

**ASSESSING THE RELATIONSHIP BETWEEN CROP CHOICE AND LAND
USE CHANGE USING A MARKOV MODEL**

Wei Hua
Ohio State University
Contact: Hua.22@osu.edu

Diane Hite, Auburn University
Brent Sohngen, Ohio State University

Abstract

There is widespread concern among environmental and agricultural interest that land use change will affect the future productivity of the agricultural industry by utilizing highly productive land for development. This paper considers the links between land use change and crop choices in order to analyze whether land use change is influencing crop choices. In order to account for potential endogeneity between crop choices and land use choices, we develop a Markov Model that allows us to capture potential endogeneity between these two choices (land use and crop choice). The Markov model is developed for the 12 Midwestern U.S. States using USDA NRI data at the county level.

*Selected Paper prepared for presentation at the American Agricultural Economics
Association Annual Meeting, Providence, Rhode Island, July 24-27, 2005*

*Copyright 2004 by Wei Hua, Diane Hite, and Brent Sohngen. All rights reserved.
Readers may make verbatim copies of this document for non-commercial purposes by any
means, provided that this copyright notice appears on all such copies.*

Acknowledgement: Funding for this research has been obtained from the USDA National Research Initiative. The authors thank participants in research seminars at Ohio State University for their comments.

ASSESSING THE RELATIONSHIP BETWEEN CROP CHOICE AND LAND USE CHANGE USING A MARKOV MODEL

Introduction

Over the years, a number of authors have explored land use change decisions (Hardie and Parks, 1997; Plantinga, 1999; Choi et al., 2001; Hsieh, 2000; Bhattarai et al., 2004; Templeton, 2004), and many acreage response models have been estimated to analyze crop choice decisions (Wu and Adams, 2001; Wu and Segerson, 1995; Lichtenberg, 1989) among the main field crops like corn, soybeans, wheat, etc. While these papers provide important insights, they do not consider the interactions between land use change decisions and crop choice decisions, although it is likely that these decisions influence each other in different ways. On one hand, land use change and crop choice decisions are determined simultaneously. In each time period, a farm landowner may decide how much land to sell for development, while at the same time, he/she may decide how much land to grow crops, and how much land to grow livestock, and how much land for timber harvest, and more specifically, he/she may decide what kinds of crops to grow and allocate his/her land among all of those land use categories. Therefore, the shares of developed land and the shares of different crop choices are determined simultaneously. On the other hand, there are recursive relationships between the two decisions. For example, there is evidence that current crop choice decisions affect future soil productivity (Orazem and Miranowski, 1994) and that development uses the least productive land first when it occurs in rural areas (Hsieh et al., 2000; Hite et al., 2001, Choi et al., 2001). At the same time, land use changes in rural areas appear to influence crop choices as landowners adjust capital and labor usage and potentially shift remaining agricultural lands towards less labor intensive uses (i.e., conservation tillage),

or higher value crops. The changing crop choices may lead to different profitability of farming operations, which may then influence land conversion themselves.

This study models the crop choices and land use change decisions in one framework, assumes that crop choices are endogenous with the land use change decision, and seeks to identify how crop choices affect land use change decisions, and vice versa. In particular, this study begins to explore the process or histories of land use change, and answer the question: what succession of land uses ultimately leads to loss of agricultural land? This objective will be accomplished by building a Markov Chain model of cropping and land use decisions that focuses on major grain and vegetable crops in the Midwest, as well as forest, grazing, and urban uses. Using the Markov Chain model, equilibrium land use shares are predicted, which have policy implications for future land use planning.

