Journal of Economic Geography **2** (2002) pp. 31–54

Interacting agents, spatial externalities and the evolution of residential land use patterns

Elena G. Irwin and Nancy E. Bockstael***

Abstract

We develop a model of land use conversion that incorporates local spillover effects among spatially distributed agents. The model is used to test the hypothesis that fragmented patterns of development in rural-urban fringe areas could be due to negative externalities that create a 'repelling' effect among residential land parcels. Identification of the hypothesized interaction effect is complicated by unobserved, spatially correlated heterogeneity. Using an identification strategy that bounds the interaction effect from above, we find empirical evidence that is consistent with a theory of negative interactions among recently developed residential subdivisions in exurban Maryland. The result offers an alternative explanation for low density sprawl to that which is frequently posited in the economics literature and one with potentially quite different efficiency implications.

Keywords: land use pattern, spatial externalities, interactions-based models, sprawl
JEL classifications: R14, C29

1. Introduction

The evolution of urban spatial structure is generally understood to be the result of forces that spur the spatial concentration of economic activity. Recent changes in urban land use patterns, however, are characterized not only by the formation of new 'edge cities' around traditional urban centers, but also by scattered residential development in outer suburban and urban-rural fringe areas.[1] Rates of conversion to residential land use have far exceeded population growth rates in these areas, leading to a low-density, *non-contiguous*, and land-intensive development pattern that many refer to as 'sprawl'.

Surprisingly, relatively few theoretical explanations for this discontinuous development pattern have been proposed in the economics literature.[2] One of the best known economic theoretical explanations for this type of sprawl can be found in Mills (1981), who expanded on the earlier idea of Ohls and Pines (1975) that land inside the urban fringe may purposefully be withheld from early development as the result of optimal intertemporal decisionmaking in a growing urban area. Using a dynamic monocentric

*Department of Agricultural, Environmental, and Development Economics, Ohio State University, 2120 Fyffe Rd, Columbus, OH 43210, USA. *email* <irwin.78@osu.edu>
**Department of Agricultural and Resource Economics, University of Maryland, 2100 Symons Hall, College Park, MD 20740, USA. *email* <nancyb@arec.umd.edu>

1 We use the term 'exurban' areas to refer to regions that are outside of, but within commuting distance to, one or more urbanized areas.

2 We treat the discontinuity of sprawl as its defining feature and therefore focus our discussion on work that has sought to explain discontinuous development. This pattern has also been referred to as leapfrog or scattered development.

model, Mills extends this idea to explain the existence of both developed and undeveloped land located within the same distance to the urban center by introducing uncertainty and heterogeneity in expectations among land developers. He shows that the presence of uncertainty results in a sequential decision making process in which too much land within the urban fringe may be reserved for future development. In the case of heterogeneous expectations, this can result in 'transition zones' consisting of permanent scattered residential development. Bar-Ilan and Strange (1996) offer a slightly different explanation. In their model, lags between the decision to develop and the completion of development can lead to leapfrogging of development when uncertainty over returns increases with distance from the urban center. Other models portray sprawl, as defined by leapfrog development, as the consequence of a dynamically efficient market process (Fujita, 1976; Wheaton, 1982; Braid, 1991).

In this paper, we develop and attempt to empirically test an alternative hypothesis about the evolution of sprawl in exurban areas that draws on the notion that urban spatial structure is determined by interdependencies among spatially distributed agents. This theme has been developed in the literature by Anas (1992), Anas and Kim (1996), Arthur (1988), Fujita (1988), Krugman (1991, 1996), Page (1999), Papageorgiou and Smith (1983), Zhang (1993), and others.[3] While these models vary in terms of the hypothesized sources and specification of agent interdependence, they adopt the common theme that urban spatial structure evolves from a 'tug-of-war' between attracting and repelling forces that result from economic interactions among agents.[4] The relative magnitudes of these interactions determine individual agents' location decisions and hence, the evolution of urban land use patterns. Because the interactions both influence future location decisions and are a function of past location decisions, the spatial distribution of agents across the landscape is endogenously determined. Although the focus of much of this research has been on explaining the emergence of multiple urban clusters, the underlying principle is sufficiently robust to explain a variety of spatial patterns, including single cluster, multiple clusters, and dispersion, and thus presents a potential alternative to explaining sprawl at the urban-rural fringe.

Empirical evidence of any type of agent interactions in the evolution of urban spatial structure has thus far been absent. This is a challenging task, in part because of the many heterogeneous features of the landscape that are likely to influence location decisions (e.g. roads, zoning, natural features). Such landscape heterogeneity is ignored by the theoretical interaction models. While these simplified but tractable analytical models demonstrate the potential role of interactions, they do not offer a means of identifying these effects using real world data. Empirically, the challenge is to separate the effects of endogenous interactions from spatially correlated exogenous landscape features, which may invoke land use patterns that are observationally equivalent. Because it is difficult to measure such interactions, separating these effects from unobserved exogenous heterogeneity is possible only for limited cases. The challenge of econometric identification has been outlined in a separate literature on empirical

3 For a review of some of this work, see Anas et al. (1998).

4 The notion of spatial interdependence leading to agglomeration is not new. Beckman (1976) explored the effect of social interactions among spatially distributed households on urban spatial structure. More recently, these interactions have been modeled as arising indirectly through market forces, e.g. transportation costs and pecuniary externalities (Krugman, 1991), as well as directly through agents' preferences (Page, 1999) or spatial externalities, e.g. knowledge spillovers (Zhang, 1993).

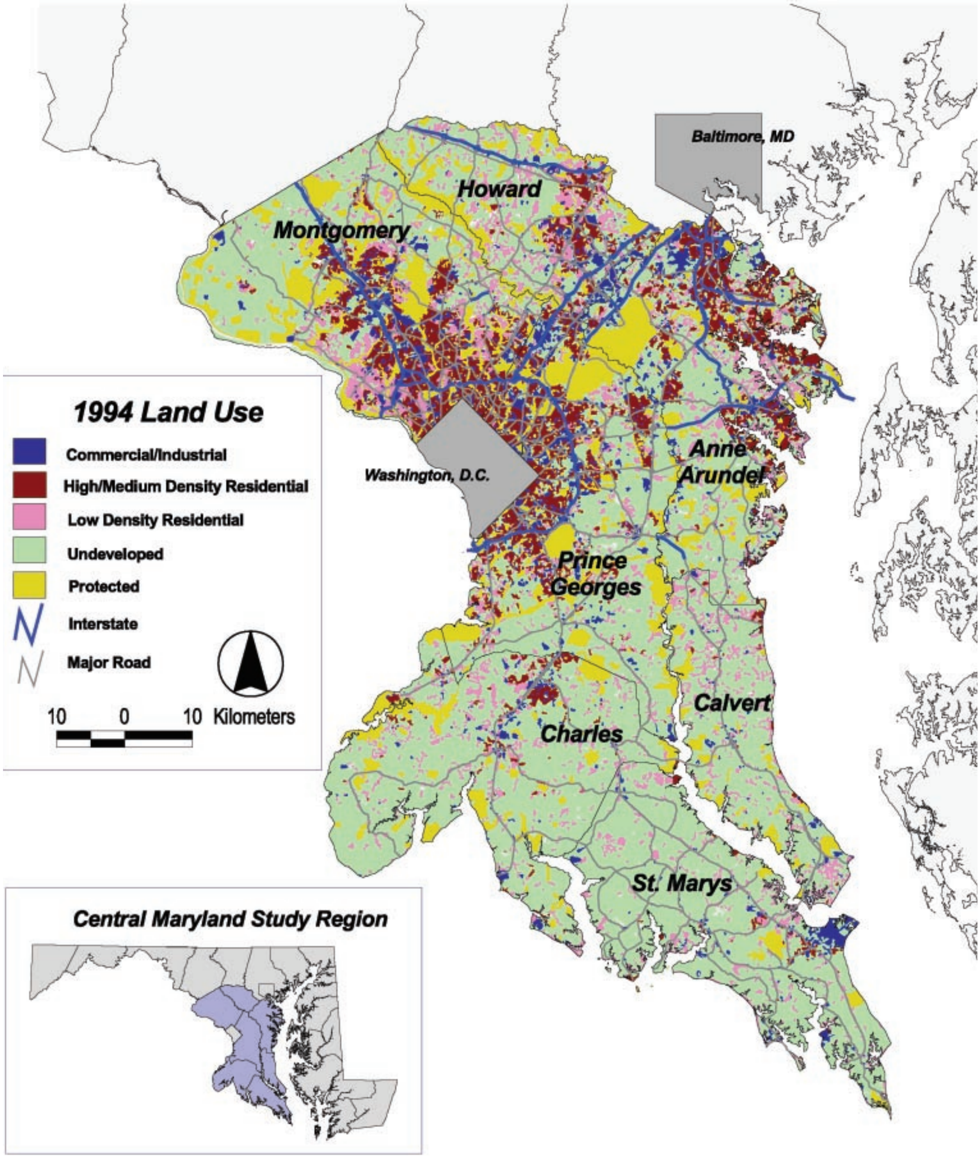
models of social interactions, most notably by Manski (1993) and Brock and Durlauf (2001).

