiveness:

$$\eta \equiv \frac{\epsilon \gamma}{\omega \left\langle A^0 \right\rangle} \tag{14}$$

where the brackets $\langle . \rangle$ denote (here and in what follows) the spatial average over the city,

$$\left\langle A^0 \right\rangle \equiv \frac{1}{L^2} \sum_{X \in \Omega} A^0(X). \tag{15}$$

_____

[2]Remark: for the multi-agent simulations it is convenient to define quantities per location, whereas for the mathematical analysis with continuous limits, we consider densities.

The origin of this parameter η can be understood by looking at the dynamics (17) and noting that the maximum possible value of $v_{k_>}(X,t)$ is controlled by the incoming rate, Γ.

The analysis done in this paper is for an arbitrary choice of A^0, and specified for particular cases, as previously explained.

What really matters in the present model are (1) the size of the domain over which A^0 is significantly non-zero (determining the part of the city on which transactions concentrate (see the Appendix, section A.3)), and (2) the spatial heterogeneity in A^0 values. This shows up in the results of the following analysis of the stationary regime, where variables are found to depend on ratios such as $\frac{A^0(X)}{\langle A^0 \rangle}$ or $\frac{A^0_{max}}{\langle A^0 \rangle}$.

As already mentioned, different social preferences could be studied, by appropriate choice of the social term $\Phi_k(X,t)$. In particular, all the results in this paper straightforwardly generalize to the case where the agents weight higher categories differently:

$$\Phi_k(X,t) = \epsilon \sum_{k' \geq k} w_{k'-k}\, v_{k'}(X,t) \tag{16}$$

with weights $w_j \geq 0$ (see in particular the Appendix, Sections A.1 and A.4).

Finally, we note that the other parameters are "harmless", that is their values are not essential in controlling the model behaviour, provided they satisfy some global condition such as being 'small' or 'large': the time scale δt (small in the mathematical analysis); the number of locations L^2, assumed to be large; the (large) city size $\mathcal{D}$ (or equivalently the small space scale $a = \mathcal{D}/L$ between two nearby locations); the number K of WTP categories (arbitrary, but it may be convenient to consider the large K limit as providing a fine discretization of any continuous distribution of incomes); the set of WTP values (the P_ks) and the price parameters (P^0, P^1, β), which can be chosen so as to set price scales for comparison with any given empirical data; the total number $\mathcal{N}$ of goods in the city, with a number per location, $N = \mathcal{N}/L^2$, assumed to be large in the mathematical analysis; the parameter λ entering the probability to choose a location, Eq. (9), assumed to be small (compared to the typical values of attractiveness), in which case it plays no role, except for providing smoothness properties with occasionally extremely high values of attractiveness.

Section 4 now mathematically characterizes the stationary regime that can be reached for a large range of parameter values.

4. Theoretical analysis: dynamics and equilibrium

4.1. Continuous time dynamics

For the mathematical analysis, the evolution of the above agent-based model is further formalized through partial differential equations, taking the continuous time limit (that is $\delta t \to 0$).

In this continuous limit, the updating rule of the attractiveness $A_k(X, t)$ of a location X seen by a k-agent, Equ. (3), gives:

$$\partial_t A_k(X, t) = \omega \left(A^0(X) - A_k(X, t) \right) + \epsilon \, v_{k_>}(X, t) \tag{17}$$

Given the outgoing rate α (the fraction of agents leaving the inside per unit of time), and $v_k(X, t)$ being the density of newly-housed k-agents, the density of insiders with willingness to pay P_k, $u_k(X, t)$, satisfies the following differential equation:

$$(1 - \alpha)\partial_t u_k(X, t) = v_k(X, t) - \alpha \, u_k(X, t) \tag{18}$$

The density of outsiders $\rho_k(X, t)$ is written as:

$$\rho_k(X, t) = v_k(X, t) + \bar{v}_k(X, t), \tag{19}$$

with $\bar{v}_k(X, t)$ denoting the density of agents who did not succeed in buying a good at location X and time t. Given the rules of the dynamics, the evolution of the density of outsiders $\rho_k(X, t)$ can then be written as:

$$\partial_t \rho_k(X, t) = -\rho_k(X, t) \, + \, \frac{\gamma}{K} L^2 \, \pi_k(X, t) \, + \, \pi_k(X, t) \sum_{X' \in \Omega} \bar{v}_k(X', t) \tag{20}$$

In the above equation, $\frac{\Gamma}{K}\pi_k(X, t)$ is the rate at which new buyers visit X at time t,

$\pi_k(X, t)$ being given by (9). The last term on the right-hand side of (20) is the fraction of agents, among those who did not buy a good at time t, who decided to search for a flat at location X. Remark: note that the scaling gives a well-defined limit as $a \to 0$:

$$\partial_t \rho_k(X, t) = -\rho_k(X, t) \, + \, \frac{\Gamma}{K} \, f_k(X, t) \, + \, f_k(X, t) \int \bar{v}_k(X, t) \, d^2 X \tag{21}$$

where $f_k(X, t) = \pi_k(X, t)/a^2$ is the probability density.

14

4.2. Equilibrium

4.2.1. Stationary state

We define the *equilibrium* as the stationary state (whenever it exists) of the dynamics, i.e., the variables of the model - $\rho_k(X,t)$, $A_k(X,t)$, $u_k(X,t)$ and $v_k(X,t)$ - become constant in time. All the variables in the stationary state are denoted with an asterisk *.

Writing $\partial_t u_k = 0$, $\partial_t A_k = 0$ and $\partial_t \rho_k = 0$, in equations (18), (17) and (20) respectively, we obtain for the stationary state:

$$v_k^*(X) = \alpha u_k^*(X), \tag{22}$$

$$A_k^*(X) = A^0(X) + \frac{\epsilon}{\omega} v_{k_>}^*(X) \tag{23}$$

$$\rho_k^*(X) = \pi_k^*(X) \sum_{X'\in\Omega} \rho_k^*(X'). \tag{24}$$

with

$$\pi_k^*(X) = \frac{1 - \exp(-\lambda A_k^*(X))}{\sum_{X'\in\Omega} 1 - \exp(-\lambda A_k^*(X'))} \tag{25}$$

or, in the small λ limit (see (10),

$$\pi_k^*(X) = \frac{A_k^*(X)}{L^2 \langle A_k^* \rangle} \tag{26}$$

with $\langle A_k^* \rangle = (1/L^2) \sum_{X'\in\Omega} A_k^*(X')$.

Summing equation (20) on the whole space, we conclude that:

$$\frac{1}{L^2} \sum_{X\in\Omega} v_k^*(X) = \alpha \frac{1}{L^2} \sum_{X\in\Omega} u_k^*(X) = \frac{\gamma}{K}. \tag{27}$$

Consequently, in the stationary state, the mean density of k-housed agents on the whole space is equal to $\frac{\gamma}{\alpha K}$. The total rate of k-transactions is $\frac{\gamma}{K}$. Thus, the total number of housed agents is the same for all the k values. In the following, we restrict the analysis to equilibria in a non-saturated regime, as defined below.

Definition: non-saturated equilibrium. *A non-saturated equilibrium is defined as an equilibrium where, for any given $k \in \{0, ..., K-1\}$, at any location $X \in \Omega$, either the WTP of the k-agents are too low, so that none of them can buy a dwelling in this location, hence $v_k^*(X) = 0$, or the k-agents*

15

can afford to buy in this location, and in this case all k-demand is satisfied, that is:

$$\bar{v}_k^*(X) = 0. \tag{28}$$

We denote by Ω_k the sub-set of locations X where all the k-agents' demand is satisfied:

$$\Omega_k \equiv \{X \in \Omega | \bar{v}_k^*(X) = 0\} \tag{29}$$

Ω_k is the set of locations (possibly empty, possibly identical to Ω) where the k-agents are able to buy homes, and the complement set, $\Omega \setminus \Omega_k$, is the set of locations where they cannot.

In what follows, the step-by-step analysis of the non-saturated equilibrium implies the unicity of such an equilibrium, whenever it exists. The existence conditions on the parameter values are the condition on the ratio of incoming to outgoing rates compared to the number of available goods (see Section 3), the criterion (12), and the conditions discussed in Section 5.2 concerning the offer and demand prices. As shown in Section 6, the above definition defines an equilibrium which is a fairly good approximation of what is observed in the numerical simulations. Results would be different on a market with rationing (saturated case), as we will explain later.

4.2.2. The conditions of segregation emergence: the WTP threshold

If the fraction of housed k-agents corresponds to the fraction of k-agents in the external reservoir and the (k-independent) intrinsic attractiveness, it means that no segregation has emerged from the market dynamics. Let us now characterize the subsets Ω_k of location possibilities for each category of k-agents. The agents with the highest WTPs have access to any location, so that for them, $\Omega_k = \Omega$. If a k-agent can afford a good at a given location, it is clear that k'-agents with $P_{k'} > P_k$ can as well: we must have $\Omega_k \subseteq \Omega_{k+1} \subseteq \Omega$. Hence, in a stationary regime, either $\Omega_k = \Omega$ for every k (in which case there is no segregation), or there exists a critical value $\bar{k}$ such that for $k < \bar{k}$, the demand of the k-agents is not satisfied on some of the locations, $\Omega_k \subset \Omega$ (and thus some socio-spatial segregation occurs). This means equivalently that there is a *WTP threshold* $P_c^* = P_{\bar{k}}$ so that only agents with WTP at least equal to this threshold can buy a good anywhere in the city.

