J Geograph Syst (2002) 4:251–274

Journal of Geographical Systems

Toward micro-scale spatial modeling of gentrification

David O'Sullivan

Department of Geography, The Pennsylvania State University, University Park, PA 16802, USA (e-mail: dbo2@psu.edu)

Received: 1 November 2001 / Accepted: 12 April 2002

Abstract. A simple preliminary model of gentrification is presented. The model is based on an irregular cellular automaton architecture drawing on the concept of proximal space, which is well suited to the spatial externalities present in housing markets at the local scale. The rent gap hypothesis on which the model's cell transition rules are based is discussed. The model's transition rules are described in detail. Practical difficulties in configuring and initializing the model are described and its typical behavior reported. Prospects for further development of the model are discussed. The current model structure, while inadequate, is well suited to further elaboration and the incorporation of other interesting and relevant effects.

Key words: gentrification, spatial models, micro-simulation, cellular automata, urban simulation

JEL classification: C51, C63, 021, R00

1 Introduction

This paper reports on progress toward the development of a preliminary, theoretically-founded model of gentrification. Gentrification has been a major theme in urban geography for at least three decades, its significance recently confirmed (in the academy at any rate) by the appearance of well-received books from two of the more prominent names in the literature (Ley 1996, Smith 1996). Nevertheless, the phenomenon remains largely untouched by model- or simulation-based approaches. Perhaps it is the wider political, social and cultural, aspects of gentrification that have immunized it

Town centres data were made available by kind permission of the UK Department of the Environment, Transport and the Regions.

against the efforts of the more methodologically inclined in geography. Of course, the explanation for such lack of interest may be more prosaic – simply that the level of detail required to model gentrification processes has been beyond anything that could readily be implemented. This is beginning to change, a fact attested to by recent contributions from Portugali et al. (1997) and Bernard (1999), and also by earlier large-scale simulations from Kain and Apgar (1985).

It must be emphasized from the outset that the model presented in this paper is preliminary and exploratory in nature, and is intended primarily as a starting point for exploration of the gentrification process using contemporary modeling techniques and ideas. The model described is certainly *not* a 'finished' article. Its major aim is to establish a viable architecture that can, in time, be developed into a more complete simulation. In this context it is important for the reader to concentrate on the description of the overall model structure presented in Sect. 3, rather than on the specifics of the transition rules presented in Sect. 4. While the transition rules are certainly important, at the present stage of development it is more important to establish that within this architectural framework, a set of transition rules can be implemented, which lead to interesting outcomes, and thus to demonstrate the promise of the present approach.

In the next section parts of the urban geography literature on gentrification relevant to the approach adopted in this paper are discussed, particularly Smith's rent gap hypothesis (Smith 1979). In Sect. 3 issues raised by this discussion are identified and the overall model structure is described. A detailed description of the 'rules' that govern the model's dynamics is presented in Sect. 4. Running the model is decribed in Sect. 5, in terms of both the data required to initialize the model, and in terms of typical output. Avenues opened up for further exploration by this framework are considered in concluding remarks in Sect. 6.

2 Gentrification and the rent gap hypothesis

Since the term's first appearance (Glass 1964) gentrification has occupied an important position in the urban geography literature (Hamnett 1991, Smith 1996; Redfern 1997; Lees 2000; provide extensive bibliographies). In particular, the introduction of the *rent gap hypothesis* (Smith 1979) was controversial, leading to bad-tempered exchanges between proponents of various approaches (Ley 1987; Smith 1987a, 1992; Bourassa 1990; Badcock 1990; Hamnett 1991, 1992; Clark 1992). More recently the dust has settled, and a more pluralistic discussion has ensued (Rose 1984; Bondi 1991, 1999; Warde 1991; Bridge 1994; Shaw 2000; Robson and Butler 2001). A recent paper by Lees (2000) exemplifies this trend. This article could not possibly examine thoroughly all of these debates. Instead, I focus on the rent gap hypothesis as a relatively simple conceptual model whose spatial aspects can be explored in a dynamic model.

The essentials of the rent gap hypothesis are easily explained. The thrust of Smith's (1979) paper is that whereas previous work had emphasized 'demand-side' explanations of gentrification, "[a] broader theory [...] must take the role of producers as well as consumers into account." Furthermore "when this is done, it appears that the needs of production – in particular the

need to earn a profit – are a more decisive initiative behind gentrification than consumer preference." (Smith 1979, page 540) In short:

"Consumer sovereignty explanations took for granted the availability of areas ripe for gentrification when this was precisely what had to be explained." (Smith 1979, pages 540–541)

Smith's explanation of the genesis of 'gentrifiable' areas lies in the dynamics of residential property values. The price of a residential site is made up of two components: house value and capitalized ground rent. House value is the value of the raw materials and labor used in construction, minus subsequent depreciation due to wear and tear, and plus any improvements. Capitalized ground rent is the "actual quantity of ground rent that is appropriated by the landowner, given the present land use" (Smith 1979, page 543).

Since rent is a flow and not an amount, this value is the expected discounted cash flow from rents associated with the site, assuming that the rent remains at or around the current level. Potential ground rent is the rent that might be realized, given the site location, under its "highest and best use" (Smith 1979, page 543). Potential ground rent must also be regarded as a discounted cash flow. Further, we must think of the house as providing a flow of services (shelter, heat, light, amenity), to clarify the rent concept in the owner-occupier case. The fact that the sale price of a residential site is equal to the house value plus the capitalized ground rent means that an owner-occupier stands to benefit from any improvements that raise the capitalized ground rent of the property, since these will be recouped when the property is eventually sold.

Smith relates these concepts to the life cycle of residential properties in inner urban neighborhoods, illustrated in Fig. 1. When a building is first constructed, it is well suited to its site, and in good condition, so that the owner can maximize the rental income. Capitalized ground rent matches potential ground rent and the *rent gap* is zero. Over time it is likely that house value will change. This depends on the extent to which regular maintenance tasks are carried out, and also whether 'upgrades' in the form of additions,

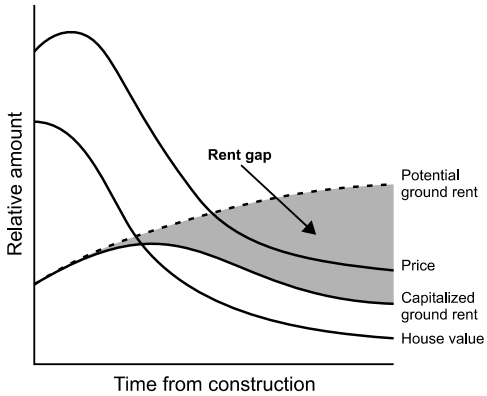

Fig. 1. The rent gap hypothesis. Changes in rents and values in an inner city neighborhood, with the rent gap shown as a shaded region (*based on*: Smith 1979)

re-plumbing, re-wiring and so forth, are performed to keep the property up to date. Smith argues for a Marxian theory of value, where the value of a property is dependent on the amount of raw materials and labor power in the initial construction, combined with subsequent injections of raw materials and labor power. The mechanism by which property value conceived in this way falls is technological progress in the construction industry, which means that new houses provide the same amenity, for a lower input of raw materials and labor power.

Notwithstanding the complexities of any theory of value, more rapid depreciation inevitably occurs in some neighborhoods, because more affluent households take up the better opportunities offered by modern housing stock in the suburbs. Such neighborhoods are likely to see a move toward higher rates of tenancy. It is a rational economic response for landlords to under-maintain their properties in these circumstances. Such landlord action further reduces the value of the stock, prompts the out-migration of more affluent households, and sees the neighborhood's decline accelerate further. Other institutional practices may further accelerate disinvestment. Financial institutions may be reluctant to provide capital for owners in a neighborhood at this stage in the cycle. Whether deliberate *redlining* policies are pursued or not, the effect is further decline, which may prompt landlords to subdivide properties to increase returns, since few other options are open to them. The final stage of this process is abandonment by landlords who can no longer profitably fill buildings. In the United States, this stage has seen catastrophic destruction of urban fabric, sometimes by deliberate arson aimed at recovering at least some of the property value from insurance (Tabb 1982).