Literature Review

There has been considerable research into the causes or determinants of land use change (Hardie and Parks, 1997; Plantinga et al., 1999; Choi et al., 2001; Hsieh, 2000; Bhattarai et al., 2004; Templeton, 2004). Most studies consider general land use categories: agriculture, forest, urban uses; while, some only focus on urban use (e.g. Hsieh, 2000; Templeton, 2004). There are also many studies on optimal timing of development (e.g. Hite, et al., 2003; Irwin and Bockstael, 2002). Besides economic research, there have been a number of geographic studies on land use change (e.g. Qihao Weng, 2002), whose analysis are done in dynamic framework. Their analysis, however, are statistically based, relying on satellite remote sensing techniques and Geographic

Information System (GIS), and do not explore the underlying economic reasons behind the phenomenon of land use change. Factors that have been found to influence land use change include land quality, land rents, population density, growth rates of population, per capita income, transportation, accessibility to urban centers, demographic characteristics such as age and education, and policy controls, etc.

Specifically, Plantinga (1999) uses a multinomial logit model and finds that population density does not have a significant effect on the ratio of agricultural to forestland, but it significantly influences the ratio of urban to forestland. Counties with lower average land quality and higher transportation costs tend to have less agricultural land relative to forestland. The results from a modified multinomial logit model estimated by Hardie and Parks (1997) indicate that population density and per capita income have a significantly negative influence on farmland and forestland, and share of high quality land is positively related to farmland, but negatively related to forestland. Landowner's age is positively related to forestland, but negatively related to farmland. The results from a spatial model by Choi et al., (2001) suggest that urbanization trends, measured by population density and distance from the nearest city, affect mainly the levels of change to urban from forest and agricultural land, but the decision between agriculture and forestland depends mainly on land rents and land quality. Using GIS data for west Georgia, Bhattarai et al. (2004) suggest that agricultural use is negatively related to per capita income but positively related to education level. Travel time to work has a positive effect on developed land use, but negatively affects agriculture and forestry.

While these papers provide important insights, they do not consider the interactions between land use change decisions and crop choice decisions, although it is

likely that these decisions influence each other in different ways. On one hand, land use change and crop choice decisions are determined simultaneously. In each time period, a farm landowner may make land allocation decisions among different land use categories: selling for development, growing crops, raising livestock, timber harvesting, and more specifically, what kinds of crops to grow. Therefore, the shares of developed land and the shares of different crop choices are determined simultaneously. On the other hand, there are recursive relationships between the two decisions. First, there is evidence that current crop choice decisions affect future soil productivity (Orazem and Miranowski, 1994). The resulting soil productivity could influence the course of land development, and potentially shift development towards the least productive land. There is evidence that when development occurs, it uses the least productive land among alternatives (Hsieh et al., 2000; Sohngen et al., 2001, Choi et al., 2001), suggesting that farmers with the least productive land are the first to sell for development. This is probably because it is cheaper to build on the land with least productivity, due to lower land prices. However, it may also be the case that less productive land is hillier, which could increase construction costs, thereby slowing urban development somewhat. Further, it is also possible that landowners with more productive land are more able to take advantage of new opportunities that arise as development occurs in once rural areas. Second, land use changes in rural areas appear to influence crop choices as landowners adjust capital and labor usage and potentially shift remaining agricultural lands towards less labor intensive uses (i.e., conservation tillage), or higher value crops. Besides, the land use conversion process rarely involves a discrete step from farmer-owner to a house or development project. Instead, it often involves a number of steps, from farmer-owner, to non-farm

owner (potentially a developer) who may rent the land, to developer, to urban use. Along this conversion pathway, land tenure issues could influence crop choices. The changing crop choices may lead to different levels of profitability of farming operations, which may then influence land conversion decisions.

Different analytical methods have been applied to study land use change. Some are performed in a static framework (Hardie and Parks, 1997; Plantinga et al., 1999; Choi et al., 2001; Hsieh, 2000; Bhattarai et al., 2004; Templeton, 2004), while others are dynamic studies on optimal timing of development (e.g. Hite, et al., 2003; Irwin and Bockstael, 2002). Besides economic research, there have been a number of geographic studies on land use change (e.g. Qihao Weng, 2002), whose analysis are done in dynamic framework. Their analysis, however, is statistically based, relying on satellite remote sensing techniques and Geographic Information System (GIS), and do not explore the underlying economic reasons behind the phenomenon of land use change. This study differs from the previous ones by investigating the process of land use change in a dynamic framework using the Markov Chain model, which can accommodate both the simultaneous and recursive relationships between land use change and crop choices.

