Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Land Use Pattern

by

Elena G. Irwin and Nancy E. Bockstael

WP 98-24

Department of Agricultural and Resource Economics
The University of Maryland, College Park

DRAFT 7/31/98

Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Land Use Pattern[1,2]

Elena G. Irwin and Nancy E. Bockstael

Working Paper 98-24

Department of Agricultural and Resource Economics
The University of Maryland, College Park

[1] This research has been supported by U.S. Environmental Protection Agency under Cooperative Agreement # CR821925010.

[2] This version of the paper was presented at the NBER Summer Institute 1998, Workshop on Public Policy and the Environment, Cambridge, MA, August, 1998.

1. Introduction

There are two motivations for this paper. The first is to provide a means of answering a question of primary importance for environmental science and policy. Humans influence both the quality of marine and estuarine systems and the size and diversity of animal and plant populations in a region, largely by changing the pattern of land use in that region. The pattern of land use affects biodiversity directly by altering terrestrial habitat. It affects water quality because changes in land use bring with it changes in the amounts and types of toxics and nutrients man introduces into the ecosystem. It also alters water quality and associated stream flows by dramatically decreasing the amount of land cover that is permeable.

The models of many environmental scientists require as input the spatial distribution of land use/land cover. Since land use/land cover change is almost entirely determined by human actions, and since environmental policies generally operate by altering human actions, economic models of human land conversion decisions are key to scientists' predictions of future environmental outcomes and to an understanding of the effectiveness of policy instruments designed to alter those outcomes.

The predominant spatial model of land use evolution in the economics literature is the monocentric city model and its descendants. However robust a generalization this model has proven to be in the past, it fails to explain current trends in the changing spatial pattern of land use, which are characterized by land fragmentation brought about by residential sprawl at the rural-urban interface. For example, consider the spatial pattern of residential subdivision development located in exurban[3] areas in central Maryland (Maps 1 and 2). The observed pattern of exurban residential land use – and in particular, of residential subdivision development shown in Map 2 – is highly fragmented. Map 2 shows that in many cases, development appears to follow the location of exogenous

---

[3] We use the term "exurban" to connote areas located at the rural-urban fringe.

landscape features, e.g. roads[4] and waterfront. This map also indicates the areas in which residential development is allowed by zoning regulations. As is the case with the other exurban areas within our study area, zoning constraints influence the resulting pattern by delineating developable areas and setting minimum lot sizes, but do not fully determine the residential land use pattern. Map 2 suggests that, conditional on these exogenous features and on zoning constraints, residential subdivisions appear to be *negatively* correlated with each other – i.e. rather than clustered in a contiguous pattern, they are scattered in a noncontiguous pattern across the landscape. In this paper we explore whether a second generation model of the evolution of land use pattern can be developed, one that draws on new spatial, interacting-agent models, to explain this highly fragmented pattern of sprawl development at the rural-urban fringe.

This brings us to the second motivation for this paper. Recent work on a variety of economic and social phenomena has emphasized the role of interactions among economic agents connected to each other via local networks in space.[5] The key insight of these models is the recognition that agents' values or available information, on which individuals' decisions are based, may be influenced by others' values and actions. While the models encompass a broad range of subject matter and employ a variety of tools and techniques, their common premise is that interaction among many, possibly heterogeneous agents at a highly disaggregate level leads to the emergence of collective behaviors and patterns in social and economic systems at an aggregate level.[6] Some of these concepts have been applied to model the formation of industrial location patterns (Krugman, 1996; Arthur, 1998). For example, Krugman (1996) models the evolution of

---

[4] The relationship between public infrastructure, e.g. roads and public sewer and water lines, and residential land use patterns is an important subject in its own right. It is clear that over time, the provision of these services and infrastructure is endogenous to residential location decisions. However, because the focus of this paper is on the endogenous interactions among agents making residential location decisions, we abstract from the endogeneity of these features of the landscape and treat them as exogenous. Over a short time horizon, this is not an unreasonable assumption.

[5] While often times referring to geographical distance, space is also used in more a general sense, e.g. social distances between agents.

[6] These models fall under the general rubric of the so-called complexity perspective in economics, which emphasizes the process and emergence of economic phenomena (e.g. nonlinear patterns in financial time series, sociological pathologies, evolution of institutions). This is also sometimes referred to as the "Santa Fe" approach. For a recent overview, see Arthur, Durlauf, and Lane (1997).

a clustered industrial pattern as the result of tension between "centripetal" and "centrifugal" forces that arise from underlying demand and supply linkages among firms, which are a function of distance between firms.

In a similar spirit, we are interested in whether the observed pattern of scattered residential land use patterns in our study area can in part be explained as the consequence of interaction effects among spatially distributed land owners. If spatial externalities from neighboring land uses influence owners' land use decisions, then individual land owners' conversion decisions are not independent and changes in land use pattern are, in part, endogenous. Clearly a host of exogenous factors influence land use conversion – e.g. accessibility to urban centers, shopping areas, and other destinations; fixed features of the landscape such as waterfront, slope, and soil type; government policies that directly constrain land use or indirectly influence the returns to land in alternative uses; and factors that influence economic activity and growth in the region.[7] The central question posed in our research is whether the observed evolution of land use patterns is fully explained by these exogenous features and if not, whether a model that incorporates endogenous interaction effects among developers provides a robust explanation of these patterns.

This paper is organized as follows. First, we briefly review models of residential location patterns found in the urban economics and explain why these approaches are limited for our purposes. We then turn to agent-based interaction models and discuss their applicability to modeling the process of land use conversion. This approach is adopted to develop a dynamic model of a land owner's conversion decision, which incorporates interaction effects that capture the influence of neighboring agents' land use decisions. Simulation of this model suggests testable hypotheses regarding the interaction effects by illustrating the conditions under which a correspondence exists between interaction effects and the evolution of a "sprawl" land use pattern. Finally, the

[7] For recent empirical analyses using data from our study area of these factors, including empirical evidence of spatial externalities from neighboring land uses, see Bell and Bockstael (1998), Bockstael (1996), Geoghahan, Wainger, and Bockstael (1997), and Leggett and Bockstael (1998).

interaction hypothesis is tested using micro-level data on residential land use conversion over a seven year time period within an exurban region in central Maryland. In simultaneous social interaction models, Manski's (1993) "reflection problem" makes identification of social interaction problems impossible except in restrictive cases. In our model the spatial process is most appropriately characterized as recursive, but, as we will explain, it remains difficult to separate the effect of the spatial interaction ("true state dependence" in Heckman's terminology or "endogenous effects" in Manski's terms) from other effects (the contextual or correlated effects of Manski).[8] The paper concludes with thoughts on the usefulness of such an approach to characterizing the dynamic spatial process of land use conversion and the possible policy implications of interaction effects.

2. Models of Land Use Pattern Formation

The formation of urban land use patterns can be understood as a dynamic system comprised of many individual land owners distributed in space, each of whom owns a land parcel and makes profit-maximizing choices regarding the use of the parcel. The collective result of these actions over time and space determines the dynamic evolution of a regional land use pattern. In the next section we consider a model of agents' land use choices where decisions are not independent of each other due to the existence of spatial externalities.

The approach we propose to modeling land use conversion differs significantly from the monocentric model of residential location, and its descendants, found in the urban economics literature.[9] Monocentric models (Alonso, 1964; Muth, 1969; Mills,

[8] In this draft, we estimate what we term a "naïve" model, in which we do not fully account for identification problems. We are currently working on implementing an estimation strategy that addresses this issue.

[9] A small number of dynamic models that treat an individual's residential location decision as endogenous to other agents' location decisions are found in the regional economics literature (see Haag, 1989, for a review). They focus on the influence of aggregate migration flows across regions on individual location decisions among regions, and they abstract from agent-to-agent interaction on a micro-level. These studies typically attempt to provide an analytical characterization of the steady state and therefore do not attempt to characterize the out-of-equilibrium dynamics (e.g. Ben-Akiva and de Palma, 1986). In addition, they ignore possible heterogeneity among individuals at a micro-level (i.e. agents located within the same region) and allow only for a "one-with-all" interaction effect.

1972) describe a long run spatial equilibrium pattern of land uses given the existence of an exogenously determined central business district to which residents commute. Because the location of the city center is not endogenous to residential location, these models are essentially a static representation of an equilibrium land use pattern for either a fixed population or utility level. Many extensions to the basic monocentric model have been developed, including policentric models[10] and dynamic models that consider the process of urban growth by incorporating the durability of housing and intertemporally optimal decision making. In particular, Mills (1981) develops a dynamic monocentric model, in which temporary "leapfrog" development results due to the speculative behavior of heterogeneous developers. However, over time this type of model predicts an infill of development rather than an increase in the fragmentation of the development pattern. In any evern, the dominant elements of the monocentric city model and its antecedents are a) the role of a (set of) exogenous landscape feature(s) in determining the value of a parcel in development, and b) an interest only in the equilibrium land use pattern.

So far we have referred to the observed pattern of exurban residential development in broad descriptive terms – scattered, noncontiguous, sprawl – and argued that this does not "look like" it could be generated by a monocentric city type process. We now turn to spatial statistic methods in order to refine our understanding of the actual pattern of development in our central Maryland exurban study area. In particular, these methods are used to explore whether the observed pattern differs from either (1) a fully random distribution of developed parcels and (2) a hypothetical distribution of development under the monocentric hypothesis. To do so, we follow methods outlined in the spatial statistics literature (e.g. Cressie, 1993; Diggle, 1983) that are intended as

[10] These models allow for more than one urban center (see Richardson, 1988, for a discussion of these models). While these models are clearly a more realistic representation of today's metropolitan areas, they share some of the same limitations as the monocentric model. For example, most retain assumptions about exogenously located centers. In addition, accessibility to urban centers remains the determining factor in land rents.

preliminary "tests" for spatial pattern. The pattern that we test is a point pattern, in which each point represents a centroid of a residential lot.[11]

The first hypothesis that we would like to test is the possibility that the spatial pattern of development is fully "randomly" distributed. In terms of spatial correlation, a randomly distributed point pattern is one that exhibits a mix of both positive and negative correlation among points in space. This is in contrast to a clustered pattern, characterized by positive correlation, or an "inhibited" or "checkerboard" pattern that is negatively correlated. For clarification, Figure 1 provides a simple example of each type of pattern. A priori, we might expect our pattern of residential development to exhibit higher levels of positive correlation than a fully random pattern for a couple of reasons. First, a large majority of the residential lots in our study area are located within subdivisions and therefore, by definition, are clustered together. Second, we know from economic intuition and by visual inspections of the land use maps that residential development clusters around exogenous features of the landscape, e.g. roads, towns, and waterfront. However, as we have argued, there also appears to be a good deal of scatteredness among the "clumps" of residential development. A test for randomness is one way in which we can investigate the degree to which the known influences that result in positive correlation may be offset by an opposing effect that generates negative correlation.

A "test" for randomness is carried out by simulating multiple patterns of development under the assumption that the N points that comprise the pattern are independently and uniformly distributed over the region.[12] A spatial statistic is used to summarize the pattern of N points. In this case, we use an inter-point distance measure. The statistic is constructed as a count variable that tallies the number of paired points whose inter-point distance falls within each of several increasing distance ranges. The counts are normalized by N(N-1) and the ranges are cumulative. For any range, d, the statistic is given by:

[11] Here we take the residential lot (i.e. a smaller scale than the residential subdivision) as our unit of observation.

$$H(d) = \sum_{ij} I(d_{ij} \leq d)/N(N-1),$$

where $I(\cdot)$ is an indicator variable such that $I(d_{ij} \leq d) = 1$ if points i and j are within distance d of each other and zero otherwise. The distance interval d ranges from 0 to the extent of the region so that for $d = d_{max}$, all pairs of points are counted in H(d); the statistic ranges from 0 to 1.