Smith argues that the passage of a neighborhood through this life cycle is manifest as a substantial rent gap between potential and capitalized ground rents. The rent gap represents an investment opportunity for would-be 'gentrifiers', thus setting the stage for gentrification. Since any gentrifier, whether owner-occupier or property developer, requires finance capital to proceed, gentrification can thus be conceived as "a back to the city movement by capital, not people" – to borrow the subtitle of Smith's paper.

Smith claims that the rent gap is "the essential centerpiece to any theory of gentrification" (Smith 1987b, page 165). While his contribution is undeniably important, efforts to verify the rent gap hypothesis empirically have met with mixed success. Studies by Badcock (1989) and Clark (1987) broadly support Smith's hypothesis. The most recent study by Yung and King (1998) provides more equivocal support. An investigation by David Ley (1986) suggested that other factors were more relevant, although this work has been heavily criticized (Smith 1987a; Clark 1988; Bourassa 1993). A major difficulty for all such research is that concepts in Smith's theory are not directly available in any public records, since they are only indirectly related to observable quantities – such as sale prices or property taxes.

3 Developing a spatial model of gentrification

Few theories – if any – in the social sciences enjoy unequivocal empirical support, and many would argue that this is to be expected, given the complex and open nature of social phenomena (Sayer 1992). The rent gap hypothesis is an important contribution to the geography literature, whatever its

limitations. In the present paper, it is a point of departure for the development of a model that examines underdeveloped spatial aspects of the hypothesis, highlighted by Hammel (1999). Hammel argues, agreeing with Badcock (1990), that, in practice, land valuations are based on sales of comparable parcels nearby. Thus potential ground rent is a *neighborhood scale* phenomenon, rather than an intrinsic characteristic of individual parcels. Capitalized ground rent, on the other hand is strongly dependent on individual building characteristics at the *parcel scale*. This refinement of the rent gap concept is embedded in the present model in the relationship between 'local' (parcel scale) and 'global' (neighborhood scale) model parameters (see Sect. 4.2 below).

This argument may also be related to criticism of Smith's theory (Bourassa 1993), that it is fine to explain *how the opportunity for reinvestment occurs*, but the rent gap fails to explain *why reinvestment becomes profitable*: what is it that changes perceptions of a declining neighborhood so that reinvestment becomes viable? After all, neighborhoods where gentrification occurs have not suddenly been relocated nearer to the downtown – they have always been there! The origin of perceived changes in value is critical. The rent gap hypothesis, by concentrating on a property's value as 'intrinsic' and related to the physical state of a property, notwithstanding the relational character of rent, fails to identify why it suddenly becomes apparent to potential gentrifiers or finance companies that a profitable rent gap exists. The answer is, of course, a spatial one and lies in the realtors' mantra: 'location, location, location'. The present model develops this perspective, by adopting a relational or *proximal* view of space where nearby locations are a key feature of each location's site, and thus the basis for decisions about its future development.

3.1 Proximal space and graph-based cellular automata

An invaluable distinction between the notions of *site* and *situation* is made by Helen Couclelis. She identifies site with simple geographical location, whereas situation includes the proximal properties of the site – put simply the other locations nearby (Couclelis 1991). The latter is a convenient way of thinking about realtors' 'location, location, location'. The concept of situation has been formalized in *proximal space* (Couclelis 1997) and an accompanying *Geo-Algebra* (Takeyama and Couclelis 1997). This approach also has roots in proposals for proximal databases to enable exploration of spatial dependence (see Getis 1994). More recently, an alternative representation of proximal space has been suggested, the *graph-based cellular automaton* (O'Sullivan 2001b), which is now described.

In a cellular automaton (CA) space is partitioned into a lattice of identical locations, usually, but not necessarily, a grid of square cells. Each cell in the lattice may be in one of a number of available discrete states – in the simplest case 'on' or 'off', but in urban simulations a variety of states are possible: undeveloped, residential at low and high densities, commercial, light and heavy industrial, and so on. Model dynamics are encoded as a set of *rules* governing cell state transitions. The model's global state evolves in a series of discrete time steps, when each cell's neighbors in the lattice are examined and each cell's next state is determined in accordance with the rules. For example,

a rule might say that a cell in the 'on' state with more than three neighbors also in the 'on' state will switch to the 'off' state at the next time step. In an urban simulation, this formulation is more complex, with rules such as "an undeveloped cell with neighboring residential and retail cells will become low density residential". Usually, rules have a stochastic element to escape the determinism inherent in simple CA. The basic CA formulation with modifications has been widely applied to urban geography (recently examples include Clarke et al. 1997; Batty et al. 1999; Webster and Wu 1999a,b; Li and Yeh 2000; Ward et al. 2000; White and Engelen 2000).

In a graph-based cellular automaton, one aspect of the CA formalism is changed: cells are no longer identical and located at evenly spaced sites across a lattice. Instead, cells may represent *any* geographical entities of interest. Furthermore, the neighborhood relations defining which cells affect the development of which others are represented as an irregular lattice, that is, as a *graph* or network. This approach immediately enables application of the CA approach to simulation of urban settings at a much finer resolution than hitherto, because cells may represent entities that are clearly related to one another in complex spatial patterns, not adequately represented by any regular lattice (O'Sullivan and Torrens 2001). How this works in practice is shown in Fig. 2, where a fragment of the present model lattice is shown. Each cell in the model represents a building or parcel. Edges connecting adjacent cells are shown, and these define the cell neighborhoods considered during application of the model rules. Thus the proximal properties of each site in the model (its situation) are immediately accessible via its neighborhood in the graph. This may be regarded as an attempt to operationalize the importance of 'location, location, location'.

Naturally, the question arises of what the graph edges between buildings *represent*. The neighbors of a building in the model are those held to have an effect on that location's value. In light of the preceding discussion it is clear that the graph structure therefore represents generally perceived relations between buildings in the urban space, since these are precisely the relations

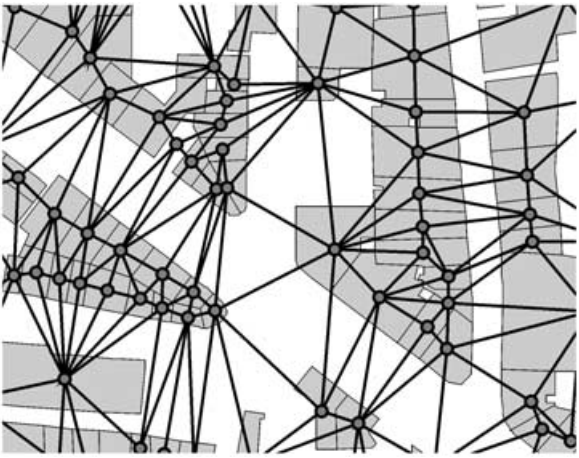

Fig. 2. A fragment of the model's spatial structure. Buildings are represented as vertices in a graph

that will affect a property's perceived value – its realtor's location, or property assessor's value.

The graph edges in Fig. 2 are based on a Delaunay triangulation of building centroids. This triangulation is the geometric dual of a proximal or Voronoi polygon tesselation of the study area, where each building centroid is associated with a region of space that is closer to it than to any other building centroid in the space. The Delaunay triangulation is a convenient basis for constructing the irregular spatial relations required by the model. This concept of neighborhood finds support at a general level in the literature (Gold 1992; Edwards 1993), and is also relatively simple to compute (Okabe et al. 2000). It also appears in recent contributions proposing Voronoi-based cellular automaton models (Semboloni 2000; Shi and Pang 2000). In fact, the graph-based CA structure presented here may be considered a generalization of the Voronoi-CA idea.