The Markov Chain model has been applied in other areas of economics. Early economic applications have included describing and predicting structures of industry or market (Adelman, 1958; Ethridge et al., 1985; Mellor, 1984; Disney et al., 1988); modeling of consumer brand choice behavior (Telser 1962); and labor force analysis (Heckman and Willis, 1977). More recently, Markov modeling has been used in recreation demand (Haab 2001); financial applications (Elliott et al, 2001), and business cycle applications (Kontolemis, 2001), while Skaggs and Ghosh (1999) employ a Markov

analysis to investigate changes in soil erosion. Markov modeling has also been used in geography for land use change analysis, together with remote sensing and GIS technologies (e.g., Weng, 2002). However, there have been few economic applications of Markov modeling on land use change. The only known study is McMillen and McDonald (1991), who used individual tract data from Chicago to estimate a Markov Chain Model of land use zoning in an urban setting. This study differs from McMillan and McDonald by using a Markov Chain model with aggregate data at county level, which employs a different estimation process.

There are several merits of using the Markov framework to model land use change. First, the Markov Chain model can capture the state dependence the property of land use change process, or rigidities in the zoning process, i.e. the influence of the past on the present. For example, consider a county with 20% urban land, and 80% agriculture land. And suppose the land is worth more in urban use than in agricultural use. The entire county might eventually be converted to urban use in a competitive market, but the adjustment likely takes place over a number of time periods rather than immediately, due to zoning restrictions or because of speculative land hold up at the urban rural fringe, for example. Therefore, if a county has a higher percentage of land in agriculture use in any given year t , then one would expect a higher percentage of agriculture land in year $t+1$. The diagonal elements of the Markov model can measure the level of dependency between time periods. Second, the Markov Chain model can take care of the endogenous relationship between crop choice and land use change decisions. In the Markov model, a share of one kind of land use in the current period depends on not only its own share in the previous period, but also the shares of all the other kinds of land use in the previous

period. Those lagged shares, therefore, serve as instrumental variables, which can help eliminate endogeneity among land use choices within the same period. In addition, the lagged shares serve as instruments for analyzing the recursive relationship between land use change and crop choices. Third, with the Markov framework, a richer set of factors that affect land use change can be modeled, compared to static analysis using a multinomial logit model. Specifically, rather than simply analyzing change from agricultural to urban uses, we extend the literature by investigating intermediate land use states that precede the ultimate change to urban use. For example, we are able to infer how changes in proportions from owned to rented farmland affect the speed of urbanization.

Econometric Model

It is difficult to collect detailed data tracing the movement of individual tracts or farms among land use categories for the 12 states midwestern region, and only the aggregate proportion of land in each category is available. With aggregate data, the general Markov Model is given by

$$\begin{bmatrix} y_1^i(t) \\ y_2^i(t) \\ \dots \\ y_j^i(t) \\ y_J^i(t) \end{bmatrix} = \begin{bmatrix} P_{11}^i(t) & P_{21}^i(t) & \dots & P_{k1}^i(t) & \dots & P_{J1}^i(t) \\ P_{12}^i(t) & P_{22}^i(t) & \dots & P_{k2}^i(t) & \dots & P_{J2}^i(t) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ P_{1j}^i(t) & P_{2j}^i(t) & \dots & P_{kj}^i(t) & \dots & P_{Jj}^i(t) \\ P_{1J}^i(t) & P_{2J}^i(t) & \dots & P_{kJ}^i(t) & \dots & P_{JJ}^i(t) \end{bmatrix} \begin{bmatrix} y_1^i(t-1) \\ y_2^i(t-1) \\ \dots \\ y_j^i(t-1) \\ y_J^i(t-1) \end{bmatrix} + \begin{bmatrix} u_1^i(t) \\ u_2^i(t) \\ \dots \\ u_j^i(t) \\ u_J^i(t) \end{bmatrix} \quad (1)$$

