Empirical Methods for Modeling Landscape Change, Ecosystem Services, and Biodiversity

David J. Lewis and Ralph J. Alig

The widespread development of land has been identified as a primary driver of global ecological change (Foley et al. 2005), affecting both terrestrial (Armstrong et al. 2004) and aquatic systems (Naiman and Turner 2000). It is increasingly recognized that understanding the effects of land-use change on ecosystems requires the integration of the social and natural sciences, a primary emphasis in the scientific literature now known as land change science (Turner et al. 2007). The discipline of economics has much to offer to land change science. On privately-owned landscapes, decisions regarding the use of land – whether to develop houses, grow trees, or plant a particular crop – are made within the context of local land markets and global commodity markets as well as local and regional regulatory institutions. As such, quantitative econometric land-use models are useful in accounting for such markets in estimation, especially if the modeling goal is the analysis of policies aimed at altering market incentives. However, because landscape pattern is fundamentally important for understanding many ecological processes, land-use models must confront the challenges associated with modeling not just aggregate land-use shares, but the spatial pattern of land use. Because landscape change is fundamentally shaped by the spatial pattern of landowner decisions, modeling landscape change must begin by modeling the decisions of individual landowners.

The purpose of this paper is to synthesize recent economics research aimed at integrating discrete-choice econometric models of land-use change with spatially-explicit landscape simulations and quantitative ecology. This research explicitly models changes in the spatial pattern of landscapes in two steps: 1) econometric estimation of parcel-scale transition probabilities from observed data; and 2) spatially-explicit simulations of landscape change that use the estimated transition probabilities as decision rules that guide land-use change. This paper will focus on the recent literature that examines the conservation of private land – an important issue for ecological change in both the western and eastern United States. Advances occurring in the past ten years regarding spatial data availability, micro-econometric modeling, spatially-explicit landscape simulation approaches, and policy applications are highlighted. The article closes by identifying multiple research challenges in need of attention by environmental and resource economists, including: 1) an increased focus on causal identification strategies; 2) improved accounting of unobserved heterogeneity in estimation; 3) the construction of new spatial-panel datasets; and 4) the development of fully coupled econometric-ecological models of landscape change.

Distinguishing Economic Methods for Modeling Land-Use Change from Other Disciplines

Many researchers from a variety of disciplines have developed methods for modeling changes in land use. Demographers have recently teamed with landscape ecologists and geographers...
to develop methods to project housing growth for the United States, typically using decennial Census data as a basis for estimation (e.g., Hammer et al. 2004; Gustafson et al. 2005; Theobald 2005). Similarly, the USDA Forest Service has developed watershed-scale projections of housing growth using projection-oriented methods for the forty-eight contiguous states to highlight regions of concern for deforestation or more houses in forests (Stein et al. 2005). Geographers offer a substantial literature that focuses on modeling land-use and land-cover change, as opposed to just housing growth (Veldkamp and Lambin 2001; Lambin et al. 2001; Walker 2003). Further, geographers have made advances in simulation-based methods applied to landscape change (e.g., Wu 2002; Allen and Lu 2003), most recently under the framework of Agent-Based Models (e.g., see Parker et al. 2003). The simulation methods devised in the geography literature have been particularly influential in pushing the economics literature to integrate landscape simulations and econometric modeling (e.g., Lewis and Plantinga 2007).

Economic methods used to analyze land-use change employ a modeling structure derived from economic theory. It captures the motivation of the landowner to convert land from one use to another. In particular, most modern micro-econometric models emphasize the importance of land rent in specifying the individual landowner's decision, and spend considerable energy on estimating causal effects. In contrast, geographic and demographic methods do not impose economic structure and typically use prior estimates of rates of land-use change to project future land-use. Although both methods can be used to project large-scale changes in land-use, a primary advantage of the economic approach is the ability to examine landscape change as a function of policy-relevant scenarios. For example, because geographic and demographic methods do not model land-use as a function of the net returns to land, such methods cannot be used to understand the effects of changes in commodity prices or incentive-based policies (e.g., Conservation Reserve Program) on land-use change. Further, the focus of many geographic and demographic methods on housing density misses other important transitions, such as the recent increase in forestland of three million acres in the U.S. southeast (Alig and Plantinga 2004). Given the importance of forestland for ecosystem service provision, transitions between agriculture and forestry are particularly important for policies aimed at enhancing ecosystem services.