Clearly, the choice of any particular graph structure is open to question, and is effectively a modeling assumption. Other approaches to the definition of neighborhoods are possible, such as including as each building's neighbors all those within some specified distance, or the k nearest other buildings. These are both well-established notions of neighborhood in quantitative geography, and are commonly used in building spatial weights matrices for use in autocorrelation studies or similar (see, for example, Anselin 1995). Less obvious approaches, sensitive to the complex geometry of urban space might be based on mutual visibility between buildings, or on shared street addresses – where buildings on the same street, or segment of street are considered neighbors. Some of these possibilities have been considered in more detail elsewhere (O'Sullivan 2000).

Whatever method is used to construct cell neighborhoods or, equivalently, the graph structure of the model, the important point is that it is now impossible to think of assessing the merits of an individual building without reference to the attributes of neighboring buildings. All change in the model takes place at the level of cell neighborhoods, and therefore takes account of the situational properties of a building.

4 A graph-based cellular automaton model of gentrification

4.1 Cell state variables

In the model, cells representing individual buildings may be in one of four discrete states at any time step, as summarized in Table 1.

A property may be either owner-occupied or landlord-owned. An owner-occupied location not currently for sale is in the NOT FOR SALE state.

Table 1. The four discrete states allowed at each location

State name	Description
NOT FOR SALE	Owner occupied
FOR SALE	For sale irrespective of current tenure type
SEEKING TENANTS	Owned by landlord and currently vacant
RENTED	Owned by landlord and currently let

Current owners who decide to leave the neighborhood determine a suitable price for the property and its state changes to FOR SALE. A property FOR SALE may be bought by new owner-occupiers, when its state reverts back to NOT FOR SALE. Alternatively it is bought by a landlord, who starts to look for a tenant, and the SEEKING TENANT state is entered. A property may remain in this state for several time periods, when the landlord may decide to resell, so reverting to the FOR SALE state. Alternatively, if a tenant is found, the location enters the RENTED state. From the RENTED state, the location may revert to either the FOR SALE state if the landlord decides to sell, or to the SEEKING TENANT state if the tenant decides to move out. The allowed transitions between discrete model states are summarized in Fig. 3.

Underlying discrete cell state transitions are two numerical cell state variables whose values in neighboring cells determine which discrete state transitions occur. The state a_i of a building v_i at time t may be described in terms of these variables according to

$$a_i(t) = \langle X_i(t), C_i(t), I_i(t) \rangle \tag{1}$$

where $X_i(t)$ represents the discrete state discussed above, $C_i(t)$ is the property's *current physical condition* and $I_i(t)$ is the *income of the current occupants*. For convenience, C_i and I_i are restricted to the range 0.0 to 1.0, and are continuously variable within this range. With these and other variables, the (t) notation denoting model time is omitted where the sense is clear.

Some comments on the physical condition and income state variables are appropriate. The condition C_i should be thought of as standardized for the property's other physical attributes such as floor space, and number of rooms, and may be considered as an 'intrinsic value per square meter'. The notion of intrinsic value summarized by the property's current physical condition is representative of Smith's Marxian concept of value, and any expenditure on the property increases this quantity directly. Household income I_i is also on a standardized per (adult) capita basis. These abstract representations of physical condition and income mean that we try to treat a single 'yuppie' in a studio flat as similar to an established professional family with two children and a dual income in a four bedroom house, since both would have high incomes

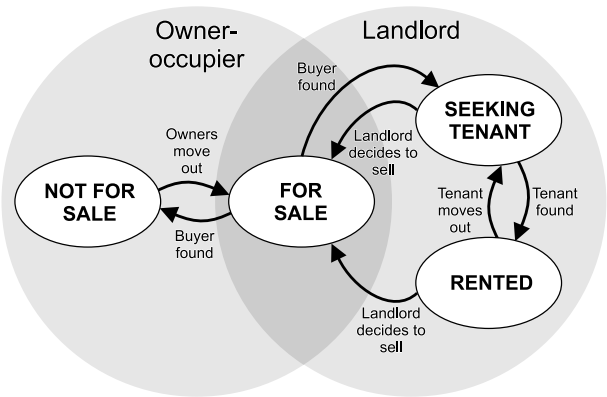

Fig. 3. Location discrete states and the allowed state transitions. The relationship of these states to property tenure is also shown

and, if the buildings were in similar condition, regardless of size, they would also have similar physical condition values. This approach obviously glosses over many of the details of the gentrification process, but as has been emphasized, the present purpose is to develop a framework that can be extended in time to cope with many of the details it currently lacks.

At this point, it is useful to look ahead to where the description of the model transition rules is going, by reviewing Fig. 4, which summarizes in schematic form the overall determination of model state transitions. Figure 4 draws attention to two *temporary* numerical state variables, associated with two of the discrete states only: *price* and *rent*. When a location is entering the FOR SALE state a sale price is set for the location. This may subsequently be adjusted downwards if no buyer is immediately found. Similarly, when a location enters the SEEKING TENANT state, a rent is set, although if a tenant is not found the property may be put up FOR SALE. In effect, a property's sale price or monthly rent—its 'value' in the market—is a 'virtual' phenomenon, only actualized at the point of sale and dependent on the buyer or prospective tenant's assessment of the property's location and prospects. This is thus related to the property's physical condition, but not in any simple way, since neighboring properties also affect value.

The overall state at an individual location is thus rather complex. Taken together, four discrete states, NOT FOR SALE, FOR SALE, SEEKING TENANT and RENTED, two permanent numerical variables, *physical condition* and *household income*, and two temporary numerical variables, *price* and *rent*, are used.

4.2 Local and global model parameters

Before describing the model transition rules in detail, we introduce various model parameters. These are derived either locally at each cell location based on a cell and its neighbors' states, or globally for the whole model. The local and global parameters are summarized in Tables 2 and 3. As will become clear, values of many of these parameters can only be sensibly interpreted with a temporal scale for the model. Effectively, each time step represents a month of actual time.

It is these parameters in combination with an individual cell's state variable values and stochastic variation that determine changes in cell states, as described in detail in the next section and schematically illustrated in Fig. 4. Two local parameters are determined for each location at each time step, as follows.

Local 'rent gap' is defined as

$$G_i^{(C)}(t) = \begin{cases} \overline{C}_N(t) - C_i(t) & \text{if } \overline{C}_N(t) - C_i(t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

that is, the difference between the mean value of the condition variable in a building's neighborhood and the value of the condition variable for that building, with the added proviso that only positive values are considered. If a building is in better condition than the local mean, then the rent gap is zero. The rent gap is an important determinant of whether or not owner-occupiers obtain finance to upgrade their properties.

