BORROWING FROM THE DEMOGRAPHER'S TOOLBOX:

LONGITUDINAL METHODS IN REGIONAL SCIENCE

by

Verena McClain, Brigitte S. Waldorf

Working Paper #17-6

September 2017

Dept. of Agricultural Economics

Purdue University

BORROWING FROM THE DEMOGRAPHER'S TOOLBOX: LONGITUDINAL METHODS IN REGIONAL SCIENCE

by
Verena McClain, Brigitte S. Waldorf
Dept. of Agricultural Economics, Purdue University
West Lafayette, Indiana 47907-2056
vmeckle@purdue.edu, bwaldorf@purdue.edu

Working Paper # 17-6
September 2017

Abstract

This paper provides a review of the regional science research employing longitudinal models. Three groups of studies are distinguished. The first group includes studies modelling variations in distance rather than duration. The second group includes studies that focus on spatial behavior in an event history setting. The last group is still in its infancy and casts regional change in a longitudinal perspective. We recommend that methodological advances should focus on designing space-time models that synthesize longitudinal with spatial econometric techniques.

Keywords: longitudinal methods, survival analysis, spatial hazard models

JEL Codes: C41; J61; R14

Introduction

Longitudinal methods and models undoubtedly belong to the most prominent demographic tools. Starting with a rudimentary life table — compiled by John Graunt (1620-1674) and based on deaths in 17th century London — demographers have since developed increasingly sophisticated models. State-of-the art modelling techniques capable of analyzing the timing of events range from hazard models with time-varying co-variates to multiple decrement life tables. The role of these techniques in regional science research and spatial data analysis has been surprisingly limited so far. Three broad sets of applications can be distinguished. First, there have been various papers that utilize the conceptual equivalence of distance and time as non-negative random variables to design spatial hazard models. Another set of applications is comprised of several papers estimating hazard functions for spatial behavior, primarily focusing on migration decisions in event history analyses. A third application area—casting regional or neighborhood change in a longitudinal model—is still in its infancy. This is quite astounding given that conceptual models of regional change, as for example the tipping point model, include an explicit temporal component. Moreover, regional science also seeks to design regional typologies whereby Markovian transition probabilities that describe the risk of switching between states or types are ideally suited to capture the timing of a region's transfer. Examples are a region switching from rural to urban, a region's exit from persistent poverty, a region becoming gentrified, or a region entering a downward spiral of population decline. Analyzing regional change in a longitudinal modelling frame is also most suitable to be augmented with spatial econometric techniques. Ultimately, the merger of the time-oriented longitudinal research with more traditional, spatially oriented techniques of regional science research will give rise to innovative space-time oriented paths in regional science research.

The paper is divided into six sections. Following this introduction, the second section briefly introduces the key concepts of longitudinal methods. Sections 3 to 5 are dedicated to the three application types distinguished above. The paper concludes with a critical assessment of longitudinal techniques for regional science research.

Longitudinal Models – A Brief Overview

Longitudinal models were designed to capture variations in the random variable T that measures duration, or the length of time elapsing until the occurrence of an event. As with any random variable, the distribution of T can be expressed via its probability density function $f(t)$ and its cumulative distribution $F(t)$. In the longitudinal setting, the survivor function $S(t) = 1\text{-}F(t)$ is a preferred representation of the variable's distribution. S(t) is the probability that the length of time elapsed is at least $T = t$, or:

$$S(t) = Prob(T \geq t).$$

$S(t)$ is a monotonically declining function with $S(0) = 1$, that is, it is certain that the event has not happened at time t=0, and $\lim_{t\to\infty} S(t) = 0$ indicating that the probability that the event has not yet occurred at time t approaches 0 as time goes to infinity.[1]

Unique for longitudinal settings is the representation of T's distribution via the hazard function, $h(t)$. The hazard function specifies the instantaneous rate of the event's occurrence:

$$h(t) = \lim_{t\to\infty} \frac{Prob(t \leq T < t + dt | T \geq t)}{dt} = \frac{f(t)}{S(t)}.$$

If T is discrete, then $h(t)$ is simply the conditional probability that the event occurs at time T=t, given that it has not occurred prior to T=t. Note that the four characterizations of T's distribution are mathematically equivalent. In the context of longitudinal methods and models, however, the survival function and the hazard function are particularly useful.

From a statistical point of view, the survivor and hazard functions allow us to also include censored observations when estimating the parameters of T's distribution by specifying the likelihood function as $L = \prod_i f(t_i) \prod_j S(t_j)$ whereby the first product refers to the observations i that have already experienced the event and the second product includes the information of censored observations j. If the population of interest is homogenous, then the distribution of the duration variable T can be fitted using a parametric approach whereby a specific functional form is *a priori* assumed. Examples are the exponential function (yielding a constant hazard), the Weibull distribution (monotonically increasing or decreasing hazard), and the gamma distribution (non-monotonic hazards).

