Coming Undone: A Spatial Hazard Analysis of Urban Form in American Regions

John I. Carruthers ★ U.S. Department of Housing and Urban Development, Office of Policy Development and Research; University of Maryland, National Center for Smart Growth Research and Education; e-mail: *john.i.carruthers@hud.gov*

Selma Lewis U.S. Department of Housing and Urban Development, Office of Policy Development and Research; University of Maryland, National Center for Smart Growth Research and Education; e-mail: *selma.lewis@hud.gov*

Gerrit-Jan Knaap University of Maryland, National Center for Smart Growth Research and Education; e-mail: *gknaap@umd.edu*

Robert N. Renner U.S. Department of Housing and Urban Development, Office of Policy Development and Research; e-mail: *robert.n.renner@hud.gov*

★ Corresponding author

U.S. Department of Housing and Urban Development Working Paper # REP 08-03; October 2008

This paper was prepared for presentation at the 2008 meetings of the Spatial Econometrics Association In New York, New York and the 2009 meetings of the Western Regional Science Association in Napa, California. Special thanks to Alastair McFarlane and Peter Gordon for their very helpful comments and feedback.

Abstract. This paper explores the viability of using proportional hazard models to study spatial point patterns generated by urbanization. The analysis demonstrates that the "spatial hazard" framework is not only viable for studying urban form, but is extremely promising: the models do an excellent job of characterizing very different patterns of development, and they lend themselves directly to the kind of probative analysis needed to guide urban and regional policy. Compared to more traditional approaches to characterizing urban form — namely, density gradients — hazard models rest on a probabilistic worldview, and, so, they portray the built environment as a quantum-like froth of stochastic transitions through which urban form unfolds in an irregular fashion until it at last comes undone. Several general conclusions and directions for future research follow from these findings.

1. Introduction

Regional science has always been eminently concerned with urban form — what Isard (1956, page 11) called the "spatial physiognomy" of urbanization. The field produced the first general theory of urban land use and has since continued to deliver increasingly sophisticated explanations for why the built environment looks as it does (Fujita 1987). Along the way, empirical research has also focused on analyzing urban form in a manner that follows directly from the socioeconomic processes that shape it. In particular, estimated density gradients (see McDonald 1988) reveal an urban fabric that is tightly knit around various points of attraction but progressively comes undone, or, as some say, "sprawls" (Bruegmann 2005) with distance from these locations. This basic pattern of land use, whether monocentric or polycentric overall, is observed worldwide and throughout history (Anas et al 1998). With its powerful theoretical models and deep empirical insights, regional science has been fundamental to establishing urban form as both a positive and normative concept.

This paper continues that tradition by exploring the viability of using proportional hazard models — a class of duration, or failure time, models normally used for analyzing lifecycles (Kiefer 1988; Odland and Ellis 1992; Lawless 2002; Waldorf 2003; Cleves et al 2004) — to study spatial point patterns generated by urbanization. The analysis has three specific objectives: (*i*) to review pertinent theoretical and empirical research on urban form and explain how the spatial hazard framework compares to traditional approaches; (*ii*) to estimate a series of spatial hazard models characterizing the built environment in the 25 largest core based statistical areas of the United States and consider the results; and (*iii*) to illustrate how spatial hazard models may be used to evaluate smart growth, or growth management, policies aimed at shaping the outcome of urbanization. The analysis, which involves over 76,000 census block groups containing nearly 42.5 million housing units, or about a third of the nation's total, demonstrates that the spatial hazard framework is not only viable for studying urban form, but is extremely promising. On the positive front, the models do an excellent job of characterizing very different patterns of regional development, and, on the normative front, they lend themselves in a natural way to the kind of analysis needed to make sound public policy.

2. Background Discussion

2.1 Theoretical Models of Urban Form

Modern economic models of urban form are rooted in von Thünen's (1826) theory of agricultural land use. The traditional framework — jointly credited to Alonso (1964), Muth (1969), and Mills

(1972) — describes a locational rent gradient that monotonically falls away from its peak around a business district at the geometric center of a circular region that is situated on an otherwise flat, featureless plain. At market equilibrium, all households, which are assumed to be identical, attain the same level of utility, meaning that the rent gradient perfectly reflects the tradeoff between location and the cost of travel to and from the central business district. Meanwhile, a corresponding density gradient emerges as a result of households occupying more-and-more space toward the urban fringe. Land is a normal good, so, in order to maintain a fixed level of utility, households consume greater amounts of it through substitution with other goods as rent declines. The density gradient and, with it, urbanization, end altogether once locational rent reaches zero and the highest and best use of land is for agriculture or some other natural resource oriented activity. Together, these basic precepts make up the so-called "monocentric model" of urban form, and they readily generalize more elaborate polycentric models having multiple business districts.

More formally, the theory underpinning economic models of urban form assumes that households have a common utility function, $U(z,s)$, which contains a composite good, or numeraire, z, and urban space, or land, s (see Fujita 1987 for a complete exposition of the material outlined in this and the following paragraph). An individual household's budgetary constraint is determined by its income, y, less the cost of travel, k, between the central business district and its location at radial distance d: $y - k(d) = z + r(d)s$, where $k(d)$ is a continuous function that increases with d and $r(d)$ is the locational rent per unit of space at d. Within this framework, households maximize their utility by choosing some combination of the numeraire and land, subject to their particular — spatially explicit — budgetary constraint:

$$\max_{d,z,s} U(z,s) \;\Big|\; z + r(d)s = y - k(d). \tag{1}$$

The outcome of this utility maximization problem is a household's bid rent, $\rho(d,\ u)$, which is the maximum price they are able to pay for land at distance d while maintaining a fixed level of utility, u:

$$\rho(d,\ u) = \max_{z,s} \left\{ \frac{y - k(d) - z}{s} \;\Big|\; U(z,s) = u \right\}. \tag{2}$$

In plain terms, bid rent is the most that a household is willing to pay per unit of space in order to secure the right to occupy their location of choice, given that they derive happiness from both land and other forms of consumption. In addition to land prices, bid rent yields a household's optimal amount of land consumption, or lot size, $\varsigma(d,u)$, which is what ultimately gives the built environment its character.

Figure 1 illustrates how the bidding process just described translates into a physical pattern of urbanization. It displays the marginal rate of substitution, described by an indifference curve (the arc) for a fixed level of utility, u, between the numeraire, z, and land, s, plus the budget constraints (the dashed lines) and corresponding consumption bundles (the dotted lines) associated with two households located at distances d_1 and d_2 from the central business district, where $d_1 < d_2$. Because the cost of travel to and from the central business district, $k(d)$, is lower at d_1 than it is at d_2, the net income of the household located at d_1 is greater than the net income of the household at d_2, or $y - k(d_1) > y - k(d_2)$. The two budget constraints, which must be tangent to the indifference curve in order for their respective households to each achieve utility level u, show that: (*i*) the bid rent, which is equivalent to the slope of the budget constraint, for the household located at d_1 is greater than the bid rent for the household located at d_2, or $\rho(d_1, u) > \rho(d_2, u)$; and (*ii*) the optimal lot size for the household located at d_1 is less than the optimal lot size for the household located at d_2, or $\varsigma(d_1,u) < \varsigma(d_2,u)$. In short, households located closer to the center of the region pay a higher price per unit of land and, so, consume less of it by substituting more of the numeraire than households located further out. This bidding process and the tradeoffs it involves are what shape the outcome of urbanization.

But, as powerful as the traditional model of urban form is — and decades of productive research attest to its ability to characterize both monocentric and polycentric regions — it has a certain practical flaw: it implicitly requires the built environment to be continually torn down and redeveloped over time in order to accommodate new development (Brueckner 1987). Population growth from period-to-period simply produces greater locational rent and higher density across the entire region and an expanded regional footprint. These results are not realistic, though, because they mean that all existing structures have somehow been instantaneously replaced when, in practice, the built environment is durable and generally remains more-or-less the same for decades or even centuries (Glaeser and Gyourko 2005).[1] An alternative approach is a dynamic, or "vintage," model of urban form that that accounts for the fact that, in practice, structures are torn down and rebuilt over time according to their age and contemporary market conditions (see Brueckner 2000 for an overview). Compared to the traditional framework, which always yields perfectly smooth density gradients, the vintage framework retains all of the same core relationships — but produces jagged patterns of urbanization that transition across geographic space in irregular ways. By bringing the age of the built environment into the mix, vintage models

[1] Some other reasons for the persistence of the built environment include land use controls, like zoning and historic preservation districts, and redevelopment costs, which, when subtracted from any anticipated profit from redevelopment, may indicate that the highest and best use of the property in question is its existing density — even if the market would call for a higher density if it were just being developed for the first time.

of urban form significantly enhance the correspondence between theoretical and actual land use outcomes. For present purposes, the key insight is that it may work particularly well to model observed patterns of urbanization in a way that accommodates their seemingly stochastic nature.

2.2 From Density Gradients to Spatial Hazards

More than a decade before economic models of urban form had been formalized, Clark (1951) found that regional population densities tend to decline along a negative exponential path:

$$\delta(d) = \delta_0 \cdot \exp(-\gamma \cdot d)\,. \tag{3}$$

In this equation, $\delta(d)$ represents residential density at radial distance d from the central business district; δ_0 represents population density at the center, where $d = 0$; and γ represents the density gradient, which registers the rate of decline per unit of distance. With data on small patches of a region — parcels, census tracts, or some other spatial unit — equation (3) can easily be expanded to include factors other than location and, after taking the natural logarithm of both sides, estimated via ordinary least squares:

$$\ln\delta(d_i) = \ln\delta_0 - \gamma \cdot d_i + \alpha_1 \cdot \ln x_{i1} ... + \alpha_k \cdot \ln x_{ik} + \upsilon_i, \tag{5}$$

where the notation is essentially the same as before, except that $x_{i1} ... x_{ik}$ are measures of relevant explanatory variables, including income and the cost of travel; the αs are estimable parameters; and υ_i is a random error term. This technique is straightforward and extremely flexible, so it has been used to analyze urbanization worldwide over the past half-century — it consistently produces patterns that that line up well with those predicted by various theoretical models of urban form (McDonald 1988).

Though nonetheless informative, the limitations of density gradients are several. To begin with, as already noted, in reality urban form rarely unfolds along a perfectly smooth, monotonic path: many, if not most, contemporary regions have a polycentric spatial structure — even if they nonetheless also have a prime center of gravity, as is often the case — and the durability of the built environment ensures that period-specific market conditions persist long after their time has passed. Moreover, from a theoretical standpoint, even the negative exponential form itself is only technically justifiable under strong assumptions about the nature of housing production, household preferences, and commuting costs (see Brueckner 1982). So, although, overall, density nearly always declines with distance from the regional center of gravity, a number of studies have shown that negative exponential density gradients may oversimplify, or even outright mischaracterize, actual development patterns (Kau and Lee 1976a, 1976b, 1977; Johnson and Kau 1980; Kau et al 1983). While these and other findings indicate that the exponential functional form is not always best for describing urbanization, there is no real consensus on an alternative general approach — or even on whether one exists at all. Nevertheless, the density gradient represents the cornerstone of nearly all of the empirical research that economists have conducted on urban form since formal models of land use were first developed.