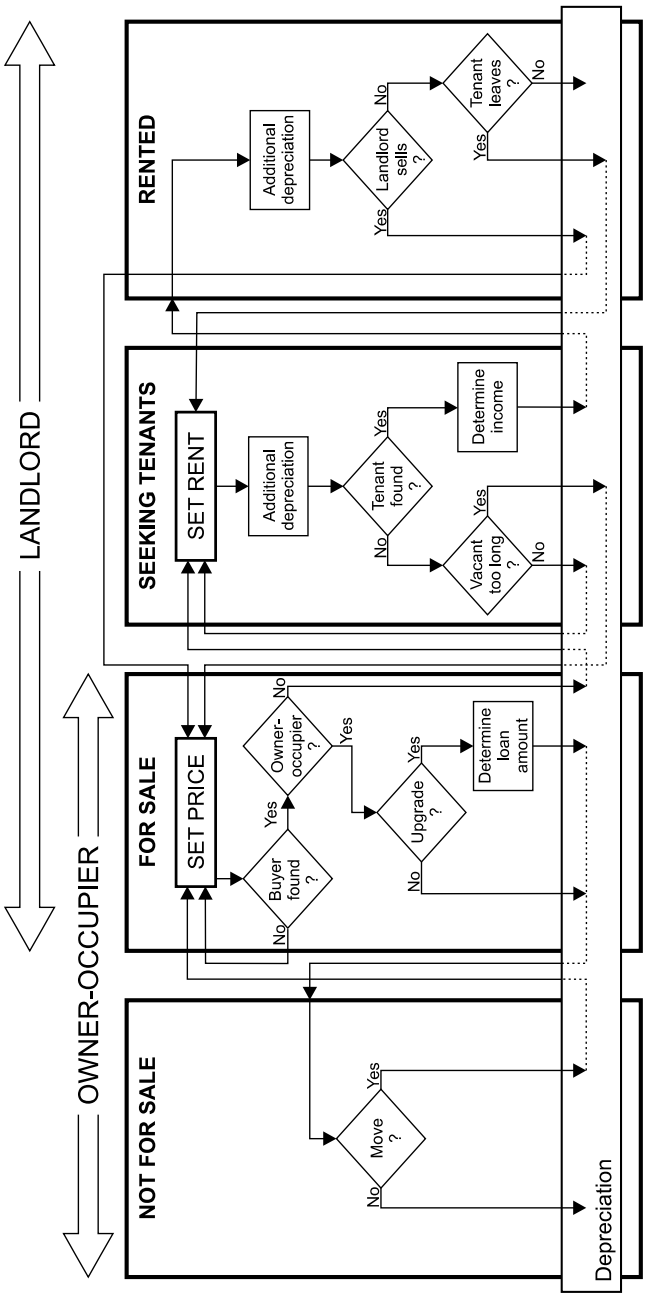

Fig. 4. Block diagram showing the relationships between location states. Note (i) depreciation occurs *every* time step, and (ii) where no arrow exits a state the cell remains in that state at the next time step

Local income gap. G_I is similarly defined as $I_i - \overline{I}_N$ and is also only allowed to adopt positive values. Thus

$$G_i^{(I)}(t) = \begin{cases} I_i(t) - \overline{I}_N(t) & \text{if } I_i(t) - \overline{I}_N(t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

In addition, the model relies on a number of global variables applicable to the whole model at each time step. Most of these are user-set parameters whose values remain constant through the duration of a single run of the model. However, one parameter—the *neighborhood status*—is updated at each time step, so that it varies as the model runs. Global parameters are summarized in Table 3 and are as follows:

Neighborhood status. S is the current standing of the whole urban neighborhood represented in the model and varies from time step to time step as the model runs. The value of S is in the range 0–1, and is an important determinant of the incomes of prospective new buyers or tenants for properties which are FOR SALE or SEEKING TENANTS. This parameter directly addresses Hammel's (1999) idea that neighborhood scale is the most important determinant of the potential value of a site.

Status variability. S_δ is a factor affecting variability in the neighborhood status. Most variation in S is due to changes in income and condition in the model, a small random variabiation is also included and this parameter controls the size of this effect, according to Eq. (19).

Abandonment factor. p_{A} is a fixed parameter whose value is effectively the probability at any particular time step that an owner-occupier household will choose to move out. The quoted typical value 0.0125 is equivalent to an annualized probability of moving out of about 0.14.

Tenant mobility. p_{TM} is a fixed parameter, whose value is equivalent to the probability that a tenant household will move out in a particular time step. Normally, p_{TM} should be set to a higher value than p_{A} reflecting the greater

Table 2. The model's two local parameters

Parameter	Description	Range
$G_i^{(C)}$	'Rent gap'	0.0–1.0
$G_i^{(I)}$	Income gap	0.0–1.0

Table 3. The model's six global parameters

Parameter	Description	Range	Typical value
S	Neighborhood (i.e. whole model) status	0.0–1.0	0.5
S_δ	Status variability	0.0–0.1	0.025
p_{A}	'Abandonment' factor	0.0–0.1	0.0125
p_{TM}	Tenant probability to move	0.0–0.5	0.0561
r_{D}	Depreciation rate	0.0–01	0.0028
n_{V}	Tolerable vacant months	1–6	3.0

transience of population in rented accommodation. The typical value 0.0561 is equivalent to an annual probability of moving of 0.5.

Depreciation rate. r_D is the loss in value applied to the physical condition variable of every location every time step, and reflecting the simple fact of physical wear and tear on property. A typical value of r_D is 0.0028 so that a period of 360 months—30 years—would see the collapse of a property in perfect physical condition ($C_i = 1.0$), to complete decay ($C_i = 0.0$). This could only occur in the absence of other effects, particularly the provision of home improvement loans, and if a property is never in the rental sector, which speeds up the depreciation process.

Tolerable vacant months. n_V is the number of months (model time steps) that a landlord will allow a location to remain vacant, that is in the SEEKING TENANTS state. After failing to let a property over this number of months a landlord sells the property, so that its state changes to FOR SALE.

4.3 Stochastic effects

One implementation detail deserves mention, before a detailed description of the model rules. A number of pseudo-random numbers are generated while running the model, and play an important role in household decisions to move out or stay in a neighborhood, in the determination of changes in household incomes, in the determination of potential buyer and tenant incomes, and in determining the size of home improvement loans. Two aspects are critical here:

Model repeatability. It is important that random number draws are carried out as far as possible to ensure the repeatability of model behavior. Therefore all five numbers that might be required at each model location are generated every time step, regardless of whether or not they will be used. This ensures that given the same pseudo-random number generator seed value at the start of a model run, an identical sequence of values is drawn at each location in each run of the model, thus allowing direct comparison of model runs with otherwise different parameter settings.

Random numbers in the model are bounded. Many random number draws in the model are taken from the normal distribution, which might produce extreme values. To avoid the associated problems, *bounded normal distributions* are used simply by making repeated draws from the normal distribution until a value inside the required bounds is obtained. This mechanism means that the desired central tendency characteristic of the normal distribution is retained. A side effect is that the repeatability discussed in the previous paragraph is compromised—although only between model runs that are beginning to diverge widely anyway.

4.4 Putting it all together—transition rules

As in a standard CA model, the state of a vertex or location at the next time step is dependent on the current states of cells in its neighborhood in the lattice:

$$a_i(t+1) = f(\{a_j(t) : v_j \in N(v_i)\}) \tag{4}$$

where a is the cell state, v_i and v_j are arbitrary cells in the lattice and $N(v_i)$ represents the neighborhood of the vertex v_i in the graph. Given that a is actually represented by a complex combination of discrete and numeric variables, it is easier to understand, develop, and describe the graph-CA process rules in terms of decisions made by individuals represented in the model dependent on the current discrete state at each model location. The diagram in Fig. 4 is helpful in following the description below.

NOT FOR SALE. Each month owner-occupier households whose property is NOT FOR SALE consider moving out, and do so with probability

$$p_{\mathrm{A}}(1.5 - S + G^{(I)}) \tag{5}$$

Thus, low neighborhood status, or high household income relative to neighboring properties is more likely to lead to a household moving out. If the household decides to move out then the discrete location state changes to FOR SALE and the sale price is set to

$$\text{Price} = 0.5(C_i + \overline{I}_N) \tag{6}$$

or the mean of the property's condition and average neighboring income values. Thus sale price is dependent on the property situation, not simply on the property characteristics itself.