The goal of this paper is to investigate whether an interactions hypothesis offers an alternative to the existing explanations of sprawl that are based on the traditional bid-rent model of urban economics. To motivate the discussion and to portray the exurban sprawl phenomenon that is the focus of this paper, land use data are mapped from a seven county region of central Maryland located between the two historical city centers of Washington, DC and Baltimore, MD. These data are then manipulated to highlight the difference between the current phenomenon and the spatial land use pattern that would be predicted by the most basic of monocentric models. Map 1 portrays the actual spatial arrangement of commercial/industrial development, and high, medium and low density residential development in the seven counties in 1994. Holding the total amount of land use in each type of development at these actual levels, we spatially reallocate the land use totals according to the predictions of a duo-centric model, in which the locations of both Washington, DC and Baltimore are treated as exogenously given and accessibility is measured as commuting distance along the major roads network.[5] Well-known results from the basic monocentric model predict that non-residential urban land will locate closest to the city center, given that firms will outbid households for these central locations, and that residential land will be characterized by a density gradient, declining with distance from the central city. Map 2 portrays what the landscape of the seven countries would have looked like in 1994 if the spatial pattern had followed the duo-centric city model predictions.[6]

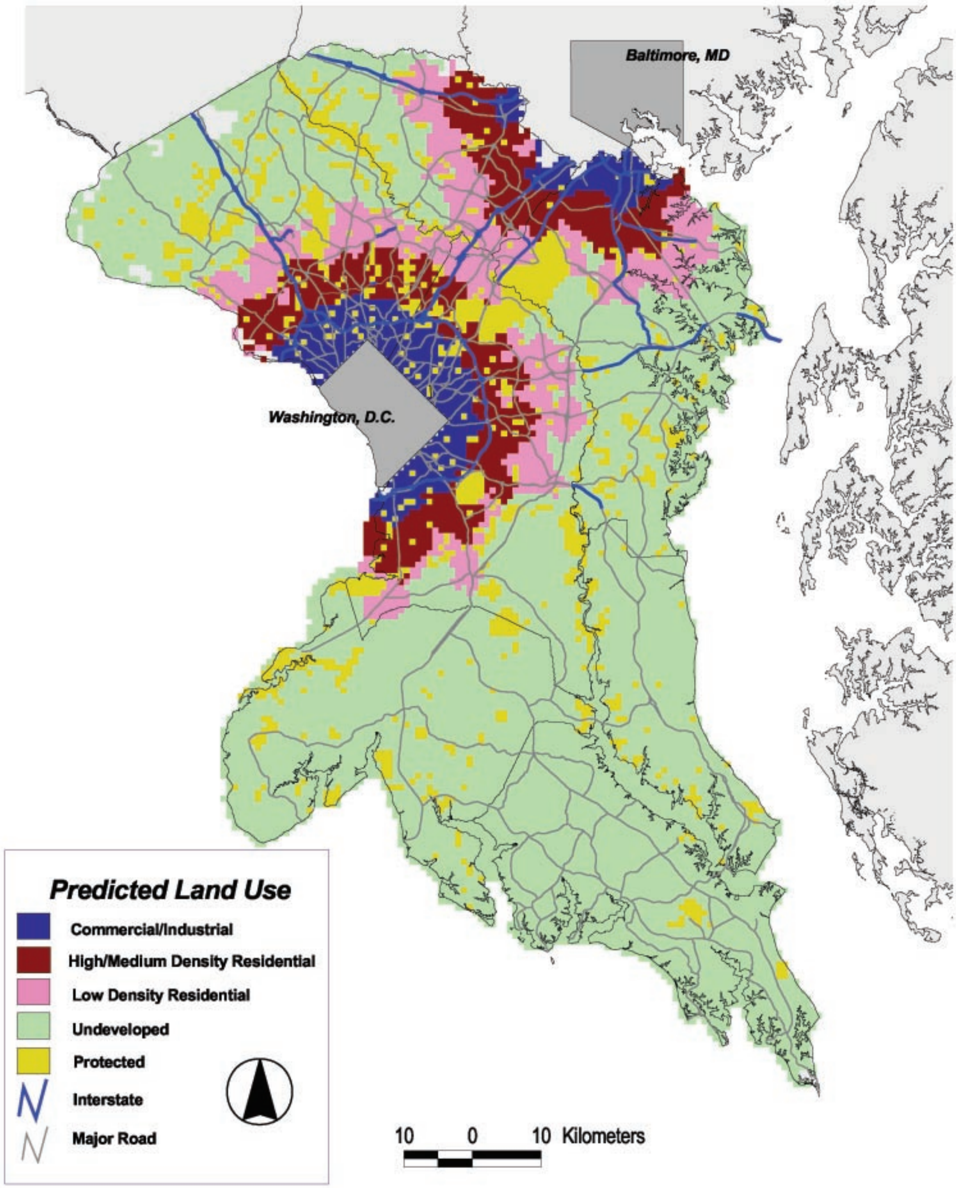
The similarity of the central city areas depicted in Maps 1 and 2 provides evidence of the power of the city-center model in describing the past evolution of land use patterns. Yet it fails to capture the exurban sprawl pattern characterized by dispersed residential development in the outlying counties. To give this model every possible benefit, we further constrain the land use allocation by current zoning restrictions. This reallocation is biased in favor of the duo-centric model, since zoning neither entirely predates the formation of development patterns nor is truly exogenous in the land development process. In Map 3, residential land use patterns are somewhat less

5 The land market of the Washington, DC metropolitan area also includes parts of northern Virginia, but spatially explicit land use data are not available for these counties.

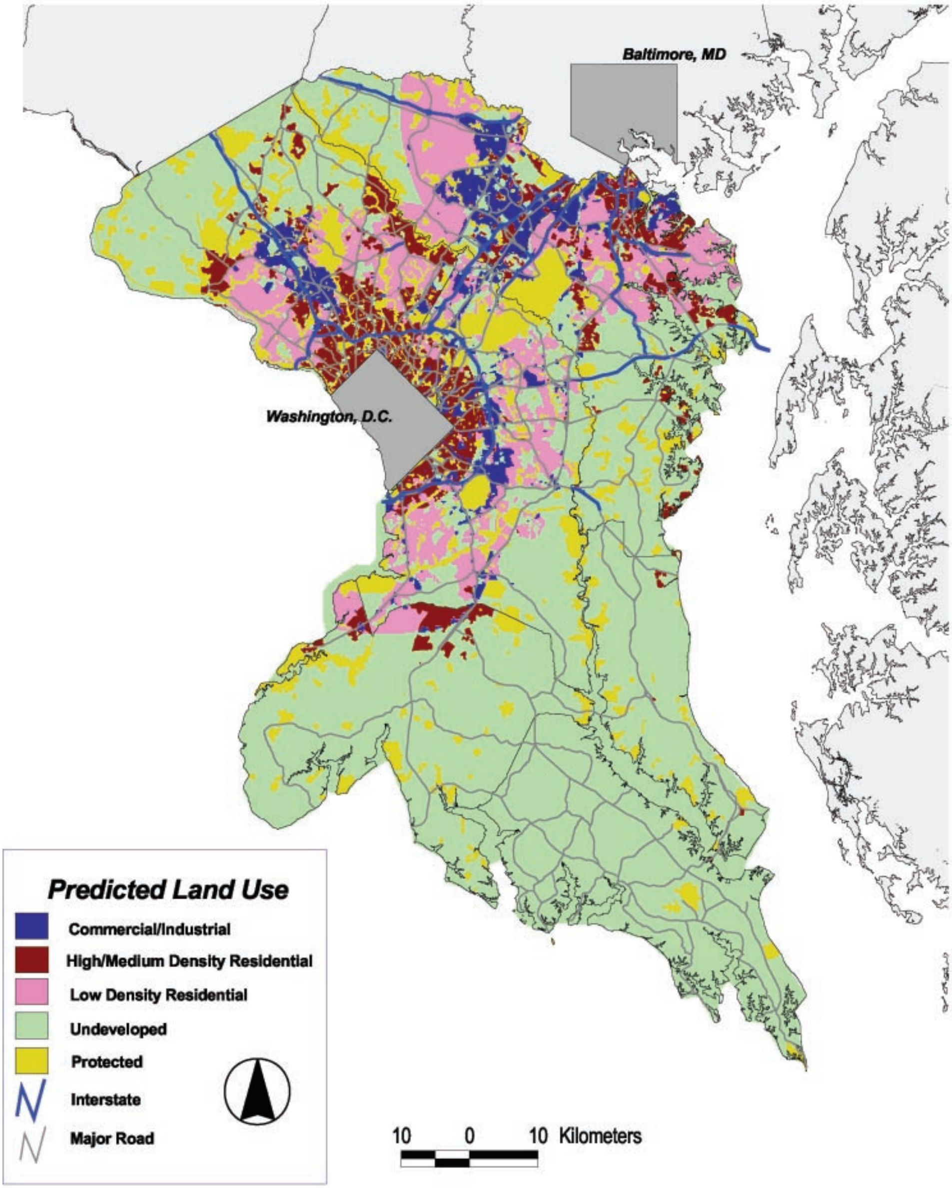
6 Maps 2 and 3 were generated by first calculating the total 1994 land area in each of the following land uses: commercial/industrial, high, medium, and low density residential, protected, and undeveloped. The spatial distribution of the land uses was then reallocated holding constant the total amount of land area in each of the developed land uses. This was done by gridding the landscape into 800 square meter cells and sequentially allocating the total area of each land use, beginning with commercial/industrial land uses, across cells based on their proximity via the major roads network to Washington, DC and Baltimore, MD. Cells with the highest accessibility to either of the two major urban centers were assigned on a cell-by-cell basis to commercial/industrial use, followed by high, medium, and low density residential use respectively. Cells falling within actual protected areas were exempted from the reallocation and the remaining cells not allocated to any urban use were categorized as undeveloped. Map 3 was generated by considering the additional constraint of zoning. Specifically, a cell was assigned a land use only if it was within an area actually zoned for this land use. For cells within residentially zoned areas, the maximum allowable density of development was used as an additional constraint to identify cells that could potentially be assigned to high, medium, or low density residential uses. Once the allowable cells were identified for each land use, the land uses were reallocated to cells based on the same proximity criterion outlined above. Both maps were created in ArcInfo Grid using a routine programmed with arc macro language.