Econometric Methods for Modeling Land-Use Change (Step One)

Beginning with Bockstael's (1996) seminal analysis, the economics literature has emphasized the estimation of discrete land-use decisions at the parcel or plot scale, e.g., develop or not. A parcel is defined by ownership boundaries, while a plot is defined as a piece of homogeneous land, typically within an individual ownership boundary. Although land-use shares models estimated at the state or county scale (e.g. Alig 1986) are useful for examining the quantity of land-use change in a region or nationwide, parcel or plot scale models are required for analyzing the spatial configuration of landscape change.

Econometric models of land-use change emphasize the discounted stream of expected land rents in their specification of a profit-maximizing landowner choosing how to allocate a

---

2 In the western U.S., the USDA Forest Service projects the largest loss in forest as occurring in portions of the Sierra foothills of California and in northern Washington state.

3 For an example of the importance of market returns on land-use, note that Lubowski et al. (2008) find that the most important factor driving the increase in U.S. forests between 1982 and 1997 was timber rents.
homogenous plot of land – see Alig (1986), Capozza and Helsley (1989), and Stavins and Jaffe (1990) for theoretical underpinnings. If rental values and conversion costs are linear in parcel size – as is typically assumed – a landowner’s decision on allocating a parcel of heterogeneous quality can be treated as the sum of land-use choices on constituent homogeneous quality plots (Lubowski et al. 2006). Econometric estimation of discrete-choice decisions typically specify Probit models of the binary development decision (Bockstael 1996; Carion-Flores and Irwin 2004); Logit models of a set of multiple land-use choices involving agriculture, forest, and development (Nelson et al. 2001; Lubowski et al. 2006; Newburn and Berck 2006; Lewis and Plantinga 2007; Langpap et al. 2008); gravity models of urbanization (e.g., Kline et al. 2001); duration models of the time to conversion (Irwin and Bockstael 2002; Irwin et al. 2003; Towe et al. 2008); and jointly estimated Probit-Poisson models of the decision to develop and the decision of how many new lots to create (Lewis et al. 2009b; Lewis 2009).

Estimating the discrete-choice land-use decision requires spatial data for at least two points in time, and such data have been derived from a number of sources. Plot-level land-use data have been derived from repeated surveys by the USDA’s National Resources Inventory (NRI) (Schatzki 2003; Lubowski et al. 2006; Lewis and Plantinga 2007; Langpap et al. 2008) and the USDA Forest Service Forest Inventory and Analysis (FIA) dataset (e.g., Kline et al. 2001). Parcel-scale land-use data have been primarily obtained from local tax assessor or land information offices (Irwin and Bockstael 2002, 2004; Carion-Flores and Irwin 2004; Newburn and Berck 2006; Towe et al. 2008), and digitized paper plat maps linked with tax assessor data (Lewis et al. 2009b). Figure 1a gives an example of parcel data. To date, parcel-level datasets have primarily been used for models of urban development, while the plot-level datasets have been used in applications involving a broader land base to model conversions between agriculture, forestry, and urban development.