Section 5 below sets out the existence conditions of such a WTP threshold P_c^* associated with a critical category $\bar{k}$, together with the characterization of the set Ω_k of locations where the $k-$agents with $k < \bar{k}$ can afford to buy

a good. We first determine the densities of housed agents in a non-saturated equilibrium whenever $\bar{k} \leq K - 1$, distinguishing between categories above the threshold, $k \geq \bar{k}$, and categories below the threshold, $k < \bar{k}$.

4.2.3. Housed densities

For a non-saturated equilibrium, we determine the density of housed agents in terms of the intrinsic attractiveness and of the sets Ω_k introduced above. We consider here the small λ limit: we take (26) for the probability of choosing a location.

Proposition: In a non-saturated equilibrium, whenever $P_c^* \leq P_{K-1}$, that is $\bar{k} \leq K - 1$:

- for any WTP level $k \geq \bar{k}$, the density of housed k-agents does not depend on the WTP level, but only on the intrinsic attractiveness of the location, according to:

$$\forall k \geq \bar{k} \ \ \forall X \in \Omega, \ u_k^*(X) = \frac{\gamma}{K\alpha} \frac{A^0(X)}{\langle A^0 \rangle} \tag{30}$$

  where $\langle A^0 \rangle$ is the average of A^0 as defined in (15).

- if $\bar{k} > 0$ (that is $P_c^* > P_0$), for any WTP category $0 \leq k < \bar{k}$, that is with P_k smaller than the WTP threshold P_c^*, the density of housed k-agents depends on the intrinsic attractiveness of the location and on the set $\Omega_k \subset \Omega$ of locations in which the k-agents can afford to buy a good, according to:

$$u_k^*(X) = \frac{\gamma}{K\alpha} \frac{A^0(X)}{\langle A^0 \rangle_k} \text{ if } X \in \Omega_k \tag{31}$$
$$= 0 \text{ otherwise} \tag{32}$$

  with

$$\langle A^0 \rangle_k \equiv \frac{1}{L^2} \sum_{X' \in \Omega_k} A^0(X'). \tag{33}$$

The proof of this Proposition, demonstrated by recurrence on k starting from the highest WTP value, is given in the Appendix, Section A.1. This Proposition can also be shown to apply when the agents weight higher categories differently, as in Eq. (16).

17

Eq. (30) shows that for agents with high enough WTP ($k > \bar{k}$), the social influence parameter does not show up in the number of housed agents. An intuitive explanation is that in a non-saturated regime, for "rich" agents not constrained by their income, the probability of settling at a given location depends on the attractiveness of this location relative to that of all the other locations.

Eq. (31) is an intuitive generalization of Eq. (30): the normalization, that is the denominator in (31), is given by the sum of A^0 over locations where the considered k-agents can locate. For $k \geq \bar{k}$, by definition $\Omega_k = \Omega$, and (31) reduces to (30).

5. Thresholds and segregation

Let us now determine the conditions for having the WTP thresholds within the range $[P_0, \; P_{K-1}]$, i.e., determine where people can be housed depending on their WTP category. In other words, let us characterize the sets Ω_k.

The absence of segregation in the stationary regime would mean that, at any location, there are transactions involving any WTP category. If, on the contrary, at some locations the market dynamics make the distribution of offer prices higher than the distribution of demand prices, the poorest people cannot be housed, and thus some segregation occurs. In the worst case, the offer prices become so high that the intersection between the two distributions is null: this would mean that the system does not settle in a stationary state, but reaches a "frozen" state where housed agents cannot sell their good to any other agent. We determine here the conditions on the parameters that separate the regime with non-segregated equilibrium from the one with segregated equilibrium, and the regime with segregated equilibrium from the frozen state. This will also determine the thresholds in the intermediate parameter regime where segregation occurs.

5.1. Stationarity criterion: non-profitable arbitrage

The existence of a stationary regime requires the absence of profitable arbitrage. Indeed, if at a given location, all the poorest inhabitants were selling their goods with a positive profit, this lowest category of income would eventually disappear from this location. We thus write the stationarity condition, in other words, at each location, for each WTP level k present in this location, the condition for the realization of transactions between two agents of

18

the same category. From the definition of the offer prices, Eq. (5), we get the condition $P_k \geq P^0 + (1 - \exp(-\xi \bar{A}^*(X)))P_k$, which we can write as:

$$\xi \bar{A}^*(X) \leq \log \frac{P_k}{P^0} \tag{34}$$

where we recall that $\bar{A}^*(X) = \frac{1}{K} \sum_k A_k^*(X)$. As explained below, this condition enables us to identify the WTP category threshold $\bar{k}$.

5.2. The frozen state

At the limit, nobody can be housed, and the attractiveness decreases towards the intrinsic value A^0. Then, a necessary condition for avoiding the frozen state is to have the criterion (34) satisfied for the highest income category, $k = K - 1$, when the attractiveness is equal to the maximum value that can be taken by the *intrinsic* attractiveness. This gives the condition:

$$\xi A^0_{max} < \log \frac{P_{K-1}}{P^0}. \tag{35}$$

A frozen market could be linked to a market bubble. This sort of crisis could be attenuated by having adaptive agents with idiosyncratic P^0 and ξ values: in order to sell their good, agents could lower at least one of these parameters, so that they would accept lower prices. Such a scenario would lead to an endogenous determination of some of the parameters. This clearly goes beyond the scope of this paper. We simply assume here that the P^0 and ξ values are the same for every agent and within a range that allows to reach a stationary regime.

5.3. The WTP Threshold

We now write the conditions for the existence of a critical WTP level, $\bar{k} \in \{1, ..., K - 1\}$. Assuming that such a threshold exists, we get the expression of $\bar{A}^*(X)$ (see the Appendix, Section A.2):

$$\bar{A}^*(X) = A^0(X) \left(1 + \frac{\eta}{2} [\frac{K+1}{K} - \frac{(\bar{k}+1)\bar{k}}{K^2}] \right) \tag{36}$$

where $\eta = \frac{\epsilon \gamma}{\omega \langle A^0 \rangle}$, as defined in (14). The category threshold $\bar{k}$, and the associated WTP threshold $P_c^* = P_{\bar{k}}$, are obtained by writing the inequality (34) for locations X where A^0 reaches its maximum value, A^0_{max}:

$$\xi A^0_{max} \left(1 + \frac{\eta}{2} [\frac{K+1}{K} - \frac{(k+1)k}{K^2}] \right) \leq \ln \frac{P_k}{P^0} \tag{37}$$

and $\bar{k}$ is the smallest value of k for which this inequality is true.

Large K limit. For illustrative purposes, let us here assume that the WTP values are evenly distributed between two extreme values, according to Eq. (1), that is $P_k = P_0 + \Delta \frac{k}{K-1}$, $k = 0, ..., K - 1$ (see the Appendix, Section A.4, for the case of an arbitrary distribution). In the large K limit, writing that the threshold corresponds to the marginal case where the inequality becomes an equality, we obtain the WTP threshold P_c^* as the solution of:

$$\log \frac{P_c^*}{P^0} - \widetilde{\xi} = \eta \frac{\widetilde{\xi}}{2} \left(1 - \left(\frac{P_c^* - P_0}{\Delta} \right)^2 \right) \tag{38}$$

where $\widetilde{\xi} = \xi A_{max}^0$, as defined by Eq. (11) Section 3, is the effective parameter controlling the attractiveness dependency of the offer prices.

Since $\eta = \frac{\epsilon \gamma}{\omega \langle A^0 \rangle}$, as defined in (14), note that the product $\eta \widetilde{\xi}$ on the right-hand side of (38) combines five key control parameters (see Section 3): the inflow of agents γ, the social strength ϵ, the relaxation time scale $1/\omega$, the offer prices parameter ξ, and a measure of the heterogeneity of intrinsic attractiveness, $\langle A^0 \rangle / A_{max}^0$.

5.4. The emergence of segregation

Let us study the existence and solutions of Eq. (38). Analysis of the equation shows that there are three parameter regimes to consider, as detailed below.

- First, there is no solution for $\widetilde{\xi} > \log(P_{K-1}/P^0)$: this corresponds to the frozen state, where no transactions occur.

- Second, there exists a parameter domain where segregation occurs even in the absence of social influence. This is the case for $\log P_0/P^0 < \widetilde{\xi} \leq \log(P_{K-1}/P^0)$. In this case, a solution exists defining a critical threshold larger than P_0. Such a condition on the parameters is reached for large enough values of $\widetilde{\xi}$, that is if the offer prices are high, even if this is only due to the intrinsic attractiveness. For large values of $\widetilde{\xi}$, Eq. (5) implies that the offer prices do not even depend on the attractiveness: $P_k^o = P^0 + P_k$.