Map 1. 1994 land use pattern, central Maryland region.



Map 2. Land use allocation under duocentric assumption with distance measured via the roads network.



Map 3. Land use reallocation under duocentric assumption with exogenous zoning.

contiguous, but the pattern of scattered development in the exurban areas still remains unexplained. While additional complications are possible, such as assumptions of a more polycentric structure, there appears no obvious way to account for the degree of fragmentation in exurban areas solely as a function of distance from exogenously defined urban centers.

Clearly additional heterogeneity of the landscape will influence the location of residential development. Both manmade features, such as public sewer and water lines, and natural landscape attributes, such as proximity to a coastline, are likely to influence the pattern of development. In this paper, we explore whether such exogenous features can fully explain the recent pattern of urban-rural fringe development or whether a model that also incorporates an interactions component—in the form of land use externalities—can better account for the observed scattered pattern of development. In doing so, our focus is not on explaining the rate or amount of growth occurring in these areas, but rather on explaining the *spatial pattern* of this development. Growth in exurban areas may be explained by a number of factors, including decreased commuting costs (e.g. due to road improvements and low gas prices), declining urban services, and socio-demographic changes. We abstract from the reasons for this growth and focus instead on the question of spatial pattern: given that new development is occurring in these exurban regions, why is it happening in such a fragmented pattern?

The rest of the paper is organized as follows. A model of an agent's land use conversion decision is developed in which interactions among neighboring agents arise from land use externalities. We argue that, in theory, a fragmented development pattern may arise as the result of sufficiently strong negative interaction effects. However, in devising an empirical test of these interactions, the effect of real world heterogeneity of the landscape must also be considered. The principal challenge is to statistically identify the interaction effects by distinguishing them from the influence of unobserved but spatially correlated landscape features. We employ an identification strategy suggested by Heckman and Singer (1985) in which the direction of the interaction effect is bounded from above. Due to the positive spatial correlation of exogenous factors, the sign of the interaction effect is identified, and the existence of the interaction effect assured, only if the estimated interaction parameter is negative. Empirical results using parcel-level data on residential subdivision conversion from an exurban region of central Maryland provide support for a hypothesis of negative interaction effects for residential conversion in recent years. Simulations of predicted changes in the development pattern demonstrate the magnitude of these estimated effects. Evidence that these negative externalities could be causing sprawl brings into question the claim that a sprawl development pattern is dynamically efficient.

2. The conceptual framework: optimal timing of development

The evolution of large-scale spatial land use patterns results from decisions of micro-level agents about individual land parcels. We begin by formulating the decision of an agent concerning the conversion of his parcel from an undeveloped to a developed state. A parcel is considered 'undeveloped' if its current use is in agriculture or another resource producing activity such as commercial forestry, as well as if it is in a natural state. The developed use of interest is residential and therefore we focus only on those parcels that could be converted to residential use as determined by current zoning regulations.[7]

In our problem the behavior of interest is the decision of a parcel owner to subdivide his land for development. Approximately 90% of all subdivided parcels are built on within three years, so this subdivision decision is a reliable precursor of development and a good indicator of the owner's intentions. The decision may be made by the original owner of a parcel, or it may be made by a developer after purchasing the property. In either case, this recorded activity reflects the original owner's intention to alter the use of his land. However, it is not necessarily a market transaction since no residential buyer is required at this stage. Consequently the decision is driven by expectations rather than current market realities. In this study, the undeveloped parcel is treated as the decision unit and conversion implies the subdivision of a previously farmed or forested parcel into multiple residential lots.

The region of study is one that has experienced considerable growth in incomes and population over the last several decades and is predicted to face continued growth into the foreseeable future. As such, it is likely that most agents expect that if their parcels are not already profitable candidates for development, they will become so at some future time. Therefore, the relevant decision from the agent's perspective is the optimal timing of this development.[8] Development is viewed as irreversible, both because reversing development is exceedingly costly and because, once developed, ownership of the parcel passes to many small lot owners.

To represent the choice problem, first define $A(i, t)$ as the returns to the original, unsubdivided parcel (denoted i) in the undeveloped use in time period t. We will refer to this as agriculture, broadly defined to include any uses of the land in an undeveloped state. Conversion of parcel i at time T will require the agent to incur costs to reap expected one-time gross returns. Costs include the provision of subdivision infrastructure, as well as permitting and other administrative fees. Gross returns equal the sum of the expected sales prices of the subdivided residential lots. We denote δ as the discount factor, defined as $1/(1+r)$ where r is the interest rate, and the one-time returns from development minus costs of conversion in time T as $V(i, T)$. Then the net returns from developing parcel i in time period T equals the one time net returns minus the present value of foregone agricultural returns and is given by:

$$V(i, T) - \sum_{t=0}^{\infty} A(i, T+t)\delta^t. \tag{1a}$$

The net returns from keeping parcel i in an agricultural use in period T and developing it in time period $T+1$, discounted to time T, are:

7 Residential use encompasses 73% of the developed land in the seven county area depicted in Map 1 and accounts for 78% of the changes in land use in that region between 1985–94. More importantly, it accounts for over 90% of conversion in the exurban areas during this time period. The form of residential development is dicated by zoning, but zoning does not greatly constrain the location of residential development as it does commercial and industrial development.

8 The optimal timing of development has been analyzed by a number of authors, most notably Arnott and Lewis (1979) and Capozza and Helsely (1989). They develop their model of the present value of land in terms of continuous time and an infinite stream of rents from development, but otherwise our approach is similar.

$$A(i,T) + \delta V(i,T+1) - \sum_{t=0}^{\infty} A(i,T+t)\delta^t. \tag{1b}$$

The optimal development time will occur in period T only if (1a) is positive and if (1a) is greater than or equal to (1b). Thus, development will occur in T if

$$V(i,T) - \sum_{t=0}^{\infty} A(i,T+t)\delta^{T+t} > 0 \tag{2a}$$

and

$$V(i,T) - A(i,T) \geqslant \delta V(i,T+1). \tag{2b}$$

The agent develops in period T only if (a) the net value of conversion is positive and (b)

$$\frac{V(i,T+1) - \{V(i,T) - A(i,T)\}}{V(i,T) - A(i,T)} < r,$$

where r is the interest rate.[9]

The smallest T that satisfies (2) is the optimal time to develop if certain assumptions about the time path of the functions hold. First, we expect that in our region of study the gross returns to development are rising over time. This will occur if population and/or income per capita continue to rise in the face of a declining stock of undeveloped land. As we have no particular prior expectations on the time path of $A(i, t)$, we assume that returns to the undeveloped use remain constant over time. With only these assumptions, all parcels will eventually meet the development criteria in (2a). The condition specified in (2b) will be met if the tendency for agents to postpone development indefinitely in the hopes of greater future returns is mitigated by residential prices that increase at a decreasing rate and/or development costs that increase over time. The latter may be true, for example, if agents fear growth controls will ultimately be instituted to pass on more and more of the public sector costs of infrastructure, such as roads and schools, to the developer. Because (2a) is likely to be met well before (2b), (2b) is the crucial condition in subsequent development of the model.