Figure 2 shows the results of a test for randomness using a sub-area of Calvert County, one of the exurban counties in our Maryland study area.[13] The mean statistic from the multiple simulations of the pattern under the random hypothesis is plotted against the empirical statistic (or empirical distribution function, EDF, as it is sometimes called in the literature). In addition, the statistic from each of the simulated patterns is plotted against the mean to generate a rough indication of a confidence interval. Interestingly, the observed pattern of development mimics a random pattern of development almost exactly. This is evidenced by the roughly linear plot of the EDF vs. the mean random statistic. The result is interesting because of what we know to be true about underlying influences that cause development to be clustered, most obvious of which is the nature of residential subdivisions and the spatial correlation of exogenous features of the landscape. If there were only these positive influences and randomness, the resulting plot would, at least for some distance ranges, lie above the 45° line, indicating higher levels of positive correlation than in a fully random pattern. Because this is not the case, it is at least suggestive of another process – one which generates negative correlation among the observed *clusters* of development – that offsets the degree of positive correlation in the data.

The second hypothesis that we "test" is the monocentric hypothesis, which is represented by a simulated pattern of development generated as the result of maximizing

---

[12] Here we use the same area and total number of points found in the observed pattern of development within the region in order to compare the simulated patterns with the actual pattern.

[13] Computational limitations necessitated analysis of a smaller region.

accessibility to the central employment district. In this case, we again use the sub-sample from Calvert County and take distance to Washington D.C. as an inverse measure of accessibility to the CBD. Again, a spatial statistic is used to represent the pattern generated under the null hypothesis, in this case the monocentric case, and we plot the monocentric statistic against the EDF. Figure 3 shows the result of this exercise. While it is not possible to generate something akin to confidence intervals in this case, it is clear that the actual data depart quite dramatically from the monocentric pattern. Specifically, the EDF lies well below the 45° line, indicating that the observed data is far more negatively correlated than it would be in the monocentric case. This is consistent with our earlier intuition that the monocentric model is a poor predictor of the observed pattern of exurban development.

3. Models of Interacting Agents

In developing a spatially disaggregate model of land use conversion with interacting land developers, we draw from interacting agent models that have recently been applied in economics to a wide range of economic and social phenomena, e.g. business cycles (Durlauf, 1991), aggregate inventory stocks (Scheinkman and Woodford, 1994), social pathologies (Glaeser, Sacerdote, and Scheinkman, 1996; Brock and Durlauf, 1995, 1997), Durlauf (1997), and employment status (Topa, 1996). In addition, Krugman (1996) and Arthur (1988) have developed models of industrial location with interaction effects.[14] While the range of topics is broad, the common underlying premise of these models is that interaction among many individual agents at a highly disaggregate level leads to the emergence of collective behaviors and patterns in social and economic systems. The principle challenge in modeling these interaction processes is the simultaneous feedback among agents' choices, which determines the aggregate choice level of the group and therefore, the equilibrium outcome. This interdependence among agents' choices leads to both theoretical and empirical challenges. From a theoretical

[14]Related to these models are agent-based simulation models, which have been developed to study the evolution of a variety of human social phenomena, including trade, migration, group formation, interaction with the environment, cultural transmission, and disease propagation (see Epstein and Axtell, 1996, for a review).

standpoint, transition equations that characterize the relationship among agents' choices can not be readily solved for equilibrium solutions, and therefore, analytical characterization of the dynamic process at an aggregate level is very difficult.[15] Empirically, the interdependence among agents' choices leads to an identification problem, referred to by Manski (1993, 1995) as the reflection problem. Namely, in trying to infer whether average group behavior influences the behavior of individuals belonging to the group, the cases are limited in which endogenous social effects can be distinguished from observed and unobserved exogenous effects that are for one reason or another correlated across the group.

The influence of interactions on aggregate outcomes can be made tractable by assigning structure to the interaction effect, a task sometimes accomplished by borrowing from models of interacting particle systems developed in the physical sciences. In particular, many of the economic models draw from models originally developed in statistical mechanics, a branch of physics that seeks to explain and predict the macroscopic properties and behavior of systems comprised of very large numbers of microscopic processes (e.g. the movement of molecules or atoms within a gas or liquid). The very large number of particles at the microscopic level make it impossible to represent these systems exactly and therefore statistical methods are used to model discrete microscopic processes in a probabilistic framework. The overall goal of statistical mechanics is to explain observable physical and chemical properties of a system in terms of these stochastic microscopic processes.

For this reason, modeling techniques from statistical mechanics and theories of interacting particle systems have proven useful in modeling the evolution of economic systems comprised of interacting, heterogeneous agents and in deriving testable hypotheses. For example, Brock and Durlauf (1995) and Durlauf (1997) adapt a mean field theory model from statistical mechanics to derive the implications of global

[15] In fact, many complexity theorists argue that analytical, equilibrium-based approaches are limited and that simulation offers a broader understanding of the dynamic process (e.g. see Arthur, Durlauf , and Lane, 1997; Epstein and Axtell, 1996).

interactions (i.e. interactions among all agents within a population) for equilibrium group behavior. They show that for a critical range of the positive interaction parameter, exogenous shocks or small changes in private utility may generate large changes in aggregate behavior due to the presence of multiple equilibria. In a subsequent paper (Brock and Durlauf, 1997), this theoretical result is used to formulate a testable hypothesis for the presence of positive social interactions. Topa (1996) models local informational spillovers among agents seeking employment by adapting a model of the contact process, which has been extensively studied in physics and mathematics. In doing so, the interaction is assumed to be a local effect between an agent and her four nearest neighbors and is governed by an "infection" or contagion parameter that determines the strength of the interaction effect among an agent and her neighbors. Theoretical results are used to illustrate the aggregate consequences of the hypothesized positive information spillovers, which are shown to result in positive spatial correlations among employed and unemployed agents in a steady state equilibrium. Glaeser, Sacerdote, and Scheinkman (1996) develop a local interaction model based on voter models, which also originate from the physical sciences, to interpret their empirical findings regarding the variation in crime rates across cities.

The interacting particle system approach is particularly appealing for modeling land use conversion for several reasons. First, these models are explicitly spatial in the relative sense that each agent is identified with a set of neighbor agents. Second, the process is viewed as stochastic and interaction among individual agents is a function of the discrete choices (i.e. the land use conversion decision) made by agents. The characterization of the problem makes explicit the similarity to standard economic models of discrete choice. The central difference here is the inclusion of a "social interaction" effect in the individual's objective function. Third, the focus is on understanding the consequences of a specified interaction among agents for the spatial distribution of agents at an aggregate level. Structure is assigned to the interaction effect depending on the maintained assumptions about the underlying economic process. For example, Brock and Durlauf (1995, 1997) specify a global interaction, in which an individual's decision is influenced by their expectation over the average behavior of a group comprised of very many individuals. On the other hand, Topa specifies a local interaction effect, in which the individual is positively influenced by the proportion of nearest neighbors that are in an employed state. As will become clear in Section 5, the maintained assumptions about the interaction effect are critical in determining whether the endogenous interaction effect can be identified empirically.

4. Dynamic Model of Residential Land Use Conversion

In developing a model of land use conversion, we start from the viewpoint of a profit-maximizing agent who owns an undeveloped[16] land parcel and makes a discrete choice in every period regarding the subdivision of the parcel for residential use.[17] Conditional on the parcel being undeveloped in the present period, the agent's decision is simplified to a binary choice of converting her parcel to residential use or keeping her parcel in an undeveloped use, such that the present discounted sum of all future expected returns from the land are maximized. Once converted, developers supply residential lots to households, who make location decisions by choosing a bundle of attributes associated with a particular location to maximize utility. Therefore, the developer-agent faces a dynamic optimization problem in which she will choose to convert the parcel to residential use in period t if the present discounted value of the parcel in residential use net of conversion costs and opportunity costs is maximized over an infinite time horizon.

To develop this model, first define two possible states for a parcel of land, $s(i,t) \in \{-1,1\}$, where -1 represents an undeveloped state and 1 represents a residential (developed) state. Viewed at the beginning of period T and given an undeveloped parcel in time T-1, the expected net present value to the developer *as of period T* of converting parcel i to residential use in the upcoming period is given by $\pi_{1|-1}(i,T)$, where

[16] Here, undeveloped uses include agricultural and other resource production-oriented uses of the land, e.g. commercial forestry, as well as land in natural states.

[17] In the reminder of the paper, we treat the undeveloped parcel as the unit of observation and therefore, the decision that is modeled is the developer's decision to subdivide her parcel into multiple residential lots or

(1) $\pi_{1|-1}(i,T) = R(i,T) - C(i,T)$

and R(i,T) is the expected sales price of the subdivided residential lots and C(i,T) is the cost of conversion. In addition, define the expected net present value, at the beginning of period T, of keeping the parcel in the undeveloped state in perpetuity as $\pi_{-1|-1}(i,T)$. Here,

(2) $$\pi_{-1|-1}(i,T) = \sum_{t=0}^{\infty} A(i,T+t)\delta^{t}$$

where A(i,T) is the expected returns from an alternative, undeveloped use of the land in period T[18] and δ is the discount rate.

The basic optimization problem can then be specified as:[19]

(3) $$\Psi(i,T) = \max_{s(i,T)\in\{1,-1\}} \{V_{s|-1}(i,T)\}$$

where

(4) $$V_{1|-1}(i,T) = \pi_{1|-1}(i,T) - \pi_{-1|-1}(i,T)$$

(5) $$V_{-1|-1}(i,T) = \delta\Psi(i,T+1)$$

In this formulation, $V_{1|-1}(i,T)$ equals the returns from converting in period T minus the opportunity cost – the foregone returns from the undeveloped state in perpetuity. $V_{-1|-1}(i,T)$ is the expected net present value of the best option in the future, if development is

---

to keep her parcel in an undeveloped use. We do not explicitly deal with commercial development which makes up a very small proportion of developed land use and tends to follow residential sprawl.

[18] While the non-developed use is referred to as "undeveloped," we have in mind some productive use of the parcel in a non-developed state, e.g. agriculture or commercial forestry.

[19] This set-up is analogous to models of stochastic dynamic decision making found in the dynamic programming literature (e.g. see Provincher, 1995).

not undertaken this period.

Land is treated as a bundle of heterogeneous attributes. The expected values of land in both the developed and undeveloped uses are functions of subsets of these attributes. In addition, the developer is likely to form expectations over future changes in residential land prices by taking exogenous growth pressures into consideration (e.g. future population and income changes). Here we simplify the consideration of growth effects by assuming that increases in regional population lead to an outward shift in the demand for regional housing in each period. Developers are assumed to have homogeneous expectations over these future increases in demand.

In specifying the expected value functions, $\pi_s(i,T)$, a key distinction is made between the influence of landscape features that are treated as exogenous characteristics of the landscape and spatial externalities that are generated by the surrounding pattern of land uses within a defined neighborhood of parcel i. Because these externalities are generated by the neighboring land use pattern, these effects are clearly endogenous to the land use conversion process. Given this distinction, the developer's expected value of parcel i in land use s in period T, $\pi_s(i,T)$, is specified as a function of the following components: (a) exogenous parcel attributes, $H_s(i,T)$, which includes features specific to parcel i, e.g. size, vegetation, soil type, slope, services available to the parcel, and exogenously determined location features, e.g. distances to cities, markets, and waterfront. These exogenous variables are likely to be constant over time, but spatially correlated over parcels; (b) the net influence of spatial externalities generated from the land use of parcels located within a defined neighborhood of parcel i, which may either increase or decrease the value of parcel i in land use s in period t, $I_s(i,T)$; and (c) a random component, $\varepsilon_s(i,T)$, which include unobservable parcel characteristics which also are likely to be spatially correlated over parcels and which may not change much over time. Taken together, these imply that the developer's expected values for the π's are given by:

$$(6) \qquad \pi_1(i,T) = R(i,T) - C(i,T) = W[H_1(i,T), I_1(i,T), \gamma_1(T), \varepsilon_1(i,T)]$$

and

$$\pi_{-1}(i,T) = \sum_{t=0}^{T} A(i,T+t)\,\delta^t = \sum_{t=0}^{T} A[H_{-1}(i,T+t), L_1(i,T+t), \gamma_{-1}(T+t), \varepsilon_{-1}(i,T+t)]\delta^t$$

where $\gamma_1(T)$ is the expected real change in residential prices due to growth pressure in period T relative to T-1 and $\gamma_{-1}(T)$ is the expected real change in prices of resource products produced by the alternative land use.