FOR SALE. Each month that a property is FOR SALE, whether or not it is sold is determined by comparing its price to that of a randomly generated potential buyer's income. The potential buyer's income is drawn from a normal distribution

$$\text{Buyer income} = \mathbf{N}(\mu = 0.5(S + \text{Price}), \sigma = 0.1) \tag{7}$$

bounded by repeated draws to the range

$$[0.25(S + \text{Price}), \min(1, 0.75(S + \text{Price}))] \tag{8}$$

This potential buyer of the property is assessed by comparing their income to the property sale price. If the buyer's income is greater than the sale price then a buyer is considered to have been found.

If the buyer's income is also higher than the mean income of neighboring properties, then the buyer is considered to be a new owner-occupier, the discrete state is set to NOT FOR SALE, and the income variable I_i is set to the buyer's income. The new owners also consider the possibility of upgrading. If the local rent gap is greater than zero

$$G_i^{(C)} > 0 \tag{9}$$

and the household income is greater than the property's condition value

$$I_i > C_i \tag{10}$$

then the household is considered to have been successful in obtaining a home improvement loan. The value of the loan provided is randomly drawn from a normal distribution. This loan size is added to the location's current physical condition variable,

$$C_i'(t) = C(t) + \mathbf{N}(\mu = S - C_i(t), \sigma = 0.1) \tag{11}$$

where $C_i'(t)$ denotes a new, intermediate value of the physical condition variable, which may still change subject to depreciation effects (see equation 18 below). The normally distributed loan size is bounded by repeated draws to the range $[0.0, 0.5]$, and the resulting value of C_i' may not exceed 1. Loan size is therefore dependent on how far below the neighborhood status a property's current physical condition is.

If a buyer's income is less than the local mean income, then the buyer is considered to be a landlord, the discrete state is set to SEEKING TENANTS and I_i is set to 0. This means that incoming new owner-occupiers will have relatively high incomes, whereas the income of incoming tenants remains to be determined by the rental market. The rent for the property is set to

$$\text{Rent} = 0.5(C_i + \min(I_N)) \tag{12}$$

where I_N is the set of incomes of neighboring properties. Note that the minimum neighboring income could be 0, if neighboring properties are in the SEEKING TENANTS state and have so far failed to find tenants.

If no buyer is found, because the potential buyer's income is less than the property sale price, then the property remains in the FOR SALE state, but its price is adjusted downward to

$$\text{Reduced price} = 0.5(\text{Price} + \text{Buyer income}) \tag{13}$$

so that other things being equal, it is more likely a buyer will be found in the next time step.

SEEKING TENANTS. First, additional depreciation due to the property being in the rental sector is simulated by subtracting r_{D} from the physical condition variable.

Then, similarly to the FOR SALE case, a potential tenant income is determined by random draw from a normal distribution

$$\text{Tenant income} = \mathbf{N}(\mu = 0.5(S + \text{Rent}), \sigma = 0.1) \tag{14}$$

bounded by repeated draws to the range

$$[0.25(S + \text{Rent}), \min(1, 0.75(S + \text{Rent}))] \tag{15}$$

Determination of whether or not a new tenant has been found proceeds in a similar manner to the FOR SALE case. If the potential tenant's income is higher than the rent, then the discrete state is set to RENTED and the income variable I_i is set to the new tenant's income.

If no tenant has been found, then a counter recording the number of months that the property has been vacant is incremented. Additional depreciation equal to the current value of this count multiplied by the depreciation parameter r_{D} is then applied. If the current number of vacant months is greater than the tolerable vacant months parameter n_{V}, then the discrete cell state is set to FOR SALE and a sale price is set according to equation (6) above, just as in the owner-occupier case.

RENTED. Regardless of other effects, the property's physical condition variable is reduced by additional depreciation r_{D} due to its being in the rental sector.

For a RENTED property, the landlord may decide to sell doing so with probability

$$S - \text{Rent} \tag{16}$$

so that landlords are more likely to try to sell when the neighborhood is high status, or where they are only getting low rents. If the landord does decide to sell, then the discrete state is set to FOR SALE and the sale price is set according to equation (6) as before. While the landlord is attempting to sell the current tenants are considered to remain in occupation, so that household income is unchanged.

If the landlord does not choose to sell, then the current tenants may decide to move out, which occurs with probability

$$p_{\mathrm{TM}} + G_i^{(I)} \tag{17}$$

so that higher income tenants are most likely to leave. If the tenants decide to move out then the discrete state is set to SEEKING TENANTS, and a rent is set according to Eq. (12) as before.

Depreciation and changes in neighborhood status

As indicated in Fig. 4 depreciation occurs regardless of any other changes that may have happened. Each time step the property physical condition variable is adjusted by subtracting the depreciation parameter value r_{D}:

$$C_i(t+1) = C_i'(t) - r_{\mathrm{D}} \tag{18}$$

Additional depreciation applied to rented properties as discussed above reflects lower levels of maintenance carried out by landlords relative to owner-occupiers.

Note that as is conventional in synchronous CA, all the changes described above are considered to occur simultaneously. Thus all the above rules operate on every location and its neighbors' states at time t to determine the next state at time $(t+1)$. Once the determination has been made for all locations, all location states are changed to the newly determined state simultaneously.

Finally, every time step, the global neighborhood status parameter S is adjusted by adding the mean value of all changes in the physical condition and income variables determined at all locations in the model, together with a random factor:

$$\begin{aligned} S(t+1) = S(t) + \tfrac{1}{n}\left[\textstyle\sum_{i=1}^{n}[C_i(t+1) - C_i(t)] + \sum_{i=1}^{n}[I_i(t+1) - I_i(t)]\right] \\ + \mathbf{N}(\mu = 0, \sigma = S_\delta) \end{aligned} \tag{19}$$

Since the status parameter is important in determining the likely incomes of new buyers and tenants, this is an important positive feedback mechanism in the model. Generally rising incomes and improving physical conditions will tend to draw higher income buyers and tenants. Conversely, falling incomes and deteriorating stock will draw lower income buyers and tenants.

5 Running the model

The model described has been built and run for a part of Hoxton in inner East London in the United Kingdom (UK). Hoxton is an area that has seen rapid gentrification in recent years, and has garnered a good deal of media attention in the process (see Jennings 2000, for example). Due to the abstract nature of the model, this choice was made more for convenience rather than from a desire to simulate the dynamics of change in this particular neighborhood.

5.1 Data sources

5.1.1 Spatial data. In the UK, the Ordnance Survey's *Landline* digital data sets are the best primary source for spatial data at the resolution required for this simulation. Unfortunately, these data are primarily intended to support the production of map line work so that building footprints and plot boundaries are stored as collections of line segments rather than as polygons. Each line is associated with a feature code so that it is possible to select all lines relating to buildings and plots. Even so, it remains non-trivial to further process these data to represent buildings or plots as polygon objects in a geographical information system (GIS). Many blocks in the urban fragment are not subdivided, even where it is clear on the ground that the block is not a single address.

As a result, considerable amendment and editing of the raw *Landline* data was undertaken to arrive at a set of polygons that is a reasonable representation of the study area morphology. It is not possible to proceed far with building an 'accurate' representation of the urban fabric before facing the issue of multi-occupancy buildings, and the relations internal to them. This leads immediately to the complexities of three-dimensional representation of built forms, which raises the issue of representing redevelopment in three-dimensional detail, as for example, when a developer breaks a former factory up into small units, or when a new owner-occupier knocks former flats or apartments into a single family dwelling. These issues are well beyond the scope of the present work, and for this rather abstract model such complexities have been ignored in favor of a representative two-dimensional view. Having built a two-dimensional map of the study area with buildings identified as 514 distinct polygons, the Delaunay triangulation structure shown in Fig. 2 was constructed as discussed in Sect. 3.1.