An alternative, geographically oriented, way of looking at the “fine structure” of urban form (Anas et al 1998, page 1431) is via point pattern analysis, a type of spatial analysis that is used extensively in the natural and social sciences alike (Boots and Getis 1988; Diggle 2003). Broadly defined, point pattern analysis is concerned with finding the degree of order or randomness in the arrangement and dispersion of mapped events, like the distribution of settlements and the spread of illness during an epidemic, in order to learn something about the spatial process that generated them. The main idea is simple: observed patterns are mathematically tested against theoretical patterns and/or patterns that result from a completely random point generating process. The technique is especially well suited for evaluating economic theories of location (see Duranton and Overman 2008 for a recent example) and it has been applied to urban form since at least the 1960s. Getis used point pattern analysis early on to examine commercial and residential land use succession in the Lansing, Michigan region (1964), and, later, to identify population clusters in the Chicago, Illinois region (1983). In the context of urbanization, the spacing of points, such as structures or population centers, is obviously related to density, but it is different in the sense that it also reflects something more qualitative about the structure of the built environment — the degree of what Lessinger (1961) famously (in urban planning circles) referred to as the “scatteration” of development.

The problem with point pattern analysis is that it is not a true behavioral approach to hypothesis testing (Odland and Ellis 1992). Its strength lies in its ability to produce inferences that are unambiguously tied to a particular spatial distribution: empirical evidence either confirms or rejects the existence of a hypothesized pattern and, by association, all of its implications. Beyond this, point pattern analysis reveals nothing about alternative patterns or, perhaps more critically, the role that specific socioeconomic factors play in contributing to the observed outcome. In the case of urban form, which is most often characterized along some sort of continuum — like from compact to sprawling — discrete descriptors typically have only limited value, especially because so much rides on whatever implications the pattern is assumed to hold. So, though excellent for qualitative description, point pattern analysis lacks the kind of behavioral nuance needed to recover detailed information on the kaleidoscope of variables that shape urban form in contemporary regions. It can be used measure degrees of compactness versus sprawl, for example, but it cannot in-and-of-itself be used explain how a particular pattern evolved, or to identify how to alter its course.

Given the offsetting strengths and weaknesses that density gradients and point pattern analysis bring to the challenge of characterizing urbanization, it seems desirable to develop a methodology that combines the best of what the two have to offer. On the upsides: (*i*) density gradients are an econometric approach that enables urban form to be modeled under the auspices of the very powerful behavioral framework outlined above; and (*ii*) point pattern analysis is an approach that accommodates the qualitative imperfections of actual, as opposed to theoretical, modes of urbanization. An effective way of integrating the two approaches — originally proposed by Odland and Ellis (1992) and persuasively elaborated upon by Waldorf (2003) — is by adapting proportional hazard models, also called accelerated failure time models, to the geographic realm. These models hold tremendous potential for the study of urban form because they are probabilistic in nature and therefore, by design, are able to accommodate the seemingly stochastic character of the built environment.

2.3 A Spatial Hazard Model of Urbanization

Hazard models are longitudinal models designed to estimate the conditional probability of a timeframe coming to an end (see Kiefer 1988; Lawless 2002; Cleves et al 2004). They were first developed in engineering to examine the lifespan of products, like light bulbs and ball-bearings, and have since been applied to a wide array of temporal problems in economics, epidemiology, regional science, and other fields. Just like time, distance, D, is a nonnegative random variable that terminates at a given point conditional on the probability of having made it to that point in the first place. At the core of this conceptualization is the spatial hazard function, which, following Waldorf (2003), is:

$$h(d) = \lim_{\Delta d \to 0} \frac{\Pr(D \in [d, d + \Delta d] \mid D \geq d)}{\Delta d} \in (0, \infty). \quad (7)$$

This function describes the instantaneous rate at which distances between spatial points — settlements or outbreaks of illness, to recycle examples from above — terminate at distance d, conditioned on the probability that they already extend that far in the first place. It shows that the hazard function can have either positive or negative spatial dependence, meaning, respectively, that the probability of terminating increases or decreases with distance. Intuitively, one might expect the hazard function for the spacing of settlements to exhibit positive dependence and the hazard function for the spread of an illness to exhibit negative dependence — in other words, it seems reasonable to expect that the probability of the distance between settlements terminating increases with distance, and that the probability of the distance an illness travels terminating decreases with distance. Zero dependence corresponds to complete spatial randomness, and estimating the form of dependence (positive, negative, or zero) in the spatial hazard function is generally analogous to the kind of hypothesis testing carried out in more traditional forms of point pattern analysis.

A behavioral model of a particular point generating process is achieved by choosing an appropriate statistical distribution for the baseline hazard, plus a set of independent explanatory variables that accelerate and/or decelerate the rate at which distances between spatial points terminate. The Weibull distribution is apparently the most widely used distribution in survival analysis (other common distributions include the exponential, log-logistic, and Gamma) and it works particularly well because it gives the hazard a flexible shape:

$$h(d \mid X) = h_0(d) \cdot \exp(X \cdot \Phi). \qquad (8)$$

In this spatial hazard model, the hazard function consists of two components: (1) a baseline hazard, $h_0(d) = \lambda d^{\lambda - 1}$, described by λ, a shape parameter, which gives the rate at which the distances between spatial points terminate when $X = 0$; and (2) an exponential scale parameter, Φ, which accelerates or decelerates the baseline hazard, depending on how the independent factors in the vector X influence the termination rate. Both the shape and scale parameters have to be estimated via maximum likelihood. While still not common, variations on the spatial hazard approach have been applied to an array of geographic phenomena ranging from the spacing of settlements (Odland and Ellis 1992); to the separation between parents and their adult children (Rogerson et al 1993); to the reach of market areas (Esparza and Krmenec 1994, 1996); to the adoption of agricultural technology (Pellegrini and Reader 1996); to the spread of disease (Reader 2000).

Last, in spatial hazard analysis, like all spatial analysis, the key trick is to structure the experimental setting in a way that lines up not only with theory, but also the logic of the models themselves. In the present instance, economic theory clearly indicates that the hazard function for distance separating the spatial points, whether structures, small-area population centers, or something else, that make up the urban fabric is bound to exhibit positive spatial dependence — but, also, that the hazard decelerates with distance from the interior of the region. Urbanization is, by definition, tightly woven, but, as illustrated in Figure 1, it gradually comes undone as a result of households consuming more-and-more space toward the fringe. Based on this theoretical framework, a Weibull distributed spatial hazard model of urban form is operationalized as follows:

$$h(d_{ij} \mid X_{ik}) = h_0(d_{ij}) \cdot \exp(\phi_{d_{ic}} \cdot x_{d_{ic}} + X_{ik} \cdot \Phi_k). \qquad (9)$$

Here, $h(d_{ij} \mid X_{ik})$ indicates that the baseline hazard for distance between nearest neighbors *i* and *j*, $h_0(d_{ij})$, is scaled by X_{ik}, a vector of *k* independent variables, including $x_{d_{ic}}$, the distance from *i* to the regional center; and Φ_k (including $\phi_{d_{ic}}$) measures the influence the vector of independent variables has on the conditional probability of distance between nearest neighbors terminating. To reiterate, the two hypotheses at the heart of this model, both of which flow directly from urban economic theory (Fujita 1987), are: (*i*) the conditional probability of distance between nearest neighbors terminating increases with distance; and (*ii*) the probability of terminating decelerates with distance from the center of the region. Note, too, that, because the model is probabilistic in nature, it is highly flexible in the sense that there is no requirement that these spatial transitions play out smoothly the way there is most density gradients. Moreover, there is even no requirement that urbanization proceeds consistently around the circumference of the region in question — in reality, regions may be monocentric or polycentric and, either way, regularly shaped or irregularly shaped, and the hazard framework should accommodate any combination of these.

3. Empirical Analysis

3.1 Data and Econometric Specification

The empirical analysis is set in the 25 largest core based statistical areas (CBSAs) of the United States. In cases where the CBSA is composed of two or more divisions — the New York City CBSA, for example, is made up of four separate divisions — the divisions themselves are used, so, counting all of these, the actual number of areas considered is 43. The units of analysis are census block groups (2000 definition) and the data itself comes from two sources: (*i*) a special nationwide count of housing units at the census block level in 2006, which was provided to the Department of Housing and Urban Development by the Census Bureau;[2] and (*ii*) Census Summary File 3 (SF-3), from the 2000 census of the population. The next paragraphs outline the development of the spatial data needed for the analysis and explain the full econometric specification in turn.

Spatial point patterns representing the urban fabric of the regions engaged in the analysis, plus relevant distance measurements, were developed from the 2006 housing unit counts using a geographic information system (GIS) via a five-step process (see Renner et al 2008 for an in-depth explanation of the data preparation process). In the first step, a base map consisting of all

[2] This count represents the universe for the American Community Survey, an annual survey of about three million households that is set to replace the so-called "long form" of the decennial census, which will eventually yield census tract level data on an annual basis.

census blocks in the continental United States — there are 8,205,582 — was created and the number of housing units in them was used to generate a housing unit weighted center for each of the 208,643 block-groups that make up the country. This so-called "mean center" is a point that marks where housing units are concentrated in the block group, as opposed to the block group's geometric center. In the second step, the same routine was run to generate housing unit weighted centers the 939 CBSAs in the country; here again, the points produced by this process mark the mean center of the regions. This step yields the regional center of gravity, which may be different from the central business district, especially in regions that have a strong polycentric structure. In the third step, each block group's mean center point was assigned to a CBSA mean center point, whether it officially belongs there or not, via a nearest neighbor routine. In the fourth and fifth steps, respectively, the GIS was used to generate a set of rays measuring the distances separating block groups from their regional center and another set of rays measuring the distance separating nearest neighbor block groups from one another. The results of these final two steps are shown in Figure 2, a map of CBSAs and their spheres of influence with the 43 regions that are the focus of this analysis shown in dark gray, and Figure 3, a map of spatial point patterns in the Chicago, IL, Dallas, TX, Los Angeles, CA, and New York, NY regions. In the latter figure, both the rays connecting block groups to their regional center and the rays connecting nearest neighbor block groups are visible.

The machinations just described yielded the two variables most important for estimating equation (9), d_{ij} and $x_{d_{ic}}$. The remaining independent variables that fill out the vector X_{ik} were identified from the economic models of urban form outlined in the background discussion. In particular, the framework holds three implications that are essential for understanding how the built environments of different regions — and, for that matter, neighborhoods within them — compare to each other: (*i*) land is a normal good, so household income positively affects the optimal lot size, meaning that income is expected to decelerate the hazard of the distance between points terminating; (*ii*) commuting costs determine the budgetary constraint, so time spent traveling to work, which is also a general measure of accessibility, is expected to either accelerate or decelerate the hazard of the distance between points terminating depending on region-specific conditions; and (*iii*) because of vintage effects, aged development, which is often of a lower density than contemporary market conditions call for, is expected to decelerate the hazard of the distance between points terminating. Beyond these factors, the number of housing units is included in order to control for the simple fact that, other things being equal, larger block groups will encompass a larger area; this variable is expected to decelerate the hazard of the distance between points terminating. Table 1 gives the specific definition and source of each variable and

Table 2, which orders the 25 CBSAs from largest to smallest, lists the local sample size, plus descriptive statistics by CBSA.