- The third case corresponds to $\widetilde{\xi} < \log P_0/P^0$, illustrated in Fig. 1. As the figure shows, and as one can verify in Eq. (38), the threshold P_c^* is an increasing function of η. This means that P_c^* increases if the social factor ϵ increases, or if the inflow of agents increases; whereas P_c^* decreases if the relaxation of the attractiveness towards the intrinsic value A^0 is faster (larger ω), or if the domain of strongly attractive locations increases (larger $\langle A^0 \rangle$): and this is consistent with the economic intuition.
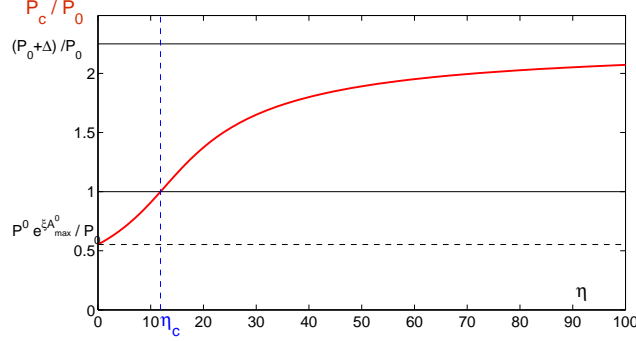
20

Figure 1: Willingness to Pay Threshold P_c^* (normalized by P_0) as function of the (effective) social strength η, illustrated for $\Delta/P_0 = 1.25$ and $P^0 e^{\xi A_{max}^0}/P_0 = 0.55$. Segregation occurs for η larger than the critical value η_c given by condition $P_c/P_0 = 1$ (here $\eta_c = 11.86$).

As η goes from 0 to $+\infty$, the critical threshold P_c^* increases from $P^0 e^{\widetilde{\xi}}$ to $P_0 + \Delta$. There is thus a critical value η_c of η:

$$\eta_c = 2 \, (\frac{1}{\widetilde{\xi}} \, \log \frac{P_0}{P^0} \, - \, 1) \; > 0, \tag{39}$$

below which P_c^* is smaller than P_0, and thus every agent can locate at the center: there is no segregation. It is only when the social influence, $\eta > \eta_c$ is strong enough, that there exists a WTP threshold that segregates the richest away from the others. This is one of the most important qualitative results: the heterogeneity in A^0 values is not enough to lead to income segregation; in a non-saturated equilibrium, segregation emerges from the strengthening of such heterogeneities, provided the strength of social influence is large enough.

For a given value of η, the condition for the emergence of segregation, $\eta > \eta_c$, is a condition on the parameters P_0, and $P^0 e^{\widetilde{\xi}}$, which characterize the smallest WTP value and the smallest offer value. Note that the threshold η_c does not depend on WTP distribution, but only on the minimum value of WTP in the population, whereas the critical WTP value does depend on WTP distribution.

5.5. Local and category thresholds

From the stationarity condition (34), we can also deduce for each location X the poorest housed category $k_c(X)$ (that is the smallest value of k such that

21

the k-agents are present on X), and for each category $k < \bar{k}$ the maximum intrinsic attractiveness level, $A^0_{c,k}$, which defines the locations X where the k-agents can be housed: $\Omega_k = \{X | A^0(X) \leq A^0_{c,k}\}$.

Consider $k < \bar{k}$. Since $\langle A^0 \rangle_k \leq \langle A^0 \rangle$, from (34) and (A-16) we obtain the inequality:

$$\xi A^0(X) \leq \frac{\ln \frac{P_c^*(X)}{P^0}}{1 + \frac{\eta}{2}\left[\frac{K+1}{K} - \frac{(k_c(X)+1)k_c(X)}{K^2}\right]} \tag{40}$$

with $P_c^*(X) = P_{k_c(X)}$. This inequality can also be read as giving the maximum intrinsic attractiveness level $A^0_{c,k}$:

$$\xi A^0_{c,k} \leq \frac{\ln \frac{P_k}{P^0}}{1 + \frac{\eta}{2}\left[\frac{K+1}{K} - \frac{(k+1)k}{K^2}\right]} \tag{41}$$

Since the inequality (40) has been derived by replacing $\langle A^0 \rangle_k$ by $\langle A^0 \rangle$, we expect it to be strict. We get a lower bound on $k_c(X)$ (and a lower bound on $P_c^*(X)$, an upper bound on $A^0_{c,k}$) by writing that the category threshold $k_c(X)$ is the smallest value of k for which this inequality is true - in a continuous limit, it is the marginal value for which the inequality becomes an equality. As will be seen with the numerical simulations, these bounds are very good approximations.

5.6. Case of a monocentric city: critical distances

Let us now apply the above results to the case of a monocentric city, with an intrinsic attractiveness monotonically decreasing from the center. Since the intrinsic attractiveness decreases from the center, the condition (41) means that for each k, the restricted set Ω_k is characterized by a critical distance $d_c^*(k)$ from the center, beyond which the $k-$agents' demand is positive. For illustrative purposes, let us assume A^0 has a two-dimensional Gaussian shape:

$$A^0(X) = A^0_{max} \exp\left(-\frac{D(X)^2}{R^2}\right) \tag{42}$$

where A^0_{max} is the maximum intrinsic attractiveness, $D(X)$ is the Euclidean distance from the center, and R determines the distance from the center at which the intrinsic attractiveness is still significantly non-zero. Then, the critical distance $d_c^*(k)$ satisfies:

$$d_c^*(k) \geq R\sqrt{-\ln\left[\frac{\frac{1}{\xi A^0_{max}} \ln \frac{P_k}{P^0}}{1 + \frac{\eta}{2}\left[\frac{K+1}{K} - \frac{(k+1)k}{K^2}\right]}\right]} \tag{43}$$

22

The above equation is valid whenever the argument of the logarithm is smaller or equal to 1, and otherwise $d_c^*(k) = 0$. This is the case for all k levels above the WTP threshold, and the critical WTP level $\bar{k}$ is the smallest value of k for which $d_c^*(k) = 0$. The distance $d_c^*(\bar{k} - 1) > 0$ gives the limit of the central domain where only the richest can buy a dwelling.

5.7. Summary

Let us briefly summarize the main results of this section and comment on the consequences in terms of space organization (social mix or segregation). For the stationary state, we have shown that if and only if the social influence is strong enough, there emerges a WTP threshold which segregates the richest from the others. The spatial distribution of agents with a willingness to pay higher than this critical value does not depend on the individual income level but only on the intrinsic attractiveness of the locations. In other words, there is a social mix between agents with an income above the WTP threshold. Agents with income below this threshold can only buy goods in a restricted set of locations, which depends both on the intrinsic attractiveness and on their income. In such locations, however, some social mix remains.

6. Numerical simulations

We now simulate the agent-based model as presented in Section 2, illustrating the theoretical results in the case of a pure monocentric city.

6.1. Simulation specifications

Following the rationale presented in Section 3, the model parameters are chosen as follows. We consider a square lattice of linear size $L = 100$, hence with $L \times L = 10000$ sites (locations). Each location X is characterized by its Cartesian coordinates (x, y) in the frame whose origin is the center of the square lattice. Without loss of generality, time and space scales are arbitrarily given by the time step $\delta t = 1$ and the lattice spacing $a = 1$ respectively. The intrinsic attractiveness $A^0(X)$ is chosen to decrease with the distance from the center, as in (42): we express it as a two-dimensional Gaussian function,

$$A^0(X) = A^0_{max} \exp(-\frac{D(X)^2}{R^2}) \tag{44}$$

where A^0_{max} is the maximum intrinsic attractiveness, $D(X) = (x^2 + y^2)^{1/2}$ is the distance from the center, and R determines the distance from the center

23

at which the intrinsic attractiveness is still significant. We set $A^0_{max} = 1$ and $R = 10$.

Most simulations are driven with $K = 10$ categories, but we also present some results for $K = 50$, which is large enough for comparison with the theoretical results obtained in the large K limit. The total number N of goods at each location is set equal to 200. This is large enough to prevent the system from saturating, i.e., to avoid having the number of assets for sale falling to 0 on some sites.

The WTP values are set following Eq. (1), with $P_0 = 10^5$ and $\Delta = 225000$. The offer price parameters are $P^0 = 9\,x10^4$, $P^1 = 2\,x10^5$ and $\xi = 0.1$.

The other parameters are: the outgoing rate $\alpha = 0.1$, the inflow $\frac{\Gamma}{K} = 1000$, the social influence strength $\epsilon = 0.022$, the attractiveness relaxation constant $\omega = \frac{1}{15}$, the transaction price parameter $\beta = 0.1$, and finally $\lambda = 0.01$.

With these parameter values, the effective social strength parameter is $\eta = 10.504$ for $K = 10$, and $\eta = 52.521$ for $K = 50$. In both cases, the critical value as given by Eq. (39) is $\eta_c = 0.107$, so that segregation is expected to occur. Finally, if one considers the distance unit as km and the price unit as euro, the results are of the same order of magnitude as what we observe for the Paris market (this is discussed in more detail in the next section).

The agent-based dynamics, detailed in Section 2.7,
is briefly recalled here. At $t = 0$, the simulated city is empty, and the initial attractiveness of a location X seen by every k-agent is set to the intrinsic attractiveness: $A_k(X, t = 0) = A^0(X)$. At each time step, a randomly chosen fraction α of housed agents put their good on the market. On the demand side, a number Γ/K of new k-agents arrive from the external reservoir and join the agents who failed to buy at the previous time step, if any. Every one of these agents chooses a location X where he would like to buy, the choice depending on the attractiveness of the locations according to Eq. (9). In each location, available offers are sold according to the chosen matching rule. Two sets of simulations are done, testing two versions of the matching rule (see Section 2.7):

(a) the goods are sold in a random order, meaning that one buyer is picked at random for a randomly-chosen offer, then another buyer is picked a random for a second randomly-chosen offer, and so on.