If the population is not homogenous, then the survivor and hazard functions need to account for these heterogeneities that affect the duration T. Let $\boldsymbol{X}$ be a vector of covariates influencing T. The first approach—accelerated life models—consists of a semi-log model as $logT = X\beta + \varepsilon$ whereby $e^{X\beta}$ represents the multiplicative effect on duration T.[2][3] The second approach is comprised of various hazard models (Cox 1972) of which the proportional hazard model is the most widely known:

$$h(t|X) = h_o(t)\exp(X\beta).$$

The hazard at time t is expressed as a baseline hazard h_o that is proportionally shifted according to $\exp(X\beta)$. Variations of the model include those that make specific assumptions about the baseline hazard (for example, assuming a constant baseline hazard—i.e., an exponential distribution). Moreover, the model can be extended to include time-varying covariates $X(t)$ and time-dependent effects $\beta(t)$. If T is discrete, then the hazard (or conditional probability) can be modelled as a logit model.

---

[1] $S(t)$ gives rise to survival curves that are well known from demographic and medical research. The time T=t_m with $S(t_m)$=0.5 is the median survival time, that is, the time at with exactly half the population has experienced the event. The average survival time or life expectancy is $\int_0^\infty S(t)dt$.

[2] If X is a dummy variable, $\beta > 0$ then the multiplicative effects mean that the duration for the group with X=1 is e^{β} longer than for the other group.

[3] If the error is assumed to be normally distributed, then the accelerated life model is equivalent to the Tobit model.

Spatial Duration

During the 1990s, several researchers adopted longitudinal models to capture variations in distances between points, rather than time. Prompted by these early studies, Waldorf (2003) provided a detailed discussion of the mathematical and conceptual equivalence of time and distance. This equivalence is based on the fact that both time and distance can be characterized by a nonnegative random variable. She argues that the transfer is conceptually sound for spatial processes that spread continuously through space. Examples include wildfires and pollution. When applied to point patterns, spatial duration models have descriptive purposes, including the handling of edge effects, and are particularly useful for analyzing point-generating processes.

Odland and Ellis (1992) were the first to use hazard models in a spatial application. They investigated the distance between settlements and their nearest neighbors in Nebraska using a proportional hazard model and find that the distance between settlements increases from east to west. Using the 1987 National Survey of Families and Households, Rogerson, Weng and Lin (1993) investigated how demographic and socioeconomic characteristics influence the distance between the residential locations of parents and their adult children. They used a spatial survival function to estimate probabilities of settling in a location conditional on having settled at a smaller distance. They found that the frequency of distances between parents and children decreases quickly within a radius of 15 to 35 miles but then decreases more slowly at distances beyond that. Esparza and Krmenec (1994) described the distances between producer services and their clients via a Weibull distribution. Pellegrini and Reader (1996) applied proportional hazard models to a study on the spatial patterns of an agricultural innovation adoption by farmers in Canada. They drew attention to the lack of a natural order in a spatial setting, censoring as a means to deal with edge effects, heterogeneity due to locational characteristics and context influencing the conduciveness for events happening at particular places, and spatial dependence of event occurrences. Reader (2000) applied survival analysis to investigate spatial point patterns in the context of spatial epidemiology, and a means to test the ‘random labeling’ hypothesis.

Carruthers et al. (2010) used spatial hazards to characterize urban form and discuss the implications for urban policies like smart growth and growth management. For the 25 largest core based statistical areas (CBSAs) of the United States and using data for 2006, they estimated spatial hazard models of nearest neighbor distances at varying distances from the center. Carruthers et al. (2012) extended the previous study for the years 1990 and 2000. They point out that, by applying spatial hazard models to the same study areas for different points in time, it is possible to address both the timing and location of regional change.

Finally, more recent studies used spatial hazard models in transportation research and spatial choice-set delineation. Anastasopulos et al. (2012) investigated factors that determine activity-based travel distance in Athens, Greece, including demographic and socioeconomic characteristics, the purpose of the trip, the travel mode and time of travel. Jian et al. (2016) studied a car sharing system and used a spatial hazard model to explore the factors that influence users’ behavior regarding the use and selection of car share vehicles. Two studies on housing search in Seattle (Rashidi et al. 2012, Rashidi and Mohammadian 2015) and a study on housing search in

Chicago (Amini et al. 2014) tackled the problem of choice set formation by using hazard models for the travel-to-work distances to delineate the choice sets.

Timing of Spatial Behavior (Mobility and Migration)

Within the migration literature, longitudinal methods and models have been utilized by regional scientists in two distinct ways. The first type includes studies specifically addressing the duration dependence of migration propensities, that is, the time that elapses before a person or household moves. The second type is comprised of studies that use proportional hazard models to investigate associations between migration propensities, personal characteristics, other life time events, and spatial characteristics.