3.2 Estimation Results — Stochastic Transitions

The maximum likelihood estimation results for the 43 spatial hazard models, which were generated using the *streg* command in *Stata*, are given by region in Table 3. Every region's shape parameter, λ, is positive and statistically significant at well over a 99% level of confidence, indicating that the baseline hazard for each exhibits positive spatial dependence, or, in other words, the greater the distance between nearest neighbor block groups, the greater the conditional probability of that distance terminating. The shape parameters further indicate that the urban fabric of these regions exhibits genuine, probabilistic order — there is no evidence of the kind of complete spatial randomness that some commentators use to characterize sprawl. Indeed, the models roundly illustrate that urban form transitions across geographic space exactly as theory predicts, even if those transitions unfold stochastically.

Moving on, the first parameter estimate under the ϕ header, for $x_{d_{ic}}$, the distance from the CBSA center, is negative and statistically significant in all but three instances: Edison, NY; Ft. Lauderdale, FL; and Rockingham County-Stratford County, NH. The significant parameters indicate that the baseline hazard decelerates — meaning that the housing unit weighted block group centers become further spaced — toward the exterior of regions. The insignificant parameters, two of which are for purely suburban areas, indicate that the baseline hazard is spatially invariant. The next parameter estimate, for household income, is always significant and negative, except in the instance of Tacoma, WA, where it is not statistically significant. Income positively influences the ability of households to consume land, and, for that reason, it decelerates the baseline hazard function and causes urbanization to be more spread out. The third parameter estimate, for travel cost, measured as time after holding income constant, is almost always significant but its sign varies by region. In those regions having a positive sign, the cost of travel is associated with a more compact pattern of urbanization whereas, in those regions having a negative sign, it is, perhaps, associated with more sprawl. The fourth parameter estimate, for the age of the built environment, is also nearly always significant and, when significant, it is negative in all but four cases: Wilmington, DE; Ft. Lauderdale; Bethesda-Frederick-Gaithersburg, MD; and Tacoma. The negative signs on this variable indicate that it decelerates the baseline hazard due to vintage effects — older development is commonly less dense than contemporary market conditions call for, so the built environment exhibits corresponding age-related idiosyncrasies. Last, as a control having to do with the geographic size of the block groups, the number of housing units is nearly always significant and is negative, as expected, in all but one instance: Bethesda-Frederick-Gaithersburg. The positive case in interesting because the area in question is the northern suburb-exurb of Washington, DC — so it may be indicative of some sort of compacting effect owed to its various town centers. All of these results are logical and consistent with expectations.

Turning now to aggregate patterns of urbanization, the real object of this analysis, the 43 spatial hazard models must be evaluated vis-à-vis the two hypotheses that lie at the heart experimental framework — namely, that: (*i*) the conditional probability of distance between nearest neighbors terminating increases with distance; and (*ii*) the probability of terminating decelerates with distance from the regional center of gravity. Recasting the hazard functions as survival functions and then mapping out the survival functions at interesting values of distance from the CBSA center accomplishes this (see Lawless 2003; Cleves et al 2005). As their names imply, hazard functions and survival functions are just opposite ways of expressing the same thing: mathematically, $H(d_{ij}) = \Pr(D < d_{ij}) \Leftrightarrow S(d_{ij}) = 1 - H(d_{ij}) = \Pr(D \geq d_{ij})$. So, whereas the hazard function, $H(d)$, expresses the conditional probability of the distance between nearest neighbor block groups terminating, the survival function, $S(d)$, expresses the conditional probability of the distance between nearest neighbor block groups extending. Although the models were estimated as hazard functions, in this case, survival functions have the advantage of being more intuitive to interpret graphically.

Thinking in terms of survival functions just turns the analytical framework around. To restate the central premise: (*i*) the conditional probability of distance between nearest neighbors extending decreases with distance; and (*ii*) the probability of extending accelerates with distance from the regional center of gravity. These hypotheses are evaluated by drawing out survival curves from the parameter estimates shown in Table 3 — that is, $\hat{\lambda}$, $\hat{\phi}_{d_c}$, and $\hat{\Phi}_k$ — while varying $x_{d_{ic}}$, distance from the CBSA center, and holding X_{ik} constant at the mean, $\overline{X}_{ik}$. In order to do this in a way that enables interregional comparison, radial distances, say $\xi_{d_{ic}}$, capturing 5%, 15%, 25%, 35%, 45%, 55%, 65%, 75%, 85%, and 95% of each region's total number of housing units were calculated and these region-specific values were used as values of $x_{d_{ic}}$. The relevant distances for each region are listed in Table 4 and they were applied by plugging relevant values into equation (9):

$$h(d_{ij} \mid X_{ik}) = \hat{\lambda} \cdot \exp(\hat{\phi}_{d_c} \cdot \xi_{d_{ic}} + \overline{X}_{ik} \cdot \hat{\Phi}_k) \,. \qquad (10)$$

All notation is the same as before, except that the hats denote estimated shape and scale parameters, the bar denotes mean values of the vector X, and $\xi_{d_{ic}} \in [d_{ic} \approx 5\%, ..., 95\%]$, where the percentages refer to the distance from the CBSA's housing unit weighted center to capture approximately that proportion of the regions housing units.

The resulting survival curves, which were drawn using the *stcurve* command in *Stata*, are shown region-by-region in alphabetical order in Figure 3. These curves describe the conditional probability of the distance between nearest neighbor block groups extending past a particular distance at relevant locations within the regions. In the graphs, the *x*-axis, which registers distance between nearest neighbors, ranges from zero to 2,500 meters, and the *y*-axis, which registers the conditional probability that d_{ij} extends, ranges from zero to one. Going from left to right, the 10 curves drawn in each of the graphs correspond to the distance from the CBSA center, $\xi_{d_{ic}}$, that captures $\approx 5\%, ..., 95\%$ of the region's housing units; in this sense, the graphs are consistent and directly comparable to one another.

As a set, the survival functions are compelling. Foremost, unlike density gradients, they to do not describe urban form by way of certain values — rather, they express probabilities, or, to be exact, the conditional probability that nearest neighbors are separated by a particular distance. With this in mind, a review of the graphs quickly reveals that the various survival functions line up extremely well with the character of the regions they represent. In particular, in regions known for their high-density, compact patterns of urbanization, like Chicago, IL and New York, NY, the estimated survival functions are steeply sloped and tightly bunched together — the probability of distance between nearest neighbors extending very far is small and the basic pattern is for the most part consistent across the entire region. Meanwhile, in regions known for their low-density, sprawling patterns of urbanization, like Atlanta, GA and Phoenix, AZ, the survival functions are more flatly sloped, especially at their tops, and spread out. In regions with high-density, compact core areas and low-density, sprawling outlying areas, like Baltimore, MD and Pittsburgh, PA, the innermost survival curves are steeply sloped and tightly bunched and the outermost survival curves flatly sloped and dispersed. Last, some regions, like Ft. Lauderdale, FL and Oakland, CA, exhibit little or no internal variation in their patterns of urbanization. Whatever the particular case, the graphs in Figure 3 expose the manner in which urban form changes going outward from the regional center of gravity. The urban form of every region ultimately comes undone, but, as the figure demonstrates, they reach that end point by way of highly disparate transitions in land use — transitions that are reflected in this inherently stochastic framework.

3.3 Detecting Urban Growth Boundaries

Having established that spatial hazard models can characterize very different patterns of regional development, the remaining step of the analysis is to determine weather or not they may be used to evaluate public policies aimed at shaping the outcome of urbanization. The regions selected for this exercise are Seattle, WA and Portland-Vancouver, OR-WA — both are surrounded by urban growth boundaries (UGBs) aimed at limiting sprawl by promoting high-density, compact land use. In the simplest of terms, UGBs are supply constraints that shape urban form by raising the price of land, a factor of production in the housing market, in order to reduce household land consumption (see Knaap 1985; Knaap and Nelson 1992; Ding et al 1999; Knaap and Hopkins 2001). Going back to Figure 1, higher land prices mean that, all else being equal, the budget constraint (the dashed line) is steeper, so it intersects the x-axis closer to the origin, where the optimal lot size, ς, is smaller. For the purpose of estimating this impact, the implication is that the housing unit weighted centers of block groups are expected to be closer together inside of the Seattle and Portland UGBs than outside. If so, the spatial hazard model shown in (9) will register the UGBs as factors that accelerate the baseline hazard function for distance between nearest neighbors.

In order to explore this proposition, shape-files of the Seattle and Portland-Vancouver UGBs were obtained from the the Puget Sound Regional Council and Metro, respectively. After obtaining the data, the GIS was used to code those block group points located inside of the UGBs with a one and those located outside of them with a zero. This dummy variable was then added to the vector X in equation (9) and the models for the two regions were re-estimated. The estimates are shown in Table 5 and survival curves, drawn at the same 10 distance bands as before, but this time twice for each region, with the UGB variables first set to one (inside the UGB) and then to zero (outside the UGB), are shown in Figure 4. As anticipated, both models register the presence of the UGBs: the variable is positive and highly significant in each case, meaning that it is associated with shorter distances between nearest neighbors. Further, Figure 4 shows that, in both regions, the survival curves are steeper and more tightly bunched together inside of the UGBs than outside. The curves corresponding to locations inside of the UGBs reflect comparatively compact, high-density development and the curves corresponding to locations outside the UGBs reflect comparatively sprawling, low-density development. These findings are clear-cut because they are exactly what is expected if the UGBs are meaningful policy instruments. Though it is important to be upfront about the fact that the positive parameters do not mean that the UGBs caused the difference — and they certainly did not, because they were established long after most of the area within them was built out — they clearly do illustrate that the pattern of urbanization is different on either side. Going forward, an important, and yet relatively unexplored, policy question spatial hazard models may help to address is: do the UGBs help to maintain this difference over time?

4. Summary and Conclusion

The introduction to this paper set out three specific objectives: (*i*) to review pertinent theoretical and empirical research on urban form and explain how the spatial hazard framework compares to traditional approaches; (*ii*) to estimate a series of spatial hazard models characterizing the built environment in the 25 largest core based statistical areas of the United States and consider the results; and (*iii*) to illustrate how spatial hazard models may be used to evaluate smart growth, or growth management, policies aimed at shaping the outcome of urbanization. Having accomplished these objectives, the remaining comments focus on some implications and directions for future research.