(b) the goods are sold in order of increasing offer price, meaning a buyer is picked at random for the least expensive offer, then a second buyer

24

is picked at random for the second least expensive offer, and so on.

Once all the transactions have been completed, the attractiveness is updated according to Eq. (3). The process is repeated until a stationary regime is reached.

In all cases, we find that the system reaches a stationary state characterized by a constant number of housed agents on the lattice. In what follows, we analyze these stationary states. Focusing on the income and space distribution of housed agents, we compare the results of the numerical simulations with the results of the theoretical analysis when the matching rule is rule (a).

6.2. Price distribution

Once a stationary regime has been reached, we measure the mean transaction prices and the associated variances. For the price distributions, we find no noticeable differences when rule (a) or rule (b) is selected, and Fig. 2 presents the results for rule (a). As expected from the theoretical results, transaction prices decrease from the center to the periphery (Fig. 2, left). Looking at the standard deviations (Fig. 2, middle), we can distinguish two parts: in the city center, a large distribution of the prices shows an active dynamics, whereas in the near peripheral area, we observe a trend towards price homogenization. The interesting fact that prices present a larger variance in the central area than at the periphery, i.e., the variance is higher for higher mean prices (Fig. 2, right), is consistent with the theoretical analysis in Section 4.2: in locations where only the richest can afford to buy a good, there is no segregation amongst these rich agents. Hence there is a large variation of transaction prices in these locations.

6.3. Socio-spatial segregation and income mix

Fig. 3 shows the occupancy ratio for the different WTP in the stationary state with respect to the distance from the center. People are located inside the zone where the intrinsic attractiveness is significantly non-zero, in keeping with the analysis of Section A.3.

Figure 3 compares the distribution of housed people with two different search rules: in case (a), the density of k-agents excluded from the center jumps discontinuously to a non-zero value at a critical distance from the center, whereas in case (b), the transition is smoother. However, the overall distribution is the same, confirming the theoretical predictions obtained for
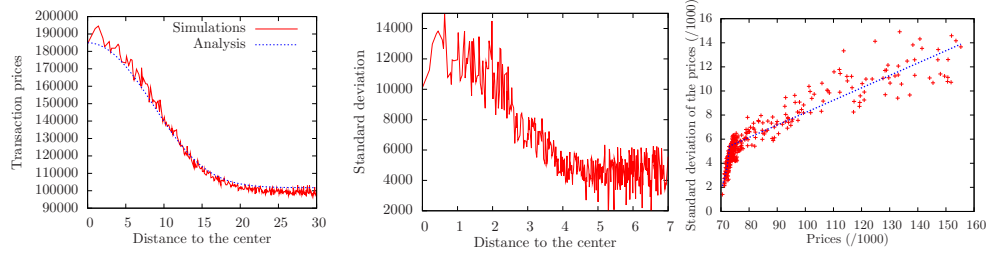
Figure 2: Left: average prices (euros) as a function of the distance from the center, for $K{=}50$: the analytical result (solid curve) fits the numerical simulations (dashed curve) well. Middle: standard deviation of transaction prices (euros) versus distance from the center; Right: standard deviation of transaction prices versus mean transaction prices (thousands of euros). The straight lines are linear regressions: (low prices) slope 0.9, intercept -62.000 correlation coeff. -1.0; (large prices) slope 0.1, intercept -2.000, correlation coeff. -0.97.
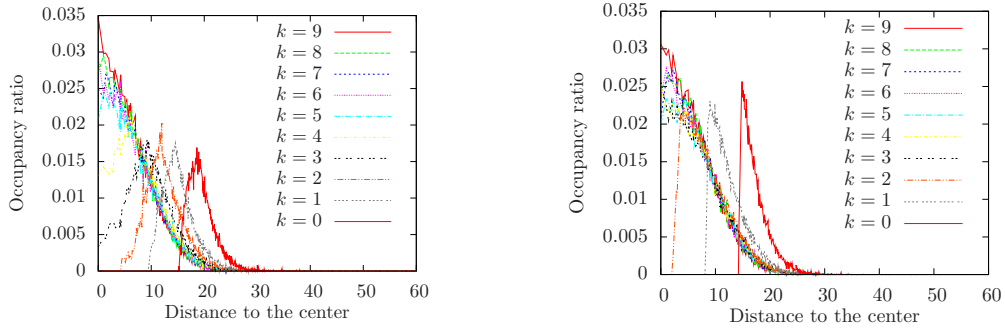
Figure 3: Occupancy ratio per WTP versus the distance from the center in the stationary state. $K = 10$ income categories are present on the lattice. The distance from the center is in arbitrary units. The offers to be sold are randomly chosen (case (a), right panel), or in order of increasing prices (case (b), left panel).
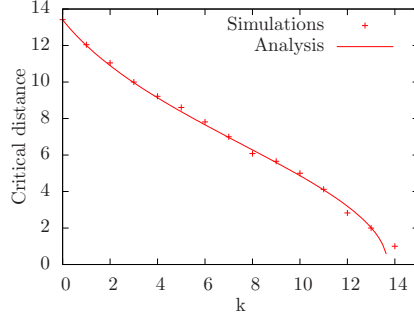
Figure 4: Critical distance with respect to the WTP category k ($K = 50$): the points are obtained from the simulations, the solid curve is the analytical bound (the right-hand side of (43)).

case (a): two zones of segregation, the rich and the poor, and in between an intermediary zone with some social mix.

The critical WTP value $\bar{k}$ is found to be around 15, and the agents with a lower WTP are distinguished by their income: the lower the WTP, the higher the distance from the center. Fig. 4 shows how the critical distance increases as the WTP level decreases. Fig. 5 represents the number of agents housed in a given area, on the left for one category of rich agents, and on the right for one category of poor agents. Both figures exhibit a good correlation between the analytical results and the numerical simulations.

To summarize, when the offers are randomly selected - rule (a) -, the numerical results fit the non-saturated equilibrium defined in the analytical part: at some distance specific to the agent's WTP, one has a sharp transition from no housing ($v(X) = 0$) to complete housing (all potential buyers become housed, $\bar{v}(X) = 0$). In the case where the less expensive assets are sold first - rule (b) -, there exists a small domain of locations where both $v(X)$ and $\bar{v}(X)$ are non-null. We recall, however, that for the distribution of transaction prices, the two rules give the same results.

6.4. Social mix index

Now we will use quantitative measures of segregation to estimate the level of social diversity. There is a large literature on such measures: here we consider two measures particularly adapted to our study. First, we introduce
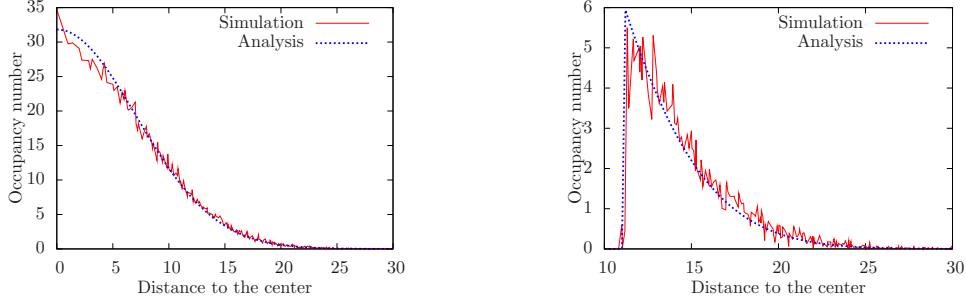
Figure 5: Mean number of housed agents with respect to the distance from the center ($K = 10$). Left: for rich agents (income level $k = 9$). Right: for poor people (income level $k = 2$). The solid curves correspond to the simulations, while the dotted ones are from the theoretical analysis: Left, Equ. (30), right: Eq. (31), with $d_c^*(k)$ replaced by its lower bound, the right-hand side of (43).

a multigroup social mix index derived from the mean relative deviation index proposed by Reardon & Firebaugh (2002):

$$ID(X) = \sum_{k=0}^{K-1} |\nu_k(X) - \frac{1}{K}| \tag{45}$$

where $\nu_k(X)$ is the relative density of k-agents:

$$\nu_k(X) \equiv \frac{u_k(X)}{\sum_{k'=0}^{K-1} u_{k'}(X)} \tag{46}$$

This gives the difference between the uniform distribution and the observed distribution: the larger this index, the greater the segregation at this location.

Second, we use the entropy (or Shannon information) associated with the income distribution at each location as a measure of segregation.

$$H(X) = -\sum_{k=0}^{K-1} \nu_k(X) \, \log \nu_k(X). \tag{47}$$

The larger the entropy, the weaker the segregation (with a maximum possible value of $H = \log K$ when all categories are equally present, and with $H = 0$ as the minimum value when a single category is found at the considered location). The entropy has unique mathematical properties that are appropriate

28

for a measure of heterogeneity[3] - for a discussion in the economics context, see Theil (1967).