or simply

$$(2) \quad y_j^i(t) = \sum_k y_k^i(t-1)P_{kj}^i(t) + u_j^i(t), \quad j = 1, \dots, J$$

where $y_j^i(t)$ and $y_j^i(t-1)$ are the proportions of land in county i in use j at time t and at time $t-1$ respectively. Generally, there are J kinds of land use categories. $P_{kj}^i(t)$ is the transition probability that a parcel of land in the i^{th} county is in use j at time t , given that the same parcel is in use k at time $t-1$, and $u_j^i(t)$ is the associated error term. The matrix of $P_{kj}^i(t)$ terms represents the Markov matrix. Two restrictions need to be satisfied in estimating the Markov model:

R1: Every element of $P^i(t)$ is bounded by zero and one, i.e., $0 \leq P_{kj}^i(t) \leq 1, \forall k, j$

R2: Each column of $P^i(t)$ must sum to one, i.e., $\sum_{j=1}^J P_{kj}^i(t) = 1, \forall k$.

Intuitively, the model states that the proportion of land in a given use at time t depends on its use at time $t-1$, and on the transitional probabilities given by the Markov matrix. The transitional probabilities may depend on various characteristics of a county; let $X^i(t)$ be the vector comprising these explanatory variables. Therefore, $P_{kj}^i(t)$ is a function of $X^i(t)$. In many previous Markov studies, $P_{kj}^i(t)$ has been assumed to be a linear function of $X^i(t)$ (Mellor, 1984; Disney et al., 1988; McMillen and McDonald, 1991). Mellor (1984) and Disney et al. (1988) use linear regression, while McMillen and McDonald (1991) use probit regression to obtain parameter estimates. However, their model estimation are all based on linear probability assumption, which cannot guarantee that the predicted probabilities satisfy the first restriction (R1), not to say the second one. In this study, it is assumed that $P_{kj}^i(t)$ is a multinomial logit function of $X^i(t)$, as in MacRae (1977), such that

$$(3) \quad P_{kj}^i(t) = \frac{\exp(X^i(t)\beta_{kj})}{1 + \sum_{m=1}^{J-1} \exp(X^i(t)\beta_{km})}, \quad k=1, \dots, J \text{ and } j=1, \dots, J-1$$

and

$$(4) \quad P_{kj}^i(t) = \frac{1}{1 + \sum_{m=1}^{J-1} \exp(X^i(t)\beta_{km})} \quad k=1, \dots, J \text{ and } j=J$$

β_{kj} is a vector of unobserved parameters to be estimated, β_{kJ} is normalized to zero (Green 2000, 859-860). The logistic specification restricts the transitional probabilities to satisfy the two conditions R1 and R2.

Plugging equations (3) and (4) into (2), then equation (2) becomes

$$(5) \quad y_j^i(t) = F(y_k^i(t-1), X^i(t), \beta_{kj}) + u_j^i(t), \quad j=1, \dots, J.$$

where $y_j^i(t)$ is now a nonlinear function of $y_j^i(t-1)$, $X^i(t)$, and β_{kj} . Since $\sum_{j=1}^J P_{kj}^i(t) = 1$

for all k , one only need to estimate the first $J-1$ equations of (5). Suppose $X^i(t)$ is a

$1 \times K$ vector, then the number of parameters that need to be estimated is

$(J-1) \times J \times (K+1)$, which requires a large amount of computational operations. Using

Nonlinear Least Square iteration on the first $J-1$ equations of (5), one can obtain the

parameter estimates $\hat{\beta}_{kj}$. In this study, we use NRI data, which is survey data suffering

from sampling error. According to MacRae (1977), least squares estimation applied to

imperfectly observed data will give inconsistent estimates, which can be remedied by

using instrumental variable estimation or limited information maximum likelihood

estimation. But as MacRae (1977) also mentions, the inconsistency problem may not be

serious if the sample size is large. NRI data consist of several hundreds of sample plots in