Foremost, over the course of this paper, spatial hazard models have emerged a highly flexible approach that holds great promise for studying urban form. As discussed throughout, urbanization is inherently stochastic — in reality, it rarely, if ever unfolds in the kind of smooth, monotonic way described by negative exponential density gradients. There are many reasons for this, including market imperfections, land use regulation, topography, and more, but chief among them is the fact that the built environment is highly durable. Development waxes and wanes through time, but it is never completely demolished and replaced all at once, no matter how rapidly a region grows. None of this is to say that spatial hazard models render the density gradient approach obsolete, just that it is important to continually explore complementary ways of analyzing patterns of urbanization. Indeed, just as Newton's laws have yielded centuries of insight in the physical sciences and continue to do so, even as special and general relativity, quantum mechanics, and, most recently, strings have challenged them, density gradients and the theory that gives rise to them remain fundamental for characterizing and understanding urban form, which has always been difficult to describe in concrete terms (Talen 2003; Song and Knaap 2007). But, to extend the physics analogy, the spatial hazard framework offers a fresh way of looking at long-familiar forces: like quantum theory, it rests on a probabilistic worldview — one that never attempts to deliver an exact measure of (in this case) urbanization. This strikes the present authors as exciting and, more important, very promising from both theoretical and empirical standpoints.

With respect to applied policy analysis, it is worth underscoring the fact that, in practice, proportional hazard models are regularly used to examine consequential problems pertaining to lifecycles. For example, economists are interested in the duration of unemployment spells; epidemiologists are interested in the survival of people having fatal diseases; and engineers are interested in the longevity of critical machine components. In all of these cases, the main concern is often on how some independent factor/s — say, employment counseling, a new medication, a particular material — adds to or takes away from the lifecycle under study. Such experiments are conceptually similar to the kind of policy experiments that are of interest to planners and others responsible for managing urban and regional development: for example, do UGBs matter to urban form? The evidence developed here suggests that they do matter — though, again, it is important to be clear that there is nothing to suggest that they are in any way a causal factor, only that they are apparently meaningful demarcations between different patterns of urbanization. In order to know that UGBs and other smart growth, or growth management, policies are sound, analysts require objective and rigorous ways of evaluating them on an ongoing basis, and, epistemologically, the proportional hazard framework seems ideal for the purpose. If the approach, when done right, is good enough to weigh on matters of life and death, then it is certainly good enough to weigh on matters of land use policy, no matter how complex and/or controversial they may be.

Finally, the spatial hazard framework is still very new. Odland and Ellis (1992) put the idea less forth than 20 years ago and Waldorf (2003) only recently developed further; in between, it has been applied to various spatial questions, but only a few (Rogerson et al 1993; Esparza and Krmenec 1994, 1996; Pellegrini and Reader 1996; Reader 2000). Further research is certainly merited and it should focus on strengthening the connection of the approach to specific spatial processes by developing it within the context of established theoretical frameworks, like, as done here, economic models of urbanization. That said, effort should also be put into adapting some of those same theoretical frameworks to line up with the spatial hazard models' probabilistic worldview and, especially, into making the benefits of that perspective explicit. In terms of urbanization, future work could go in several directions: (*i*) looking at the spacing of actual structures or some other smaller units of analysis; (*ii*) tracking land use change through time; (*iii*) comparing urban, suburban, exurban, and rural patterns of development; (*iv*) determining whether estimated shape and/or scale parameters are themselves useful as land use metrics; and (*v*) examining the utility of spatial hazard models for evaluating additional forms of land use policy. Each of these directions and more would be a worthwhile extension of the research presented here.

References

Alonso W (1964) *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press: Cambridge, MA.

Anas A, Arnott R, Small KA (1998) Urban Spatial Structure. *Journal of Economic Literature*, 36: 1426 – 1464.

Boots BN, Getis A (1988) *Point Pattern Analysis*. Sage: Newbury Park, CA.

Brueckner JK (2000) Urban Growth Models with Durable Housing: An Overview. In Hurion JM, Thisse JF (eds.) *Economics of Cities: Theoretical Perspectives*. The Cambridge University Press, Cambridge, UK.

Brueckner JK (1987) The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model. In Mills ES (ed.) *Handbook of Regional and Urban Economics*, Vol. II, pages 821 – 845. North-Holland: Amsterdam.

Brueckner JK (1982) A Note on the Sufficient Conditions for Negative Exponential Population Densities. *Journal of Regional Science*, 22: 353 – 359.

Clark C (1951) Urban Population Densities. *Journal of the Royal statistical Society*, 114: 490 – 494.

Cleves MA, Gould WW, Guitierrez RG (2005) *An Introduction to Survival Analysis Using Stata*. Stata Press: College Station, TX.

Diggle PJ (2003) *Statistical Analysis of Spatial Point Patterns*. Arnold: New York, NY

Ding C, Knaap GJ, Hopkins LD (1999) Managing Urban Growth with Urban Growth Boundaries: A Theoretical Analysis. *Journal of Urban Economics*, 46: 53 – 68.

Duranton G, Overman, HG (2008) Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Microgeographic Data. *Journal of Regional Science*, 48: 213 – 243.

Esparza AX, Krmenec A (1996) The Spatial Extent of Producer Service Markets: Hierarchical Models of Interaction Revisited. *Papers in Regional Science*, 75: 375 – 395.

Esparza AX, Krmenec A (1994) Business Services in the Space Economy: A Model of Spatial Interaction. *Papers in Regional Science*, 73: 55 – 72.

Fujita M (1987) *Urban Economic Theory: Land Use and City Size*. The Cambridge University Press, Cambridge, UK.

Getis A (1983) Second-order Analysis of Point patterns: The Case of Chicago as a Multi-center Urban Region. *Professional Geographer*, 35: 73 – 80.

Getis A (1964) Temporal Land Use Pattern Analysis with the Use of Nearest Neighbor and Quadrat Methods. *Annals of the American Association of Geographers*, 54: 391 – 399.

Glaeser EL, Gyourko J (2005) Urban Decline and Durable Housing. *Journal of Political Economy*, 113: 345 – 375.

Isard, W (1956) *Location and Space Economy*. The MIT Press: Cambridge, MA.

Johnson SR, Kau JB (1980) Urban Spatial Structure: An Analysis with a Varying Coefficient Model. *Journal of Urban Economics*, 7: 141 – 154.

Kau JB, Lee CF (1977) A Random Coefficient Model To Estimate a Stochastic Density Gradient. *Regional Science and Urban Economics*, 7: 169 – 177.

Kau JB, Lee CF (1976a) The Functional Form in Estimating the Density Gradient: An Empirical Investigation. *Journal of the American Statistical Association*, 71: 326 – 327.

Kau JB, Lee CF (1976b) Functional Form, Density Gradient, and Price Elasticity of Demand for Housing. *Urban Studies*, 13: 193 – 200.

Kau JB, Lee CF, Chen RC (1983) Structural Shifts in Urban Population Density Gradients: An Empirical Investigation. *Journal of Urban Economics*, 13: 364 – 377.

Kiefer NM (1988) Economic Duration Data and Hazard Functions. *Journal of Economic Literature*, 26: 646 – 679.

Knaap GJ (1985) The Price Effects of Urban Growth Boundaries in Metropolitan Portland, Oregon. *Land Economics*, 61: 26 – 35.

Knaap GJ Hopkins L (2001) An Inventory Approach to Urban Growth Boundaries. *Journal of the American Planning Association*, 67: 314 – 326.

Knaap GJ, Nelson AC (1992) *The Regulated Landscape: Lessons on State Land Use Planning From Oregon*. Lincoln Institute of Land Policy: Cambridge, MA.

Lawless JF (2003) *Statistical Models and Methods for Lifetime Data*. Wiley-Interscience, Hoboken, NJ.

Lessinger J (1961) The Case for Scatteration: Some Reflections on the National Capitol Region. *Journal of the American Institute of Planners*, 27: 159 – 169.

Lynch K (1983) *Good City Form*. The MIT Press: Cambridge, MA.

McDonald JF (1988) Econometric Studies of Urban Population Density: A Survey. *Journal of Urban Economics*, 26: 361 – 385.

Mills ES (1971) *Studies in the Structure of the Urban Economy*. Johns Hopkins University Press: Baltimore, MD.

Muth RF (1969) *Cities and Housing*. University of Chicago Press: Chicago, IL.

Odland J, Ellis, M (1992) Variations in the Spatial Pattern of Settlement Locations: An Analysis Based on Proportional Hazards Models. *Geographical Analysis*, 24: 97 – 109.

Pasquale PA, Reader S (1996) Duration Modeling of Spatial Point Patterns. *Geographical Analysis*, 28: 217 – 243.

Reader S (2000) Using Survival Analysis to Study Spatial Point patterns in Geographical Epidemiology. *Social Science and Medicine*, 50: 985 – 1000.

Renner RN, Lewis S, Carruthers JI, Knaap GJ (2008) A Note on Data Preparation Procedures for a Nationwide Analysis of Urban Form and Settlement Patterns. *Cityscape*, forthcoming.

Rogerson P, Weng R, Lin G (1993) The Spatial Separation of Parents and Their Adult Children. *Annals of the American Association of Geographers*, 83: 656 – 671.

Song Y, Knaap GJ (2007) Quantitative Classification of Neighborhoods: The Neighborhoods of New Single-family Homes in the Portland Metropolitan Area. *Journal of Urban Design*, 12: 1 – 24.

Talen E (2003) Measuring Urbanism: Issues in Smart Growth Research. *Journal of Urban Design*, 8: 195 – 215.

von Thünen JH (1826) *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalekonomie*. Hamburg, Germany.

Waldorf BS (2003) Spatial Point patterns in a Longitudinal Framework. *International Regional Science Review*, 26: 269 – 288.

Table 1. Data Definitions and Sources

	Definition	Source
Distance from Nearest Neighbor	Distance from population (housing unit) weighted center to the population weighted center of the nearest block group, 2006	Authors' calculations, U.S. Census Bureau — from 2006 count of housing units
Distance from CBSA	Distance from population (housing unit) weighted center to the population weighted center of the nearest CBSA, 2006	Authors' calculations, from U.S. Census Bureau — 2006 count of housing units
Household Income	Median household income, 1999	U.S. Census Bureau — SF-3, Table P68
Travel Cost	Average duration of journey to work, 2000	Author's calculations, from U.S. Census Bureau — SF-3, Tables P31 and P33
Age of Housing Units	Median age of housing units, 2000	U.S. Census Bureau — SF-3, Table H35
Number of Housing Units	Count of housing units, 2006	U.S. Census Bureau — 2006 count of housing units

Note: All data is at the level of census block groups.