5.1.2 Property and income data. Clearly, initializing even a small model with realistic empirical data on property condition and occupant incomes would be a major undertaking. The physical condition variable is particularly problematic—recalling the difficulties many have had testing the rent gap hypothesis. In fact, *any* state variable based on property values, however conceived faces difficulties. As has been remarked, the monetary value of a property is a virtual phenomenon, only actualized when it changes hands at a price set by the market. Even when market turnover is high it is rare for more than (say) 10% of the properties in an area to change hands in one year so that complete data on the current sale price of properties is unavailable. Any estimation of missing values requires a method for normalizing prices for variations in the characteristics of properties that did change hands. Only limited information about the numbers of rooms and their sizes, or the physical state of buildings is generally available. In any case the relationship between sale price and value is a contingent one, as has been discussed.

Since one of the aims of abstracting to a simple physical condition variable was to reduce data requirements, these complexities have been avoided by synthesizing values from a proxy variable. Property condition figures were obtained by 'drilling down' into the UK Department of the Environment Transport and the Regions (DETR) *Index of Town Centredness* (ITC) surface data set (Thurstain-Goodwin and Unwin 2000). This is a composite

data set developed on the basis of detailed, confidential UK government data sets intended for use in defining town centers for statistical monitoring purposes. The ITC takes into account parameters such as the amount of employment and turnover at different locations in urban areas. By 'drilling down' into a surface is meant assigning to each building the value on the surface that attains at the building centroid. For the model study area, values of the ITC are in the range 8.6–25.8. These numbers were rescaled to a more appropriate range for the present purpose, giving condition variable values in the range 0.258–0.774.

Using census data, per capita household income at an *aggregate* level is easier to determine (Dorling 1999, comments on particular difficulties in the UK). However, considerable work is required to disaggregate data to produce realistic data for simulation purposes (Clarke and Holm 1987). Assigning the individual synthetic households produced by disaggregation to particular street addresses is also difficult. Although some have suggested that the more detailed information provided by marketing surveys may represent a way out of this difficulty (Openshaw and Turton 1998), others are more wary (Longley and Harris 1999) and there seem to be important ethical issues to be considered in developing individualized datasets at the household level.

Again, a pragmatic approach has been adopted, given the abstract nature of the present model. Income values were simply set equal to the physical condition values summed with a normally distributed random offset ($\mu = 0.0$, $\sigma = 0.025$). This means that the model start from a situation where value and income are approximately matched, and has the desirable side-effect that no large disparities between income and property value are likely to strongly influence initial development of the model dynamics. Additionally, given the smooth spatial variation in the physical condition variable, no strong local 'gaps' are likely.

The resulting overall physical condition distribution in the study area is shown in Fig. 5. Incomes closely match the mapped distribution, and the

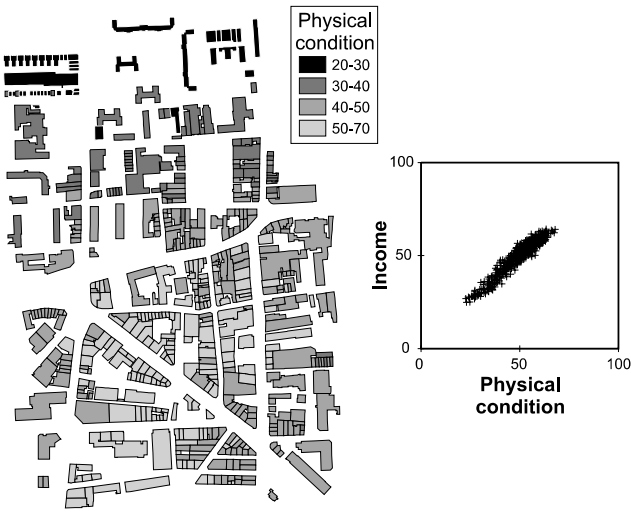

Fig. 5. The distribution of initial values used in the model

initial system configuration is also shown as a scatter plot of income against physical condition to demonstrate the point.

5.2 Results

It is not intended to present extensive details of results for different runs of the model. Instead, a single typical run is presented, to demonstrate that model dynamics are potentially interesting. This is followed by a discussion of model behavior and consideration of the potential of the present approach in the concluding section.

As has been discussed, the model was initialized to the state shown in Fig. 5. With the parameter settings recorded in the 'Typical value' column of Table 3 the model was run for 720 time steps, equivalent to a real time period of 60 years. The time series shown in Fig. 6 record the overall dynamics.

As can be seen, over this 60 year period, there is a period of decline in the neighborhood extending over about 20 years from the 15th to the 35th years after the model starts running. Decline occurs in two phases, the initial period running from about year 15 to year 22, and seeing increased numbers of rented properties and larger numbers of properties for sale for extended periods. After this initial decline there is partial recovery in the neighborhood's fortunes, although the next few years are very different in character from the early phase of stable owner-occupation, with larger numbers of properties for sale than previously, indicating larger turnover of households. Notably, during this period, although there is some recovery in the mean household income, the mean physical condition of properties falls continuously.

From the 30th to the 35th years the neighborhood is very unstable with large numbers of rental properties, large numbers of sales, and continuing decline in mean household income and in the physical condition of properties. At the end of this period, properties start to return to owner occupation and the fall in both incomes and the condition of buildings is halted. These effects combine to see the neighborhood's status sharply rise, and mean household incomes subsequently rise sharply. Within a matter of only 3 years virtually all properties are back in owner-occupation. Over a period of 5 years neighborhood status increases from its minimum to its maximum value, and a further 20 year period of stable owner-occupation has started. By the end of the 60 year period there are signs that decline may be setting in again.

Clearly, even these summary data are relatively complex. Furthermore, observation of repeated runs of the model shows that there is no readily discernible pattern to the duration of cycles, although they do tend to be shorter with more aggressive settings of the depreciation factor r_D (see Eq. 18) and of status variability S_δ (see Eq. 19). The partial recovery seen in this example is not observed in every case. Furthermore, from the given initial conditions, with various settings of the random number generator seed values, time series that see an immediate decline in neighborhood fortunes are observed, as well as time series where the neighborhood status initially rises – as in the illustrated case – but remains stable for longer than in this example.

What is clear is the important role of the global neighborhood status parameter to the model dynamics. Without this positive feedback mecha-