Duration Dependence

The migration literature has a rich tradition of addressing questions regarding who moves, where people move, and how people's relocation affects origin and destination. The question of *when* people move was long neglected. In the late 1960s, the sociologist McGinnis (1968) suggested that the standard Markov model—because of its stationarity assumption—is insufficient to capture the temporal intricacies of the mobility process. Instead, he argued, the probability of leaving the current location decreases monotonically with the length of stay in that location. This idea of cumulative inertia or—in the terminology of longitudinal models—the negative duration dependence of migration propensities was picked up and empirically tested in research conducted by demographers (Morrison 1967, Land 1969).

Early on, regional scientists—especially those with a disciplinary background in geography—made important contributions to this line of research. Clark and Huff (1977), for example, used individual household data to empirically identify such a cumulative inertia effect. They concluded that cumulative inertia is, at best, a weak effect. Ginsberg (1979a, 1979b) developed a semi-Markov model for individuals' residential histories. Subsequently, Pickles and his colleagues published a series of articles that critically investigate the identification of duration dependence (Pickles, Davies and Crouchley 1982; Pickles 1983; Pickles and Davies 1984). Emphasizing the conceptual foundation of duration dependence, Huff and Clark (1978) and Clark, Huff and Burt (1979) juxtaposed the cumulative inertia effect with a residential stress effect that increases over time. In the context of residential search and relocation behaviors, they consequently conceptualize the propensity to move as the result of the two opposing forces.[4] This was an important contribution to the literature because—in the absence of a cumulative inertia effect—the residential stress component leads to a positive duration dependence of the relocation probability. Similarly, Waldorf and Esparza (1991) postulated that two opposing forces influence immigrants' decisions to return to their country of origin: attachment to the home country and assimilation into the host country.[5]

[4] Using data for homeowners in Kansas City, Preston (1984) models the duration dependence of residential stress and cumulative inertia and finds that the residential duration effect is weak.
[5] Waldorf (1994) investigates the duration dependence of immigrants' attachment and assimilation with data for guestworkers in Germany. She finds that attachment levels decrease as guestworkers prolong their stay in Germany whereas assimilation increases at a decreasing rate. The net effects are declining return rates with guestworkers' increasing length of stay.

The implications of this broader conceptualization of duration dependence for empirical studies are substantial. Studies that model duration dependence of migration propensities or migration rates are confronted with theoretical justifications for both positive and negative effects, or weak and insignificant duration effects. Pickles, Davies and Crouchley (1982), for example, analyzed migration data from Wisconsin and concluded that duration-of-stay effects are weak. Waldorf and Esparza (1991) derived a generalized gamma function to capture the duration-of-stay dependence. In their empirical analysis of guestworkers' return from Germany, they found that the pattern of return hazards is a sequence of first decreasing hazards, followed by increasing hazards and finally decreasing hazards. In a follow-up study, Waldorf (1998) linked the temporal changes of the return hazard to age-dependent mortality in a three-dimensional life table. This design allowed her to derive cumulative years spent abroad for age-heterogeneous immigrant cohorts. Empirical studies employing hazard models (see section 4.2) can choose a flexible functional form such as a Weibull or gamma distribution for the baseline hazard. Interestingly, rarely do these studies justify their choice of a baseline hazard.

Heterogeneities, life course and migration

Starting in the 1980s, two developments prompted social scientists to increasingly adopt a life course perspective. First, several longitudinal data sets became more widely available. Examples from the US are the Panel Study on Income Dynamics (PSID) of the University of Michigan, the National Longitudinal Survey of Youth (NLSY) of the US Bureau of Labor Statistics, and the Survey of Income and Program Participation (SIPP) of the US Census Bureau. Second, statistical software for the analysis of longitudinal data became available, especially software that was capable of dealing with censored observations.

A small group of regional scientists enthusiastically participated in this shift from a cross-sectional to a longitudinal perspective. In an attempt to demonstrate the superiority of longitudinal over cross-sectional models, Davies and Pickles (1985) use a simulation to numerically show that cross-sectional research produces misleading, biased results. They are so dismissive of cross-sectional research that Clark (1992) felt compelled to point out the relative advantages of the two approaches. The longitudinal perspective emphasizes that changing location is part of a person's migration history which consists of a sequence of residential spells at distinct locations. The migration history itself is embedded in the person's life course. The life course provides a trajectory or sequence of other life events—such as a new job, changes in marital status, and birth of a child—that may trigger, or be induced, by relocation (Clark and Withers 1999).