Table 2. Descriptive Statistics by CBSA

	n	Dist. from Nearest Neighbor		Dist. from CBSA		Household Income		Travel Cost		Age of Housing Units		Number of Housing Units	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1. New York, NY-NJ-PA													
Edison	1,206	793	562	23,347	8,606	61,818	26,707	31.88	8.21	41.56	113.90	538.07	367.44
Nassau-Suffolk	1,624	671	498	19,375	10,243	71,735	22,121	31.91	5.69	45.36	97.63	437.13	232.85
Newark-Union	2,000	687	568	20,147	7,999	66,686	32,815	28.83	6.59	52.48	131.57	475.12	299.69
New York-Wayne-White Plains	8,448	285	198	14,998	7,804	49,086	29,673	36.28	10.46	64.19	165.47	520.37	426.71
2. Los Angeles, CA													
Los Angeles-Long Beach	5,695	472	416	19,795	12,130	47,500	26,619	27.86	6.62	52.23	144.49	513.33	350.58
Santa Ana-Anaheim-Irvine	2,652	548	408	20,897	11,371	61,282	27,034	27.00	5.63	36.50	115.21	535.75	273.14
3. Chicago, IL-IN-WI													
Chicago-Naperville-Joliet	4,780	492	391	19,939	8,922	51,316	25,612	31.43	9.35	59.99	171.97	516.55	435.03
Gary	786	878	952	19,597	8,824	44,066	16,540	28.86	7.26	44.46	99.67	530.72	321.79
Lake-Kenosha	652	1,026	659	20,847	9,778	73,776	33,056	29.03	6.18	29.65	13.33	622.90	409.30
4. Philadelphia, PA-MD-NJ-DE													
Camden	913	750	526	13,931	6,299	52,289	21,506	26.70	6.13	47.98	130.49	510.24	299.89
Philadelphia	2,824	454	413	13,673	6,291	43,153	25,953	29.85	9.73	61.39	151.46	435.45	283.58
Wilmington	608	1,174	1,093	17,808	10,879	54,586	22,462	23.78	4.93	39.36	81.18	576.25	327.40
5. Dallas-Ft. Worth, TX													
Dallas-Plano-Irving	2,086	782	742	17,540	9,634	54,319	31,343	26.28	6.57	36.82	130.03	565.44	438.55
Ft. Worth-Arlington	1,187	977	957	16,023	9,242	48,202	23,639	25.51	5.57	30.71	59.09	559.63	371.58
6. Miami- Ft. Lauderdale, FL													
Ft. Lauderdale	773	752	331	12,975	6,339	46,156	21,020	25.62	7.68	26.42	10.82	1,064.42	851.45
Miami-Miami Beach	1,207	673	1,239	13,688	10,237	40,745	27,685	28.43	7.26	48.07	169.67	692.00	640.16
West Palm Beach-Boca Raton	511	932	586	15,495	9,151	49,274	26,469	23.70	5.89	27.20	88.18	947.89	670.44
7. Washington, DC-VA-MD-WV													
Bethesda-Frederick-Gaithersburg	758	1,033	968	18,270	10,018	77,399	30,961	30.86	4.63	27.73	13.36	650.76	405.94
Washington-Arlington	1,925	748	658	17,756	9,248	69,350	33,234	31.12	6.54	39.74	101.33	643.10	402.13
8. Houston, TX	2,469	987	926	25,184	17,275	47,880	27,435	27.19	6.40	34.94	112.84	668.30	449.39

Table 2. Descriptive Statistics by CBSA (cont…)

		Dist. from Nearest Neighbor		Dist. from CBSA		Household Income		Travel Cost		Age of Housing Units		Number of Housing Units	
	n	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
9. Detroit-Warren-Livonia, MI													
Detroit	2,222	487	226	13,292	6,187	42,005	20,920	25.55	6.75	56.69	137.54	398.13	234.07
Warren-Farmington Hills	1,416	1,020	949	22,895	16,745	60,945	26,087	25.44	5.42	37.43	91.44	537.78	311.36
10. Boston-Cambridge, MA-NH													
Boston-Quincy	1,223	718	594	14,942	7,756	54,672	23,018	29.43	5.96	49.48	57.58	490.67	260.63
Cambridge	1,218	632	484	14,709	5,517	64,834	30,380	26.27	4.68	50.97	57.40	505.16	240.58
Essex County	562	821	623	16,270	5,033	55,914	23,875	25.84	4.65	47.35	14.27	529.81	236.81
Rockingham-Strafford	238	1,923	1,236	22,322	7,858	52,812	14,162	24.76	5.52	42.43	128.29	646.59	323.70
11. Atlanta, GA	1,543	1,249	810	21,782	11,781	55,256	27,554	29.86	6.93	28.62	88.06	852.66	614.03
12. San Francisco-Oakland, CA													
Oakland	1,267	552	343	17,239	6,767	63,350	29,767	29.81	5.36	42.80	79.21	584.15	399.25
San Francisco	1,195	442	385	13,045	9,547	66,391	29,030	28.34	5.63	50.62	80.93	594.08	360.35
13. Riverside-San Bernardino, CA	1,552	1,255	1,801	36,246	23,006	42,302	18,312	28.93	8.43	34.74	132.88	644.19	524.52
14. Phoenix-Mesa-Scottsdale, AZ	2,186	915	1,579	20,554	14,636	47,405	23,646	24.48	6.74	38.83	179.17	602.15	480.48
15. Seattle-Tacoma-Bellevue, WA													
Seattle	1,686	691	728	16,412	8,863	60,344	22,994	25.37	5.22	35.45	69.51	464.61	236.04
Tacoma	776	987	1,383	15,225	8,896	49,020	16,138	27.61	5.87	28.55	14.70	515.06	253.51
16. Minneapolis-St. Paul, MN-WI	2,091	1,003	1,115	18,776	14,295	56,855	22,938	22.31	4.59	37.57	63.24	522.40	305.17
17. San Diego, CA	1,755	731	976	22,187	11,027	51,817	24,611	24.35	5.15	33.79	94.73	591.89	390.19
18. St. Louis, MO-IL	1,861	1,061	1,260	20,452	15,459	43,016	22,285	24.61	6.50	49.90	136.87	565.17	348.61
19. Baltimore, MD	1,723	758	707	14,261	11,103	48,171	24,465	28.84	7.66	52.34	156.98	548.06	340.97
20. Pittsburgh, PA	1,570	905	766	17,476	11,606	40,215	18,786	24.79	5.08	48.67	51.03	533.68	236.96
21. Tampa-St. Petersburg, FL	1,411	893	666	22,034	9,917	40,707	18,592	24.17	6.06	38.30	139.17	735.05	636.03
22. Denver, CO	1,551	828	1,069	14,146	8,624	56,609	27,636	25.16	5.33	37.91	123.38	534.55	283.93
23. Cleveland, OH	1,556	676	568	15,389	10,121	43,708	23,123	23.79	5.53	55.40	131.46	517.44	350.26
24. Cincinnati, OH-KY-IN	1,332	1,063	1,088	16,726	10,781	46,273	21,597	23.26	5.41	41.89	77.81	547.72	331.51
25. Portland-Vancouver, OR-WA	1,175	1,031	1,072	15,113	9,485	49,374	17,561	23.20	4.43	32.96	17.18	634.68	399.45

Table 3. Estimated Spatial Hazard Functions — Distance from Nearest Neighbor

			ϕ										
	λ		Distance from CBSA		Household Income		Travel Cost		Age of Housing Units		Number of Housing Units		
	Est.	z	Est.	z	Est.	z	Est.	z	Est.	z	Est.	z	LL
1. New York, NY-NJ-PA													
Edison	1.70 ***	29.11	1.00000 n/s	–0.48	0.99999 ***	–9.62	0.99706 n/s	–0.72	0.99859 ***	–6.89	0.99954 ***	–5.56	-1,082
Nassau-Suffolk	1.97 ***	43.51	0.99997 ***	–12.08	0.99999 ***	–10.58	1.04569 ***	11.74	0.99916 ***	–3.85	0.99829 ***	–17.03	–1,219
Newark-Union	2.03 ***	45.68	0.99993 ***	–26.74	0.99997 ***	–26.51	1.01247 ***	3.46	0.99846 ***	–9.82	0.99878 ***	–13.58	–1,543
New York-Wayne-White Plains	2.52 ***	130.69	0.99991 ***	–52.14	0.99998 ***	–35.95	1.02121 ***	27.64	0.99849 ***	–23.47	0.99969 ***	–12.52	–4,584
2. Los Angeles, CA													
Los Angeles-Long Beach	2.36 ***	100.14	0.99993 ***	–50.95	0.99999 ***	–23.15	0.98234 ***	–8.99	0.99898 ***	–10.60	0.99968 ***	–8.33	–3,245
Santa Ana-Anaheim-Irvine	2.04 ***	57.91	0.99996 ***	–26.38	0.99999 ***	–14.73	1.02514 ***	8.29	0.99983 n/s	–0.97	0.99957 ***	–6.71	–1,864
3. Chicago, IL-IN-WI													
Chicago-Naperville-Joliet	2.24 ***	81.98	0.99995 ***	–35.35	0.99998 ***	–29.91	1.02002 ***	18.15	0.99960 ***	–5.00	0.99924 ***	–17.56	–3,201
Gary	1.54 ***	17.62	0.99996 ***	–10.91	0.99997 ***	–10.78	1.01478 ***	2.59	0.99997 n/s	–0.10	0.99912 ***	–7.08	–808
Lake-Kenosha	1.88 ***	24.01	0.99999 **	–2.32	1.00000 ***	–3.39	0.97083 ***	–3.87	1.03070 ***	8.39	0.99996 n/s	–0.39	–542
4. Philadelphia, PA-MD-NJ-DE													
Camden	2.05 ***	31.27	0.99992 ***	–13.63	0.99997 ***	–15.27	1.01183 ***	2.61	1.00006 n/s	0.29	0.99905 ***	–7.93	–694
Philadelphia	2.41 ***	68.13	0.99992 ***	–27.74	0.99996 ***	–36.64	1.01182 ***	7.88	0.99854 ***	–12.46	0.99847 ***	–19.47	–1,697
Wilmington	1.63 ***	16.59	0.99994 ***	–11.59	0.99998 ***	–8.07	0.96836 ***	–3.67	1.00086 ***	2.35	0.99952 ***	–3.59	–618
5. Dallas-Ft. Worth, TX													
Dallas-Plano-Irving	2.26 ***	55.86	0.99990 ***	–33.48	0.99999 ***	–12.98	0.98727 ***	–3.53	0.99935 ***	–3.85	0.99949 ***	–9.19	–1,344
Ft. Worth-Arlington	2.09 ***	37.28	0.99991 ***	–20.52	1.00000 **	–2.09	0.96436 ***	–5.75	1.00010 n/s	0.26	0.99949 ***	–5.77	–871
6. Miami- Ft. Lauderdale, FL													
Ft. Lauderdale	3.16 ***	46.04	0.99999 n/s	–0.82	0.99998 ***	–10.20	1.00092 n/s	0.13	1.03310 ***	7.29	0.99961 ***	–7.53	–266
Miami-Miami Beach	2.08 ***	36.07	0.99990 ***	–22.94	0.99999 ***	–6.02	1.04039 ***	10.35	1.00002 n/s	0.11	0.99976 ***	–5.26	–888
West Palm Beach-Boca Raton	2.61 ***	30.67	0.99996 ***	–8.02	0.99998 ***	–10.09	0.91906 ***	–10.07	0.99887 ***	–2.93	0.99956 ***	–6.18	–273
7. Washington, DC-VA-MD-WV													
Bethesda-Frederick-Gaithersburg	1.71 ***	21.33	0.99994 ***	–15.60	0.99998 ***	–10.21	0.98365 *	–1.85	1.03082 ***	7.83	1.00028 ***	2.88	–715
Washington-Arlington	1.91 ***	41.16	0.99993 ***	–25.09	0.99999 ***	–17.95	0.97489 ***	–6.73	0.99937 ***	–3.49	0.99982 ***	–3.21	–1,597
8. Houston, TX	1.93 ***	45.14	0.99995 ***	–29.14	1.00000 ***	–4.42	0.96199 ***	–12.43	0.99914 ***	–5.47	0.99985 ***	–3.24	–2,100

Notes: All models estimated via Weibull regression; LL the is log-likelihood; *** denotes two-tailed hypothesis test significant at $p < 0.01$; ** denotes two-tailed hypothesis test significant at $p < 0.05$; * denotes two-tailed hypothesis test significant at $p < 0.10$; n/s denotes two-tailed hypothesis test not significant.