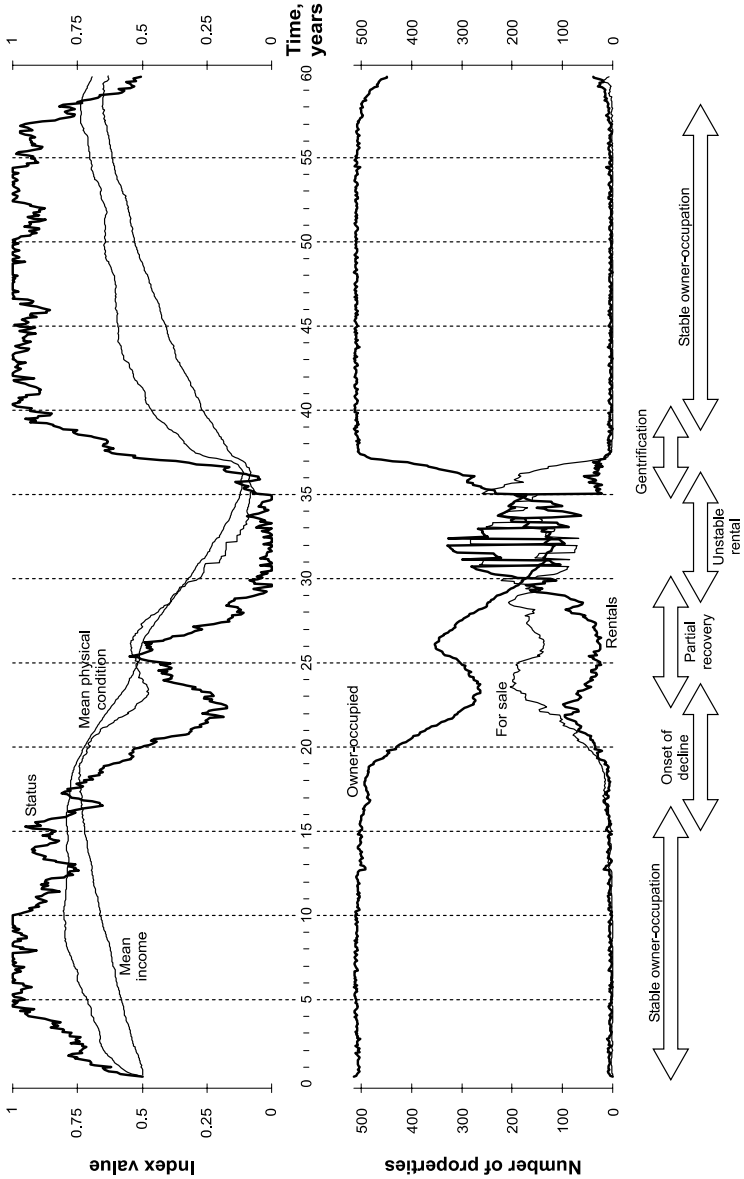

Fig. 6. Typical time series data generated by the model over a 60 year period

nism, a high or low income neighborhood tends to stay that way. At high levels of household income properties are maintained in good condition, and similarly rich buyers are quickly found for properties that go on sale. This situation is disrupted when neighborhood status starts to fall, because incoming households tend to have lower incomes (see the sections above on the FOR SALE and SEEKING TENANTS states). Similarly, without the status variable, at low income levels many properties move into the rental sector where they are undermaintained, incomes remain low and the physical condition of properties declines and remains low. Richer incoming households also tend not to stay long under these circumstances. For any sustained recovery from this state to occur the neighborhood status must be increasing, 'pulling up' the incomes of newly arrived households, and creating the cycle of improvements that further increases both incomes and property values. It is important to note that the *random* component of neighborhood status changes (see Eq. 19) is therefore critical to the dynamics.

6 Conclusions

This paper has outlined a theoretical and conceptual basis on which a detailed model of the residential and property market dynamics underlying gentrification might be built. The behavior of the resulting model has been described in outline. It seems clear that the model structure outlined is rich enough to merit further investigation, although many difficult challenges remain.

As has been emphasized, the behavior of the neighborhood status parameter is critical to the global behavior of this model. The question of its interpretation is therefore an important one. Although changes in the parameter value are calculated internal to the model, it is apparent that it would be much better to treat this parameter as exogenous to the model, because it reflects the standing of the neighborhood represented relative to other neighborhoods. A much more satisfactory implementation would embed the neighborhood presented here in a larger model of the urban system. The neighborhood status parameter would then be adjusted contingent on varying perceptions of the relative merit of other neighborhoods in the wider urban system in a manner consistent with scale-based developments of rent gap theory (Hammel 1999). A further level of embedding the urban system in a network of competing cities might also be envisaged. The graph-CA model architecture is readily extendible, so that each level of a scale-based hierarchy could be implemented in a similar way (O'Sullivan 2001b). Somewhat similar nested hierarchies of models have been discussed by White and Engelen (1997).

Other parameters raise fewer broad questions about the overall architecture of the model. However, it is interesting to note that an earlier version of the same model (O'Sullivan 2000) attempted to operate using a 'purer' cellular automaton with only two state variables – income and value – at each location. However, it proved difficult within such a limited framework to produce the cyclical behavior of the present model. Also, the distinct maintenance and depreciation behaviors of owner-occupiers and landlords in the present model could not be represented in that framework. It is interesting to speculate that *any* model of a human system represented at this scale must inevitably explicitly represent aspects of the social system. This is

very different from how many more conventional urban CA models developed on grids are conceptualized, where state transitions are based on probablistic rules for different landuse changes. In such cases, human agency is in danger of becoming a 'ghost in the machine', rather than the primary driver of urban change—whether the agents involved are individual householders or more diffuse corporate actors, such as mortgage lenders. This observation is consistent with much that has been written about gentrification itself.

This also suggests that an explicitly agent-based model may be a more appropriate starting point than any cellular framework (Portugali et al. 1997; Benenson 1998). Indeed, in their most recent work Portugali (2000) and his colleagues at Tel Aviv University adopt an agent-based approach, in a similarly graph or network-structured representation of the built environment. For agent models in a network structured space it may even be appropriate for different agents to use different criteria for determining the cell neighborhoods that they consider in making different decisions, and hence to operate in a variety of network spaces. In the present model, for example, landlords might respond to events in larger cell neighborhoods than owner-occupiers. This would represent a significant increase in the complexity of the model. In any case, as discussed in Sect. 3.1, other graph structures may be more defensible than the Delaunay triangulation, which effectively assumes spatial uniformity. Given that the graph structure is likely to affect the overall dynamics (O'Sullivan 2001a) decisions on this aspect are critical to any model of this or similar arhitecture.

Finally, this work has deliberately and self-consciously attempted to make use of *geographical* theory of gentrification rather than economic theory. Although this choice has led to some difficulties, particularly in the definition of key state variables, I feel that it has been justified. First, it has demonstrated that current modeling techniques are capable of formally representing quantitative and qualitative aspects of a complex phenomenon such as gentrification. This holds out the promise of future productive encounters between human geographical theory and contemporary modeling techniques. Second, it suggests that the opportunity now exists to harness the riches of theory to the task of developing urban simulation models for decision support that move beyond the narrow world-view of economics. Certainly, it seems clear that any successful model of gentrification must find ways to represent the rich complexity of the phenomenon. This paper has described the beginnings of a framework that may help to make that formidable task possible.

References

Anselin L (1995) Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2): 93–115

Badcock B (1989) An Australian view of the rent gap hypothesis. *Annals of the Association of American Geographers*, 79(1):125–145

Badcock B (1990) On the nonexistence of the rent gap a reply. *Annals of the Association of American Geographers*, 80:459–461

Batty M, Xie Y, Sun Z (1999) Modelling urban dynamics through GIS-based cellular automata. *Computers Environment and Urban Systems*, 23:205–233

Benenson I (1998) Multi-agent simulations of residential dynamics in the city. *Computers Environment and Urban Systems*, 22(1):25–42

Bernard RN (1999) Using adaptive agent-based simulation models to assist planners in policy development: the case of rent control. *Working Paper 99-07-052, Santa Fe Institute*

Bondi L (1991) Gender divisions and gentrification: a critique. *Transactions of the Institute of British Geographers* NS16:190–198

Bondi L (1999) Gender class and gentrification: enriching the debate. *Environment and Planning, D: Society & Space*, 17(3):261–282

Bourassa SC (1990) On 'An Australian view of the rent gap hypothesis' by Badcock'. *Annals of the Association of American Geographers*, 80:458–459

Bourassa SC (1993) The rent gap debunked. *Urban Studies*, 10:1731–1744

Bridge G (1994) Gentrification: a reappraisal. *Environment and Planning*, *D: Society and Space*, 12:31–51

Clark E (1987) The Rent Gap and Urban Change: Case studies in Malmö 1860–1985. *Meddelanden från Lunds Universitets Geografiska Institutioner*, Lund University Press, Lund Sweden, Vol 101

Clark E (1988) The rent gap and transformation of the built environment: case studies in Malmö 1860–1985. *Geografiska Annaler B*, 70(2):241–254