In selecting between the national plot-level datasets such as the NRI, or the parcel-level data obtained by local governments, the prior literature has identified several tradeoffs. First, a clear advantage of the NRI and FIA is consistent nationwide data, while digitized parcel records tend to only be available at very local scales. Second, since agricultural and forestry prices typically exhibit little variation within regions such as counties, it is difficult to estimate the effects of agricultural or forestry rents on land-use change with localized data, and the NRI has been usefully deployed to expand the geographic scope for estimation in prior analyses interested in such effects (Lewis and Plantinga 2007). Third, the NRI and FIA data do not typically disclose the exact location of plots due to confidentiality concerns, thereby reducing the usefulness of these datasets for analyses interested in the effects of fine-scale factors on land-use change. Examples of fine-scale factors of interest to land-use change modelers include minimum lot zoning policies (e.g., McMillen and McDonald 1989) and spatial externalities across land-uses (e.g., Irwin and Bockstael 2002), among others. Fourth, the NRI and FIA data do not provide information on the density of urban developments, and so models derived from these data are limited in their ability to differentiate land-use change by development density. Given its basis in ownership boundaries, parcel data have been shown to be particularly suitable for modeling the determinants of low-density development (Irwin and Bockstael 2007).

4 Parcel-scale data has become more widely available in recent years.

5 To further highlight the importance of data sources, it should be noted that conclusions regarding the extent of urban sprawl derived from satellite-based land cover maps (Burchfield et al. 2006) have differed substantially from results derived with parcel data – see Irwin and Bockstael (2007) for the comparison.
Numerous econometric challenges arise with discrete-choice models that the literature has attempted to address – though much effort has been focused on the treatment of unobserved heterogeneity in estimation. For example, the question of whether a parcel is more likely to develop if its neighbor is developed has been examined in urban-rural fringe settings, and is challenging due to the presence of unobserved spatial heterogeneity (Irwin and Bockstael 2002, 2004; Newburn and Berck 2006). In particular, difficulties in estimating the effects of spatial externalities arises from the endogeneity of measures of neighboring development, and, although there is evidence that spatial externalities are important (Irwin and Bockstael 2002), this issue remains an active area of research. Another challenge from unobserved heterogeneity arises when researchers are interested in quantifying the determinants of development density, e.g. average lot size. Using data on observed developments to estimate the effects of a variety of covariates on density (e.g., McConnell et al. 2006) is subject to sample selection bias if there are unobservables correlated across both the decision to develop and the density decision. Recent work has specified jointly estimated Probit-Poisson models to account for such selection bias (Lewis et al. 2009b). One common feature of some of the most recent parcel-scale models is the use of simulation in the estimation stage as a means of accounting for important unobserved heterogeneity (e.g., Newburn and Berck 2006; Lewis et al. 2009b; Lewis et al. 2009c). Recent advances in micro-econometric estimation techniques (e.g., Train 2003) and computational speed have greatly facilitated simulation-based estimators and allow for parcel-scale econometric models with a less restrictive set of assumptions regarding the statistical independence of land-use decisions across parcels.

**Linking Econometric Land-Use Models with GIS and Ecological Models (Step Two)**

Parcel-scale econometric models of land-use change generate estimates of the transition probability of a parcel of land. Such transition probabilities are a function of parcel-scale covariates such as lot size and physical characteristics (Irwin and Bockstael 2002), aggregated measures of net returns to various land uses (Lewis and Plantinga 2007; Langpap et al. 2008), or land policy variables such as zoning (Newburn and Berck 2006; Lewis et al. 2009b). Spatial data on each covariate found in the econometric model are then used to link the estimated transition probabilities to particular points on a landscape. Digitized GIS data from tax assessor databases typically provide the foundation necessary for landscape simulations when econometric estimation is derived from local tax assessor data. However, other sources of spatial GIS data must be used for landscape simulations when the land-use model is derived from NRI data.

---

6 The endogeneity of measures of neighboring development in land-use change models arises due to the presence of spatially-correlated unobservables that influence multiple development decisions. For example, suppose parcel A develops because its development rents are bid up due to its location next to a scenic bluff. If neighboring parcel B is observed to develop after A, and the presence of the scenic bluff is not accounted for, then the effects of A’s development status on B’s decision is confounded by the presence of the bluff (an unobservable that is spatially-correlated across both A’s and B’s development decisions).