Given the stochastic nature of the agent's decision making process, the land use conversion process can be described in terms of the transition probabilities of land parcels. Based on (3), (4), and (5), the probability that agent i will convert her parcel from an undeveloped to developed state is:

(7) $\Pr ob\{s(i,T) = 1 \mid s(i,T-1) = -1\} = \Pr ob\{\pi_{1|-1}(i,T) - \pi_{-1|-1}(i,T) > \delta\Psi(i,T+1)\}$

Development is viewed as economically irreversible due to exceedingly high costs of reconversion from a developed to an undeveloped state. Reversion is treated in the model as a random shock of very small or zero magnitude:

(8) $\mathrm{Prob}\{s(i,T) = -1 \mid s(i,T-1) = 1\} = \omega$ where $\omega << 1$

The assumptions of increasing growth pressures in the region, combined with the inclusion of spatial externality effects in the expected value functions, have some interesting implications for the developer's expectations formation over future returns from conversion. On the one hand, the increasing growth pressure effect implies that developers will find it optimal to postpone development in period T, *ceteris paribus*, even if the returns from conversion are greater than the returns from keeping the parcel in an undeveloped state. However, other factors may induce developers to convert their parcels earlier. For example, congestion externalities from increasing levels of neighboring development may decrease the residential value of parcel i, in which case the developer may find it optimal to convert her parcel in an earlier period. In addition, government land use controls often become more restrictive as regional development increases, e.g. developers may be required to pay development fees or cover other costs of development, which would also prompt developers to develop earlier. While the interplay of these effects may be quite complex, we assume that the developer's expectations over the time paths of these forces is such as to make the following arbitrage condition between converting in T and converting in T+1 as the decision rule:

$$\text{(9)} \qquad \text{Prob}\{s(i,T)=1; s(i,T-1)= -1\} = \text{Prob}\{\pi_1(i,T) - \pi_{-1}(i,T) > \delta\pi_1(i,T+1) - \delta\pi_{-1}(i,T+1)\}$$

where $\pi_s(i,T+1)$ is defined as the present value of state s evaluated at time T+1.

Assuming an additive error and recognizing the implicit definitions of the π's as given in (6), equation (9) can be rewritten as:

$$\text{(10)} \qquad \text{Prob}\{s(i,T)=1; s(i,T-1)= -1\} = \text{Prob}\{W(i,T) - \delta W(i,T+1) - A(i,T) > \varepsilon(i,T) - \varepsilon(i,T+1)\}.$$

After discussing the possible nature of the interaction terms imbedded in W and A, we consider first what the form of this model implies for the evolution of land use using simulation experiments and then we consider the difficulties of empirically testing whether the proposed interaction effects actually exist.

Interaction Effects

The important dynamic feature of this model is the inclusion of an interaction effect, $I_s(i,T)$, which introduces the possibility of interdependence among developers' residential conversion decisions over time. In specifying the interaction effect generated by spatial externalities, we adopt a variant of the Ising model from statistical mechanics (Bar Yam, 1997). This model of pattern formation posits a "spin-spin" interaction among particles i and j as a function of distance between the two particles. The term "spin-spin" refers to the interaction of two particles, such that the direction of spin of particle i is influenced by the direction of spin of particle j and vice versa. Because pair-wise interaction between particles is treated as a function of the state of each particle (one of two alternatives) and the distance between them, this model generalizes in a straightforward manner to the interaction among parcels, given that each parcel can be in one of two land use states and the interaction effect is likely a function of distance between the parcels.

A general expression for an Ising-type interaction between particle i and particle i's neighbor j in any period T, can be written as:

$$(11) \qquad I_s(i,T) = \sum_{j \in N_i} J_{s(i)}(d_{ij}) s(i,T) s(j,T)$$

where N_i includes all particles in the neighborhood of particle i and $J_{s(i)}$ is an interaction parameter, which is a function of d_{ij}, the distance between i and j, and is conditioned on the state of i. This specification of the interaction effect takes advantage of the {-1,1} indexing assigned to the two states. If i and j are in the same (different) states, then the product of s(i,T)s(j,T) is positive (negative). The sign of $J_{s(i)}$ determines whether particles in identical states have positive or negative interaction terms. Written in this form, the interaction effect of j on i varies in sign but not in magnitude with the state of j.

In our application, the particular interaction effect is a function of the types and magnitudes of spatial externalities between land uses, a set of interaction effects potentially more diverse than particle interaction. As such, we adapt (11) to reflect this added complication. One generalization is that we allow $J_{s(i)}$ to depend not only on the state of the "receiving" parcel i, but also on the state of the "emitting" parcel j. Thus we do not impose the constraint that $J_{s(i)}$ be equal in magnitude over different types of neighbors. In expressions (12) and (13) below, we specify two interaction effects that correspond to the total interaction effects of all neighbors on parcel i when i is in an undeveloped and a developed state respectively.

In constructing the interaction effect for parcel i given that it is in an undeveloped state, we hypothesize that other undeveloped parcels within parcel i's neighborhood increase the value of parcel i in an undeveloped state. Additionally, we propose that developed parcels within this neighborhood may decrease parcel i's value in an undeveloped state. This case can be made most convincingly if the undeveloped use is agriculture. The presence of other agricultural lands around parcel i may increase the value of parcel i in an agricultural use because they contribute to the viability of the area as an agricultural area, and therefore, services and other infrastructure that support agriculture may be more likely to be located within close proximity of parcel i. On the other hand, the fragmentation of the surrounding landscape with residential development may decrease the agricultural value of parcel i for several reasons, e.g. increased congestion on the roads from non-farm residents and restrictions on farming activities because of nuisance complaints from residences.[20]

Given these considerations, the interaction effect for parcel i, given that i is in an undeveloped use, is specified as:

$$(11)\, I_{-1}(i,T) = \sum_{j \in N_{i,-1}} \left(\frac{1-s_j}{2}\right) J_{-1}^{j=-1}(d_{ij}) s_i s_j + \left(\frac{s_j+1}{2}\right) J_{-1}^{j=+1}(d_{ij}) s_i s_j$$

For all neighbors in the same state as i (the undeveloped state), the second term drops out and the first term captures the interaction effect. Because the product $s_i s_j$ is positive, the sign of the interaction effect is determined by the sign of $J_{-1}^{j=-1}$. The effect of neighbors in the alternative developed state is captured by the second term, in which case the product of $s_i s_j$ is negative and the sign of the interaction effect will be opposite that of $J_{-1}^{j=+1}$. The case of positive spillovers from neighboring undeveloped land and negative externalities from neighboring residential parcels within the neighborhood defined by $N_{i,-1}$ suggests that both the interaction parameters will be positive, but by

---

[20] If the undeveloped state of the parcel were forest, air pollution from neighboring residential congestion might damage timber stands.

writing it in the above form we allow them to be of different magnitude.

The total effect of neighbors' externalities on parcel i's value in a residential state can be expressed in parallel form as:

$$(12) \qquad I_{+1}(i,T) = \sum_{j \in N_{i,1}} \left(\frac{1-s_j}{2}\right) J_{+1}^{j=-1}(d_{ij}) s_i s_j + \left(\frac{s_j+1}{2}\right) J_{+1}^{j=+1}(d_{ij}) s_i s_j$$

The first term is non-zero for neighbors in the undeveloped state and, because $s_i s_j$ is negative in this case, the sign of the interaction effect will be opposite that of $J_{+1}^{j=-1}$. A reasonable assumption is that neighboring parcels in the undeveloped state will generate positive "open-space" amenities for residences, suggesting that $J_{+1}^{j=-1}$ would be negative. The spatial externalities of neighbors in a developed use may be more complex, however, in that both positive and negative spillovers from neighboring developed uses are likely. For example, positive spillover effects between developed parcels may include various "community" spillover effects; people may find it desirable to live in close enough proximity to others so as to feel part of a community, to have the social benefits of neighbors. In addition, there may be positive effects associated with a critical density of residents in an area, which may be necessary to attract public and private services to the area. Negative spillover effects may occur between neighboring developments, however, due to congestion.

By making J an explicit function of distance between parcels, we signal that the strength of the interaction varies with proximity. Presumably the strength is a decreasing function of distance, but different externalities may have different rates of decay with distance. This suggests that where interaction effects are made up of both positive and negative externalities, as we hypothesized is likely with the effects of neighboring development on a parcel's value in the developed use, the net effect may change sign over the range of the relevant neighborhood around a parcel.

Figure 4 illustrates a hypothetical case for the effects of positive and negative externalities from developed parcels on the expected value of parcel i in residential use. Here positive externalities from neighboring developed parcels are posited to be stronger than negative externalities for very near parcels but are assumed to drop off quickly as distance from parcel i increases. In contrast, the influence of negative externalities from neighboring development decreases more slowly and over a longer range. As a result, the net effect of development located within radius a on parcel i's residential value is positive. At distance a, the spillover effects exactly offset each other and the net effect of development located at distance a from parcel i is zero. The net effect of development within the "doughnut" neighborhood between distances a and b is negative. In the general case, with the additional effects from neighboring undeveloped land included in the interaction effect as well, the net effect of these spillovers on the value of parcel i in residential use will depend on their sign and relative magnitudes.

Spatial Implications Of Interaction Effects

We are interested in determining the implications of the above model of interacting micro-level agents for the regional landscape pattern. In particular, we are interested in determining whether there are conditions under which the above model would predict the type of land use change – highly fragmented sprawl – that many regions in the U.S. have witnessed over the last several decades. If there are conditions under which this model is consistent with observed patterns, then the model may suggest hypotheses that can be tested empirically.

One possible approach would be to formulate the joint conversion probability of all land owners and attempt to characterize the qualitative aspects of an equilibrium solution(s) – in the sense of the social interaction theory. In these models, interaction between agents is simultaneous: i's state affects j's decision about state and j's state affects i's state decision. Equilibrium is characterized as a set of states of all agents such that taking into account all other agents' states, no individual wishes to change state. Techniques from statistical mechanics may prove useful in characterizing such equilibria, much like the approach of Brock and Durlauf (1995) and Durlauf (1997), in which they use results from mean field theory to identify a self-consistency equation that characterizes the equilibrium solution(s). However, because we are interested in capturing the influence of both attractive and repelling effects (vs. attractive interaction only) and because we assume interaction is limited to a local (vs. global) neighborhood, our problem is more complicated.[21]

However, while the analysis of an equilibrium may be of theoretical interest, it may not be very meaningful or very interesting for the empirical reality of our problem. First, we argue that the interaction process in our case is not simultaneous. The development process takes time, so there will be a temporal lag between stimulus and response. Thus the process is better characterized as a recursive rather than a simultaneous one. Second, the land use conversion process in non-stagnant regional economies is an evolving one that is best characterized as "out of equilibrium" in the social interaction terminology. There are several features of our problem that make this so. First, population growth and changing incomes introduce a continual impetus for conversion. Second, reversion (change from a developed to an undeveloped state) is generally considered economically infeasible. Given these two features, the only absorbing state for our system is a particularly uninteresting one -- one in which all the land available for development has been developed and a steady state equilibrium land use pattern is reached.

For these reasons, we focus on the out-of-equilibrium dynamics of the land use conversion process in characterizing the implications of agents' interaction for the regional land use pattern. Using the model laid out in the previous section, the evolution of an aggregate land use pattern is simulated over many periods and many agents as the

---

[21] As far as we can tell, the vast majority of theoretical results in statistical mechanics and interacting particle systems seem to concern processes with positive (i.e. attractive) interactions among particles in like states. Therefore, adapting these models for our purposes, in which we are interested in both attractive and repelling effects among parcels in the developed land use, is not straightforward. In addition, a mean field theory approximation to the interaction structure is not appropriate in our case, as agents are assumed to interact with a small subset of neighbors and therefore, the mean group effect is not independent of any one agent's choice.

cumulative result of individual profit-maximizing decisions made by developers at the parcel level. A type of cellular automaton is used to capture the explicitly spatial relationships among agents.[22] Ranges of parameter values are identified for which the evolution of a regional land use pattern is qualitatively similar to the observed sprawl pattern of land uses in exurban areas. In doing so, a correspondence between a hypothesis regarding interaction among agents and a hypothesis regarding the spatial distribution of residential patterns is drawn.

Based on the model developed in the previous section, developers' conversion decisions are represented in probabilistic terms, and therefore we need a means of translating conversion probabilities into actual land use states. The total amount of conversion in any given time period is clearly a function of regional growth pressures, e.g. population and income changes. In what follows, we simplify the regional growth effects by assuming that the regional demand for new housing is constant, and we define time periods such that one new conversion occurs in each period. Given this, we assume that the parcel with the highest probability of conversion in each time period is the parcel that is converted. Once converted, the probability of a parcel's re-conversion to an undeveloped state is assumed to be very close to zero.