To give more substance to this specification, consider the vector of parcel characteristics that the owner can expect will affect a parcel's value in residential use, its cost of conversion, and its value in the alternative use. Based on knowledge of the existing residential land market, the owner will expect that factors such as commuting distance to employment centers, provision of public services, lot size, and amenities of the landscape will affect the value in residential use. Conversion costs can be expected to vary over parcels according to excavation and clearing costs and availability of public utilities. Agricultural returns will vary with soil fertility and distance to markets.

Interactions can be incorporated into (2b) in several ways, depending on the assumed spatial structure of the agent interdependence. We follow other economic and social models of interdependent behavior that model an individual's choice as a function of either the current or past choices of neighboring individuals.[10] In our case, land use externalities are hypothesized to create interaction effects among parcels that influence

9 To see this, note that (2b) can be rewritten with $1/(1+r)$ replacing δ.

10 For a full review of these models, see Brock and Durlauf (2001).

the value of a parcel in a developed use. Because of temporal lags in the development process, this interaction is viewed as a recursive process, so that a parcel's value in a developed use is a function of the past period's land use states of parcels within a local neighborhood of each parcel. The magnitude of the net interaction effect may vary over distance and the direction of the effect may vary as well over distance, leading to potentially complex structures of interaction. Here we are interested in explaining the *scattered* pattern of exurban development. Such a pattern would result if net *negative* interactions were present, e.g. due to congestion externalities, and if these effects were sufficiently strong as to create a 'repelling' effect among residential development.[11]

Let $\lambda_s I_s(i,t)$ represent this spillover effect, where $I_s(i,t)$ is the proportion of neighboring parcels that are in a developed state at the time the development decision is made, λ_s is the interaction parameter, and s indexes the order of the spatial lag, which increases with increasing distance from parcel i. Because neighboring developed land could conceivably have positive and/or negative spillover effects, the parameter λ_s, which represents the net effect of these spillovers, could be either positive or negative at any given distance s. Positive spillovers from surrounding development may arise, for example, if people value a sense of community or if there are benefits associated with contiguous residential development such as the provision of better public services. Because neighboring development signals the loss of open space amenities, spillover effects may, instead, be negative due to congestion or aesthetic considerations.

Rewriting (1a) to incorporate the effect of interactions, the net returns from developing parcel i in time period T equals:

$$V(i,\ T) + \sum_s \lambda_s I_s(i,\ T) - \sum_{t=0}^{\infty} A(i,\ T+t)\delta^{T+t}. \tag{3}$$

Rewriting the conversion rule in (2b), development now occurs in the first period in which:

$$V(i,T) - \delta V(i,\ T+1) - A(i,\ T) + \sum_s (1-\delta)\lambda_s I_s(i,\ T) \geqslant 0, \tag{4}$$

where the value of $I(i,\ T)$ is assumed to be constant between periods T and $T+1$.[12] Because we are interested in testing whether the scattered pattern of residential development we observe could be the result of net negative development externalities, the hypothesis to be tested is that $\lambda_s < 0$ for some range of s.

3. The empirical model

Thus far we have assumed that variation over parcels in the timing of development will occur only due to variation in observable parcel characteristics and surrounding land use. But in reality, landowners are heterogeneous and owners of parcels with the same basic attributes will have different reservation prices. Some owners may be especially

11 For more discussion and simulations that illustrate this point, see Irwin (1998).

12 Since we cannot observe changes in $I(i,T)$, we assume that agents do not expect immediate neighborhood changes, so that $\Delta I = I(i,T) - I(i,T+I) = 0$. In any event, we expect the first order effects of $I(i,T)$ to dominate.

good at farming or may have high recreational and aesthetic values for their land; others may be near retirement and looking for a means to liquidate their assets. These idiosyncrasies will induce a distribution of unobservable factors, randomly distributed across the landscape, that will, in turn, induce a distribution of optimal development times conditioned on explanatory variables.

3.1. Hazard model of development

To take account of these differences across agents, define ε_i as these unobservable factors associated with the owner of parcel i. Given that ε is viewed by the researcher as a stochastic variable, the following gives the probability that parcel i, with surrounding land use pattern $\Sigma_s I_s(i, T)$, will be converted by period T:

$$Prob\left\{\varepsilon_i < \frac{1}{1-\delta}\big(V(i,T) - \delta V(i,T+1) - A(i,T)\big) + \sum_s \lambda_s I_s(i,T)\right\}. \tag{5}$$

This implies that agents with large ε's, such as those who are particularly good farmers or those that place a particularly high value on their undeveloped land as a source of direct utility, will convert later than agents with the same type of parcel but smaller values of ε.

The probability that a parcel with a given set of characteristics will be converted in period T is its hazard rate for period T, which is given by:

$$h(T) = \frac{F[\varepsilon^*(T+1)] - F[\varepsilon^*(T)]}{1 - F[\varepsilon^*(T)]}, \tag{6}$$

where F is the cumulative distribution function for ε and define ε^* as the ε that makes (5) an equality—i.e. the ε such that the owner of the parcel is just indifferent between converting and not converting in T.

Since (6) is a hazard rate, duration analysis is a convenient approach for estimating the parameters embedded in (6) and for testing hypotheses about these parameters. In this analysis we choose Cox's partial likelihood method because it can accommodate time-varying covariates—an essential element of our problem. In order to estimate the interaction parameters, we need to capture accurately the fact that the land use surrounding a parcel is changing over time. However, while the neighborhood land use pattern changes over time, we note that other parcel attributes, such as aspect, soil type, distance to urban centers, and the provision of public sewers, are unlikely to change within the short run. Therefore these other parcel features are treated as spatially varying, but temporally constant. Nonetheless, expectations on $V(i,t)$ will vary over time to general growth pressures in the area. Assuming that $V(i,t)$ is separable in these factors that are time varying, but spatially constant, we can apply Cox's model to our problem by defining a baseline hazard rate that is a function of time only. Let $\omega(T)$ represent the exponential of this baseline hazard rate and assume that the log of the hazard rate is linear in the other arguments. Then the hazard rate for parcel i is given by:

$$h(i,T) = \omega(T)^*\exp(Z\beta) \tag{7}$$

where Z is a vector of parcel i's attributes, including $I_s(i,T)$, and β is a corresponding parameter vector.

Cox's method is a semiparametric approach that relies on formulating the likelihood in such a way that the baseline hazard, $\omega(T)$, drops out and therefore specification of an error distribution is unnecessary. The resulting expression is called the partial likelihood function, which gives the conditional probability that, given an event occurs in a particular time period, it occurs to a specific individual. It is the product of N contributions to the likelihood function, where N is the number of developable parcels, and the form of the n^{th} contribution is given by:

$$L_n = \frac{h(n, T_n)}{\sum_{j=1}^{J_n} h(j, T_n)}. \tag{8}$$

By definition, T_n is the time at which the n^{th} parcel is converted. In (8), $h(n, T_n)$ is the hazard rate for the n^{th} parcel, $h(j, T_n)$ is the hazard rate for the j^{th} parcel, but evaluated at time T_n, and J_n is the set of parcels that have 'survived' in the undeveloped state until time T_n. This expression gives the ratio of the n^{th} parcel's hazard rate to the sum over the hazards of all other parcels that have not yet been developed as of time period T_n. When multiplied by all other n–1 contributions, this forms the conditional probability that, given an event occurs in a particular time period, it occurs to a specific individual.

Substituting (7) into (8), it is clear that the baseline hazard term cancels out in each contribution to the likelihood function; the estimation procedure produces information only on the β's. While in some settings this is a disadvantage, it helps us to focus attention on the exogenous heterogeneity and the interaction terms without having to make assumptions about the functional form of $\omega(T)$ or measure variables that might affect $\omega(T)$ over time. For example, it seems plausible to assume that the baseline hazard is a function of growth pressures and the interest rate, as well as other variables (e.g. development fees) that vary over time but not over parcels. In the absence of the baseline hazard, the only aspect that matters in estimating the parameters is the order of parcel conversion over time rather than the absolute time of conversion.[13] If all the attributes embedded in vector Z were time invariant, this would imply a fixed proportional hazard rate for any pair of parcels. However, because we allow the interaction measure to be time-varying, the ratio of the actual hazard rates of two parcels changes as the surrounding landscape changes.