each county, and for now, we assume that the sample size is large enough for the inconsistency to be a problem. Later, consistent estimators will be employed in the model. The obtained parameter estimates $\hat{\beta}_{kj}$ can be plugged back into (3) and (4) to get the predicted transition probabilities \hat{P}_{kj} for a representative observation. The matrix of $\hat{P}_{kj}^i(t)$ represents the one-period transition probability matrix P . The t-period transition probability matrices Q_t are simply the matrix powers of P . In equilibrium, the multiperiod transition matrices Q_t will converge to a matrix Q as t goes to infinity. In that case, all rows of Q are identical, representing the steady-state distribution, which in this case are the equilibrium land use shares. Policy simulation can also be performed with the predicted Markov matrices, and the marginal effects of explanatory variables x_m on p_{ik} (Greene, p861), where

$$(6) \quad \frac{\partial p_{ik}}{\partial x_m} = p_{ik} \left(\beta_{ikm} - \sum_{j=1}^J p_{ij} \beta_{ijm} \right)$$

Data

The study area in this paper is the 12 Midwest states (Ohio, Indiana, Illinois, Michigan, Minnesota, Wisconsin, Iowa, Missouri, Kansas, Nebraska, North Dakota, South Dakota). See Appendix I for a map of this region. A total of 1,054 counties are examined. This region is interesting because it is quickly urbanizing, but still has significant agricultural production.

Land Use Shares

Data used in this study are obtained from various sources. In particular, the county level land use proportions are obtained from National Resource Inventory (NRI)

database for the years 1982, 1987, 1992 and 1997. The NRI has been conducted by the Natural Resource Conservation Service of the U.S. Department of Agriculture every five years since 1982. It is a scientifically based, longitudinal panel survey that contains information on nearly 800,000 sample sites across the continental United States (USDA, 2000). Estimates from those sample sites are valid for only when aggregated to the county level, which we thus use in this study. A range of ten land management categories are examined in this study: high value crops, row crops (such as corn, soybeans), close grown crops (such as wheat, oats), hay land, other cropland, grazing land, forest, CRP land, other rural land, and developed land. The changes in the acreage of the ten land use categories are shown in table 1. The definitions of the land use categories based on 1997 NRI are as follows: 1) high value crops include fruit, nuts, vineyard, bush fruit, berries, flowers, vegetable and truck crops including melons; 2) row crops include corn, sorghum, soybeans, cotton, peanuts, tobacco, sugar beets, potatoes, all other row crops, sunflowers. Vegetable and truck crops are defined as row crops in the 1997 NRI, but in this study, they are grouped together with horticulture crops into the high value crops category. 3) Close grown crops include wheat, oats, rice, barley, and all other close grown crops. 4) Hay land includes grass, legumes, and mix of legumes and grass. 5) Other cropland includes summer fallow, aquaculture in a crop rotation, and other cropland not planted. 6) Grazing land include pastureland, rangeland, and grazed forestland. 7) Forest is defined as ungrazed forest land. 8) Other rural land includes farmsteads and ranch headquarters, other land in farms not associated with farmsteads¹, barren land, permanent snow and ice fields, marshland, and all other land. 9) Conservation reserve program (CRP) land is

¹ Includes agroforestry and tree plantations, airplane landing strips, commercial feedlots, duck farms, field windbreaks, greenhouses, hog facilities, mink farms, mushroom farms, nurseries, poultry facilities. (1997 NRI)

separated out from other rural land as a single land use category, because although both CRP land and other rural land as defined in category eight are out of production land, different patterns of land use change may occur between CRP land and other rural land. From table 1, we can see that over the years, other rural land has been decreasing and CRP land has been increasing. 10) Developed land includes urban and built-up areas and rural transportation land. Table 1 indicates that there has been a 20% increase in developed land in the midwestern region between 1982 and 1997. Although the total acreage of high value crops is relatively small compared to other land use categories, it almost doubled during the same period. There has been an increase in row crops land, hay land, forestland, but a decrease in close grown crops, other cropland, graze land, and other rural land.