Table 3. Estimated Spatial Hazard Functions — Distance from Nearest Neighbor (cont…)

	λ		ϕ Distance from CBSA		Household Income		Travel Cost		Age of Housing Units		Number of Housing Units		
	Est.	z	Est.	z	Est.	z	Est.	z	Est.	z	Est.	z	LL
9. Detroit-Warren-Livonia, MI													
Detroit	2.69 ***	70.60	0.99993 ***	–16.32	0.99999 ***	–11.07	1.01255 ***	4.41	0.99958 ***	–2.91	0.99901 ***	–9.98	–1,030
Warren-Farmington Hills	1.95 ***	36.01	0.99996 ***	–18.07	0.99999 ***	–10.75	0.92267 ***	–15.79	0.99854 ***	–6.34	0.99950 ***	–5.71	–1,153
10. Boston-Cambridge, MA-NH													
Boston-Quincy	1.88 ***	30.06	0.99994 ***	–17.03	0.99997 ***	–18.46	0.98649 ***	–2.48	0.99885 ***	–4.45	0.99921 ***	–7.10	–1,094
Cambridge	1.95 ***	32.74	0.99994 ***	–11.64	0.99997 ***	–21.85	1.01975 ***	3.27	0.99971 n/s	–1.31	0.99909 ***	–7.49	–1,013
Essex County	2.11 ***	26.15	0.99997 ***	–3.75	0.99997 ***	–11.48	1.01215 n/s	1.14	1.04023 ***	10.23	0.99882 ***	–6.05	–403
Rockingham-Strafford	1.96 ***	13.57	0.99999 n/s	–0.97	0.99997 ***	–4.43	0.91369 ***	–5.71	0.99941 n/s	–1.29	0.99968 n/s	–1.25	–208
11. Atlanta, GA	2.80 ***	56.93	0.99991 ***	–32.28	0.99999 ***	–8.00	0.99962 n/s	–0.09	1.00005 n/s	0.18	0.99960 ***	–8.62	–736
12. San Francisco-Oakland, CA													
Oakland	2.38 ***	46.39	0.99998 ***	–6.01	0.99998 ***	–20.42	0.97565 ***	–4.89	0.99850 ***	–5.86	0.99934 ***	–9.66	–751
San Francisco	1.91 ***	35.47	0.99993 ***	–19.58	0.99999 ***	–9.60	0.99921 n/s	–0.16	0.99875 ***	–5.14	0.99965 ***	–4.62	–925
13. Riverside-San Bernardino, CA	1.45 ***	21.97	0.99997 ***	–23.30	0.99999 ***	–4.49	0.97627 ***	–8.70	0.99887 ***	–5.80	0.99989 **	–2.34	–1,676
14. Phoenix-Mesa-Scottsdale, AZ	1.83 ***	43.25	0.99993 ***	–34.07	0.99999 ***	–8.38	1.01096 ***	3.41	0.99930 ***	–5.33	0.99971 ***	–6.82	–1,836
15. Seattle-Tacoma-Bellevue, WA													
Seattle	2.27 ***	47.89	0.99991 ***	–24.74	0.99998 ***	–17.74	0.96176 ***	–7.09	0.99907 ***	–3.73	0.99932 ***	–6.01	–1,128
Tacoma	1.80 ***	22.63	0.99992 ***	–18.34	1.00000 n/s	–0.04	0.99222 *	–1.72	1.01219 ***	4.06	0.99970 **	–2.24	–709
16. Minneapolis-St. Paul, MN-WI	2.50 ***	59.55	0.99990 ***	–35.02	0.99998 ***	–14.69	0.98192 ***	–3.28	0.99904 ***	–4.20	0.99944 ***	–7.32	–1,215
17. San Diego, CA	1.59 ***	30.92	0.99995 ***	–20.43	0.99998 ***	–15.53	0.97315 ***	–5.87	0.99922 ***	–3.62	0.99957 ***	–6.80	–1,694
18. St. Louis, MO-IL	2.04 ***	43.92	0.99992 ***	–33.55	0.99998 ***	–15.95	0.97856 ***	–4.94	0.99964 **	–2.18	0.99965 ***	–5.21	–1,451
19. Baltimore, MD	2.24 ***	48.51	0.99989 ***	–29.29	0.99998 ***	–14.95	1.01597 ***	4.67	0.99949 ***	–3.18	0.99988 *	–1.75	–1,178
20. Pittsburgh, PA	2.06 ***	39.04	0.99991 ***	–28.39	0.99998 ***	–11.87	0.96583 ***	–6.01	1.00004 n/s	0.15	0.99911 ***	–8.03	–1,251
21. Tampa-St. Petersburg, FL	1.90 ***	36.95	0.99995 ***	–17.16	0.99998 ***	–9.31	0.96294 ***	–9.40	0.99903 ***	–5.69	0.99968 ***	–7.43	–1,134
22. Denver, CO	2.36 ***	48.54	0.99987 ***	–30.40	0.99999 ***	–7.48	0.97900 ***	–3.87	0.99851 ***	–6.05	1.00001 n/s	0.11	–970
23. Cleveland, OH	2.18 ***	43.70	0.99993 ***	–21.72	0.99998 ***	–14.89	0.99246 n/s	–1.58	0.99934 ***	–3.63	0.99902 ***	–12.00	–1,109
24. Cincinnati, OH-KY-IN	2.27 ***	40.64	0.99989 ***	–26.19	0.99998 ***	–12.81	0.95693 ***	–6.74	1.00015 n/s	0.54	0.99938 ***	–7.43	–945
25. Portland-Vancouver, OR-WA	2.41 ***	42.93	0.99987 ***	–24.45	0.99998 ***	–12.94	0.96111 ***	–5.07	1.00011 n/s	0.04	0.99962 ***	–4.52	–720

Notes: All models estimated via Weibull regression; LL the is log-likelihood; *** denotes two-tailed hypothesis test significant at $p < 0.01$; ** denotes two-tailed hypothesis test significant at $p < 0.05$; * denotes two-tailed hypothesis test significant at $p < 0.10$; n/s denotes two-tailed hypothesis test not significant.

Table 4. Radial Distance from Center of CBSA — Percent of Population

	Total Housing Units	Meters to Capture Percent of Housing Units ≈ 5%	≈ 15%	≈ 25%	≈ 35%	≈ 45%	≈ 55%	≈ 65%	≈ 75%	≈ 85%	≈ 95%
1. New York, NY-NJ-PA											
Edison	648,912	7,442	15,744	18,678	21,217	23,963	25,541	27,056	29,505	33,017	40,222
Nassau-Suffolk	709,902	5,418	9,135	11,973	14,579	17,073	19,848	22,891	26,004	28,570	39,918
Newark-Union	950,239	7,730	13,045	15,466	16,996	18,520	20,154	22,375	24,912	28,515	36,334
New York-Wayne-White Plains	4,396,063	2,516	5,320	7,564	9,611	11,690	14,058	16,566	18,940	22,287	28,461
2. Los Angeles, CA											
Los Angeles-Long Beach	2,923,432	4,404	8,585	11,504	14,407	17,211	19,835	22,651	26,241	31,083	42,975
Santa Ana-Anaheim-Irvine	1,420,802	4,795	8,863	11,945	14,896	18,292	21,528	24,961	28,665	33,192	42,032
3. Chicago, IL-IN-WI											
Chicago-Naperville-Joliet	2,469,113	6,996	11,885	14,971	17,339	19,574	20,908	22,924	25,635	28,851	39,482
Gary	417,145	5,041	9,209	12,068	15,134	18,182	20,885	23,988	26,428	29,091	32,451
Lake-Kenosha	406,130	5,198	9,294	11,610	15,913	20,667	23,331	25,699	27,701	31,054	37,816
4. Philadelphia, PA-MD-NJ-DE											
Camden	465,845	4,205	6,845	8,740	10,444	12,133	14,240	16,033	18,418	20,794	24,005
Philadelphia	1,229,707	5,510	8,434	10,194	11,580	12,938	14,247	15,747	18,658	22,826	29,226
Wilmington	350,360	4,207	6,475	8,674	10,757	13,565	17,663	21,182	25,632	30,607	37,218
5. Dallas-Ft. Worth, TX											
Dallas-Plano-Irving	1,179,507	4,216	7,916	10,267	13,516	16,124	18,357	20,625	23,824	28,614	37,365
Ft. Worth-Arlington	664,275	5,099	8,796	11,696	13,740	15,479	17,143	19,466	21,981	25,713	37,384
6. Miami- Ft. Lauderdale, FL											
Ft. Lauderdale	822,795	3,689	6,704	8,561	10,711	12,449	14,390	16,707	18,489	20,291	23,482
Miami-Miami Beach	835,246	4,092	6,450	8,188	10,074	12,238	14,619	16,422	18,054	20,875	30,702
West Palm Beach-Boca Raton	484,372	3,338	6,048	8,328	10,832	13,198	15,468	17,930	22,422	24,658	34,233
7. Washington, DC-VA-MD-WV											
Bethesda-Frederick-Gaithersburg	493,274	4,035	7,695	10,086	12,752	16,368	19,633	22,192	25,462	29,523	37,080
Washington-Arlington	1,237,958	4,178	7,907	10,798	13,365	15,065	17,492	20,533	23,389	27,288	33,860
8. Houston, TX	1,650,038	5,411	9,853	13,522	16,817	20,383	24,082	28,226	33,324	40,542	56,795

Table 4. Radial Distance from Center of CBSA — Percent of Population (cont…)