3.2. The identification problem

The vector $Z(i)$ contains all attributes associated with parcel i, and the stochastic term ε_i captures the existence of idiosyncratic factors associated with agent i. In any actual empirical application, though, data on many of the important attributes of the parcel will not be available and will reside in ε_i. This raises potential problems. Consider the types of factors that are likely to affect the value of land in developed and undeveloped uses. One set of factors will be proximity to desirable and undesirable features of the man-made landscape—such as hazardous waste sites, airports,

13 This allows us to avoid identifying those regional economic factors (such as rising incomes and population) that affect the rate at which housing starts take place in our study area.

freeways, or public parks. Another may be school quality or the quality of other publicly supplied services. Yet another may be features of the natural landscape, such as topology or soil quality. In each of these cases, the factors described can be expected to be spatially correlated. If a parcel is in a district with high quality public schools, its neighbor will have a high likelihood of being in the same high quality school district. If a parcel is situated in rolling hills, a neighboring parcel is likely to be so also. To the extent that some of these attributes will be unobserved, the unobserved heterogeneity associated with individual land parcels is likely to exhibit strong positive correlation over space. The presence of unobserved but positively spatially correlated features that influence the conversion decision complicates the identification of the interaction effects.

This version of the identification problem has arisen, not surprisingly, in the social interactions literature as well. Identification of the effects of social norms and peer pressures on individual choices requires controlling for unobserved heterogeneity (Manski, 1993, 1995; Brock and Durlauf, 2001). The same problem arises in the literature on own-state dependence over time, which seeks to separate 'true' temporal state dependence (e.g. habitual effects) from 'spurious' state dependence (Heckman, 1981, 1991). The identification problem in the land use conversion model is most similar to the social interactions case, both in terms of the source of 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 interrelated, even if they are not, and therefore complicate the task of discerning spatial interactions. Although the nature of the development process implies a temporally lagged spatial interaction effect, the presence of time-invariant unobserved heterogeneity creates the same identification problem due to correlated unobservables as that which arises in the simultaneous social interaction models.

While a variety of identification strategies have been suggested in the literature, the problem in the land use conversion model is further complicated in ways that prevent ready adoption of these strategies (Irwin, 1998). As a consequence, exact identification of the interaction parameter is not possible. However, it is possible to bound the interaction term, as suggested by Heckman and Singer (1985). In our case, the spatial correlation that exists in the unobservable factors is most likely positive on net, since such factors as distances from exogenous landscape features, topography, and school and tax districts will all exhibit positive spatial correlation.[14] As a result, the empirical estimate of the interaction effect is expected to be biased in the positive direction. This implies that the estimated interaction effect bounds the true interaction effect from above. If the estimated effect is negative, then it must hold that the 'true' interaction effect is negative for at least some range of the sample and over some interval of time. If the estimated interaction effect is positive, however, the existence of a true interaction effect is indeterminate.

14 It is, in fact, difficult to imagine many natural or man-made features that would cause negative spatial autocorrelatioin among undeveloped parcels. Regularly distributed hilltops that may be sought after for their view are one of the few landscape features that potentially could be negatively correlated.

3.3. Specification and data

Data used to estimate the land use conversion model and test for interaction effects include spatially defined, micro-level data on land parcels in the exurban areas of five counties in central Maryland. The data set was constructed using the Maryland Office of Planning's geo-coded current file of land parcels and historical information from the state's tax assessment data base. Parcels were tracked backwards in time, so that the population of parcels that could be developed in residential use as of January 1991 was identified. The data set is comprised of all privately held parcels in the exurban[15] areas of Anne Arundel, Calvert, Charles, and Howard Counties that were large enough as of 1991 to accommodate a subdivision of at least five houses given current zoning and that were not otherwise restricted from development by conversion trusts or easements.[16] The year in which conversion takes place is the event date, with parcels tracked from 1991 through 1997. Censoring occurs in 1997. An event is defined as the subdivision of an undeveloped parcel into residential lots in preparation for house construction.

To represent the neighborhood interaction term, $I_s(i, T)$, a measure of neighborhood development is calculated for each developable parcel to capture the potential spillover effects of neighbors on a parcel's conversion probability. The surrounding land use variable is constructed as the percent of the neighboring land within a certain distance s that is in a developed use in the year prior to the conversion decision. Development is defined as all commercial, industrial, and residential uses for which a structure exists on the land parcel, excluding extremely low density uses (defined by a structure on more than 5 acres). Since this variable changes over time, it is updated for every year from 1991 through 1997.

Specification of the *extent* of the relevant neighborhood around the potentially developable parcel is essentially an empirical question. Choosing too wide a radius will dilute effects, while choosing too narrow a range will omit potentially important effects of the spatial externality gradient. We specify s_{max}, the maximum distance in which we expect to find interaction effects, as equal to 1600 meters (approximately 1 mile). This is admittedly arbitrary, but work by Fleming (1999) using semi-variogram analysis supports ranges of this order of magnitude for land use interactions in this area. Since different spatial externalities may have different rates of decay, it is possible that the direction of the interaction parameter may change with distance. It is important to note that from the perspective of the individual residential housing lot, these 'neighborhoods' relate to land surrounding the lot's subdivision and not land surrounding the individual lot. Within the 1 mile maximum

15 Exurban land is defined using the 1990 US Census definition of urban fringe, which is generally categorized as a contiguous territory adjacent to an urbanized area that has a density of at least 1000 persons per square mile (US Census Bureau, Appendix A: Area Classifications, STF3 Technical Documentation, 1990). We used residential zoning restrictions of minimum lot size to approximate this definition by selecting nonurbanized areas that were zoned with minimum lot sizes of 0.5 dwelling units per acre. Given that the average household size in 1990 for these four counties was 2.92, this translates into a minimum density of 934 persons per square mile.

16 All the basic data employed are available upon request, as are the programs used to produce the generated data series.

radius, we allow for changing interaction effects by specifying two non-overlapping neighborhoods, $0 < s^* \leqslant s^*$ and $s^* < \; s \leqslant s_{max}$, and we vary s^* to see if the results are sensitive to its choice.[17] The variables DEVLUSE1 and DEVLUSE2 are measures of the fraction of neighboring land in developed uses within the two non-overlapping neighborhoods.

Of the many exogenous characteristics that could affect the likelihood of a parcel's conversion, zoning is perhaps the most troublesome. For the most part residential development is allowed throughout the study area. The only exception is the narrow strip within 1000 feet of the Chesapeake Bay coastline, defined by Maryland's Critical Areas legislation, in which development is intended to be strictly controlled. This has not translated into a complete ban on development within this zone, since some developments have been approved and others had been grandfathered. As a consequence, development in this area is less likely but not precluded. We include an index variable (CRITAR) to denote parcels that fall in the Critical Areas Zone.

Zoning's most significant effect on residential land use is to dictate maximum allowable development densities. In our study area, maximum allowable densities vary from 0.05 to two dwelling units per acre. While there is some variation within county, much of the variation occurs across counties. For example, Anne Arundel zones its most rural areas at one housing unit per 20 acres while Calvert zones such areas at one dwelling unit per 5 acres and Howard and Charles Counties at one housing unit per 3 acres. The smallest minimum lot sizes within Anne Arundel and Calvert Counties are about 1 acre, while in Charles and Howard Counties the smallest minimum lot sizes are about 0.5 acre. Variation in zoning can dramatically affect the returns to development, so we include the maximum density of allowable development in our model. This variable is included in logarithmic form (lnDUPA).

In addition, we include a set of exogenous variables intended to capture some of the other major sources of spatial heterogeneity among parcels. Following the basic insight of the urban bid-rent models, the dominant exogenous factor affecting net returns to development is arguably accessibility to urban centers. Distances to Washington, DC and to Baltimore, MD are measured via the roads network and are included in logarithmic form (lnDCDIST and lnBADIST). Costs of conversion will vary over parcels for a number of reasons including the nature of the topography and soils. We define an indicator variable (COSTCON) that takes the value of 1 for parcels that have steep slopes (more than 15%) and/or poorly drained soils and 0 otherwise.[18] Ideally we would wish to measure the opportunity costs of development using cross-sectional data on farm returns, but because of confidentiality restrictions such data are not available at the spacial resolution that we require. To substitute for this, we include an indicator variable (PRIME) that takes the value of 1 for parcels that are currently employed in agricultural activity and have prime agricultural soils, and 0 otherwise.

17 Note that this is a simplification of s. Rather than treating it as a continuous distance variable, as in Equations (3) and (4), s, now indexes two discrete neighborhoods, as defined by two non-overlapping ranges of $s, 0 < s^* \leqslant s^*$ and $s^* < \; \leqslant s_{max}$. The corresponding parameters to be estimated are λ_1 and λ_2 where 1 and 2 refer to the inner and outer neighborhood respectively as defined by s.