In contrast to most of the literature, we do not consider averages of the indices over all the spatial locations (the social mix index $ID(X)$ and the entropy $H(X)$ are defined for each location X). To quantify the segregation as a function of the distance from the center, we consider the averages of the indices ID and H over locations situated at a same distance from the center. Figure 6 shows theses average values, comparing two different rates
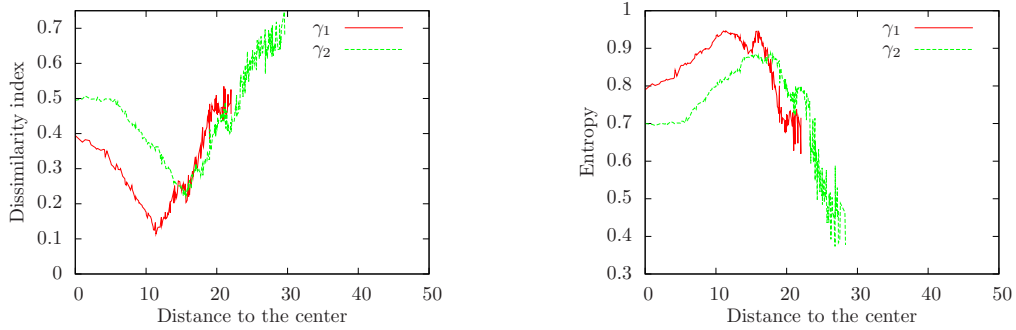
Figure 6: Social mix index ID (left) and Entropy $H(X)$ (right), as functions of the distance f,om the center, with $K = 10$, for two rates of newcomers $\gamma_1 = 1$ (that is $\Gamma/K = 1000$) and $\gamma_2 = 3$. Other parameters as for the other simulations.

of newcomers, $\gamma_1 < \gamma_2$.

The plots quantify what we inferred from Fig. 3: there is a zone near the city center with a moderate value of each index, a peripheral zone with a high index (low entropy), and an intermediate zone with a weak index (high entropy). This confirms the presence of an intermediate area of social mix. We also see that when we increase the rate of newcomers and hence the local demand, the index (resp. entropy) is on the whole higher (resp. lower), meaning there is a general decline in the social mix. The domain of social mix becomes smaller, and income segregation is thus amplified.

---

[3]In particular, it allows the coherent study of segregation at different scales *e.g.* if one wants to see how the measure changes when increasing the number of income categories.

6.5. Saturated regime

Increasing the incoming rate further eventually leads to a saturated state, i.e., with a situation of excess demand at some locations. An example is illustrated in Fig. 7. The parameters of the simulations are the same as before, except for the rate of newcomers which is higher. Comparing with
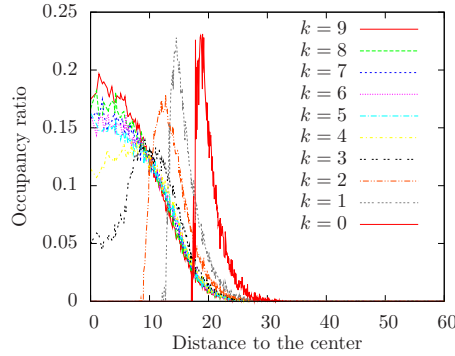

Figure 7: Saturated stationary regime: occupancy ratio per WTP versus the distance from the center (in arbitrary units). The parameters of the simulations are the same as in Fig. 3 except for $\Gamma/K = 2000$.

Fig. 3, we see that saturation increases the critical distances for the low WTP agents: the poorer agents move further away from the center.

The analytical expressions derived previously do not hold in this saturated case, for which the mathematical analysis requires a specific study beyond the scope of the present paper. Nevertheless, the simulations indicate that the distribution of the population exhibits similar features to those obtained in the non-saturated case, with increased income segregation.

7. Application to the Paris housing market

This section proposes a first comparison between the results of the model - analytically demonstrated and validated by simulations - and empirical observations. The aim here is to see how far the price distribution in Paris can be explained by the phenomena described above.

7.1. The empirical database

The main reliable information source on real-estate prices in Paris is the B.I.E.N. database, managed by the "Chambre des Notaires de Paris", which

records real-estate transactions for Paris and the Ile-de-France region. This database covers all categories of real estate, indicating, for each transaction, more than 100 different elements of information drawn from the associated legal documents. For our study, we extract the information we need, concerning the locations and transaction prices for flats. Since the data base contains information from 1990 until 2004, the long term dynamics may be studied: this will be the subject of future works, requiring additional features in the model, notably the introduction of a slow dynamics for the intrinsic attractiveness. In the present work we restrict the analysis to one particular year, namely 1994, for which the number of registered transactions is large enough. Moreover, this is an important year because it is situated before the fall in prices which affected Paris from 1995 to 2000.

During the year 1994, about 13000 transactions were recorded in Paris. The average price of a flat was 143300 euros, the standard deviation was around 90000 euros and the distribution was not normal (the normalized kurtosis coefficient is strictly positive). Since the database does not contain the incomes of the buyers, we use the transaction prices as a proxy for income distributions.

Fig. 8 and Fig. 9 give some idea of the spatial distribution of prices in Paris. Paris is not a very large city, and its geometrical form ('like a potato') endows the geographical center, identified as the historical center *Notre-Dame de Paris*, with particular importance. Looking at the map of the prices in Fig. 8, a first obvious fact is that prices are high in the center, with a marked trend of decreasing prices as one moves away from the center. Hence, the monocentric model could be seen as a first-order approximation of the Paris market. However, there are obvious deviations from this general trend. In particular, the "Rive Gauche" effect is important: at a similar distance from the centre, prices are higher on the left bank than on the right. In addition, prices are high in the 16th *arrondissement* on the very outskirts of Paris, creating a "hot-spot" of high prices far from the center. This map is obtained by averaging transaction prices over areas of 500 by 500 meters.

In Fig. 9, standard deviations are plotted with respect to the corresponding average prices used in Fig. 8, for each of the areas that contain enough transactions for the measure to be relevant. As expected, the standard deviation grows with the prices, suggesting a higher diversity when prices are high. However, it should be noted that the standard deviation is an *affine*
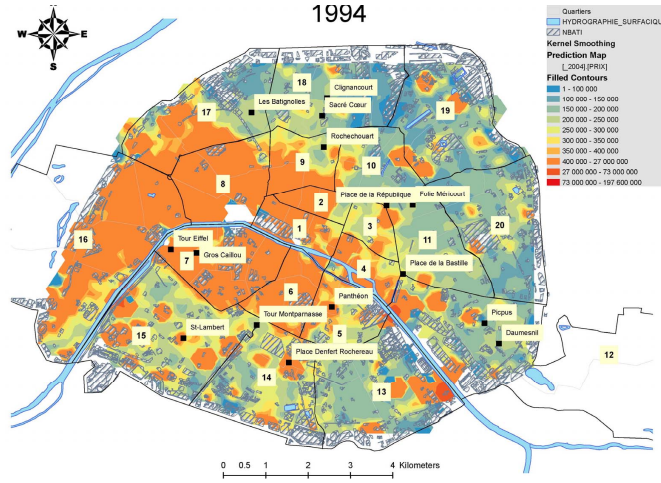
Figure 8: Transaction prices in Paris in the year 1994.

function of the average price.[4] This behaviour is remarkably similar to that obtained in the model, both qualitatively and quantitatively: compare the plot in Fig. 9 with Fig. 2, right (case of the modeled monocentric city), and the linear regression slopes in both cases.

7.2. Data-driven modeling

As explained above, a simple monocentric model would not allow to fit the price distribution in Paris. Here we simulate the Paris prices by considering a more specific model of the city, which combines a general preference for the center together with preferences for some local particularities.

First, instead of using a square lattice in the simulations, we use a stylized map of Paris. It consists of three concentric zones of radius R_1, R_2, R_3, each zone with 4,7 and 9 areas respectively representing the *arrondissements*, and the ratios of the simulated areas fitting the real ratios well. The stylized map of Paris is presented on the left of Fig. 10. Second, we assign to each one of these "*arrondissements*" a specific intrinsic attractiveness, decreasing with distance from the city center.

---

[4]This observation would have been less evident had we normalize the standard deviation by the mean.
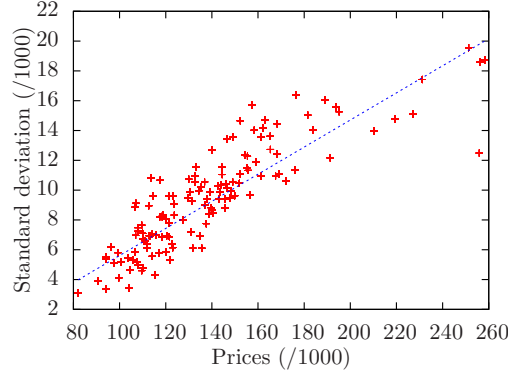
Figure 9: Paris data. Standard deviation versus transaction prices in thousands of euros. The straight line is the corresponding linear regression: slope 0.091, intercept −3.482, correlation coeff. −0.972.