	Total Housing Units	Meters to Capture Percent of Housing Units ≈ 5%	≈ 15%	≈ 25%	≈ 35%	≈ 45%	≈ 55%	≈ 65%	≈ 75%	≈ 85%	≈ 95%
9. Detroit-Warren-Livonia, MI											
Detroit	884,653	3,939	7,125	8,977	10,561	12,086	14,010	15,918	18,224	21,156	24,290
Warren-Farmington Hills	761,492	6,164	9,684	12,366	14,786	17,115	19,497	22,348	25,244	30,704	64,765
10. Boston-Cambridge, MA-NH											
Boston-Quincy	600,085	5,121	6,770	7,949	11,751	13,234	14,640	15,855	18,280	23,868	32,577
Cambridge	615,290	6,532	9,226	11,196	12,903	14,075	15,025	16,349	18,420	21,354	24,471
Essex County	297,754	7,097	10,613	12,526	14,272	16,004	17,138	17,925	19,063	21,400	25,595
Rockingham-Strafford	153,889	8,206	12,865	17,252	20,154	21,840	22,819	23,823	26,899	30,177	36,112
11. Atlanta, GA	1,315,653	4,544	9,793	13,671	16,949	20,464	23,964	27,486	31,563	35,882	45,344
12. San Francisco-Oakland, CA											
Oakland	740,116	8,623	10,695	12,426	13,993	15,271	17,153	19,425	22,381	26,672	31,954
San Francisco	709,928	2,309	4,493	6,190	7,205	8,119	9,220	15,512	22,051	25,661	31,708
13. Riverside-San Bernardino, CA	999,787	10,322	17,799	23,019	27,758	32,001	36,805	46,838	57,422	66,812	88,586
14. Phoenix-Mesa-Scottsdale, AZ	1,316,303	5,638	10,524	13,697	16,259	18,652	20,954	23,873	27,559	32,586	43,578
15. Seattle-Tacoma-Bellevue, WA											
Seattle	783,325	5,688	7,643	9,256	10,802	12,951	15,644	18,158	21,212	25,559	33,026
Tacoma	399,689	3,459	6,778	8,722	10,499	12,193	14,574	18,124	20,959	25,207	30,415
16. Minneapolis-St. Paul, MN-WI	1,092,336	3,799	7,337	10,103	12,850	15,473	19,000	21,717	25,532	31,392	45,484
17. San Diego, CA	1,038,770	7,892	11,588	14,100	15,866	18,312	20,828	24,591	29,246	34,690	43,104
18. St. Louis, MO-IL	1,051,789	4,520	8,053	10,261	12,555	16,261	19,690	24,034	31,278	36,652	55,830
19. Baltimore, MD	944,306	2,197	4,459	7,074	9,259	11,849	15,095	18,578	22,463	29,442	37,126
20. Pittsburgh, PA	837,879	3,276	5,782	8,629	10,677	13,364	16,395	20,702	25,996	32,211	38,418
21. Tampa-St. Petersburg, FL	1,037,157	6,698	12,653	15,824	18,584	21,754	24,739	27,340	30,054	32,996	38,675
22. Denver, CO	829,093	3,431	5,603	7,952	9,853	12,151	14,057	15,784	17,743	20,062	30,487
23. Cleveland, OH	805,137	3,836	6,771	8,699	10,706	12,793	15,390	18,817	22,260	28,242	37,668
24. Cincinnati, OH-KY-IN	729,564	3,780	6,700	9,046	11,671	13,608	16,356	19,577	23,043	27,641	37,097
25. Portland-Vancouver, OR-WA	745,746	3,035	6,063	8,384	10,839	13,200	14,871	16,787	19,752	23,298	33,045

Table 5. Estimated Spatial Hazard Functions for Seattle and Portland w/ UGB Dummy — Distance from Nearest Neighbor

	λ			ϕ Distance from CBSA			Household Income			Travel Cost			Age of Housing Units			Number of Housing Units			Inside UGB			LL
	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	Est.		z	
Seattle	2.36993	***	49.04	0.99992	***	–19.89	0.99998	***	–16.02	0.96905	***	–5.66	0.99907	***	–3.50	0.99924	***	–6.67	2.20734	***	9.73	–1078
Portland-Vancouver	2.48952	***	43.64	0.99990	***	–16.60	0.99998	***	–12.69	0.95354	***	–5.95	0.99997	n/s	–0.01	0.99949	***	–5.86	2.21248	***	7.10	–694

Notes: All models estimated via Weibull regression; LL the is log-likelihood; *** denotes two-tailed hypothesis test significant at $p < 0.01$; ** denotes two-tailed hypothesis test significant at $p < 0.05$; * denotes two-tailed hypothesis test significant at $p < 0.10$; n/s denotes two-tailed hypothesis test not significant.