18 These attributes are defined according to the Maryland Department of State Planning, Natural Soil Groups of Maryland, Technical Series Publication 199 (December, 1973).

4. Empirical estimation

The proportional hazards model of (7) and (8) implies that at most one event happens at a point in time. Since our data are measured annually, there are many observational ties. We use the 'exact method' for handling these ties developed by DeLong et al. (1994). This is based on the assumption that ties arise only because the data are measured in discrete rather than continuous time. Results did not vary when other methods for treating tied events were employed.

4.1. Estimation results

Estimates of the parameters and their standard errors, together with Wald tests of their joint significance, are reported in Tables 1 and 2. Two different specifications of s^* are considered ($s^* = 800$ meters and $s^* = 1,000$ meters), although the results are qualitatively similar. We begin by presenting the most parsimonious model in Table 1 that includes our neighborhood variables (DEVLUSE1 and DEVLUSE2) together with lnDUPA and CRITAR included to control for the constraints imposed

Table 1. Results from proportional hazards model of land use conversion

	$s^*=800$ m				$s^*=1000$ m			
	Parameter estimate	Standard error	Wald chi-square	Pr > chi-square	Parameter estimate	Standard error	Wald chi-square	Pr > chi-square
DEVULSE1	−1.282336	0.55179	5.40070	0.0201	−1.067893	0.60188	3.14805	0.0760
DEVLUSE2	−1.123275	0.62853	3.19389	0.0739	−1.264581	0.62039	4.15489	0.0415
InDUPA	−0.335009	0.15791	4.50086	0.0339	−0.355132	0.15750	5.08438	0.0241
CRITAR	−0.391523	0.14605	7.18602	0.0073	−0.388760	0.14583	7.10653	0.0077
	InL = −1046.32		$\chi^2 = 57.66$		InL = −1046.96		$\chi^2 = 56.38$	

Note: Binary dependent variable = conversion to residential subdivision in a given year, 1991–97. Number of observations = 4509.

Table 2. Results from proportional hazards model of land use conversion

	$s^*=800$ m				$s^*=1,000$ m			
	Parameter estimate	Standard error	Wald chi-square	Pr > chi-square	Parameter estimate	Standard error	Wald chi-square	Pr > chi-square
DEVULSE1	−1.468307	0.55927	6.89262	0.0087	−1.524438	0.62218	6.00324	0.0143
DEVLUSE2	−2.442978	0.72250	11.43308	0.0007	−2.418691	0.71119	11.56601	0.0007
InDUPA	−0.744440	0.17830	17.43308	0.0001	−0.753419	0.17805	17.90610	0.0001
CRITAR	−0.852976	0.19790	18.57760	0.0001	−0.857964	0.19786	18.80273	0.0001
InDCDIST	−3.139205	0.42871	53.61939	0.0001	−3.162272	0.42876	54.39605	0.0001
InBADIST	−0.242259	0.55428	0.19103	0.6621	−0.272435	0.55460	0.24130	0.6233
PRIME	−1.009741	0.40193	6.31146	0.0120	−0.985413	0.40208	6.00624	0.0143
COSTCON	−1.432126	0.29955	22.85704	0.0001	−1.434015	0.29962	22.90735	0.0001
	InL = −980.62		$\chi^2 = 189.07$		InL = −980.44		$\chi^2 = 189.42$	

Note: see note to Table 1.

by zoning regulations. In Table 2 we control for a larger set of explanatory variables in an attempt to increase the precision of our estimates of neighborhood effects, but also to reduce the bias that might be introduced by omitting important explanatory variables that are spatially correlated.

In the parsimonious model reported in Table 1, we find evidence of negative interaction effects at the 90% and 95% confidence level. The larger the share of development within both the inner and outer neighborhoods, the lower the hazard of development. The size of the estimated coefficient is slightly smaller and the significance somewhat weaker for the inner as compared to the outer ring when $s^* = 1000$ m and for the outer as compared to the inner ring when $s^* = 800$ m. Given the identification problem described earlier and the likelihood that omitted variables are positively spatially correlated, this provides at least preliminary evidence of negative interaction effects among parcels in the development decision.

The estimated coefficients associated with the two additional explanatory variables are both negative and significantly different from zero at the 95% confidence level. The results with regard to the critical areas dummy variable are as expected. While development is not completely precluded from this region, it is certainly discouraged. The estimated negative coefficient on the density of development variable is consistent with theory. Assuming that returns to development are increasing over time and that the housing production function is concave, a landowner faced with a binding density constraint will find it optimal to postpone development given a marginal increase in the maximum allowable density.[19]

Table 3 reports the results when additional explanatory variables are added. Parameter estimates associated with these other exogenous effects are significantly different from zero, with one exception, and are consistent in sign with intuition. As commuting distance to Washington, DC increases, the hazard of development declines. The estimated coefficient suggests that a 1% increase in distance causes an approximate

Table 3. Results from restricted proportional hazards model

	$s^* = 1000$ m			
	Parameter estimate	Standard error	Wald chi-square	Pr > chi-square
lnDCDIST	−3.292632	0.43989	56.02723	0.0001
lnBADIST	1.101146	0.57522	3.66461	0.0556
lnDUPA	−0.948285	0.16701	32.23800	0.0001
CRITAR	−0.698391	0.19372	12.99732	0.0003
PRIME	−0.953867	0.40026	5.67938	0.0172
COSTCON	−1.437265	0.30021	22.92037	0.0001
	lnL = −1000.73		$\chi^2 = 148.84$	

Note: see note to Table 1.

19 We would like to thank the Editor for pointing this out.

3% decline in the hazard of development, holding other variables constant. The estimated coefficient associated with commuting distance to Baltimore is not significantly different from zero, however. This is not surprising. Baltimore employment opportunities tend to pay lower salaries and are more heavily weighted toward blue collar jobs than in Washington, DC. Also, a large portion of the workforce lives within the city. Commuters tend to travel from the northern and eastern Baltimore suburbs (which are not in our study area) rather than compete in the Washington, DC housing market. Finally, the areas closest to Baltimore in our study area are dominated by industrial uses and the Baltimore-Washington International airport, rather than desirable residential neighborhoods.

Estimated coefficients on the dummy variables representing prime agricultural land (PRIME) and land that is costly to excavate (COSTCON) are both significantly different from zero and negative as expected. The hazard of conversion for prime agricultural land is about 35% of that for less desirable farm land, and the hazard of land conversion for parcels with steep slopes and poorly drained soils is only 25% of that for more buildable parcels. In this set of specifications, both poor soils for construction and prime soils for agriculture depress the hazard of development more than the Critical Areas designation. The hazard for parcels found in the Critical Area is about 42% of that for parcels outside this zone.

Perhaps the most interesting result emerges from a comparison of Tables 1 and 2. With the inclusion of an expanded set of explanatory variables, the estimated coefficients on the interaction effects remain negative, but are larger in absolute value and statistically more significant. This is in keeping with our expectations that the omission of important explanatory variables may bias the neighborhood effect in the positive direction because parcel attributes are likely to be positively spatially correlated. Of course, there are additional, unmeasured explanatory variables still missing from the specification, and since we have no knowledge of these, we can not be certain that they too exhibit positive spatial correlation. However, as noted earlier, it is difficult to think of parcel attributes that might be negatively correlated in space. As a consequence, these results support the notion that the true interaction parameters are likely to be negative and that the estimated parameters on the interaction variables are lower bounds (in absolute value terms) for these 'true' interaction effects. Interpreting the estimated parameters in this manner, increasing the development in the surrounding neighborhood by 1% decreases the hazard of development by 1.5 to 2.5%.

To some, the fact that surrounding development has a depressing effect on future development might seem unrealistic, since infill development is a common occurrence in the city and inner suburbs. However, our study area is the rural-urban fringe, where one of the chief attractions is open space and reduced congestion. In addition, the magnitude of the results is based on commuting distance *being held constant*. Interpreting the results in a different way, suppose that two parcels *i* and *j* are identical except that half of the land surrounding parcel *i* is already developed none of the land surrounding parcel *j* is developed. The probability that parcels *i* and *j* are developed by the same time period will be roughly the same if parcel *i* is located 18 miles from the city center (the inner edge of the rural-urban fringe in our data set) and parcel *j* is located 33 miles from the city.[20]

20 Parcels within our exurban study area extend to 70 miles from the city center.

While the above examples are illustrative, the estimated coefficients may in fact be biased in the positive direction. The above may actually be an underestimate of the depressing effect that surrounding development has on new development. In any event, these results provide support for the notion that significant negative net interactions among neighboring parcels in residential land use may exist. Of course, it is not possible to observe these externalities directly, as we do not have direct observations on individual preferences, only market outcomes. However, we find it difficult to devise competing hypotheses that would provide alternative explanations for the negative effects, given that most unobserved factors are likely to be positively spatially correlated. One alternative hypothesis might be an institutional one—that local governments purposefully or inadvertently encourage fragmentation or discourage clustering. Purposeful intervention is unlikely; fragmentation is seen as costly from a public finance perspective and local governments in Maryland are actively combating dispersed development. Regulations such as adequate public facilities moratoria that allow local governments to slow development in areas where schools are overcrowded are a possible explanation, but the school districts upon which these moratoria are based are an order of magnitude or more larger than the extent of interaction specified in our model. Other competing hypotheses include the possibility that the intervening open spaces are protected or that the region is controlled by a few large developers that find it optimal to diversify their spatial portfolio. The majority of conserved land in this region is protected as agricultural preserved land, which tends to be positively spatially correlated due to a regulation that requires newly preserved land to be contiguous to existing agricultural districts. The latter hypothesis is possible, although the development market is large and it is likely that considerable competition exists among developers.