For each *arrondissement* in the model, the intrinsic attractiveness parameters are scaled according to the average transaction price of the corresponding (real) *arrondissement* in the year 1994. More precisely, the maximum
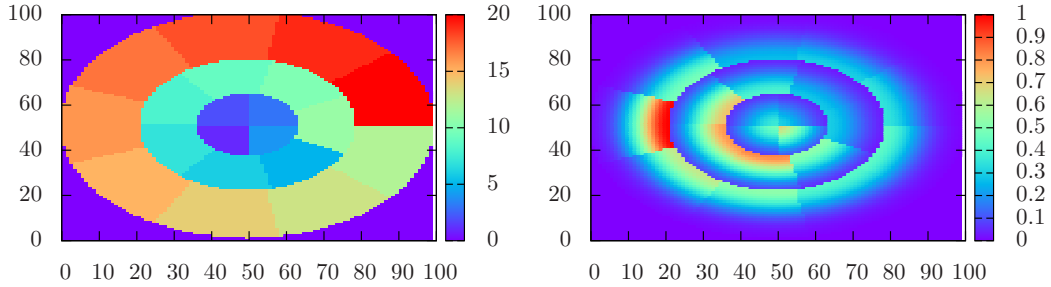

Figure 10: On the left: stylized map of Paris. On the right: map of the intrinsic attractiveness. $P_{ref}^1 = 151002$; $P_{ref}^2 = 119440$; $P_{ref}^3 = 149606$; $P_{ref}^4 = 142332$; $P_{ref}^5 = 157623$; $P_{ref}^6 = 188535$; $P_{ref}^7 = 190785$; $P_{ref}^8 = 203960$; $P_{ref}^9 = 146715$; $P_{ref}^{10} = 119515$; $P_{ref}^{11} = 121604$; $P_{ref}^{12} = 131973$; $P_{ref}^{13} = 134051$; $P_{ref}^{14} = 152234$; $P_{ref}^{15} = 162560$; $P_{ref}^{16} = 225985$; $P_{ref}^{17} = 150991$; $P_{ref}^{18} = 104776$; $P_{ref}^{19} = 115073$; $P_{ref}^{20} = 110912$; $R_1 = 13.6; R_2 = 28.9; R_3 = 50$.

attractiveness is taken as the ratio of the squared mean transaction price in the *arrondissement* studied to the squared maximum mean transaction price observed among all the *arrondissements*. Finally, the intrinsic attractiveness

$A^{0,a}$ of each *arrondissement* $a = 1, ..., 20$ is given by:

$$A^{0,a}(X) = \left(\frac{P^a_{ref}}{P^{max}_{ref}}\right)^2 exp\left(-\frac{(R^a - D(X))^2}{R^2}\right) \qquad (48)$$

where $D(X)$ is the distance from the center, R^a the shortest distance from the *arrondissement* to the center (which can take one of the three values, R_1, R_2, R_3), and P^a_{ref} the mean transaction price of the *arrondissement* calculated for the year 1994. The value of R is chosen to be of the same order of magnitude as the mean distance to the center of the lattice. The resulting map of the intrinsic attractiveness is presented in the right-hand panel of Fig. 10.

Note that firstly, the above form and parametrization of the attractiveness gives a stronger attractiveness to the more expensive *arrondissements*, even if they are far from the center of the city, and secondly, the chosen calibration only provides the overall price trend of the *arrondissements* - it does not determine how the prices vary within a given *arrondissement*.

For the year 1994, we order the transaction prices and divide the price domain into intervals such that each one includes about 1000 transaction prices. We postulate that prices belonging to a same price interval correspond to transactions made by individuals belonging to a same income class. In the model, each of these intervals is thus associated with one WTP category. This gives $K = 13$ WTP levels in the model.

According to data published by the INSEE (the French National Institute of Statistics and Economic Studies), there are about 375000 home-owners in Paris (we only consider main residences), and about 13000 transactions in one year. We therefore deduce a rate of house-moving of about 3.5%, hence $\alpha = 0.035$. Lastly, we take $\frac{\Gamma}{K} = 1000$, $\epsilon = 0.18$, $R = 10$, $\lambda = 0.005$, $\xi = 0.05$, $\Delta = 225000$, $\beta = 0.1$, $P^0 = 90000$, and $P_0 = 100000$. Assuming the price unit to be the euro, this set of parameter values allows us to obtain transaction prices of the same order as the empirical prices.

Arrondissement-specific segregation. When applied to the present case, the general theoretical results predict the emergence of segregation with critical distances specific to each *arrondissement*. The space is here decomposed into the union of $m = 20$ disjoint sub-spaces (*arrondissements*) $\Omega^a, a = 1, ..., m$, such that, on each of these sub-spaces, the intrinsic attractiveness is continuous and monotonically decreasing with the distance from the center of the city.

Each Ω_k (the subset of locations to which k-agents have access) is then the union of sub-sets Ω_k^a, where the agents have a sufficiently high WTP to exchange, $\Omega_k = \bigcup_{a=1}^{m} \Omega_k^a$ with

$$\Omega_k^a \equiv \{X \in \Omega^a | D(X) \geq d_c^a(k)\} \tag{49}$$

where $d_c^a(k)$ gives the critical distance specific to the sub-set Ω_k^a: for each space Ω^a, the $k-$agents can afford the goods in the locations $X \in \Omega^a$ for which $D(X) \geq d_c^a(k)$. Similarly, a WTP threshold $P_c^{*,a}$ can be defined for each *arrondissement*. The agents with a WTP above the threshold of a given *arrondissement* can buy a good in any location in this *arrondissement*. The WTP threshold, P_c^*, on the whole city is given by $P_c^* = max_a(P_c^{*,a})$.

Figure 11 shows two different measures of the segregation within the (real and the simulated) city of Paris. On the left, we present the degrees of exclusion for different levels of income as obtained from simulations with 50 income levels. Clearly, when the poorest people (income level 0) are not able to buy a dwelling in Paris, the richest ones (income level 45) can live just about wherever they want. It is also interesting to note that, even for a low level of income (income level 9 for example), there exist some opportunities to live in the expensive *arrondissements*. On the right-hand side of the figure, the social mix index (defined in Eq. 45 in Section 6.4) obtained in the simulations is compared with the one computed from real data, with transaction price being used as a proxy for income. The patterns are clearly similar, confirming the pertinence of our model for explaining prices in Paris.

In particular, the lowest levels of social diversity (corresponding to the highest values of the social mix index) are observed in the 8th,16th,18th,19th and 20th *arrondissements*, both in the simulations and in the data: the 8th and the 16th are the *arrondissements* where people's incomes are above the threshold, while the 18th, 19th and 20th *arrondissements* are the poorest.

8. Conclusion

Going beyond the simple Schelling segregation model, this paper studies social segregation through the spatial distribution of income resulting from dynamic price formation and agents' localization in a particular housing market model. In doing so, it actually introduces a new general framework for studying such markets and the resulting socio-spatial segregation. A specific feature of the model is the specification of the attractiveness of each location,
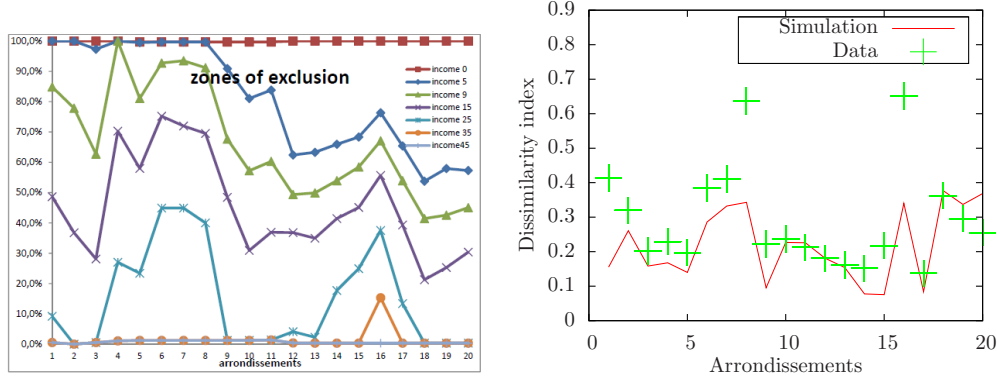
Figure 11: Left: for different WTP categories, the degree of exclusion from each arrondissement, measured by the percentage of arrondissement area where the considered category cannot buy a home (results from simulations with $K = 50$). Right: social mix index computed for the empirical data (points) and simulations (curve).

composed of both an intrinsic part and a subjective part that depends on the agents' social preferences and therefore evolves over time with the social characteristics of each neighborhood.

For the particular choices made in the present study - a preference for neighbors with similar or higher incomes -, the analysis of a non-saturated market yields the following main results. First, as expected, for a wide range of parameters, a socio-spatial segregation occurs, with richer people living in locations with higher intrinsic attractiveness. Second, such segregation occurs only if social influence is strong enough: on its own, the heterogeneity in attractiveness is not enough to provoke segregation. Third, and more original, whenever socio-spatial segregation occurs, there exists a large area of locations with intermediate values of intrinsic attractiveness where some social diversity is preserved. Finally, we prove that segregation is due to the emergence of income thresholds which divide the space between richer and poorer people. This endogenous threshold varies following the dynamic attractiveness of the rich location. This result could have important economic implications for the regulation of the housing market. For example, a policy aiming to boost demand (by subsidizing the poorer potential buyers) without constraining the supply side would lead to an increase in the distribution of sellers' prices and thus exacerbate the segregation.