7 In a binary land development model, Irwin and Bockstael (2002) argued that including variables indicating the amount of neighboring development would produce positively biased parameter estimates of such variables. Therefore, their estimate of a negative externality effect is bound from above.

8 For example, a traditional Logit model of the binary development decision requires an assumption that unobservable components of the net returns to land-use are independent across observations. Simulation-based estimators – such as Random-Parameters Logit – allow researchers to relax this independence assumption in a variety of ways.
Simulations of landscape change can be performed once the estimated transition probabilities are linked to points on a landscape. Early analyses treated the transition probabilities as deterministic rules by, for example, assuming that each parcel converts to the use with the highest estimated transition probability (Chomitz and Gray 1996; Irwin and Bockstael 2002). Although such an approach has the advantage of producing a single estimated landscape pattern, Train (2003) argues that interpreting choice probabilities as deterministic rules is “opposed to the meaning of probabilities and the purpose of specifying choice probabilities” (p.73). An alternative method is to simulate a large number of different landscapes that conform to the underlying probabilistic transition rules (Lewis and Plantinga 2007). Such simulations are performed by repeating the land-use decision for each parcel multiple times through the use of random number generators. For example, suppose a parcel’s estimated probability of converting from agriculture to development is 0.2, while its probability of remaining in agriculture is 0.8. The land-use decision can be simulated by drawing a random number \( r \) from a \( U \sim [0,1] \) distribution, where the parcel remains in agriculture if \( r \leq 0.8 \), and converts to development otherwise. Repeating this process \( S \) times for each parcel on the landscape generates \( S \) different spatial landscape patterns.

The repeated Monte Carlo simulation approach is theoretically consistent with the econometric specification, but produces many different landscape patterns that must be summarized with spatial indices or some other measure of landscape output. The measure of output chosen largely depends on the goals of the analysis, and past studies have calculated the following metrics for each simulated landscape: core forest and average forest patch size (Lewis and Plantinga 2007; Lewis et al. 2009b), environmental benefit indices (Newburn et al. 2006), watershed health indices (Langpap et al. 2008), and lakefront development density (Lewis et al. 2009a). The landscape simulations generate distributions of each output measure that can be used to estimate moments, such as the mean and variance.

In terms of explicitly integrating economic and ecological models, the two-step approach has the advantage of being able to link with any ecological model that uses landscape pattern to estimate ecological response – a common feature in the field of landscape ecology. For example, Nelson et al. (2008) and Lewis et al. (2009c) use the two-step landscape modeling approach to estimate explicit biodiversity indices, defined to represent population persistence probabilities over a set of 24 terrestrial species in the Willamette Basin of Oregon. Lohse et al. (2008) use a two-step model to forecast the spatial pattern of land development and its corresponding effects on fish populations in a California watershed conservation analysis. Lewis (2009) simulates lakeshore development patterns in northern Wisconsin and integrates an ecological population model that estimates the effect of development on the extinction probability of green frogs (figure 1b). These examples are notable for explicitly integrating behavioral land-use decision models with quantitative ecological models, and thus are a step in fulfilling the broader goals of quantitatively coupling human and natural systems.

**Policy Analysis with Landscape Models**

Contemporary land-use policy tools range from government or non-profit purchases of land and easements, other voluntary incentives such as found in the U.S. Farm Bill programs, and local regulatory approaches such as zoning. All these policies can be and are frequently used to alter the provision of the many environmental public goods that arise from landscapes. The two-step landscape models offer the ability to examine the landscape consequences of alternative policy scenarios. The policy link derives from the fact that the underlying econometric models
are functions of either net returns to land, or functions of regulatory policies such as zoning. As such, because the transition probabilities provide the foundation for landscape simulations, the effects of policy scenarios on landscape pattern can be examined. This last section looks at two example policy analyses conducted with the two-step landscape modeling framework.