Regional scientists have made substantial contributions to life course based migration research. A rich body of research has focused on the connection between migration and employment status. Bailey (1993) used proportional hazard models and NLSY data to show differences in migration propensities by type of move, educational attainment, and the timing of being unemployed. In a follow-up study (Bailey 1994) using the same data, he switched perspective, focusing on the timing of migration as the key predictor of unemployment duration. The intricate linkages between migration histories and employment histories call for their simultaneous consideration. This has not yet been accomplished in any of the empirical studies. Alternatively, researchers have tackled this challenge by differentiating various types of migration types and employment-status transitions. Bailey and Cooke (1998) resorted to logit models of employment status—disaggregated by onward migrants versus return migrants—with residential spell length as the key determinant. Detang-Dessendre and Molho (1999) focused on sojourn spells of young men in rural France, with employment status transition being the key predictors. They differentiated by migration distance making a distinction between long-distance migration—associated with contracted work—and short-distance migration—associated with speculative migration among the unemployed. Interestingly, they also paid attention to the differences in duration dependence by migration distance. Exclusively considering short-distance relocations, Clark and Withers (1999) found that job changes trigger household moves, but that there are variations by household type. In fact, several studies alluded to household type playing a central role in how the migration-employment histories are intertwined (Withers 1997, Bailey and Cooke 1998, Clark and Withers 2002).

An important subset of the literature connects the residential sojourn length to spatial characteristics. These studies recognize that where people live influences their propensity to leave and, in turn, affects the composition and characteristics of these places, especially given that migration propensities are not homogenous across residents. Odland and Bailey (1990) compared survivor functions to show that places with a high influx of migrants turn into places with high exit rates. In a very detailed study on duration-of-stay in poverty neighborhoods, Quillian (2003) used longitudinal measures—spell length, recurrence, exposure, and immobility—to draw conclusions about the interplay between relocation and neighborhood attributes and their change over time. Earlier, Glavac and Waldorf (1998) modelled the linkage between residential mobility propensities and ethnic concentrations, using longitudinal data for Vietnamese immigrants in Brisbane. Their results suggest that the immigrant composition influences the speed of neighborhood change, and that immigrants sort in such a way that dominant ethnic clusters strengthen while secondary clusters weaken.

Regional change in a longitudinal perspective

Much of the research described so far has focused its attention on either spatial patterns by applying spatial hazard models in which distance is the nonnegative random variable (section 3) or on the temporal aspects of spatial behavior (section 4). In land-use research, incorporating both the spatial and temporal dimension provides insights into *where* and *when* transitions happen, for example from rural to urban, from forest to agricultural land, or from low-income to high-income. Longitudinal methods have some unique features that make them viable complements to other econometric methods traditionally used in this field.

There have been several methodological papers that outline how longitudinal methods can be incorporated into land-use change research (An and Brown 2008, Wang et al. 2013, An et al. 2015). An and Brown (2008) proposed the use of survival analysis in land-use research and showed how it could complement other methods in this field. They identified four types of complexities that are frequent characteristics of space-time data (spatial complexities, temporal complexities, implicit dynamic information, and land-unit complexities) and argued that not all of these complexities can be sufficiently accommodated in traditional methods. Although An and Brown (2008) were not the first to use survival analysis in land-use research, their paper was the first to present a coherent framework for the use of this method in the field. They further proposed how the interpretation of hazard rates and survival probabilities can be translated in a land use context. Hazard rates can be interpreted as the average risk for land parcels to switch states at different points in time while survival probabilities may represent the proportion of land parcels that do not switch states over time. Instead of 'event', they labeled a switch of states 'development'. The authors argued that survival analysis has the additional advantage that time-dependent variables, which are often present in land use data, can be accounted for in the model. Finally, they identified three strengths of survival analysis in land–use science, that is (1) its ability to deal with censored data, (2) the concept of hazard rates and the inclusion of time-dependent variables to address the issue of implicit dynamic information, and (3) its capacity to handle problems arising due to development of land into different types.

An et al. (2011) used this approach to investigate what factors drive the timing and location of urban development in southeastern Michigan. They combined data on land-use at different points in time derived from aerial photos with data on geographic, socioeconomic, environmental and biophysical attributes as time-dependent explanatory variables. In their model, the state 'undeveloped' is terminated when a parcel of land becomes 'developed' so that the survival function describes how the probability of a parcel being undeveloped changes over time. This approach allowed them to analyze both the temporal as well as the spatial aspects of urbanization.

Wang et al. (2013) compared logistic regression and survival analysis in terms of their ability to identify spatial predictors of land-use change by applying both methods to data generated in an agent-based simulation model and running Monte Carlo experiments. They found that survival analysis performs better than logistic regression in identifying the predictors largely because of its ability to account for the effect of time-dependent variables.

Several studies have linked socioeconomic, demographic or biophysical data to satellite images of the study area and used longitudinal methods to analyze land-use changes over time. Vance and Geoghegan (2002) used a hazard model approach combined with satellite images to determine hazard rates for forest conversion in southern Mexico and Greenberg et al. (2005) used a similar approach to analyze deforestation of a tropical rainforest in Ecuador. The pixels in satellite images represent land parcels on the ground and current land-use is identified based on the pixels' characteristics. Deforestation events are then identified by comparing the color of pixels in satellite images taken at different points in time. The at-risk group in these studies are those pixels that represent forest cover in each image and the hazard rates can be interpreted as the deforestation rate at each point in time while the survival function describes the remaining forest area as a function of time. Both studies included additional explanatory variables in order to analyze what factors and characteristics of the land parcel and the people using it impact the likelihood that deforestation occurs.