Specifying the results for the case of a monocentric city (the closer to the center, the higher the intrinsic attractiveness), all the results can be

36

simply stated in terms of properties depending on the distance from the center. Whenever social influence is strong enough, there exists a critical distance isolating a central zone occupied by the richest people. Within this zone, there is no segregation between these rich agents: some are richer than others, and this shows up in the high variance in transaction prices. Poorer people find themselves on the outskirts, where social diversity is the weakest, leading to a low variance in transaction prices. At intermediate distances from the center, there is an area of moderate social mix.

Looking at the empirical distribution of prices in Paris for the year 1994, we observe that, as a first-order approximation, the simple Alonso model (prices higher at the center) is a good description: higher prices are indeed concentrated around Notre Dame cathedral. The behavior of the mean and standard deviation of price transactions with the distance from the center are fairly well reproduced by the model. However, this monocentric characterization of the housing market in Paris is only correct on average; it misses certain particularities, such as the difference between prices on the left bank (higher) and the right bank (lower). A better representation of the Paris housing market is then proposed by considering an intrinsic attractiveness specific to each of the 20 *arrondissements* of Paris. Taking the spatial distribution of transaction prices as a proxy for the spatial distribution of income, the results compare favorably with empirical data in the distinction between *arrondissements*: a low level of social diversity corresponds to *arrondissements* with either the highest or the lowest average prices, and a lower level segregation corresponds to *arrondissements* with intermediate average prices.

The mathematical analysis is limited to a non-saturated regime, as explained in Section 4.2.1. On the technical side, this allows for a complete mathematical characterization of the stationary states. On the qualitative side, it demonstrates that heterogeneity in intrinsic attractiveness alone is not enough to provoke segregation, which only emerges when the social influence is strong enough. Analytical study of the saturated regime - briefly explored numerically in Section 6.5 - will require a much more complex analysis, beyond the scope of the present paper. The present model provides a framework that is general enough to take into account other components of social preferences such as ethnic features or other population characteristics.

This study also allows for further investigations. Concerning social preferences, alternative hypotheses such as agents preferring to live in more working-class areas have yet to be considered. Preliminary results indicate that a preference to live with poorer neighbors has opposite effects to

37

those shown in the present paper, working against the emergence of segregation. Concerning the dynamics of the model, the long-term evolution will be studied by introducing a slow dynamics for the intrinsic attractiveness, corresponding to the effect of, e.g., investments in new amenities, or, on the contrary, the deterioration of a neighborhood. This should also allow for the analysis of the emergence of polycentric cities. A further step would consist in analyzing the conditions on the dynamics of subjective attractiveness for the emergence of "hot spots" - local areas of high prices unrelated to the level of intrinsic attractiveness. This could accompany extensions to the model, notably in the line of Short et al. (2008) and Berestycki & Nadal (2010), adding a diffusion term to the attractiveness dynamics: high attractiveness of a location is expected to have some positive influence on the attractiveness of nearby locations. It is this kind of extension that is likely to lead to dynamical instabilities, with the emergence of hot spots of high prices, that could, for example, explain changes in the ranking of *arrondissements* by property prices over the years. The study of these dynamical instabilities could provide new insights into the mechanisms behind gentrification.

References

Alesina, A., & Zhuravskayay, E. (2011). Segregation and the quality of government in a cross-section of countries. *American Economic Review*, *101*, 1872–1911.

Alonso, W. (1964). *Location and land use: toward a general theory of land rent*. Harvard University Press, Cambridge.

Ballester, C., & Vorsatz, M. (2010). Random-walk-based segregation measures. *FEDEA WP 2010-30*, .

Baltagi, B., & Bresson, G. (2010). Maximum likelihood estimation and lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: an application to hedonic housing prices in paris. *Journal of Urban Economics*, *forthcoming*.

Berestycki, H., & Nadal, J.-P. (2010). Self-organized critical hot spots of criminal activity. *European Journal of Applied Mathematics*, *21 Issue 4-5*, 371–399.

Berg, L. (2002). Prices on the second-hand market for the swedish family houses: Correlation, causation and determinants. *European Journal of housing policy*, *2*, 1–24.

Bernard, S., & Willer, R. (2007). A wealth and status based-model of residential segregation. *The Journal of Mathematical Sociology*, *31*, 149.

Bono, P., Gravel, N., & Trannoy, A. (2007). L'importance de la localisation dans la valorisation des quartiers marseillais. *Economie Publique/Public Economics*, *20*, 169–205.

Brueckner, J., Thisse, J., & Zenou, Y. (1999). Why is central paris rich and downtown detroit poor? - an analysis of consumer demand for location. *European Economic Review*, *vol. 43, number 1*, pp. 91–107(17).

Clapp, J., Dolde, W., & Tirtiroglu, D. (1995). Imperfect information and investor inferences from housing price dynamics. *Real estate economics*, *23*, 239–269.

Clark, A., & Oswald, A. (2002). The international library of critical writings in economics: Happiness in economics. chapter Unhappiness and Unemployment. Edward Elgar.

Cutler, D., Glaeser, E., & Jacob Vigdor (2008). When are ghettos bad? lessons from immigrant segregation in the united states. *Journal of Urban Economics*, *63*, 759–774.

Cutler, D., Glaeser, E., & Vigdor, J. (1999). The rise and decline of the american ghetto. *Journal of Political Economy*, *107*, 455–506.

Echenique, F., & Fryer, R. (2007). A measure of segregation based on social interactions. *Quarterly Journal of Economics*, *122, Issue 2*, 441–485.

Figlio, D., & Lucas, M. (2004). What's in a grade? school report cards and the housing market. *American Economic Review*, *94:3*, 591–604.

Fossett, M. A., & Senft, R. L. (2003). *Status and Housing Systems: Rationale and Methods for Generating Status Score Distributions for Ethnic Groups and Area Stratifiation in Housing Quality in the SimSeg Simulation Model*. Technical Report.

Gobillon, L., Selod, H., & Zenou, Y. (2007). The mechanisms of spatial mismatch. *Urban Studies*, *44*, 2401–2427.

Goyal, S., & Ghiglino, C. (2010). Keeping up with the neighbors: Social interaction in a market economy. *Journal of European Economic Association*, .

Holly, S., Pesaran, M., & Yamagata, T. (2010). The spatial and temporal diffusion of house prices in the uk. *Journal of Urban Economics*, *forthcoming*.

Ioannides, Y. M., & Zabel, J. E. (2003). Neighbourhood effects and housing demand. *Journal of Applied Econometrics. Chichester:*, *18 (5)*, pp 563–583.

Jenks, C., & Meyer, S. (1990). in inner city poverty in the united states. chapter Residential Segregation, Job Proximity and Black Job Opportunity. National AcademicPress, Washington, DC.

Luttmer, E. (2005). Neighbors as negatives: Relative earnings and well-being,. *Quarterly Journal of Economics*, *120,3*, 963–1002.

Meen, G. (1996). Spatial agregation , spatial dependence and predictability in the u.k. housing market. *Housing studies*, *vol 11*, pp345–372.

Oikarinen, E. (2005). The diffusion of housing price movements from centre to surrounding areas. Discussion paper n 979.

Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology*, *32*, 3367.

Rosen, S. (1974). Hedonic prices and implicit markets. *Journal of Political Economy,*, *82*, pp 34–55.

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

Seo, Y., & Simons, R. A. (2009). The effect of school quality on residential sales price. *Journal of real Estate Research*, *3 1: 3*, 307–327.

Short, M. B., D'Orsogna, M. R., Pasour, V. B., Tita, G. E., Brantingham, P. J., Bertozzi, A. L., & Chayes, L. B. (2008). A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, *18*, 1249.

Theil, H. (1967). *Economics and information theory*. North-Holland.

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64: 5*, pp 41626.

Zhang, J. (2004). A dynamic model of residential segregation. *Journal of Mathematical Sociology*, *28*, 147.

A. Appendix

A.1. Proof of the Proposition, Section 4.2.3

We give here the proof by recurrence on k.

Proof: case $k \geq \bar{k}$.

At equilibrium, $u_k^*(X) = \frac{v_k^*(X)}{\alpha} = \frac{\rho_k^*(X)}{\alpha}$ which gives:

$$u_k^*(X) = \frac{\gamma}{K\alpha} \frac{A_k^*(X)}{(1/L^2) Z_k} \tag{A-1}$$

with

$$Z_k \equiv \sum_{X \in \Omega} A_k^*(X) \tag{A-2}$$

From the fixed point equations, we get, for any $k \in \{0, ..., K-1\}$,

$$Z_k = Z^0 + \frac{\epsilon}{\omega} \sum_{X \in \Omega} \sum_{k' \geq k} v_{k'}(X) = Z^0 + L^2 \frac{\epsilon\gamma}{\omega K} (K - k) \tag{A-3}$$

In particular, for the housed $(K-1)$-agents (highest WTP):

$$u_{K-1}^*(X) = \frac{\gamma}{K\alpha} \frac{A^0(X) + \frac{\epsilon\alpha}{\omega} u_{K-1}^*(X)}{(1/L^2)Z^0 + \frac{\epsilon}{\omega}\frac{\gamma}{K}} \tag{A-4}$$

Solving this equation leads to:

$$u_{K-1}^*(X) = \frac{\gamma}{K\alpha} \frac{A^0(X)}{(1/L^2)Z^0} \tag{A-5}$$

Now assume that (30) is true for $k + 1 \leq k' \leq K - 1$ with $k \geq \bar{k}$. Then we have:

$$u_k^*(X) = \frac{\gamma}{K\alpha} \frac{A^0(X) + \frac{\epsilon\alpha}{\omega} u_k^*(X) + (K - k - 1)\frac{\epsilon\gamma}{\omega K}\frac{A^0(X)}{(1/L^2)Z^0}}{(1/L^2)Z^0 + \frac{\epsilon\gamma}{\omega K}(K - k)} \tag{A-6}$$

which gives $u_k^*(X) = \frac{\gamma}{K\alpha} \frac{A^0(X)}{(1/L^2)Z^0}$. Hence the proposition is proved by recurrence for all $k \geq \bar{k}$.