The question of where to purchase land reserves when budgets are limited has occupied the conservation biology literature for at least 20 years. However, no analyses in conservation biology account for heterogeneity in both land costs and the probability of losing different types of undeveloped land to development. Newburn et al. (2006) devise a modeling framework that combines the two-step approach discussed above with a dynamic programming algorithm to optimally target conservation funds to increase the provision of environmental public goods at least cost. The analysis is conducted using tax assessor data in Sonoma County, CA, and demonstrates how the positive correlation between land costs and the probability of land-use conversion significantly alters optimal conservation targeting strategies.

Use of voluntary payments for altering land practices is now a widely discussed policy instrument. However, using voluntary payments to achieve spatial goals is challenged by the fact that the willingness of private landowners to accept a payment is private information, and thus, agencies cannot directly control landscape pattern through voluntary payments. Lewis et al. (2009a) discuss the theoretical issues that arise when such information problems combine with the fact that the marginal benefits of forest restoration are convex due to habitat fragmentation effects. In particular, their theoretical results show that conservation targeting might optimally involve corner solutions when deciding how much of particular regions to restore as forest. A two-step empirical landscape model is developed for the coastal plain of South Carolina to demonstrate the targeting solution and compare its results to other spatial targeting strategies. Private information on the willingness of landowners to accept conservation payments is empirically accounted for in the two-step model by treating land-use decisions in a probabilistic fashion and simulating multiple realizations of landscape change.

**Concluding Thoughts**

This paper has synthesized the recent literature pertaining to models of landscape change using a two-step approach that combines parcel-scale econometric models with GIS-based landscape simulations. Given the widely acknowledged importance of spatial landscape pattern for understanding the links between land-use, ecosystem services, and biodiversity, the two-step approach provides an empirical framework to integrate rigorous economic models with ecology. As demonstrated in multiple papers, such a framework provides a tool for resource economists to engage ecologists directly in analyzing policies to enhance the provision of environmental public goods from landscapes.

Future research in the spatial land-use modeling area can be enhanced in multiple ways. First, on the estimation side, further attention should be paid to causal identification strategies through greater use of quasi-experimental methods (e.g., Greenstone and Gayer 2009). For example, the hedonic pricing literature has recently made great strides with identification strategies based on techniques such as regression discontinuity (e.g., Chay and Greenstone 2005) and difference-in-differences (e.g., Horsch and Lewis 2009). In addition, because data available for estimation are often limited, continued refinement of micro-econometric approaches can improve the manner in which unobserved heterogeneity is accounted for in estimation.
Second, the resource economics profession should continue to exploit new spatial data sources, and in some cases, develop new datasets. Many advances that have been made in this literature in recent years have been made possible by enhanced data, especially increased availability of GIS land-use datasets, including increasing availability of local digital parcel data. Linking together multiple sources of local digitized parcel data would allow the geographic scaling up of many localized models. At the national scale, there are new challenges because the last year of available plot-level data from the NRI is 1997, and it is uncertain when additional plot-level data will be available to researchers in the future. Therefore, more work is needed to explore how other sources of nationally-consistent spatial data, such as the National Land Cover Dataset, could be exploited.

Finally, policy relevance for conservation can be strengthened through greater integration with quantitative ecology. The most promising integrations are for those ecosystem goods and services that are impacted by land-use change, such as wildlife conservation and water quality. A particularly challenging task is the development of fully coupled economic-ecological models, whereby estimation accounts for feedbacks between economic and ecological systems. Such an estimation task requires detailed historical data on land-use change, along with indicators of ecosystem goods and services. Development of such a fully-coupled analysis would represent a significant contribution of economics for land change science.

References


Figure 1. Example of Two-Step Modeling Framework – Shoreline Development across Northern Wisconsin Lakes (from Lewis 2009)

Figure 1.a Input to econometric model: spatial-panel data from local land information office and plat maps

Figure 1.b Output from landscape simulations: 20-year forecast of expected extinction probabilities (mean of 1000 simulations) for green frogs across 138 lakes