Clark E (1992) On blindness centrepieces and complementarity in gentrification theory. *Transactions of the Institute of British Geographers*, NS17:358–362

Clarke KC, Hoppen S, Gaydos L (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay Area. *Environment and Planning B: Planning & Design*, 24(2):247–262

Clarke M, Holm E (1987) Micro-simulation methods in human geography: a review and further extensions. *Geografiska Annaler B*, 69:145–164

Couclelis H (1991) Requirements for planning-relevant GIS: a spatial perspective. *Papers in Regional Science*, 70:9–19

Couclelis H (1997) From cellular automata to urban models: new principles for model development and implementation. *Environment and Planning B: Planning & Design*, 24: 165–174

Dorling D (1999) Who's afraid of income equality? *Environment and Planning A*, 31(4):571–574

Edwards G (1993) The Voronoi model and cultural space: applications to the social sciences and humanities. In: Frank AU, Campari I (eds) *Spatial Information Theory: A Theoretical Basis for GIS*. Vol 716 Lecture Notes in Computer Science, Springer-Verlag, Berlin, pp 202–214

Getis A (1994) Spatial dependence and heterogeneity and proximal databases. In: Fotheringham AS, Fischer M (eds) *Spatial Analysis and GIS*. Taylor & Francis, London, pp 105–120

Glass R (1964) Introduction: aspects of change. In: Centre for Urban Studies (ed) *London: Aspects of Change*. MacGibbon & Kee, London, pp xiii–xlii

Gold CM (1992) The meaning of neighbour. In: Frank AU, Campari I, Formentini U (eds) *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, Vol 639, Lecture Notes in Computer Science, Springer-Verlag, Berlin, pp 220–235

Hammel DJ (1999) Re-establishing the rent gap: an alternative view of capitalised land rent. *Urban Studies*, 36(8):1283–1293

Hamnett C (1991) The blind men and the elephant: the explanation of gentrification. *Transactions of the Institute of British Geographers*, NS16:173–189

Hamnett C (1992) Gentrifiers or lemmings? A response to Neil Smith. *Transactions of the Institute of British Geographers*, NS17:116–119

Jennings C (2000) Lost and Found. Pages 18–19, *Space* supplement to *The Guardian* newspaper, 20 January

Kain JF, Apgar WC (1985) *Housing and Neighbourhood Dynamics: A Simulation Study*, Harvard University Press, Cambridge, MA and London

Lees L (2000) A reappraisal of gentrification: towards a 'geography of gentrification'. *Progress in Human Geography*, 24:389–408

Ley D (1986) Alternative explanations for inner-city gentrification: a Canadian assessment. *Annals of the Association of American Geographers*, 76(4):521–535

Ley D (1987) Reply: the rent gap revisited. *Annals of the Association of American Geographers*, 77(3):465–468

Ley D (1996) *The New Middle Class and the Remaking of the Central City*. Oxford Geographical and Environmental Studies, Oxford University Press, Oxford, New York

Li X, Yeh AG-O (2000) Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science*, 14(2):131–152

Longley PA, Harris RJ (1999) Towards a new digital data infrastructure for urban analysis and modelling. *Environment and Planning B: Planning & Design*, 26(6):855–878

Okabe A, Boots B, Sugihara k, Chiu SN (2000) *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley Series in Probability and Statistics, 2nd edn, John Wiley & Sons Ltd, Chichester, UK

Openshaw S, Turton I (1998) Geographical research using lifestyles databases. *RGS-IBG annual conference, 5–8 January*, Kingston University, Guilford, England

O'Sullivan D (2000) *Graph-based Cellular Automaton Models of Urban Spatial Processes*. (unpublished Ph.D. thesis), University College London, Bartlett School of Architecture and Planning

O'Sullivan D (2001a) Exploring spatial process dynamics using irregular cellular automaton models. *Geographical Analysis*, 33(1):1–18

O'Sullivan D (2001b) Graph cellular automata: a generalised discrete urban and regional model. *Environment and Planning B: Planning & Design* 28(5): 687–705

O'Sullivan D, Torrens PM (2001) Cellular models of urban systems. In: Bandini S, Worsch T (eds) *Theoretical and Practical Issues on Cellular Automata*, Proceedings of the Fourth International Conference on Cellular Automata for Research and Industry, Karlsruhe, 4–6 October, Springer-Verlag, London, pp 108–116

Portugali J (2000) *Self-Organization and the City*. Springer-Verlag, Berlin

Portugali J, Benenson I, Omer I (1997) Spatial cognitive dissonance and sociospatial emergence in a self-organizing city. *Environment and Planning B: Planning & Design*, 24(2):263–285

Redfern PA (1997) A new look at gentrificatian: 1. Gentrificatian and domestic technologies. *Environment and Planning A*, 29(7):1275–1296

Robson G, Butler T (2001) Coming to terms with London: middle-class communities in a global city. *International Journal of Urban and Regional Research*, 25(1):70–86

Rose D (1984) Rethinking gentrification: beyond the uneven development of marxist urban theory. *Environment and Planning D: Society and Space*, 1(1):47–74

Sayer A (1992) *Method in Social Science: A Realist Approach*. 2nd edn, Routledge, London

Semboloni F (2000) The growth of an urban cluster into a dynamic self-modifying spatial pattern. *Environment and Planning B: Planning & Design*, 27(4):549–564

Shaw W (2000) Ways of whiteness: Harlemising Sydney's Aboriginal Redfern. *Australian Geographical Studies*, 38(3):291–305

Shi W, Pang MYC (2000) Development of Voronoi-based cellular automata — an integrated dynamic model for geographical information systems, *International Journal of Geographical Information Science*, 14(5):455–474

Smith N (1979) Toward a theory of gentrification: a back to the city movement by capital not people. *Journal of the American Planning Association*, 45(October):538–548

Smith N (1987a) Gentrification and the rent gap. *Annals of the Association of American Geographers*, 77(3):462–465

Smith N (1987b) Of yuppies and housing: gentrification social restructuring and the urban dream. *Environment and Planning D: Society and Space*, 5:151–172

Smith N (1992) Blind man's buff or Hamnett's philosophical individualism in search of gentrification. *Transactions of the Institute of British Geographers*, NS17:110–115

Smith N (1996) *The New Urban Frontier: Gentrification and the Revanchist City*. Routledge, London, New York

Tabb WK (1982) *The Long Default: New York City and the Urban Fiscal Crisis*. Monthly Review Press, New York

Takeyama M, Couclelis, H (1997) Map dynamics: integrating cellular automata and GIS through Geo-Algebra, *International Journal of Geographical Information Science*, 11:73–91

Thurstain-Goodwin M, Unwin D (2000) Defining and delineating the central areas of towns for statistical monitoring using continuous surface representations, *Transactions in GIS*, 4:305–317

Ward DP, Murray AT, Phinn SR (2000) A stochastically constrained cellular model of urban growth, *Computers Environment and Urban Systems*, 24: 539–558

Warde A (1991) Gentrification as consumption: issues of class and gender, *Environment and Planning D: Society and Space*, 9: 223–232

Webster CJ, Wu F (1999a) Regulation land use mix and urban performance. Part 1 Performance, *Environment and Planning A*, 31(8):1433–1442

Webster CJ, Wu F (1999b) Regulation land use mix and urban performance. Part 2 Theory, *Environment and Planning A*, 31(9):1529–1547

White R, Engelen G (1997) Cellular automata as the basis of integrated dynamic regional modelling, *Environment and Planning B: Planning & Design*, 24(2):235–246

White R, Engelen G (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems, *Computers Environment and Urban Systems*, 24:383–400

Yung C-F, King RJ (1998) Some tests for the rent gap theory. *Environment and Planning A*, 30:523–542