Iovanna and Vance (2007) applied this approach to analyze urbanization in North Carolina using satellite images from five points in time between 1976 and 2001. Data from satellite images are often interval-censored, meaning that the timing of the land-use change cannot be precisely determined as satellite images are typically taken infrequently so that there is often an interval of several years between available images. Nevertheless, the use of satellite images presents an efficient and accessible way to study land use changes when the focus area of the study is relatively large as is often the case when the topic is deforestation or agricultural land-use change. In some cases, the research design may require a higher spatial resolution of the data that satellite images or aerial photos may not provide. Irwin and Bockstael (2004) analyzed how urban sprawl develops over time using a dataset of land-use at the county level that includes for example a parcel's size, zoning and current use and combine it with parcel characteristics derived from GIS data such as distance to major employment centers.

In many applications of hazard models there is only one event that ends a particular state, for instance 'death' is the only event that terminates 'being alive' in demographic models. Competing-risk duration models allow for the possibility of multiple terminating events. A study on how vehicle access impacts the residential mobility of low-income households by Dawkins, Jeon and Pendall (2015) provided an example of how a competing-risk duration model can be applied in urban studies. They investigated the dynamics of residential mobility of households that participated in a program that aimed at increasing upward social mobility by relocating participants from high-poverty neighborhoods to low-poverty neighborhoods so that they can benefit from the better infrastructure and economic opportunities associated with low-poverty neighborhoods. In their model, a spell of a household's residence in one neighborhood can end with a move into either a low-poverty neighborhood or a high-poverty neighborhood. This shows another way of how both time (duration of stay in a neighborhood) and space (location or characteristic of new neighborhood) can be included within the framework of a hazard model.

Conclusions

This review has shown that longitudinal models and methods are important but also underutilized techniques of regional science research. As of now, regional science research that employs survival models is still the exception. This is unfortunate as the longitudinal perspective emphasizes development, time before change, rather than the *faits accomplis* of cross-sectional models. And even the temporal sequencing of cross-sections in a panel setting cannot match the longitudinal perspective as an unknown number of transitions may have occurred in–between the points of observations.

Methodological advances are needed to explicitly integrate longitudinal models with the space-oriented models and techniques of regional science, for example, the spatial econometric models. The previous discussion illustrates that some progress on joint space-time research has already been made. In Section 5, we emphasized that there are several ways of including both the spatial and temporal dimension of land-use change within the framework of longitudinal models. One method that has been used is to combine satellite images or aerial photos taken at multiple points in time with data from other sources, such as GIS or census data, to estimate the impact of different factors on land-use change over time using temporal hazard models. When satellite images are either not available or do not provide the necessary spatial resolution other sources may provide data that can be used to track land-use changes over time such as census data, or records on zoning, taxation or housing from local authorities. Even when images are used, other data sources can be used to supplement and verify the accuracy of the image analysis, as done for example by An et al. (2011). Competing-risk models incorporate the spatial dimension into a temporal hazard model by allowing for multiple terminating events (e.g. moving to different locations) as in Dawkins, Jeon and Pendall (2015). Finally, Carruthers et al. (2012) show that temporal aspects can be accounted for when using spatial hazard models as described in Section 3 by repeating the models for several points in time.

As a final point we like to emphasize that—unlike spatial econometric models—longitudinal methods and models do not belong to the standard curriculum of regional science oriented graduate programs. However, this review has shown that they enrich regional science research, making their inclusion in the curriculum highly desirable.

References

Amini, B., Javanmardi, M., & Mohammadian, K. (2014). Application of Hazard Based Model for Housing Location Based on Travel Distance to Work. International Journal of Transportation Engineereing, 2(4), 267-278.

An, L., Brown, D. G., Nassauer, J. I., & Low, B. (2011). Variations in development of exurban residential landscapes: timing, location, and driving forces. *Journal of Land Use Science*, *6*(1), 13-32.

An, L., & Brown, D. G. (2008). Survival analysis in land change science: Integrating with GIScience to address temporal complexities. *Annals of the Association of American Geographers*, *98*(2), 323-344.

An, L., Tsou, M. H., Crook, S. E., Chun, Y., Spitzberg, B., Gawron, J. M., & Gupta, D. K. (2015). Space–time analysis: Concepts, quantitative methods, and future directions. *Annals of the Association of American Geographers*, *105*(5), 891-914.

Anastasopoulos, P. C., Islam, M. B., Perperidou, D., & Karlaftis, M. G. (2012). Hazard-based analysis of travel distance in urban environments: longitudinal data approach. *Journal of Urban Planning and Development*,*138*(1), 53-61.

Bailey, A. J. (1993). Migration history, migration behavior and selectivity. *The Annals of Regional Science*, *27*(4), 315-326.