Assuming that land use parcels are distributed on a 30x30 grid of equal area cells, a cellular automaton type of model[23] is used to simulate the stochastic evolution of a residential land use pattern. The effective neighborhood of influence for each parcel is locally defined, so that beyond a certain distance, the land use states of parcels i and j are independent. Interaction effects are weighted by inverse distance so that nearest neighbor interactions have the greatest effect on a parcel's value in a given land use. In

[22] This is the same type of approach used by some complexity theorists to study the emergent properties of group behavior as a function of the interactions among many heterogeneous agents, each operating in a localized, spatially explicit environment (see Epstein and Axtell, (1996) and Tesfatsion (1997) for more discussion on the use of agent-based models in modeling complex social processes).

[23] The difference between this type of simulation and the more standard cellular automata is the lack of simultaneity in the evolution process. This arises because of (1) the lagged effect among agents making land use decisions and (2) exogenous factors and growth pressures in the region, which also influence the state of a parcel. We used Grid, a program within the Geographical Information System package Arc/Info, and Arc Macro Language to perform the simulations.

addition, the following assumptions regarding the signs and neighborhood extents of the externalities were made: (1) the net influence of neighboring development on the value of parcel i in the developed use is positive for "nearest neighbor" or "first order" spatially lagged parcels, negative for "higher order" spatially lagged parcels,[24] and zero for parcels that are sufficiently far apart such that they are located outside the effective neighborhood of influence; (2) the net influence of neighboring open space parcel i's value in the developed state is positive for all spatially lagged parcels located within the effective neighborhood; and (3) the net influence of neighboring parcels in the developed (undeveloped) state is negative (positive) on the value of parcel i in the undeveloped state for all spatially lagged parcels located within the effective neighborhood.

As outlined below, the residential land use pattern is simulated for different parameter values. In all cases, the initial conditions are specified by an almost fully undeveloped region, in which one parcel is initially in a residential state. In order to compare the outcomes of the simulations, we start with the same, randomly determined initial state. Given that one conversion occurs in each period and that once converted, a parcel is rarely unconverted, the simulation reaches a "natural" end point after 900 time steps (i.e. the entire region is developed). Because the interest here is in identifying the qualitative pattern that emerges for different parameter values, it is unnecessary to simulate to this final end state. We find via trial and error that approximately 100 time steps are sufficient for this purpose. Multiple simulations were performed for several different scenarios, including the cases in which two exogenously determined features – a city center and a road that cuts through the region to the city center – differentiate the region. Each parcel's value in residential and undeveloped uses is an additive function of these two exogenously determined effects together with the interaction effects from development. We assume that the location of both the road and the city increases parcel i's residential value more than the undeveloped value.

---

[24] First order and higher order spatial lags can be used to represent different neighborhoods around parcel i. For example, if parcels are indexed in space as i-2, i-1, i, i+1, etc. then, with respect to parcel i, parcel i-1 and i+1 are first order spatial lags and parcels i-2 and i+2 are second order spatial lags.

Figures 5-6 illustrate the case in which the influence of an exogenously defined city center is considered and Figures 7-8 show results for which both a city center and road are considered. In general, varying degrees of clustered patterns emerge, depending on the relative magnitudes of the endogenous and exogenous effects. In particular, the "tension" between the attractive influences – both endogenous and exogenous – on one hand and the endogenous repelling effects on the other determines the degree to which clustering and scatteredness emerge. For example, Figure 6 illustrates the competing effects between the road and a city center, which create positive correlation among residential parcels, and the repelling interaction effects, which lead to clustering. The location of development is clearly determined by the exogenous landscape features, but the repelling effects cause development to leapfrog away from the city center to an undeveloped area that is adjacent to the road. The result is a clustering of development around the city and road. In contrast, Figure 5 shows the case in which the exogenous influences from the city fully dominate the repelling effects. This contiguous agglomeration of development around the city center is consistent with the predictions from the monocentric model.

Comparison of Figure 8 to the observed pattern of residential subdivisions in Map 2 shows a qualitative similarity between the two patterns. In both cases, the residential parcels are located in clusters, some of which are adjacent to the road. In the simulation, this pattern was generated via a mix of endogenous and exogenous effects: (1) a positive effect from proximity to the road and city and (2) relatively strong repelling effects generated from negative development externalities and positive open space effects. We conclude that the interaction hypothesis -- with relatively strong repelling effects -- is a *viable* explanation of the observed residential pattern, which provides us with a testable hypothesis regarding the actual role of interaction effects in individual's conversion decisions. It is this question to which we turn in the next section.

5. Empirical Evidence of Interaction Effects

Comparing the actual land use pattern with a simulated one based on a model of interacting agents is one thing; empirically testing for the presence of spatial interaction effects is something else. In attempting to do so, we run headlong into the difficulty of distinguishing between endogenous interaction effects from neighboring land use decisions ("true" state dependence) and exogenous heterogeneity from unobserved characteristics associated with agents or land parcels ("spurious" state dependence). This problem is related to the identification problem - known as the "reflection" problem - described by Manski (1993, 1995). This form of the identification problem has been most extensively discussed in the social/behavioral modeling literature in economics. In this literature researchers typically wish to distinguish, empirically, the effects on individual decisions of social interactions and correlated exogenous characteristics across individuals. For example, do students perform similarly on tests because each student's scholastic achievement is influenced by the achievements of her peers or, are students' achievements similar because they share common experiences, e.g. the same teachers or similar family backgrounds? As Manski shows, the conditions under which these effects can be distinguished or "identified" are limited, due to the contemporaneous interaction of the individual's choice and the "spatially lagged" endogenous social effect.

A related identification problem, but one in which the endogenous effect is specific to each individual over time, has been studied by Heckman (1981, 1991) and others (e.g. Honoré and Kyriazidou, 1998). In this case, the dependence is not over space, but rather over time and therefore the structure is recursive. For example, if an individual's unemployed state is influenced by previous states of unemployment, then this is evidence of what Heckman terms state dependence (or path dependency). While this avoids the difficulties of simultaneity that arise in Manski's reflection problem, the problem of discriminating between "true" state dependence (in this case, unemployment in past periods) and "spurious" state dependence due to heterogeneity across individuals (e.g. differences in education and ability) remains.

As we will soon explain more fully, our problem cuts across these two streams of literature. We are interested in determining whether past decisions about surrounding land use influences the land use decision of the owner of a given parcel. Specifically, we hypothesize that the land use decision about the parcel at location i in time t is potentially a function of the land use decisions in previous time periods of parcels $i \pm s, s \in N$ where N is a neighborhood of the parcel at location i. Thus, we are looking for a spatially lagged endogenous effect that is recursive (not simultaneous). However an "identification" problem arises because persistent (i.e. time invariant) but unobserved characteristics of the parcel at location i will likely be correlated with the persistent unobserved characteristics of the surrounding parcels. It will therefore be difficult to distinguish between the effects of true spatial externalities from surrounding land uses and the effects of unobserved characteristics that are spatially correlated across observations. First, let us outline some findings from the identification literature.

Strategies for Solving Identification Problems

The social interactions literature deals with the issue of contemporaneous correlation in "space"[25] among individuals' choices. Manski (1993, 1995) outlines three competing hypotheses that explain observed correlation among individuals' choices: (1) endogenous effects -- the propensity of an individual to behave in some way varies with the behavior of the group; (2) contextual effects -- the propensity of an individual to behave in some way varies with the exogenous characteristics of her reference group; and (3) correlated effects -- individuals within the same group tend to behave similarly because they have the same exogenous individual characteristics or face similar exogenous institutional environments. For a linear model with simultaneous interactions, Manski shows that the separation of these effects is not possible, due to the "reflection problem." More specifically, if an individual's expected choice varies with the mean choice of the individuals that comprise the individual's reference group, then the variable that captures the expected mean choice of an individual enters the "social equilibrium" regression equation as both a dependent and an independent variable. Manski warns that

[25] The notion of space is not necessarily limited to geographic distances, e.g. it may refer to social or economic distances among individuals.

this result may also carry over to non-linear models[26] and strong assumptions about the interaction process may be necessary in order to ensure identification.

The extent to which issues of this sort have been resolved in the empirical literature on social interactions in a discrete choice setting varies. The reflection problem has been addressed by some by assigning structure to the interaction such that the simultaneous "reflection" between an individual's mean choice and the group's mean choice is in some sense deflected. The model of global interactions developed by Durlauf (1997) and Brock and Durlauf (1995, 1997), in which the mean choice of many individuals enters an agent's utility function, is an example of this. In their "mean field theory" approach, the simultaneity problem is handled by assuming that the mean group effect is independent of any one individual's action. Alternatively, the problem can be handled by assuming that the interaction is recursive, in which case an individual's mean choice is a function of a temporally lagged social interaction effect. So long as the process is observed out of equilibrium, the reflection problem does not arise. Manski stresses the need for underlying economic justification of such assumptions in an empirical setting for social interaction phenomena.

Heckman (1981) outlines a second type of endogenous interaction model, in which the discrete choice of individuals may be influenced by their own past states. The recursive nature of the process avoids Manski's reflection problem, but identification of the endogenous effect is still an issue due to unobserved heterogeneity across individuals that leads to "spurious" state dependence. Heckman develops a "mirror image" approach to separating these effects in a time-series setting with a binary choice, in which the symmetry properties of the bivariate normal distribution are used to aid in the separation of the endogenous interaction parameter from the correlation coefficient that relates the errors across time. In this case, a panel data with at least two time periods is necessary to enable identification. However, the applicability of this approach is limited, due to some

---

[26] Manski (1993) outlines the cases under which the parameters of a model with all three effects are identified for the linear regression model, but states that the conditions under which these parameters are

strict assumptions about the nature of the process, e.g. observations in the first period for which data are available are assumed to be uncorrelated with the past. In addition, the separation strategy relies on observing probability "runs" in both directions – meaning the estimation of both (1) the joint probability of choosing A in period 1 and B in period 2 and (2) the joint probability of choosing B in period 1 and A in period 2.

A second approach to separation of true vs. spurious state dependence, which draws on Chamberlain's (1985) notion of using probability "runs," is outlined by Honoré and Kyriazidou (1998). In this case, the unobserved heterogeneity that causes correlation over time is assumed to be time invariant and specific to the individual. A strategy analogous to "first differencing" is applied to the binary logit model with a lagged dependent variable and unobserved fixed effects, where the first differencing cause the latter effects to drop out of the model. Given observations across individuals over four time periods, an identification scheme is developed for the conditional probability of a particular sequence of choices over time, given that either that particular run or the reverse run occurs. Brock and Durlauf (1997) argue that the extension of these approaches to a spatial setting is possible, given either a conditional model of interactions in which next period's utility depends on the average group choice made in the current period or the simplification of contemporaneous interactions via the mean field theory approach.

Identification Issues in the Land Use Conversion Model

We argue that our problem is not a true reflection problem because (1) interaction among agents is lagged over time and (2) the observed process of land use change is "out of equilibrium." The lag is due to the nature of the development process: the developer must incur the time and money costs of conversion, which include securing permits for subdividing and developing the land, as well as the costs associated with the physical aspects of conversion, e.g. clearing and grading the land, building residential structures,

---

identified for non-linear models, e.g. a binary discrete choice model, with contemporaneous effects have not been established.

and providing any public services (e.g. utility lines, roads) that may be required of subdivision development. The second assumption stems from the observation that growth pressures, in the form of increases in population and income, imply that the land use conversion process will be subject to continual influxes of new conversion pressure.

Our problem is best described as a "cross-section" of the two types of endogenous interaction models outlined above. Spatial externalities imply a spatially lagged dependent variable, which is also temporally lagged due to the nature of the development process. The temporal lag prevents Manski's reflection problem. However, since the heterogeneity (omitted explanatory variables) in our model are likely to be both time invariant and correlated over space, we still can not easily distinguish between true and spurious state dependence. Therefore our problem remains similar to the social interaction models, both in terms of the source of endogenous effects (i.e. associated with neighboring agents' choices) and the correlation of exogenous variables over space. Analogous to the correlated effects among individuals described by Manski, heterogeneous landscape characteristics that vary over space may generate spatial correlation among neighboring land use decisions. If unobserved, these effects will make decisions appear related, even if they are not, and therefore complicate our ability to discern true state dependence.