Land Use Rents

Land use rents are important land use change determinants. In this study, population density and population growth are used as a proxy for urban land values. Different approaches have been used in previous literature to estimate agricultural rents. Hardie and Parks (1997) use farm revenue and costs as two variables separately; Lichtenberg (1989) uses agricultural commodity prices; Plantinga et al. (1999) uses present discounted value of the stream of annual net revenues, while Choi et al. (2001) use annual revenue above variable cost. In this study, the following formula is used to estimate crop rents from county level data:

$$(7) \text{ Annual crop rent (\$/acre)} = [\text{Price (\$/unit)} - \text{variable cost (\$/unit)}] \times \text{yield (unit/acre)}$$

County level crop yield data and state level crop prices data are obtained from the USDA Agricultural Statistics Data Base (USDA-NASS). Regional level variable cost data are

from USDA Economic Research Service commodity costs and returns dataset (USDA-ERS). Because ERS employs different regional definitions and different measures of production costs before and after 1995, adjustments are made in order to make the measure of variable cost as consistent as possible across the four years. Crop rents for corn, soybeans, wheat, and oats are estimated for each county based on formula (7). Since on average, corn and soybeans account for 86% of land in row crops in the study region, county level estimates of row crop rents are then determined by weighting the rents for corn and soybeans by the number of acres in the two crops in the county for each period. Similarly, because wheat and oats occupy 91% of land in close grown crops, weighted crop rent between wheat and oats is used as the close grown crop rent for each county.

Annual forest rents is estimated as the annual share of discounted net present value of timber revenue per acre, using data from USDA Forest Inventory Mapmaker Version 1.7. Net present value of timber revenue per acre is obtained with the Faustmann formula (Johansson and Logfren, 1985):
$$NPV = \frac{p \times f(a) \times \exp(-ra) - c}{1 - \exp(-ra)},$$

where a is rotation age, p is timber price, r is interest rate, c is harvest cost. Annual forest rent is then the annual share of NPV ($r \times NPV$).

Urban land values are measured by population density (POP_DEN) and population growth rate (POP_GROWTH). Population density is measured as population per acre at county level. County level population estimates are from the Bureau of Economic Analysis (USDC – BEA), and the total area of each county is from NRI data. Other measurements of urbanization level will be included later, for example, distance from each county to the nearest major city will be calculated from ZIPFIP developed by USDA.

Other Variables

Data for soil characteristics are from NRI data. Average land capability class (ALCC) for each county is estimated to measure average land quality. The lower the value is, the higher the average land quality. The share of land within the first two capability classes (LCC_12) is also included in the regressions. Other characteristics such as water holding capacity, soil permeability can be obtained by linking NRI to the SOIL5 database developed by the NRCS (Tanaka and Wu, 2004).

Climate data are included to control for ecological reality, such as unfitness of some areas for forestry. County level seasonal precipitation and temperature data are obtained from Dr. Brent Sohngen.

Capital investment intensity is measured by two variables, one is market value of buildings and land per acre (BLDGAC), the other is market value of machines and equipment per acre (MACHAC). The cattle inventory (CATTLE_N) is included as a measure of livestock operation size. Share of owned farmland (OWN_PER) is included as a measure of land tenure. Data for all these variables are from Census of Agriculture (USDA- NASS).

Other data will need to be collected as follows: County level unemployment rate will be obtained from the Bureau of Economic Analysis; tillage practice data will be obtained from Conservation Technology Information Center (CTIC), which will be studied in combination with crop choice later. Non-agricultural wage rates need to be collected to replace the general wage rate, in order to get a better measurement of opportunity cost of working on the farm. The full data set is based on four years of observations at county level—1982, 1987, 1992, and 1997. All monetary values were

deflated for base year 2002. Specifically, current dollars were converted into constant dollars by using the formula:

$$(8) \quad \text{Year } Z \text{ constant dollar value} = \text{Year } Z \text{ current dollar value} \times \left(\frac{IPD_B}{IPD_Z} \right)$$

, where IPD_B is base year IPD number and IPD_Z is year Z IPD number.²

Based on the data currently collected, the statistics of the variables used in the regressions are shown in table 4.