Bailey, A. J. (1994). Migration and unemployment duration among young adults. *Papers in Regional Science*, *73*(3), 289-307.

Bailey, A. J., & Cooke, T. J. (1998). Family migration and employment: the importance of migration history and gender. *International Regional Science Review*, *21*(2), 99-118.

Carruthers, J. I., Lewis, S., Knaap, G. J., & Renner, R. N. (2010). Coming undone: A spatial hazard analysis of urban form in American metropolitan areas. *Papers in Regional Science*, *89*(1), 65-88.

Carruthers, J. I., Hepp, S., Knaap, G. J., & Renner, R. N. (2012). The American Way of Land Use: A Spatial Hazard Analysis of Changes Through Time. International Regional Science Review, 35(3), 267-302.

Chen, Y., Li, X., Liu, X., Ai, B., & Li, S. (2016). Capturing the varying effects of driving forces over time for the simulation of urban growth by using survival analysis and cellular automata. *Landscape and Urban Planning*, *152*, 59-71.

Clark, W. A. (1992). Comparing cross-sectional and longitudinal analyses of residential mobility and migration. *Environment and Planning A*, *24*(9), 1291-1302.

Clark, W. A., & Huff, J. O. (1977). Some empirical tests of duration-of-stay effects in intraurban migration. *Environment and Planning A*, *9*(12), 1357-1374.

Clark, W. A., Huff, J. O., & Burt, J. E. (1979). Calibrating a model of the decision to move. *Environment and Planning A*, *11*(6), 689-704.

Clark, W. A., & Davies Withers, S. (1999). Changing jobs and changing houses: mobility outcomes of employment transitions. *Journal of Regional Science*, *39*(4), 653-673.

Clark, W. A., & Withers, S. D. (2002). Disentangling the interaction of migration, mobility, and labor-force participation. *Environment and Planning A*, *34*(5), 923-945.
Crouchley, R., Davies, R. B., & Pickles, A. R. (1982). Identification of some recurrent choice processes. *Journal of Mathematical Sociology*, *9*(1), 63-73.
Davies, R. B., & Pickles, A. R. (1985). Longitudinal versus cross-sectional methods for behavioural research: a first-round knockout. *Environment and Planning A*, *17*(10), 1315-1329.
Dawkins, C., Jeon, J. S., & Pendall, R. (2015). Vehicle access and exposure to neighborhood poverty: Evidence from the moving to opportunity program. *Journal of Regional Science*, *55*(5), 687-707.
Detang-Dessendre, C., & Molho, I. (1999). Migration and changing employment status: a hazard function analysis. *Journal of Regional Science*, *39*(1), 103-123.
Huffman, W. E., & Feridhanusetyawan, T. (2007). Migration, fixed costs, and location-specific amenities: A hazard analysis for a panel of males. *American journal of agricultural economics*, *89*(2), 368-382.
Ezra, M., & Kiros, G. E. (2001). Rural out-migration in the drought prone areas of Ethiopia: A multilevel analysis. *International Migration Review*, 749-771.
Ginsberg, R. B. (1979a). Timing and duration effects in residence histories and other longitudinal data: I-stochastic and statistical models. *Regional Science and Urban Economics*, *9*(4), 311-331.
Ginsberg, R. B. (1979b). Timing and duration effects in residence histories and other longitudinal data: II—studies of duration effects in Norway, 1965–1971. *Regional Science and Urban Economics*, *9*(4), 369-392.
Glavac, S. M., & Waldorf, B. (1998). Segregation and residential mobility of Vietnamese immigrants in Brisbane, Australia. *The Professional Geographer*, *50*(3), 344-357.
Greenberg, J. A., Kefauver, S. C., Stimson, H. C., Yeaton, C. J., & Ustin, S. L. (2005). Survival analysis of a neotropical rainforest using multitemporal satellite imagery. *Remote sensing of environment*, *96*(2), 202-211.
Grube-Cavers, A., & Patterson, Z. (2014). Urban rapid rail transit and gentrification in Canadian urban centres: A survival analysis approach. *Urban Studies*, *52*(1), 178-194.
Huff, J. O., & Clark, W. A. (1978). Cumulative stress and cumulative inertia: a behavioral model of the decision to move. *Environment and Planning A*, *10*(10), 1101-1119.
Huffman, W. E., & Feridhanusetyawan, T. (2007). Migration, Fixed Costs, and Location-Specific Amenities: A Hazard Analysis for a Panel of Males. *American Journal of Agricultural Economics*, *89*(2).
Iovanna, R., & Vance, C. (2007). Modeling of continuous-time land cover change using satellite imagery: an application from North Carolina. *Journal of Land Use Science*, *2*(3), 147-166.
Irwin, E. G., & Bockstael, N. E. (2004). Land use externalities, open space preservation, and urban sprawl. *Regional science and urban economics*,*34*(6), 705-725.
Jian, S., Rashidi, T. H., Wijayaratna, K. P., & Dixit, V. V. (2016). A Spatial Hazard-Based analysis for modelling vehicle selection in station-based carsharing systems. Transportation Research Part C: Emerging Technologies, 72, 130-142.
Land, K. C. (1969). Duration of residence and prospective migration: Further evidence. *Demography*, *6*(2), 133-140.
Le Gallo, J. (2004). Space-time analysis of GDP disparities among European regions: A Markov chains approach. *International Regional Science Review*, *27*(2), 138-163.