Our problem is further complicated in three important ways. In particular, while growth pressure effects "save" us from the pitfalls of the reflection problem, they introduce additional identification complexities. Because growth effects imply that conversion is uni-directional over time (from undeveloped to developed), a parcel's neighborhood can only be increasing in the amount of development. At the same time, however, increasing population and income growth will cause the likelihood of conversion to be increasing over time, irrespective of changing states of neighboring parcels. Our empirical task is to isolate the potential interaction effect separate from the effects of *both* the spatially correlated exogenous variables and the temporal effect of increasing growth pressure. A second complication introduced by the growth pressure effect relates to some of the identification strategies mentioned above (e.g. Heckman

(1981) and Honoré and Kyriazidou (1998)). Because these strategies rely on observing both types of "switches" from state A to state B and vice versa, they do not apply in our case due to the irreversible nature of our problem.

Finally, our problem is complicated by the unknown nature of the interaction effects – they could be either positive or negative or more likely, a combination of the two that is likely to vary over distance. Therefore our problem is similar to the Manski and Heckman identification problems only if net positive externalities result. In this case, the two effects – positive interaction effects and positive correlation of exogenous variables over space – will reinforce each other and therefore, in the presence of correlated effects, the parameter estimate on the interaction term will be biased upward. However, in the case of net negative externalities, the two effects will work against each other and it is possible that the interaction effect will be washed out by the competing positive correlation effect. Because we can't be sure of either the direction or magnitudes of the hypothesized externality effects, this problem underscores the importance of controlling for possible unobserved correlation.

Empirical Approach

Spatially articulated, micro-level data on land parcels is available to us from the Maryland Office of Planning's geo-coded tax assessment data base. To estimate the naïve version of a model that would test for spatial interaction effects of the form described above, we considered the conversion decisions associated with a set of parcels in the rural-urban fringe area around Washington, D.C. Our preliminary analysis considered only one year – 1992, and three counties – Calvert, Charles and Howard. These counties have been experiencing extremely rapid growth. Our naïve model (one that does not address the potential identification problem) considered all parcels that could be subdivided as of January 1992 and, given zoning regulations, were large enough to accommodate a housing development of at least 5 houses. There were 3901 such parcels, only 2.5% of which were subdivided in 1992.

We attempted to explain the conversion decision using a binary discrete choice model as a function of factors. The first is a nonlinear function of the commuting distance from the centroid of the parcel to Washington, D.C., along the road network. The second factor is the percent of the surrounding land use in a developed rather than an undeveloped use, as of January 1992. We defined "surrounding" as a given radius around the parcel and considered a few different values for this radius. In addition, we attempted to correct for county specific policies that might alter the likelihood of conversion by including county dummy variables. Everything else is purposefully relegated to the error term.

The results suggest that, other things equal, land has a significantly higher likelihood of conversion in Calvert County than the other two counties. In addition, and not surprisingly, the probability of conversion declines with distance to Washington, but at a decreasing rate. The coefficient associated with the "lagged interaction" variable, represented here by the amount of development surrounding a parcel, is negative and significantly different from zero. This suggests that a higher amount of development around a parcel is associated with a lower probability of development. This effect is significant within a kilometer radius of the potentially developable parcel, but loses significance as the proposed radius of interaction increases.

While our sample must be investigated with more care, we are still inclined to speculate about the role of the identification problem. A priori, we expect that the omitted variables are characteristics of location, which – almost by definition – are likely to be positively spatially correlated. That suggests that we would expect to find clustering of development in the absence of any type of interaction effect. It is difficult to postulate a source of negative spatial correlation among omitted variables (e.g. scenic views from hilltops are one of the few that come to mind). As such, if we have no interaction or only positive interaction effects, then we would expect to find a positive coefficient on our spatially lagged, temporally lagged development measure. The fact that we estimate a significant negative coefficient argues for the presence of negative spatial interaction caused by negative externalities among developed land uses.

However, even if these preliminary results hold up under further scrutiny, we will need to address the identification problem to be confident that our estimated interaction effect is not biased.

6. Conclusion

We attempt in this paper to develop an economic model of land use change as a spatial process, in which the important spatial element is the interaction between agents. In contrast, the standard approach to residential location patterns, as described by the monocentric models, assumes the an agent's location relative to some exogenous entity, such as the central business district, is the determining spatial factor. Such an approach is unable to offer robust explanations of the observed changes in spatial land use patterns. While our model is a relatively simple alternative, it has the potential for offering robust explanations of a variety of different development patterns as the consequence of endogenous interaction effects and exogenous features of the landscape. Simulation exercises demonstrate that an interaction model can explain the observed evolution of *increasing* fragmentation of residential development patterns in some exurban areas as the result of growth pressures and relatively strong repelling effects, generated from positive open space effects and negative development externalities.

Empirical evidence is more difficult to come by, however, given the pervasive problem of distinguishing between correlated effects and true spatial interaction. Distinguishing between the endogenous spatial externality effects and the effects of unobserved heterogeneity is important from a policy perspective, in which the ability to predict future changes in land use pattern is key. The presence of endogenous interaction effects combined with growth pressures would imply a very different type of spatial process than that assumed by the monocentric and policentric models. The influence of spatial externalities would suggest that the land use conversion is path dependent. At any given point in time, land use pattern changes would be influenced by the types of interactions among economic agents, determined by the location of agents relative to each other. Therefore, changes in the spatial pattern of land use would influence future changes in the pattern due to the presence of spatial externalities. Add to this the fact that the conversion process is further characterized by growth pressures, which suggests that conversion is likely to be economically irreversible over time due to increasing demand for housing. As a result, rather than evolving to a unique, well-defined, long-run spatial equilibrium of land uses, many different spatial configurations of land use could be generated by a process influenced by such effects. Given the irreversible nature of the process, the long-run spatial equilibrium in this case is a trivial one -- one in which all the land available for development is developed. This underscores the importance of understanding the dynamic process of land use change when the system is out of equilibrium.

State and local governments throughout the U.S. are attempting to come to terms with new emerging patterns of land use. While the stated goal of many local and state governments is to "contain" growth in some fashion, it is unclear how effective some of the current growth management regulations are in attaining such goals. To the extent that policies may regulate development pattern without consideration of individuals' preferences, it could be that the problem of negative development spillovers is actually exacerbated. For example, this could be the case if undesirable policy prescriptions were to unintentionally create more incentives for individuals to seek variances or waivers to land use zoning and, in so doing, led to increased fragmentation of the landscape. In this case, it seems that a more effective approach would be to recognize individuals' preferences and design policies that minimize such negative development spillovers. For reasons such as these, designing effective growth management policies relies on distinguishing among true and spurious interaction effects.

References

Alonzo, W. *Location and Land Use,* Harvard University Press, 1964.

Anselin, L. *Spatial Econometrics: Methods and Models.* The Netherlands: Kluwer Academic Publishers, 1988.

Arthur, W.B. "Urban Systems and Historical Path Dependence." *Cities and Their Vital Systems.* eds. J. Asubel and R. HermanNational Academy Press, 1988.

Arthur, W.B., S. Durlauf, and D. Lane. *The Economy As an Evolving Complex System II.* Addison-Wesley, 1997.

Bar-Yam, Y. *Dynamics of Complex Systems.* Reading, MA: Addison-Wesley, 1997.

Bell, K. and N.E. Bockstael. "Applying the Generalized Method of Moments Approach to Spatial Problems Involving Micro-Level Data." Working Paper, Dept. of Agricultural and Resource Economics: University of Maryland, 1998.

Ben-Akiva, M. and A. de Palma. "Analysis of a Dynamic Residential Location Choice Model With Transaction Costs." *Journal of Regional Science* 26, no. 2 (1986): 321-41.

Bockstael, N.E. "Modeling Economics and Ecology: The Importance of a Spatial Perspective." *American Journal of Agricultural Economics* 78 (1996): 1168-80.

Brock, W. and S. Durlauf. "Discrete Choice With Social Interactions I: Theory." Working Paper 9521, Social Systems Research Institute: University of Wisconsin Madison, 1995.

———. "Discrete Choice With Social Interactions II: Econometrics." Dept. of Economics: University of Wisconsin, 1997.

Chamberlain, G. "Heterogeneity, Omitted Variable Bias, and Duration Dependence." *Longitudinal Analysis of Labor Market Data.* eds. J.J. Heckman and B. Singer, Cambridge University Press, 1985.

Cressie, N. *Statistics for Spatial Data.* New York: John Wiley and Sons, 1993.

Diggle, P. *Statistical Analysis of Spatial Point Patterns.* London: Academic Press, 1984.

Durlauf, S. "Multiple Equilibria and Persistence in Aggregate Fluctuations." *American Economic Review* 81 (1991): 70-74.

———. "Statistical Mechanics Approahces to Socioeconomic Behavior." *The Economy As an Evolving Complex System II.* W. B. Arthur S. Durlauf and D. Lane eds., 81-

104. Vol. Vol. XXVII. Addison-Wesley, 1997.

Epstein, J. and R. Axtell. *Growing Articificial Societies*. Washington, DC: Brookings Institute Press, 1996.

Geoghagan, J., L.A. Wainger, and N.E. Bockstael. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS." *Ecological Economics* 23 (1997): 251-64.

Glaeser, E., B. Sarcedote and J. Scheinkman. "Crime and Social Interactions." *Quarterly Journal of Economics* (1996).

Haag, G. *Dynamic Decision Theory: Applications to Urban and Regional Topics*. Dordrecht: Kluwer Academic Publishers, 1989.

Heckman, J. "Identifying the Hand of Past: Distinguishing State Dependence From Heterogeneity." *American Economic Review,AEA Papers and Proceedings* (1991): 75-79.

Heckman, J. "Statistical Models for Discrete Panel Data." *Structural Analysis of Discrete Data With Econometric Applications*. Manski, C. and D. McFadden, eds., 114-80. Cambridge, MA: MIT Press, 1981.

Honore, B and E. Kyriazidou. "Panel Data Discrete Choice Models With Lagged Dependent Variables." Manuscript, Dept. of Economics, Princeton University and Dept. of Economics, University of Chicago, 1998.

Krugman, P. *The Self-Organizing Economy*. Blackwell Publishers, 1996.

Leggett, C. and N.E. Bockstael. "Evidence of Effects of Water Quality on Residential Land Prices." Working Paper, Dept. of Agricultural and Resource Economics: University of Maryland, 1998.

Manski, C. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60 (1993): 531-42.

———. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press, 1995.

Mills, D. "Growth, Speculation, and Sprawl in a Monocentric City." *Journal of Urban Economics* 10 (1981): 201-26.

Mills, E.S. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *American Economic Review* 57 (1972): 197-210.

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

Scheinkman, J. and M. Woodward. "Self-Organized Criticality and Economic

Fluctuations." *American Economic Review* May (1994): 417-21.