4.2. Comparison of actual and predicted spatial patterns

To further interpret these results, we use the estimated parameters to simulate predicted changes in land use. Using a small portion of our study area, we attempt to illustrate the empirical model's ability to explain actual land use conversion pattern. Before doing so, we estimate one final model, one in which the interaction effects are restricted to zero.[21] These results are reported in Table 3. Using parameter estimates from the full model (estimated with $s^* = 1000$ m) and this restricted model, changes in the 1990 land use configuration of northeast Charles County are predicted. For each parcel that was 'developable' in 1990, the time-invariant exogenous attributes and the time-varying neighborhood land use variables are recorded. Each parcel's likelihood of conversion is then calculated using both the full and restricted models' parameter estimates. In order to translate probabilistic measures of conversion into actual conversion, a constant regional demand for new housing is assumed and the parcel with the highest probability of conversion in each time period is the parcel chosen for conversion, where a time period is arbitrarily defined as the period long enough for one conversion to take place. Neighborhood interaction effects are recalculated after each predicted conversion for the unrestricted case. Simulations using both the full and restricted models' estimated parameters are carried out for 114 rounds of development and the results are then

21 A Wald test comparing the restricted to the unrestricted model is significant at the 95% confidence level.

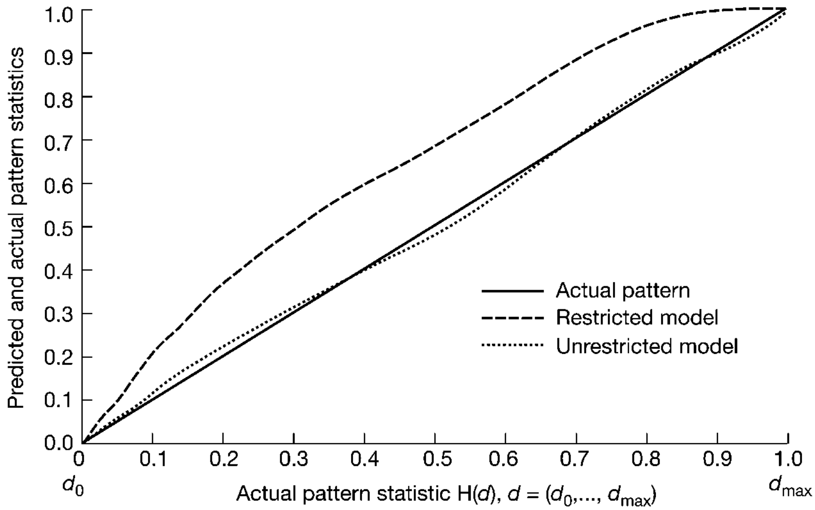


Figure 1. Spatial distribution of inter-distance point statistic $H(d)$ for actual and predicted patterns of development in Northeast Charles County, Maryland.

compared with the actual pattern of 114 subdivisions that were developed in northeast Charles County between 1991 and 1997.

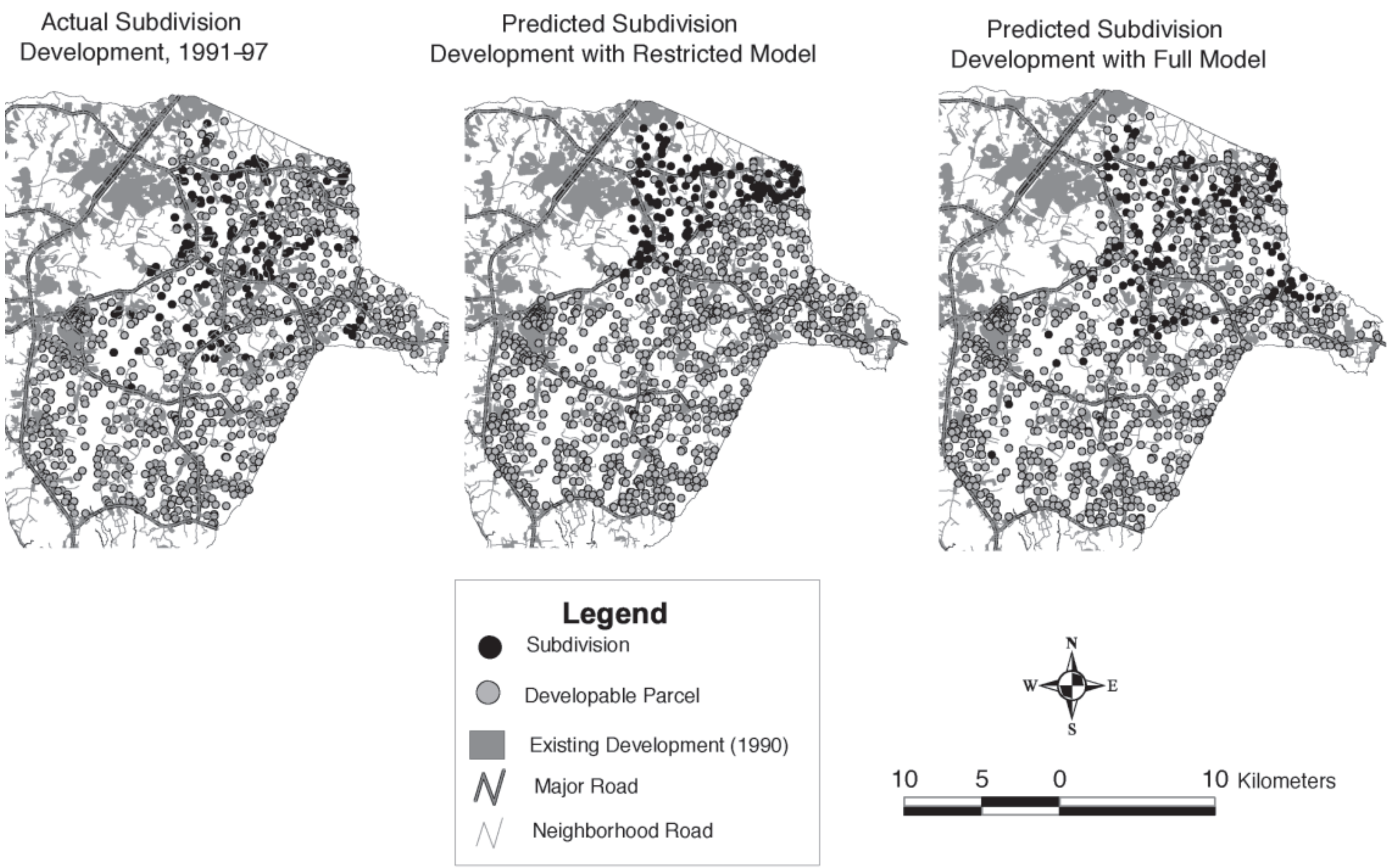
Map 4 shows the comparison of the predicted patterns with observed changes in the development pattern in northeast Charles County between January 1991 and December 1997. Each point corresponds to the centroid of a parcel that could be subdivided as of 1991. In comparing the two predicted patterns to the actual development pattern, the pattern simulated with only the exogenous effects generates a much higher degree of clustered development than the actual development pattern, as would be expected. In this case, the location of development is primarily determined by relative accessibility to Washington, DC, located to the northwest of the developable region. The inclusion of the negative interaction effects generates a pattern that is significantly more fragmented and one that appears to mimic more closely the actual pattern of residential subdivision development.