McGinnis, R. (1968). A stochastic model of social mobility. *American Sociological Review*, 712-722.
Morrison, P. A. (1967). Duration of residence and prospective migration: the evaluation of a stochastic model. *Demography*, *4*(2), 553-561.
Motamed, M. J., Florax, R. J., & Masters, W. A. (2014). Agriculture, transportation and the timing of urbanization: Global analysis at the grid cell level. *Journal of Economic Growth*, *19*(3), 339-368.
Odland, J., & Bailey, A. J. (1990). Regional Out-Migration Rates and Migration Histories: A Longitudinal Analysis. *Geographical Analysis*, *22*(2), 158-170.
Odland, J., & Ellis, M. (1992). Variations in the spatial pattern of settlement locations: an analysis based on proportional hazards models. *Geographical Analysis*, *24*(2), 97-109.
Palloni, A., Massey, D. S., Ceballos, M., Espinosa, K., & Spittel, M. (2001). Social capital and international migration: A test using information on family networks1. *American Journal of Sociology*, *106*(5), 1262-1298.
Pellegrini, P. A., & Grant, J. T. (1999). Policy coalitions in the US Congress: A spatial duration modeling approach. Geographical Analysis, 31(1), 45-66.
Pellegrini, P. A., & Reader, S. (1996). Duration modeling of spatial point patterns. *Geographical Analysis*, *28*(3), 219-243.
Pickles, A. R. (1983). The analysis of residence histories and other longitudinal panel data: A continuous time mixed markov renewal model incorporating exogeneous variables. *Regional Science and Urban Economics*, *13*(2), 271-285.
Pickles, A. R., Davies, R. B., & Crouchley, R. (1982). Heterogeneity, nonstationarity, and duration-of-stay effects in migration. *Environment and Planning A*, *14*(5), 615-622.
Pickles, A. R., & Davies, R. B. (1984). Recent developments in the analysis of movement and recurrent choice. In *Spatial statistics and models* (pp. 321-343). Springer Netherlands.
Preston, V. (1984). A path model of residential stress and inertia among older people. *Urban Geography*, *5*(2), 146-164.
Quillian, L. (2003). How long are exposures to poor neighborhoods? The long-term dynamics of entry and exit from poor neighborhoods. *Population Research and Policy Review*, *22*(3), 221-249.
Rashidi, T. H., Auld, J., & Mohammadian, A. K. (2012). A behavioral housing search model: Two-stage hazard-based and multinomial logit approach to choice-set formation and location selection. Transportation Research Part A: Policy and Practice, 46(7), 1097-1107.
Rashidi, T. H., & Mohammadian, A. K. (2015). Behavioral Housing Search Choice Set Formation A Spatial Hazard-based Screening Model. International Regional Science Review, 38(2), 151-170.
Reader, Steven. "Using survival analysis to study spatial point patterns in geographical epidemiology." *Social Science & Medicine* 50.7 (2000): 985-1000.
Rogerson, P. A., Weng, R. H., & Lin, G. (1993). The spatial separation of parents and their adult children. Annals of the association of American Geographers, 83(4), 656-671.
Sabia, J. J. (2008). There's No Place Like Home A Hazard Model Analysis of Aging in Place Among Older Homeowners in the PSID. *Research on Aging*, *30*(1), 3-35.
Vance, C., & Geoghegan, J. (2002). Temporal and spatial modelling of tropical deforestation: a survival analysis linking satellite and household survey data. *Agricultural Economics*, *27*(3), 317-332.

Waldorf, Brigitte. "Spatial patterns and processes in a longitudinal framework." *International Regional Science Review* 26.3 (2003): 269-288.
Waldorf, B. S., & Esparza, A. (1991). A parametric failure time model of international return migration. *Papers in Regional Science*, *70*(4), 419-438.
Waldorf, B. S. (1994). Assimilation and attachment in the context of international migration: The case of guestworkers in Germany. *Papers in Regional Science*, *73*(3), 241-266.
Waldorf, B. (1998). A three-dimensional life table approach to immigrants' sojourns abroad. *Papers in Regional Science*, *77*(4), 407-425.
Withers, S. D. (1997). Methodological considerations in the analysis of residential mobility: A test of duration, state dependence, and associated events. *Geographical Analysis*, *29*(4), 354-372.
Wang, N., Brown, D. G., An, L., Yang, S., & Ligmann-Zielinska, A. (2013). Comparative performance of logistic regression and survival analysis for detecting spatial predictors of land-use change. *International Journal of Geographical Information Science*, *27*(10), 1960-1982.