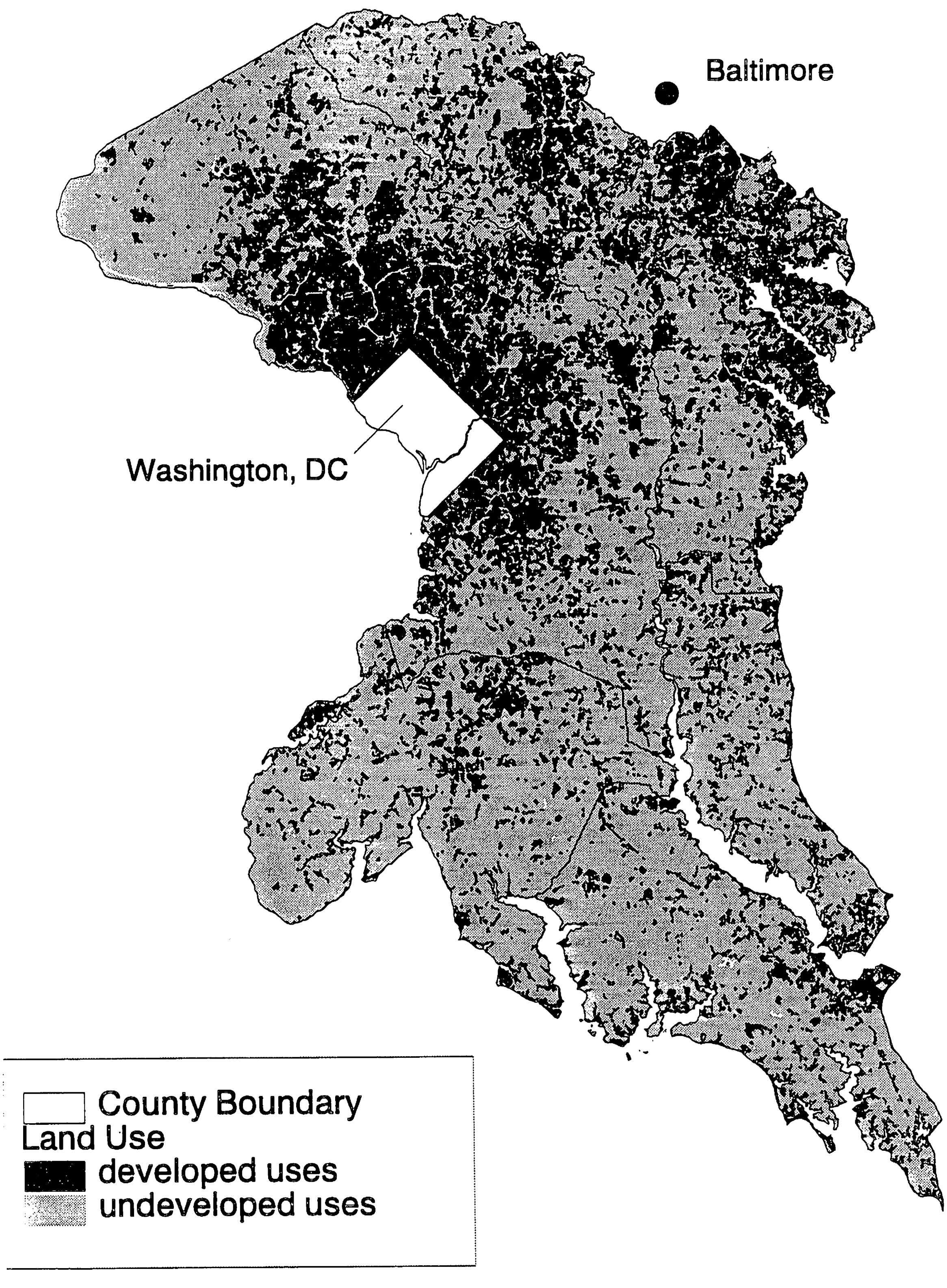
Tesfatsio, L. "How Economists Can Get ALife." *The Economy As an Evolving Complex System II.* W. B. Arthur, S. Durlauf, and D. Lane eds., 533-64. Vol. XXVII. Addison-Wesley, 1997.

Topa, G. "Social Interactions, Local Spillovers, and Unemployment." Working Paper, University of Chicago, 1996.

Werczberger, E. "A Dynamic Model of Urban Land Use With Externalities." *Regional Science and Urban Economics* 17 (1987): 391-410.

Map 1

1994 Land Use
Patuxent Watershed Counties



Map 2

Recent Residential Subdivisions

Calvert County

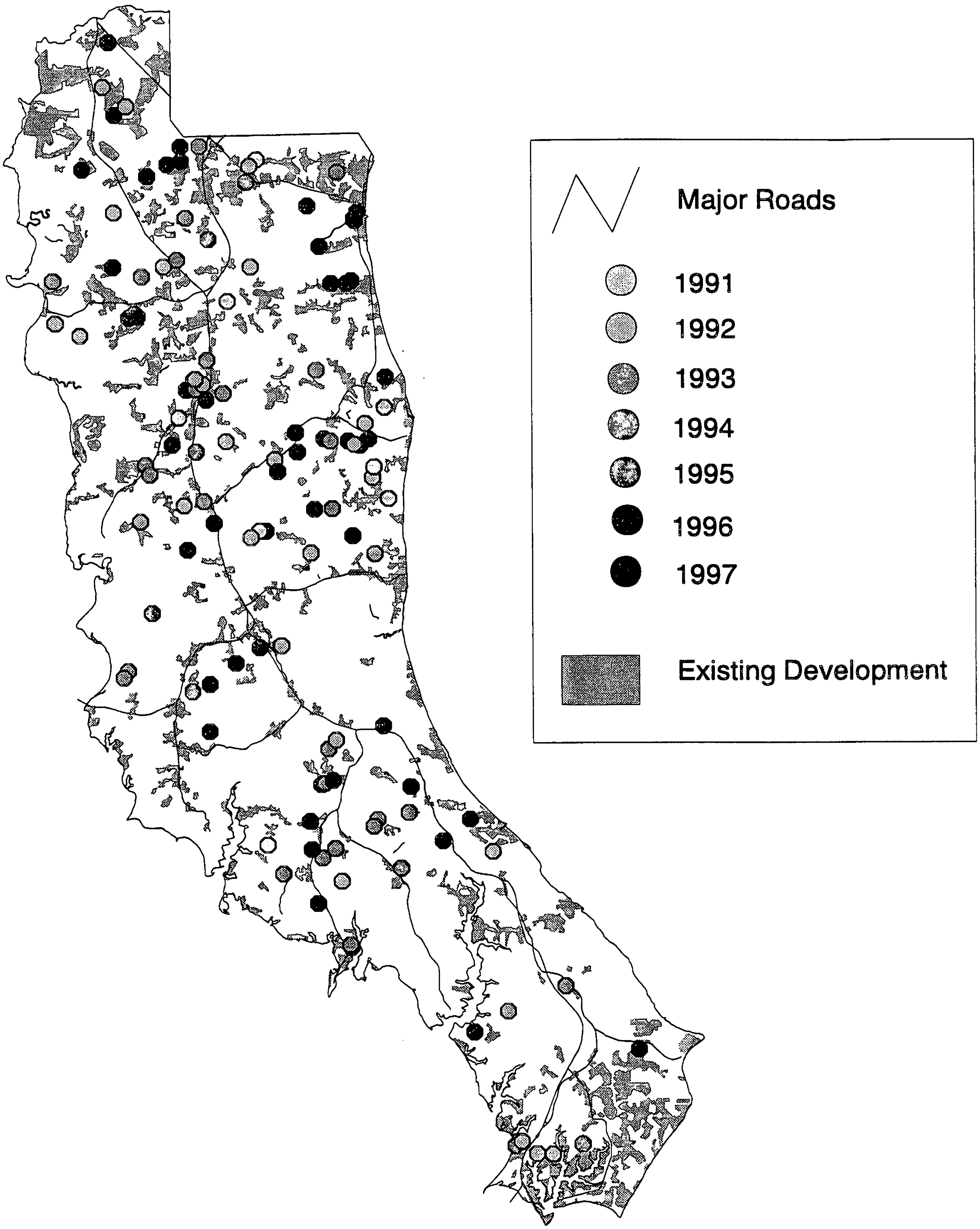
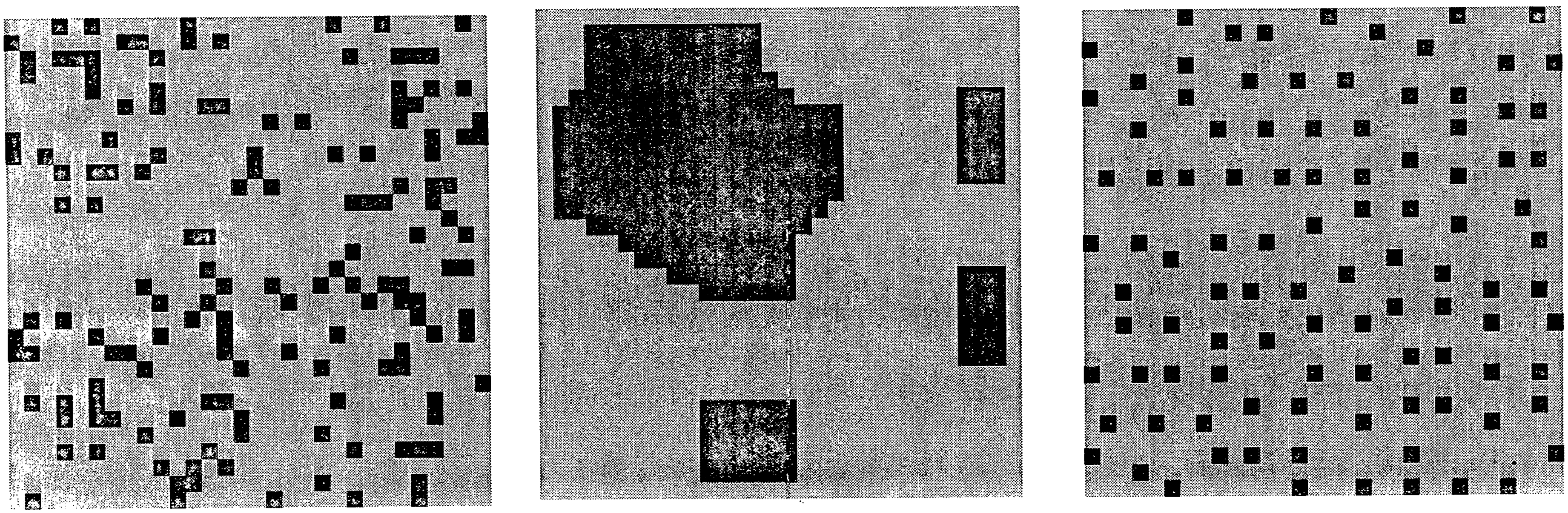


Figure 1

Spatial Patterns



Random Distribution   Positive Correlation   Negative Correlation

Figure 2

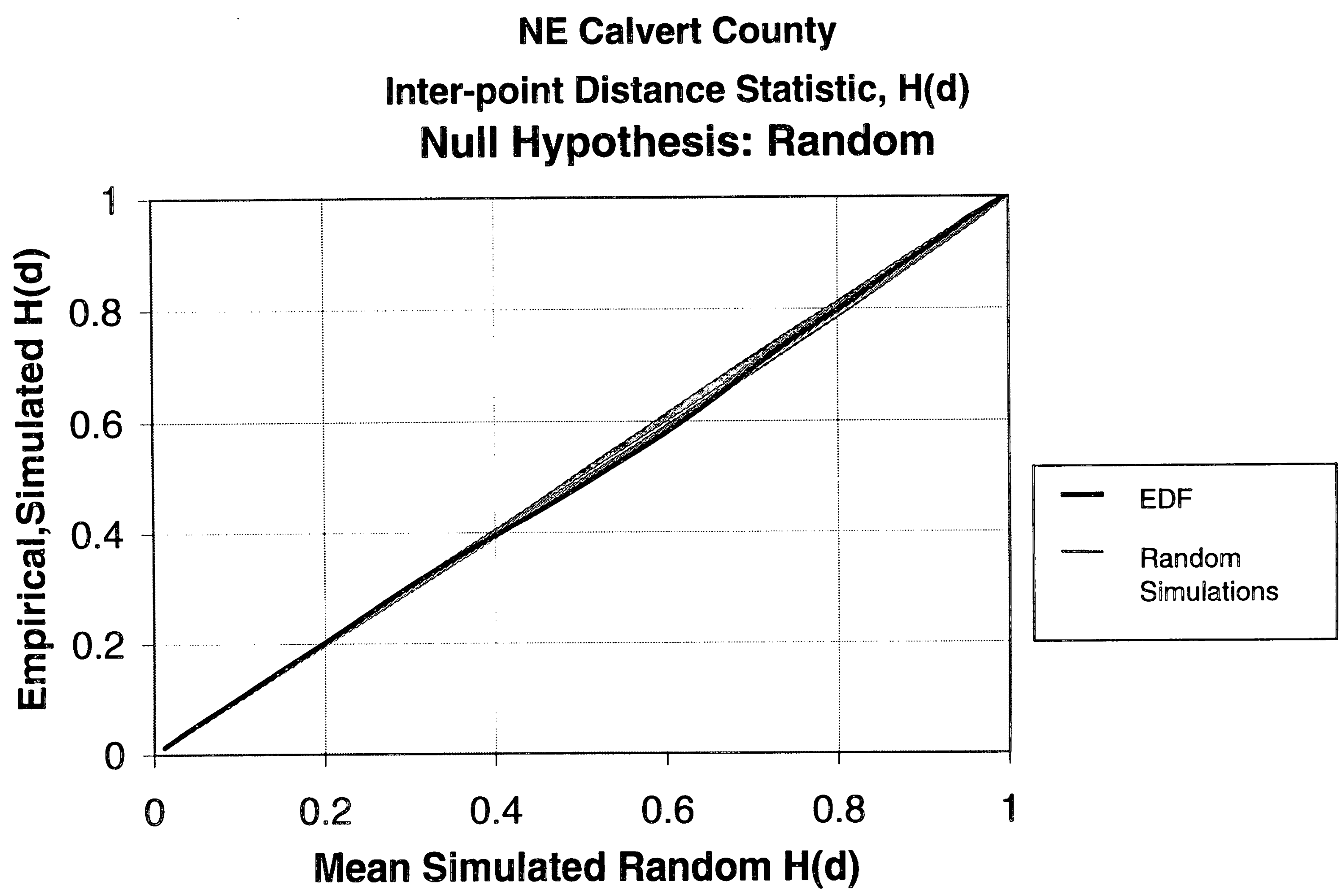

NE Calvert County
Inter-point Distance Statistic, H(d)
Null Hypothesis: Random
Empirical,Simulated H(d)
0
0.2
0.4
0.6
0.8
1
Mean Simulated Random H(d)
EDF
Random Simulations