In order to quantify the differences between the actual pattern and the two predicted patterns, an inter-distance point statistic is used to summarize each pattern of n points.[22] This statistic is a count variable that tallies the number of paired points whose inter-point distance falls within increasing distance ranges. The counts are normalized by $n(n-1)$ and the distance ranges are cumulative. For any range, d, the statistic is given by:

$$H(d) = \Sigma_{ij}\Phi(d_{ij} \leqslant d)/n(n-1).$$

where $\Phi()$ is an indicator variable such that $\Phi(d_{ij}) = 1$ if points i and j are within distance d of each other and 0 otherwise; $d \in \{0, d_{max}\}$, where d_{max} is the extent of the region; and $H(d) \in \{0, 1\}$.

22 To calculate this statistic, we adopt the methods outlined in the spatial statistics literature (Diggle, 1984; Cressie, 1993).



Map 4. Northeast Charles County, Maryland, actual vs predicted development.

To gauge the degree of difference between predicted and actual patterns, the inter-distance point statistics from the actual vs predicted patterns are calculated for the same intervals of d and plotted against each other. The statistic representing the actual development pattern is mapped against itself for the relevant distance range, $d = 0, \ldots, d_{max}$. The statistic representing the predicted pattern is mapped against the actual pattern statistic for each distance range, so that the degree of difference between the actual and predicted patterns is evidenced by the degree to which the plot of the predicted pattern statistics differs from the 45° line. This is illustrated in Figure 1 for the predicted patterns generated by the restricted and full models. The statistic corresponding to the unrestricted model lies quite close to the 45° line, suggesting that the spatial pattern predicted by this model is qualitatively similar to the actual pattern. In contrast, the statistic corresponding to the restricted model lies well above the diagonal, suggesting that this pattern has a much higher degree of positive spatial correlation than either the actual pattern or the pattern simulated with the inclusion of the interaction effect. These observations provide further support for a model that incorporates both exogenous landscape features and interaction effects.

5. Implications

Several recent theoretical papers in urban and regional economics have focused on the role of interdependencies among spatially distributed agents in the formation of urban spatial structure. In most of these, theories about the role of agent interactions have focused on the formation of urban centers. In this paper, we adopt an interactions-based model to study the influence of land use spillovers and exogenous landscape features on the evolution of land use pattern at the rural-urban fringe. By incorporating the influences of both exogenous landscape features and interaction effects, our model permits an empirical test of the interactions hypothesis. We find support for a hypothesis of negative spillovers among exurban land parcels converted to residential subdivisions. Although an unbiased estimate of the interaction effect is not possible, the results nonetheless provide the first empirical evidence of the potential role of negative interactions in the formation of sprawl development patterns at the rural-urban fringe.

Much of the theoretical economics literature to date has portrayed sprawl as a dynamically efficient outcome of the development process (Ohls and Pines, 1975; Fujita, 1976; Wheaton, 1982; Braid, 1991). Under conditions of uncertainty, it is viewed minimally as ex ante efficient: given initial uncertainty and irreversible development, developers make optimal intertemporal tradeoffs based on their expectations, which may lead to discontinuous development that is only inefficient ex post (Mills, 1981). In all of these approaches, the monocentric model is the framework within which the theoretical explanation for sprawl is developed. Because location relative to other agents is not accounted for in this model, this approach is not sufficiently general to explain sprawl as the result of another market failure—namely, externalities. Our results suggest that such externalities may be present.

If indeed externalities of the sort we hypothesize exist, then they bring into question the claim that the unregulated land market is efficient, whether under conditions of perfect foresight or uncertainty. Rather than being the result of an optimal intertemporal strategy, in which land is purposely withheld from development because of anticipated higher returns in the future, sprawl may be caused by negative externalities between developments. The implication is that undeveloped land that is adjacent to development is *less* valuable in residential use and therefore less likely, ceteris paribus, to be developed residentially in the future. A suboptimal pattern may emerge because individual landowners will fail to take into consideration the external cost of their decision to develop on the development potential of neighboring land. Whether the resulting spatial pattern of land use will be more or less than the socially optimal level of sprawl remains unanswered and will depend on how other potential external benefits and costs not considered here vary with the development pattern, e.g. public service provision costs, landscape amenities, and environmental impacts. In any case, it is impossible to conclude that an optimal outcome is necessarily obtained given the presence of the externality.

Acknowledgements

Earlier versions of this paper were presented at the 1998 NBER Summer Workshop on Public Policy and the Environment, the 1999 North American Meetings of the Regional Science Association International, and the 2000 W-133 Regional Research Project Meetings. The authors would like to thank participants at these meetings as well as William Brock, Kerry Smith, James Opaluch, Richard Arnott, and two anonymous reviewers for helpful suggestions. Support for this research was provided by the MacArthur Foundation and the Environment Protection Agency under Cooperative Agreement CR-821925010.

References

Anas, A. (1992) On the birth and growth of cities: laissez-faire and planning compared. *Regional Science and Urban Economics*, 22(2): 243–258.

Anas, A. L., Kim, I. (1996) General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. *Journal of Urban Economics*, 40(2): 232–256.

Anas, A., Arnott, R., Small, K. (1998) Urban spatial structure. *Journal of Economic Literature*, 36: 1426–1464.

Arnott, R., Lewis, F. (1979) The transition of land to urban use. *Journal of Political Economy*, 87(11): 161–169.

Arthur, W. B. (1988) Urban systems and historical path dependence. In J. Ausubel and R. Herman (eds) *Cities and Their Vital Systems*. Washington, DC. National Academy Press.

Bar-Ilan, A., Strange, W. (1996) Urban development with lags. *Journal of Urban Economics*, 39: 87–113.

Beckmann, M. (1976) Spatial equilibrium in the dispersed city. In G.J. Papageorgiou (ed.) *Mathematical Land Use*. Lexington, MA: Lexington Books.

Braid, R. (1988) Uniform spatial growth with perfect foresight and durable housing. *Journal of Urban Economics*, 23: 41–59.

Brock, W., Durlauf, S. (2001) Interactions-based models. In J. J. Heckman and E. E. Leamer (eds) *The Handbook of Economics 5*. Amsterdam: Elsevier Science.

Capozza, D., Helsley, R. (1989) The fundamentals of land prices and urban growth. *Journal of Urban Economics*, 26 (3): 295–306.

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

DeLong, D. M., Guirguis, G., So, Y. C. (1994) Efficient computation of subset selection probabilities with application to Cox regression. *Biometrika*, 81: 607–11.

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

Fleming, M. (1999) Growth controls and fragmented suburban development: the effect on land values. *Geographic Information Sciences*, 5 (2): 154–162.

Fujita, M. (1976) Spatial patterns of urban growth: optimum and market. *Journal of Urban Economics*, 3: 209–241.

Fujita, M. (1988) A monopolistic competition model of spatial agglomeration: differentiated products approach. *Regional Science and Urban Economics*, 18 (1): 87–124.

Heckman, J. J. (1981) Statistical models for discrete panel data. In C. Manski and D. McFadden (eds) *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press.

Heckman, J. J. (1991) Identifying the hand of the past: distinguishing state dependence from heterogeneity. *American Economic Review*, 81 (2): 75–79.

Heckman, J. J., Singer, B. (1985) Social science duration analysis. In J. Heckman and B. Singer (eds) *Longitudinal Analysis of Labor Market Data*. Cambridge: Cambridge University Press.

Irwin, E. (1998) Do spacial externalities matter? The role of endogenous interactions in the evolution of land use pattern. Unpublished Ph.D dissertation, University of Maryland.

Krugman, P. (1991) Increasing returns and economic geography. *Journal of Political Economy*, 99: 438–499.

Krugman, P. (1996) *The Self-Organizing Economy*. Cambridge, MA: Blackwell Publishers.

Manski, C. (1993) Identification of endogenous social effects: the reflection problem. *Review of Economic Studies*, 60: 531–542.

Maryland Department of Planning (1973) *Natural Soil Groups of Maryland, Technical Series Publication* 199, Annapolis, MD: Maryland Department of Planning.

Mills, D. (1981) Growth, speculation, and sprawl in a monocentric city. *Journal of Urban Economics*, 10:201–226.

Ohls, J. Pines, D. (1975) Discontinuous urban development and economic efficiency. *Land Economics*, 3: 224–234.

Page, S. (1999) On the emergence of cities. *Journal of Urban Economics*, 45 (1): 184–208.

Papageorgiou, Y.Y., Smith, A.R. (1983) Agglomeration as local instability in spatially uniform steady-states. *Econometrica*, 51: 1109–1119.

US Census Bureau, US Department of Commerce (1990) Appendix A: Area Classifications, STF3 Technical Documentation, US Census of Population and Housing. Washington, DC: US Department of Commerce, Bureau of the Census.

Wheaton, W. (1982) Urban residential growth under perfect foresight. *Journal of Urban Economics*, 12: 1–21.

Zhang, W. B. (1993) Location choice and land use in an isolated state: endogenous capital and knowledge accumulation. *Annals of Regional Science*, 27 (1): 23–39.