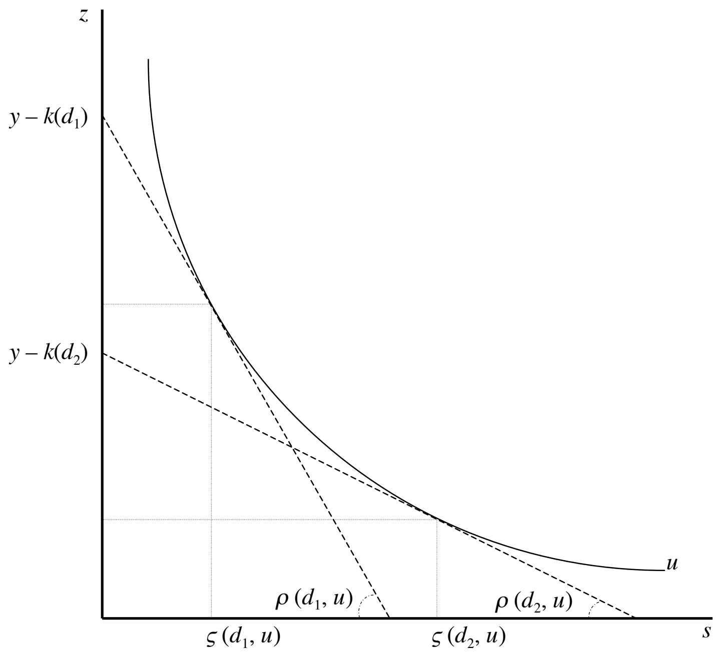


Figure 1. The Bid Rent Process and Urban Form

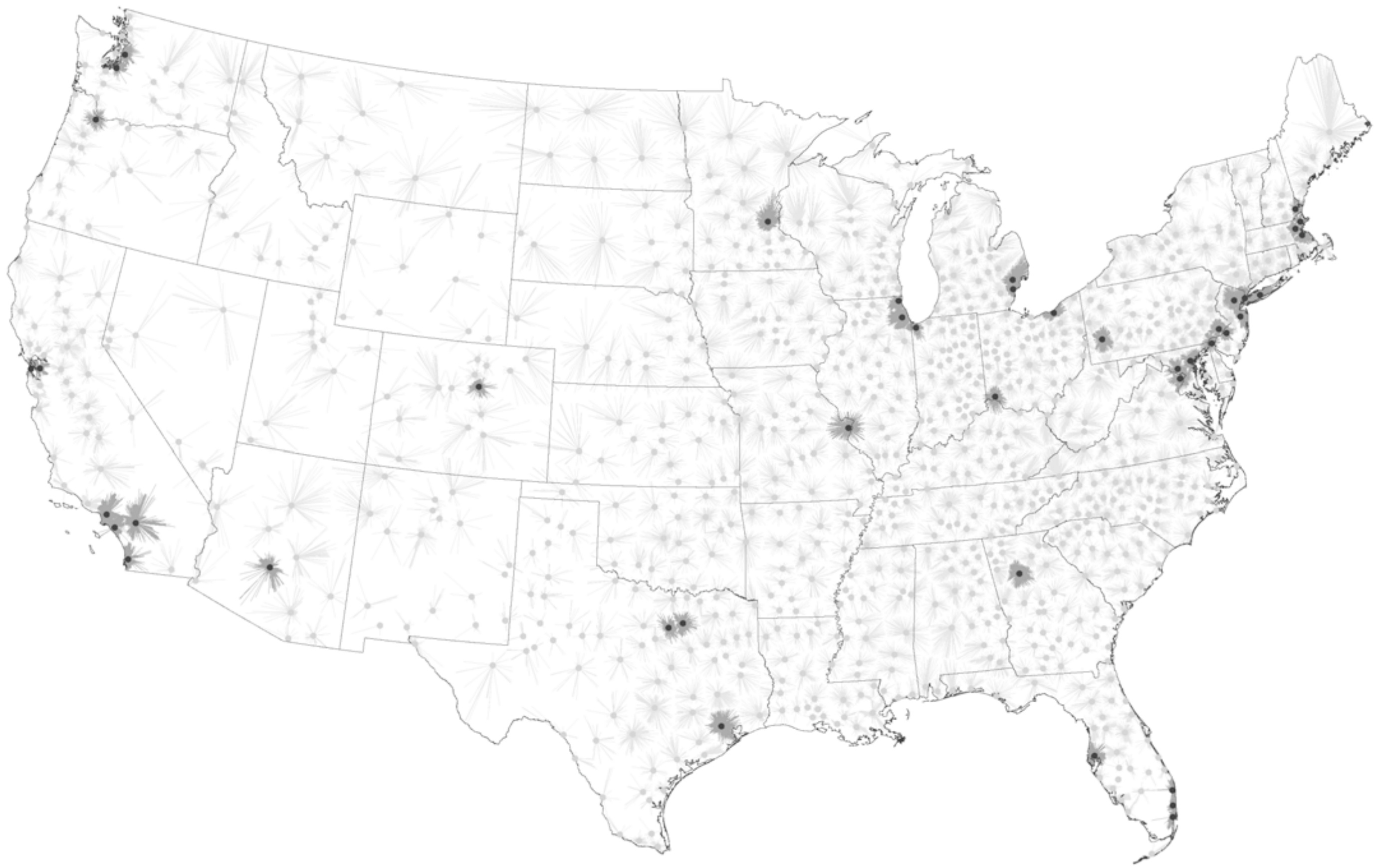


Figure 1. Core Based Statistical Areas and their Spheres of Influence

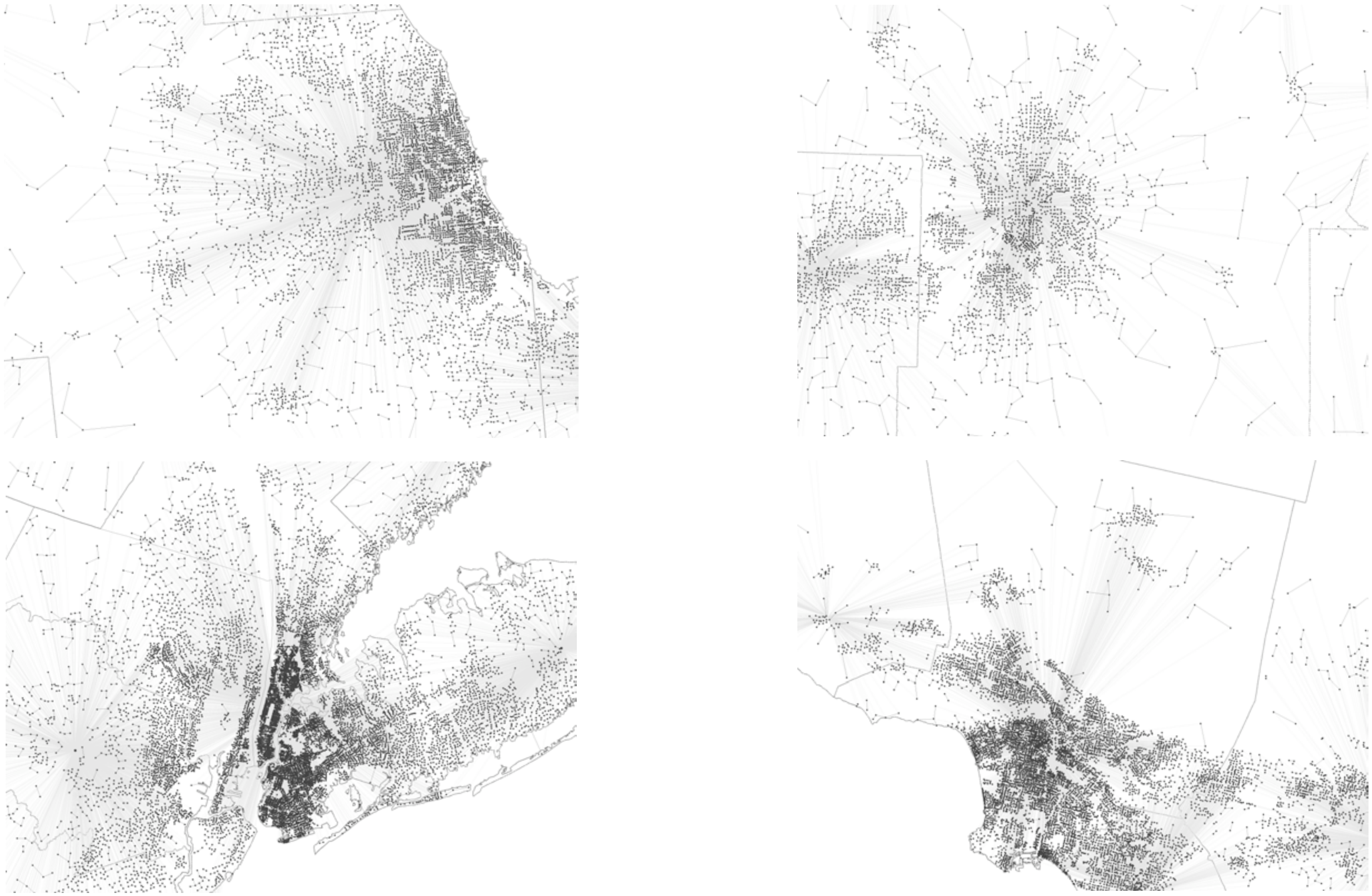


Figure 2. Spatial Point Patterns in the (clockwise from upper left) Chicago, Dallas, Los Angeles, and New York Regions

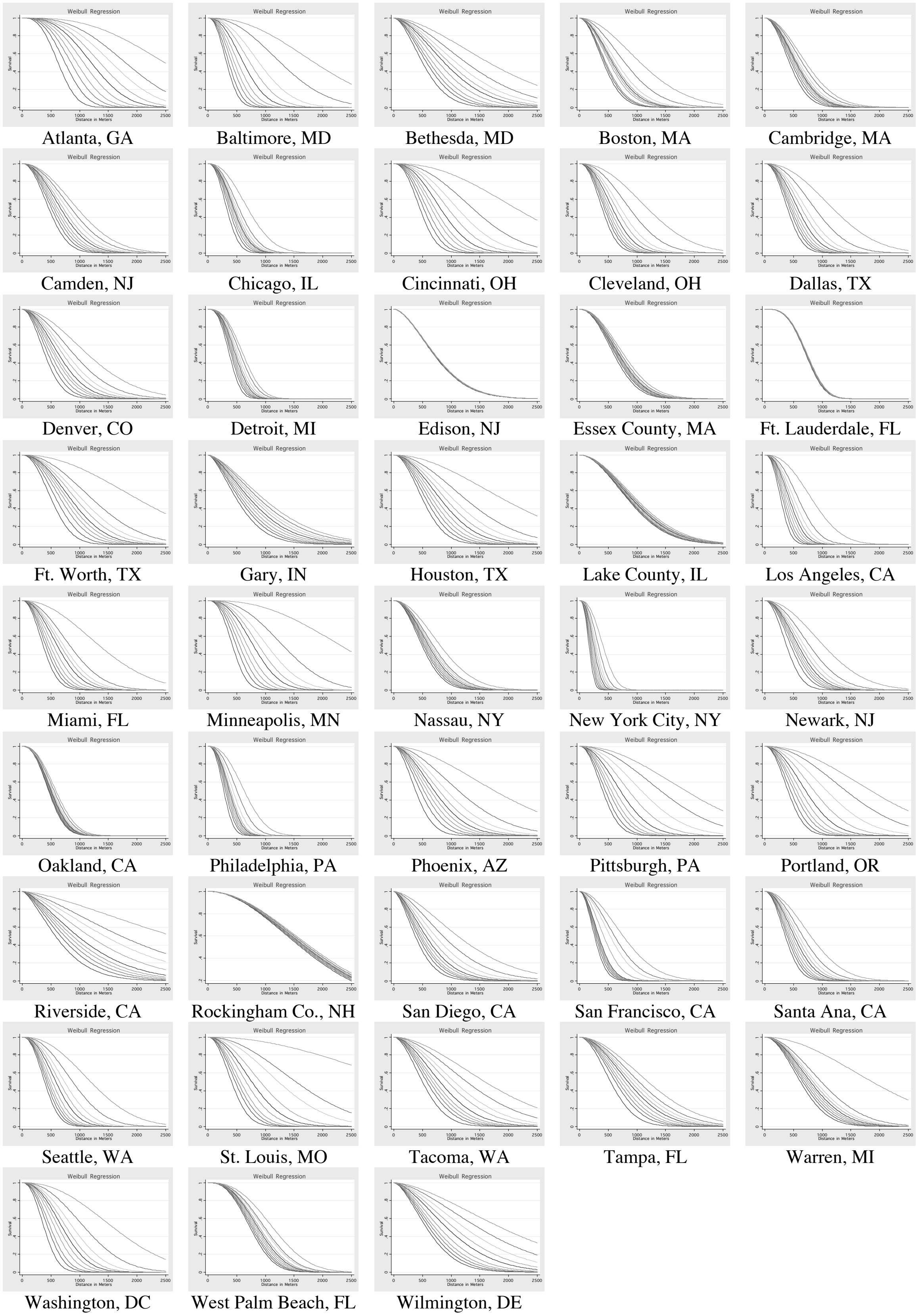
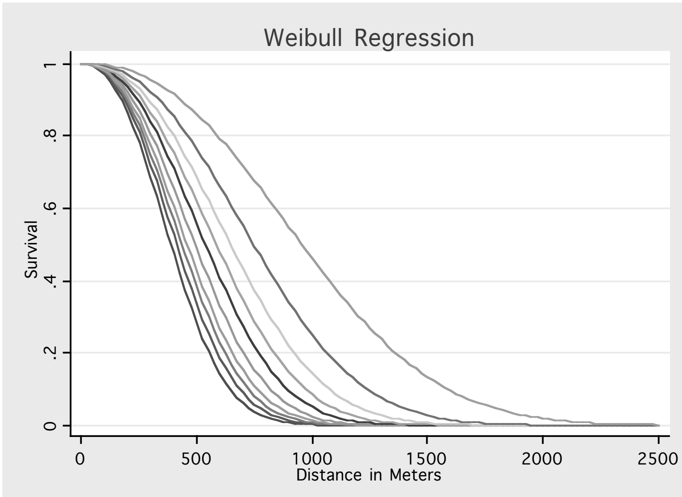
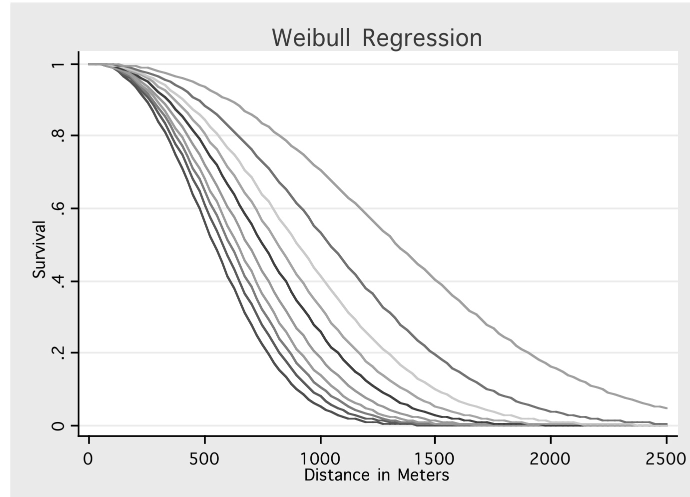


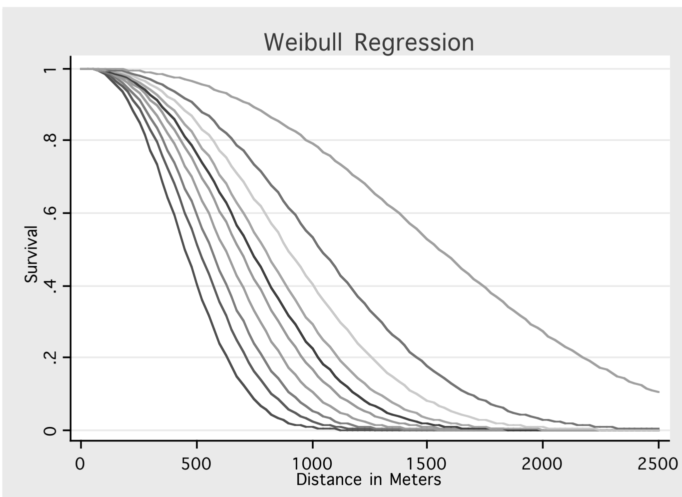
Figure 3. Estimated Survival Functions — Meters to Nearest Neighbor



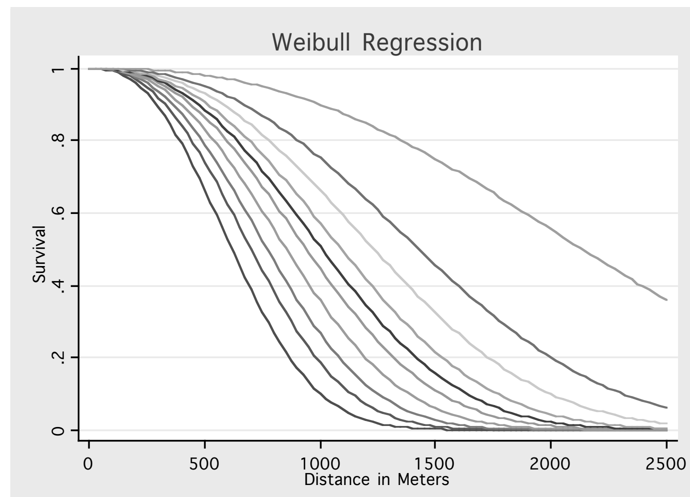
Seattle — Inside UGB



Seattle — Outside UGB



Portland-Vancouver — Inside UGB



Portland-Vancouver — Outside UGB

Figure 4. Estimated Survival for Functions for Seattle and Portland-Vancouver Regions w/ UGB— Meters to Nearest Neighbor