Appendix

Table 1: Overview of studies referenced in the article and additional examples of applications of longitudinal models in regional sciences

	Application	Reference
Group I	**Spatial Duration**	
	Housing search and commuting distance	Amini et al. (2014), Rashidi and Mohammadian (2015), and Rashidi et al. (2012)
	Travel distance in urban environments	Anastasopoulos et al. (2012)
	Urbanization in American metropolitan areas	Carruthers et al. (2010)
	Vehicle selection in station-based carsharing systems	Jian et al. (2016)
	Spatial pattern of settlement locations in Nebraska	Odland and Ellis (1992)
	Policy coalitions in the US congress	Pellegrini and Grant (1999)
	Adoption of an agricultural innovation	Pellegrini and Reader (1996)
	Investigate 'random labeling' hypothesis	Reader (2000)
	Spatial separation pf parents and their adult children	Rogerson et. Al (1993)
	Conceptual equivalence of time and distance	Waldorf (2003)
Group II	**Timing of spatial behavior (mobility and migration)**	
	Migration behavior and migration history of young adults	Bailey (1993)
	Timing of migration as the key predictor of unemployment duration	Bailey (1994)
	Residential spell length, employment status and migration	Bailey and Cooke (1998)
	Comparison of longitudinal and cross-sectional models of in migration research	Clark (1992)
	Cumulative inertia and residential stress effect as opposing forces	Clark, Huff and Burt (1979),Clark and Huff (1977), and Huff and Clark (1978)
	Changing locations as part of a person's life course	Clark and Withers (1999)

	Application	Reference
	Household types and migration-employment histories	Clark and Withers (2002)
	Duration-of-stay effects in migration using data from Wisconsin	Crouchley, Davies and Pickles (1982)
	Comparison of longitudinal and cross-sectional models in regional sceinces	Davies and Pickles (1985)
	Migration decision and employment status among young adults in France	Detang-Dessendre and Molho (1999)
	School attendance and remittances in El Salvador	Edwards and Ureta (2003)
	Drought risk and rural out-migration in Ethiopia	Ezra and Kiros (2001)
	Semi-Markov model for individuals' residential histories	Ginsberg (1979a) and Ginsberg (1979b)
	Gentrification in Canadian cities and access to rail transit	Grube-Cavers and Patterson (2014)
	Likelihood of Vietnamese immigrants in Brisbane, Australia of changing residence	Glavac and Waldorf (1998)
	Interstate migration of adult working-age males	Huffman and Feridhanusetyawan (2007)
	Empirically tests cumulative inertia	Land (1969)
	Introduced concept of cumulative inertia, negative duration dependence	McGinnis (1968)
	Empirically tests cumulative inertia	Morrison (1967)
	Migration behavior and the duration of time individuals spend in aregion	Odland and Bailey (1990)
	Social capital and international migration	Palloni et al. (2001)
	Identification of duration dependence	Pickles (1983), Pickles, Davies and Crouchley (1982), Pickles and Davies (1984)
	Empirical test of cumulative inertia and residential stress for homeowners in Kansas City	Preston (1984)
	Duration of stay in poor neighborhoods	Quillian (2003)
	Analysis of factors influencing aging homeowners' decision to age in place	Sabia (2008)
	International return migration - attachment to home country versus assimilation into host country in the case of guestworkers in Germany	Waldorf and Esparza (1991) and Waldorf (1994)
	Immigrants' sojourns abroad - linking temporal changes of return hazard to age-dependent mortality	Waldorf (1998)
	Household types and migration-employment histories	Withers (1997)
Group III	**Regional change in a longitudinal perspective**	
	Methodological discussion of incorporating survival analysis into land change science	An and Brown (2008)
	Timing, location and driving forces of urbanization in Michigan townships using satellite images and survival analysis	An et al. (2011)
	Discussion of different methods for space-time analysis	An et al. (2015)
	Urbanization in American metropolitan areas over time	Carruthers et al. (2012)
	Combination of survival analysis and cellular automata to simulate urban growth	Chen et al. (2016)
	Influence of vehicle access on residential spells and transitions to and from high and low-poverty neighborhoods	Dawkins, Jeon and Pendall (2015)

Application	Reference
Application of survival analysis and satellite image analysis to deforestation of a rainforest in Ecuador	Greenberg et al. (2005)
Using survival analysis and satellite images to study urbanization in North Carolina	Iovanna and Vance (2007)
Urban sprawl over space and time using survival analysis	Irwin and Bockstael (2004)
Evolution of GDP disparities in Europe over space and time	Le Gallo (2004)
Timing of a location's transition from rural to urban over the last 2000 years	Motamed et al. (2014)
Application of survival analysis and satellite image analysis to deforestation in southern Mexico	Vance and Geoghegan (2002)
Comparison of logistic regression and survival analysis using simulations	Wang et al. (2013)