Figure 3

Inter-Point Distance Statistic, H(d)

Null Hypothesis: Monocentric

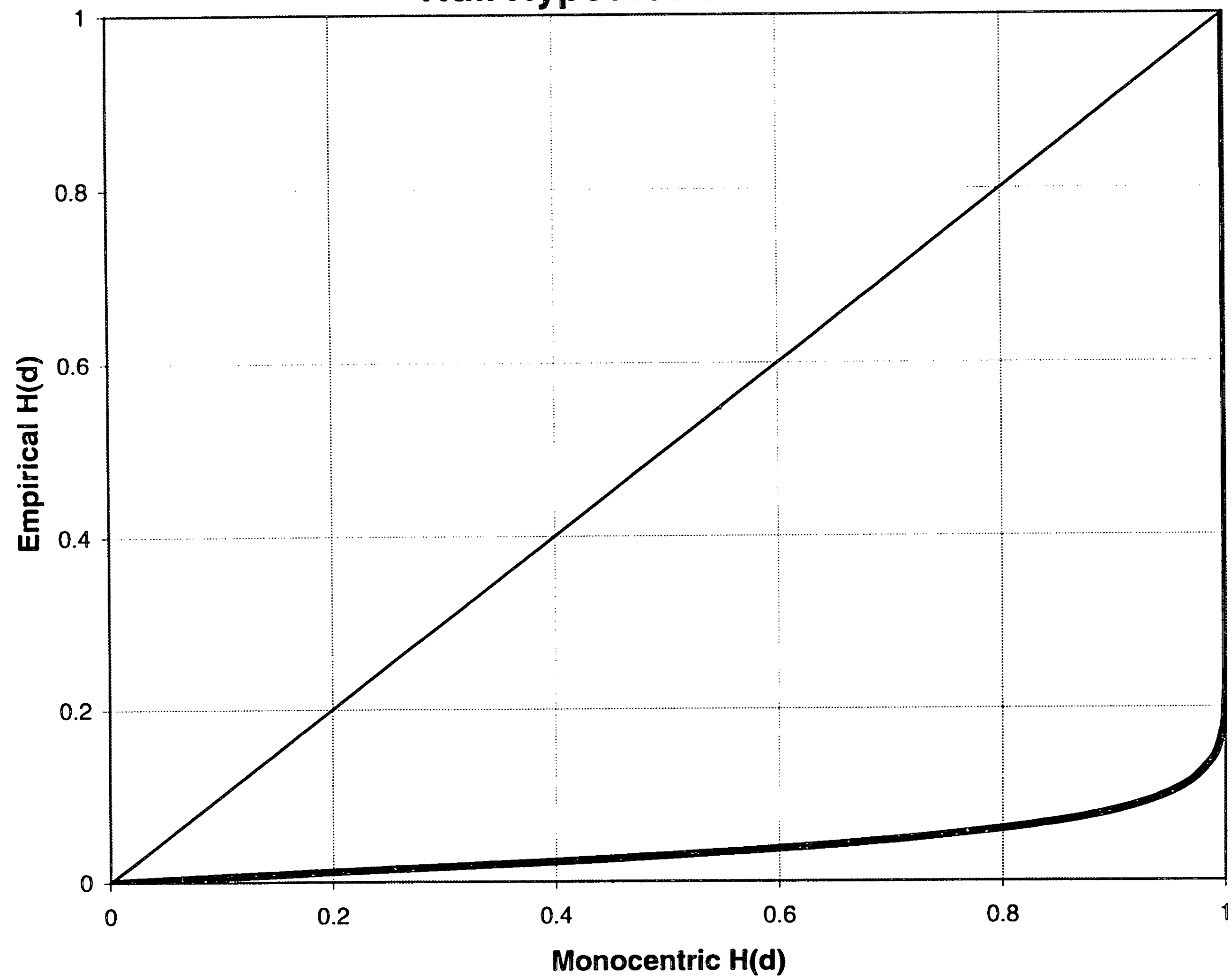
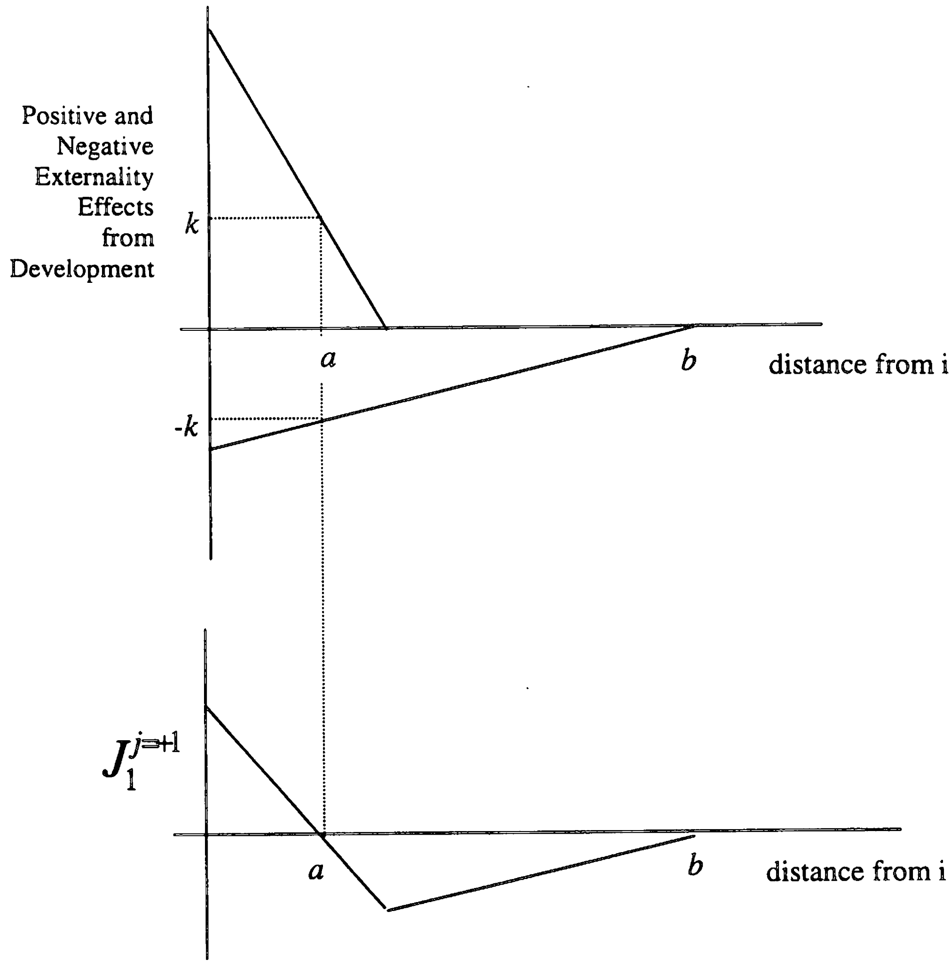