Proof: case $k < \bar{k}$ (assuming $\bar{k} > 0$).

From (24), the density of potential buyers in a location X (outsiders) is given by:

$$\rho_k^*(X) = \frac{A_k^*(X)}{\sum_{X' \in \Omega} A_k^*(X')} \sum_{X' \in \Omega} \rho_k^*(X') \tag{A-7}$$

42

For $k < \bar{k}$, $v_k^*(X) = 0$ for $X \in \Omega \setminus \Omega_k$, and $\bar{v}_k^*(X) = 0$ for $X \in \Omega_k$. Hence, since $\rho_k^*(X) = v_k^*(X) + \bar{v}_k^*(X)$, we have:

$$\sum_{X \in \Omega_k} \rho_k^*(X) = \sum_{X \in \Omega_k} v_k^*(X) = \sum_{X \in \Omega} v_k^*(X) \tag{A-8}$$

As we have seen, in the stationary state the total number of k-agents on the lattice is $\frac{\gamma L^2}{\alpha K}$, and $v_k^*(X) = \alpha u_k^*(X)$, so that finally:

$$\sum_{X \in \Omega_k} \rho_k^*(X) = L^2 \frac{\gamma}{K} \tag{A-9}$$

Then, summing Eq. (A-7) over X in Ω_k, we get:

$$\rho_k^*(X) = \frac{\gamma L^2}{K} \frac{A_k^*(X)}{\sum_{X' \in \Omega_k} A_k^*(X')} \tag{A-10}$$

Let us now consider $k = \bar{k} - 1$, that is the highest WTP category for which Ω_k is different from Ω, i.e., the highest $k-$WTP which does not allow to buy a good anywhere. The agents with a WTP greater than P_k are distributed according to the expression (30). Thus, if $X \in \Omega_k$:

$$\rho_k^*(X) = v_k^*(X) = \alpha u_k^*(X) = \frac{\gamma L^2}{K} \frac{A_k^*(X)}{\sum_{X' \in \Omega_k} A_k^*(X')} \tag{A-11}$$

$$u_k^*(X) = \frac{\gamma L^2}{K \alpha} \frac{A^0(X) + \frac{\epsilon \alpha}{\omega}[(K - k)\frac{L^2 A^0(X)}{\sum_{X' \in \Omega} A^0(X')} + u_k^*(X)]}{\sum_{X' \in \Omega_k}[A^0(X')] + \frac{\epsilon}{\omega}[(K - k)\frac{\gamma L^2}{K} \frac{\sum_{X' \in \Omega_k} A^0(X')}{\sum_{X' \in \Omega} A^0(X')} + \frac{\gamma L^2}{K}]} \tag{A-12}$$

The equation (A-12) leads to:

$$u_k^*(X) = \frac{\gamma L^2}{K \alpha} \frac{A^0(X)}{\sum_{X' \in \Omega_k} A^0(X')}. \tag{A-13}$$

By recurrence, we can generalize to any $k < \bar{k}$:

$$v_k^*(X) = \alpha u_k^*(X) = \frac{\gamma L^2}{K} \frac{A^0(X)}{\sum_{X' \in \Omega_k} A^0(X')} \text{ if } X \in \Omega_k \tag{A-14}$$

$$= 0 \text{ otherwise} \tag{A-15}$$

Note that the above proof can easily be shown to apply to the case where agents weight higher categories differently, as in Eq. (16).

A.2. Computing the mean attractiveness

Assuming $\bar{k} \in \{1, ..., K-1\}$, we obtain the expression of $\bar{A}^*(X)$ as follows. From the stationary equation for A_k^*, Eq.(23), and from the Proposition stated in Section 4.2.3, we get, in a non-saturated equilibrium:

$$A_k^*(X) = A^0(X) + \frac{\epsilon\gamma}{\omega K} \sum_{k' \geq \max(k, k_c(X))} \frac{A^0(X)}{\langle A^0 \rangle_{k'}} \tag{A-16}$$

$k_c(X)$ is the smallest value of k such that the k-agents are present on X.

First consider $k \geq \bar{k}$. In this case, $\langle A^0 \rangle_k = \langle A^0 \rangle$ so that, from (A-16), we get: $A_k^*(X) = A^0(X) + \frac{\eta \, A^0(X)}{2}[1 - \frac{k}{K}]$, where $\eta = \frac{\epsilon\gamma}{\omega \langle A^0 \rangle}$, as defined in (14). At locations X where $k_c(X) = \bar{k}$, for $k < \bar{k}$, we obtain from (A-16) $A_k^*(X) = A^0(X) + \frac{\eta \, A^0(X)}{2}[1 - \frac{\bar{k}}{K}]$. As a result, the average value $\bar{A}^*(X) = \sum_0^{K-1} A_k^*(X)$ is

$$\bar{A}^*(X) = A^0(X) + \frac{\eta \, A^0(X)}{2}[\frac{K+1}{K} - \frac{(\bar{k}+1)\bar{k}}{K^2}] \tag{A-17}$$

A.3. The residential city

The densities obtained with the partial differential equations approach can take arbitrarily small values. In the agent-based version, the number of agents at a given location obviously only take integer values: we can expect a good comparison considering that the actual density is zero whenever $a^2 \, u_k(X)$ is less than 1. For a concrete example, consider a monocentric city with $A^0(X) = A^0_{max} \exp(-D(X)^2/R^2)$, with $R < \mathcal{D}$. The diameter $\mathcal{D}_{eff}$ of the space of residential locations is obtained by writing $a^2 \sum_k u_k(X) = 1$. Taking as an approximation the expression (30) for every location, and setting a continuous limit by assuming a small, we obtain how $\mathcal{D}_{eff}$ (or equivalently, the number of active sites, $L_{eff} = \mathcal{D}_{eff}/a$) scales with the model parameters:

$$\mathcal{D}_{eff} = R\sqrt{\ln(\frac{\mathcal{D}}{R}\sqrt{\frac{\Gamma}{2\pi\alpha}})} \tag{A-18}$$

Note that the demand pressure, expressed by the ratio $\frac{\Gamma}{\alpha}$, only contributes through a logarithmic term. For the parameter values (notably $\mathcal{D} = aL = 100$, $R = 10$) used in the numerical simulations in Section 6, the above formula gives $\mathcal{D}_{eff} \simeq 28$, in keeping with what can be seen in, e.g., Fig. 3.

A.4. Extension to an arbitrary WTP distribution

In the model we have assumed that the K values of WTP are uniformly distributed among the agents in the external reservoir. One may want to start with a particular probability distribution function (PDF) of the reservation price, corresponding to the empirical income distribution of a given population, for example. Let $Q(P)$ be this PDF, with support on $[P_0, P_0 + \Delta]$. We can construct the discrete set of K levels from $Q(P)$ by requiring the same density of agents in any one of these levels. The $\{P_k, k = 0, ..., K - 1\}$ are thus defined by:

$$\int_{P_0}^{P_k} Q(p)dp \;=\; \frac{k}{K} \tag{A-19}$$

for $k = 1, ..., K$ (with $P_K \equiv P_0 + \Delta$). Of particular interest is the large K limit. In this case, we have $P_k = P(k/K)$, the function $P(z)$ on $z \in [0, 1]$ being given as solution of $\int_{P_0}^{P(z)} Q(p)dp \;=\; z$, or, equivalently by:

$$Q(P(z))\, P'(z) = 1 \tag{A-20}$$

where $P'(z) = dP/dz$, with $P(0) = P_0$ and $P(1) = P_0 + \Delta$.

Formally, all the analysis done in this paper for the uniform case, (1), can be applied to an arbitrary WTP function. In particular, the critical WTP threshold is here given by $P_c^* = P(z_c^*)$, z_c^* being obtained by solving the equation

$$\frac{1}{\tilde{\xi}}\, \log \frac{P(z_c^*)}{P^0} \;-\; 1 \;=\; \frac{\eta}{2}\left(1 - (z_c^*)^2\right). \tag{A-21}$$

The above results can be generalized to the case of agents weighting higher categories differently, as in Eq. (16), with $w_{k'-k} = w(\frac{k'-k}{K})$. We get:

$$\frac{1}{\tilde{\xi}}\, \log \frac{P(z_c^*)}{P^0} \;-\; 1 \;=\; \eta \int_{z_c^*}^{1} dt' \int_{0}^{t'} dt\; w(t' - t). \tag{A-22}$$

The critical value η_c of η is obtained by setting $z_c^* = 0$ in the above equation.