Figure 4

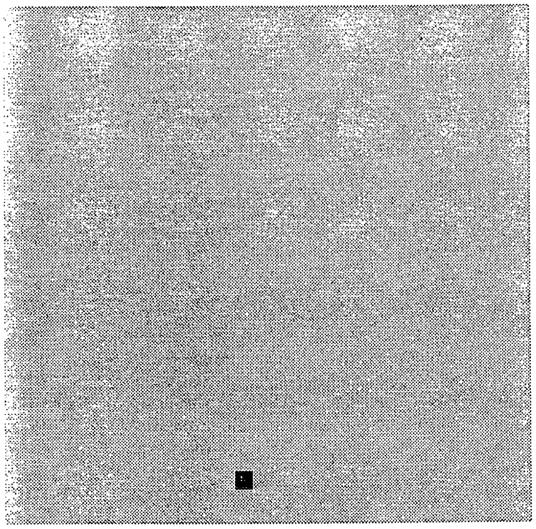


The Net Effect of Neighboring Development on Parcel i's Residential Value

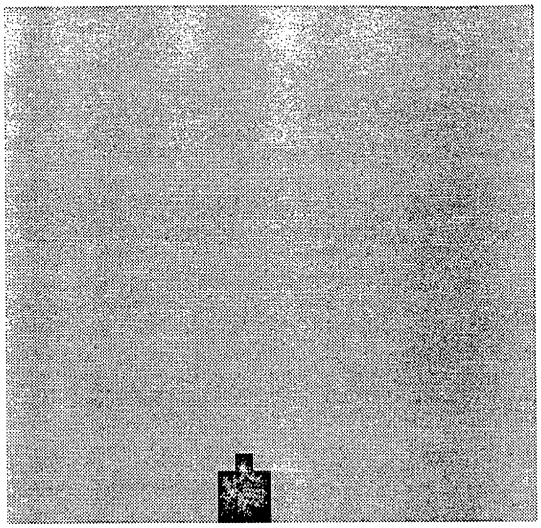
Figure 5

Monocentric Pattern

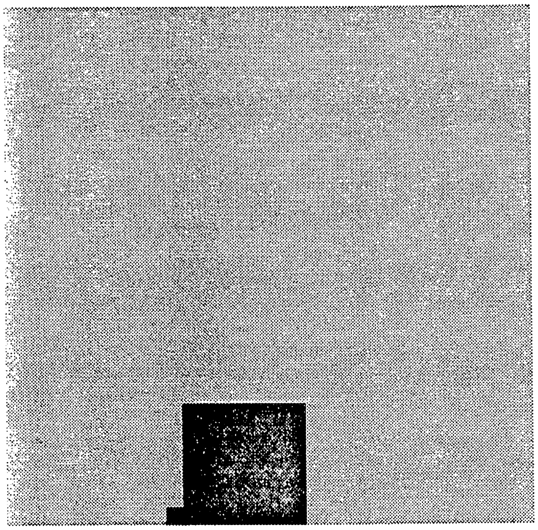
Parameters: city = 2, road = 0



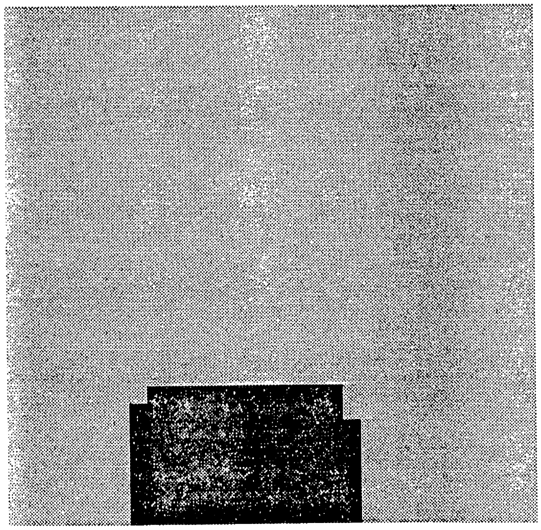
Initial conditions



Period 10



Period 50

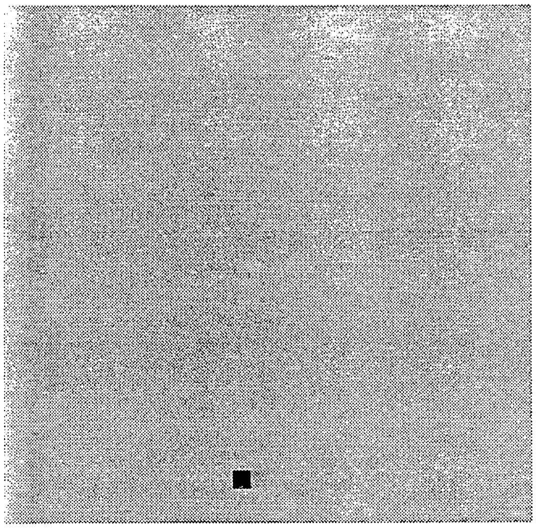


Period 100

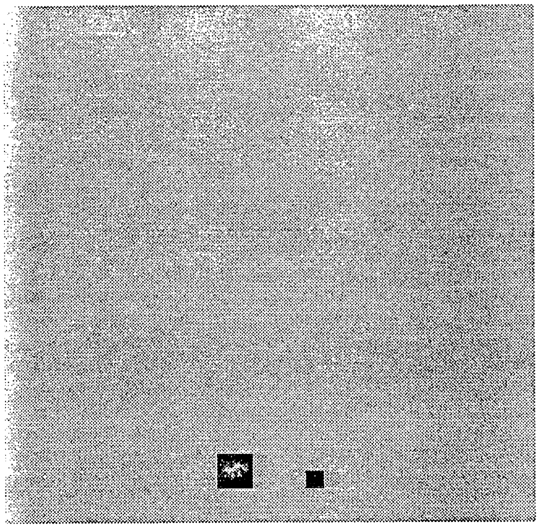
Figure 6

Clustered Monocentric

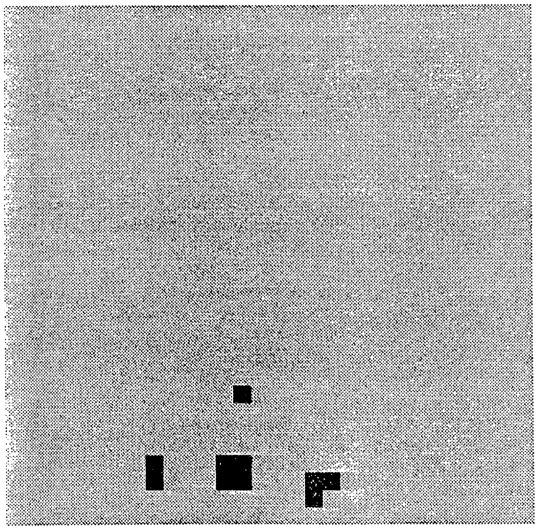
Parameters: J(1) = J(-1) = 1, J2(1) = -3, city = 2, road = 0



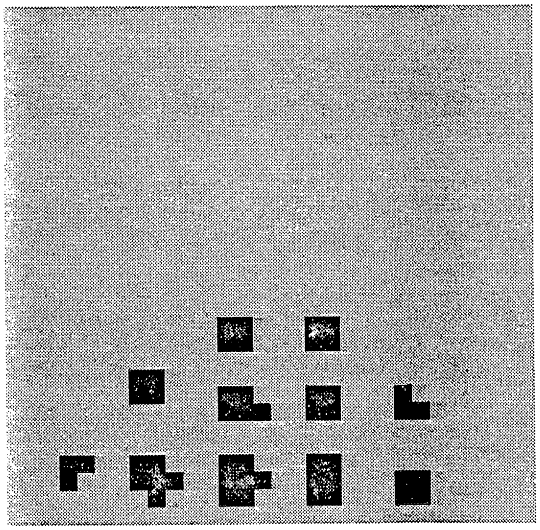
Initial conditions



Period 5



Period 10

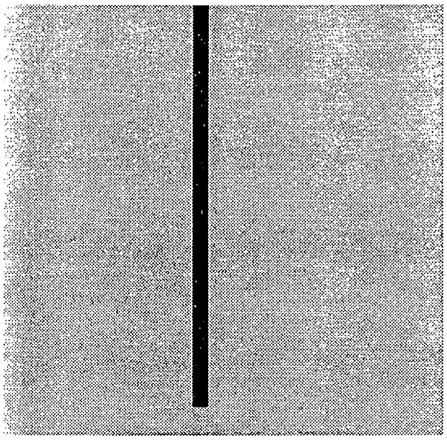


Period 50

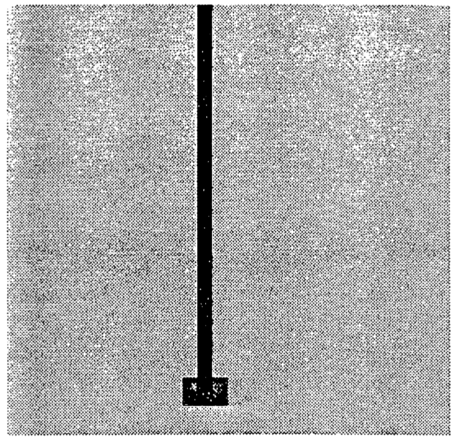
Figure 7

Road Development Pattern 1

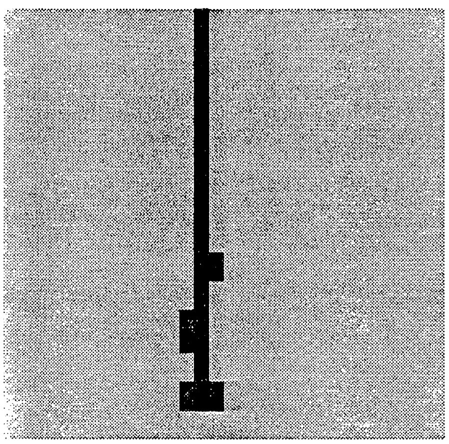
Parameters: J1(1) = J1(-1) = 1, J2(1) = -3, city = 1, road = 1



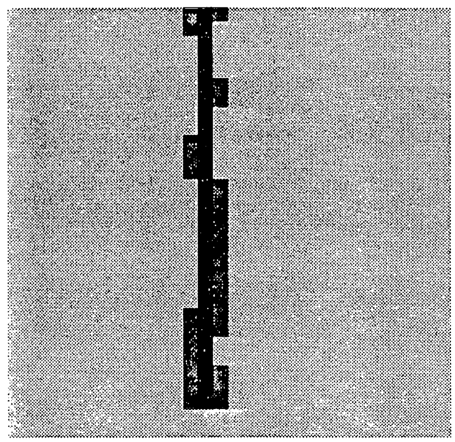
Initial conditions



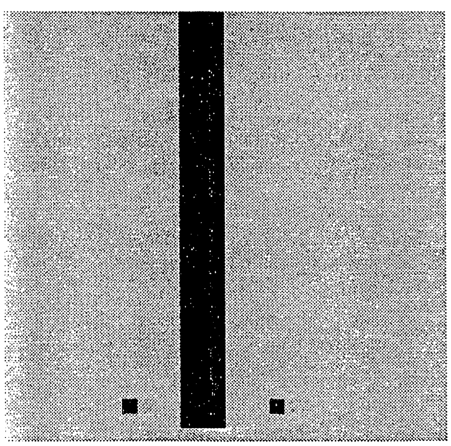
Period 5



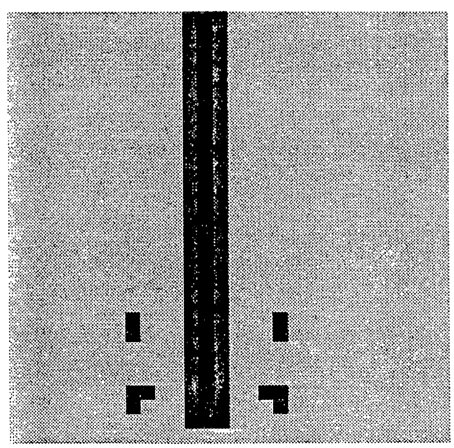
Period 10



Period 30



Period 62

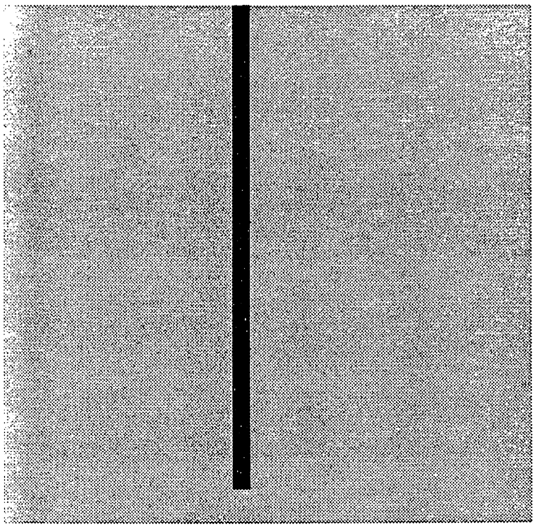


Period 70

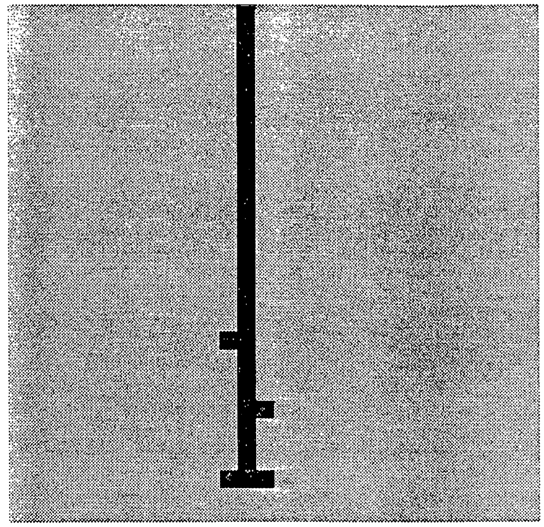
Figure 8

Road Development Pattern 2

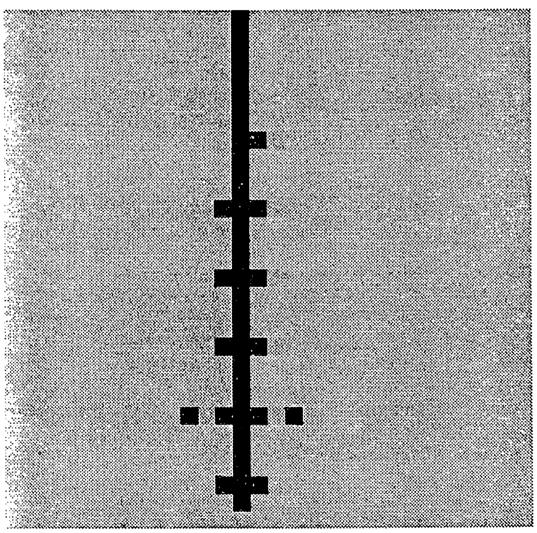
Parameters: J1(1) =1, J1(-1) = 0, J2(1) = -3, city = 1 road = 2



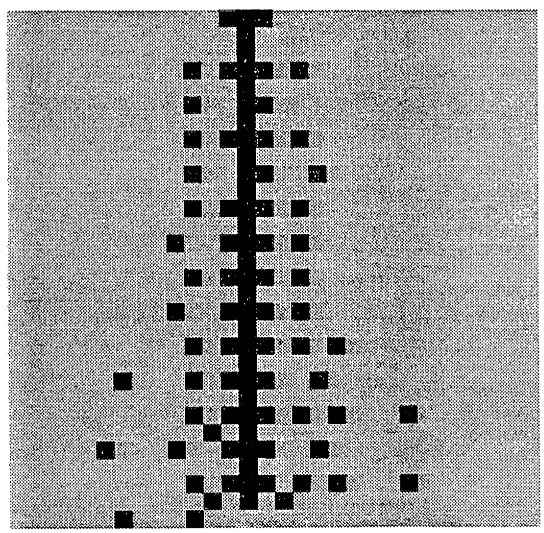
Initial conditions



Period 5



Period 15



Period 65