Empirical Results and Discussion

In order to check estimation results, a nonparametric Markov transition matrix is

estimated based on statistical frequency, with $p_{ik} = \frac{n_{ik}}{\sum_{j=1}^J n_{ij}}$, where p_{ik} is the transition

probability of land use change from i to k , and n_{ik} is the number of transition from land use i to land use k (Table 2).

For the sake of simplicity, we illustrate the full econometric model based on four land use categories: (1) agricultural land (2) forest land (3) developed land, and (4) other land. The average share of each land use is displayed in table 3. We can see that agricultural use is still the predominant use in the midwestern region, but its average land use share decreases by 5% between 1982 and 1997. Forestland increases by about 1%. Developed land increases by 1.2%, and other land uses increase by 3.2%. The descriptive statistics of variables used in the analysis are listed in table 4.

² Implicit price deflator (IPD), or GNP deflators, for year 1982, 1987, 1992, and 1997 are calculated based on quarterly IPD published by the BEA in March 2004 on the basis year of 2000.

Only three rows of the Markov matrix need to be estimated since $\sum_{j=1}^4 P_{kj}(t) = 1$.

Using the estimated model system from Equation (5), the predicted one-period transition probability matrix evaluated at the sample mean of explanatory variables is presented in table 5. As expected, the diagonal elements are the largest elements in the Markov matrix. For a representative county, the probability of land use change from agriculture to forest is 0.0032, and the probability of land use change from agriculture to developed land is 0.0036, implying that agricultural land is slightly more likely to shift to developed land than to forest land. The probability of land use change from forest to agriculture, and from forest to developed land is 0.0003 and 0.0001 respectively, implying that forest land tends to more likely to shift to agricultural land than to developed land. The probability of land use change from developed land to agriculture is 0.0036. The probability of land use change from developed land to forest and other uses are both zero. If we want to retrieve the process of land use change, i.e. examine the way agricultural land eventually converge to developed land, first look at the third line of the transition matrix, which represents the conditional probabilities that land use changes to developed land. Except p_{33} , the largest element is p_{43} , which implies that the “other uses” category is more likely to convert to developed land than agriculture land and forestland. Then we look at the fourth line. Except p_{44} , the largest element is p_{14} , which implies that agriculture land is more likely to convert to “other uses” than forestland and developed land. Based on those steps, one can infer roughly the land use change process, i.e. from agricultural land to other uses, and then to developed uses. In order to understand what kind of agricultural

land are more likely to be converted to “other uses”, I will subdivide the agricultural land into detailed crop choice categories.

The equilibrium land use shares are $\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = \begin{bmatrix} 0.0965 \\ 0.5330 \\ 0.2594 \\ 0.1111 \end{bmatrix}$, implying that in the long

run, a representative county will have 53.3% of forestland, 25.94% of developed land, 11.11% of other land, and only 9.65% of agricultural land. This is not surprising, because from the Markov matrix, we see that although the diagonal elements are the largest elements, p_{11} , the probability of remaining in agriculture, is the smallest one among the diagonal elements. In other words, agricultural land is most vulnerable, and more likely to change to other land use categories. It is important to note that the above Markov matrix and the equilibrium land use shares are obtained at the means of explanatory variables, and they can vary by county. In other words, different characteristics of different counties will lead to different transition probabilities and equilibrium land use shares. The marginal effect of explanatory variables on the transition probabilities will be calculated, which will yield more interesting results. For example, we would expect that higher population density would lead to higher p_{33} .

Conclusions and Further Research

In the demonstration example, we develop a Markov model for general land use categories, agricultural, forest, and developed land, and other land uses. By adding more detailed land management categories within agricultural land, we can take a further step to analyze the influence of crop choice on land use change, which kind of agricultural

land management is more likely to be converted to non-agricultural land, and finally to developed land. In addition, the study can be expanded by incorporate tillage information into crop choice data. Tillage practice data will be obtained from Conservation Technology Information Center (CTIC). By combining crop choice and tillage information, we can explore more in detail how crop choice and tillage practices together influence land use change process.

More explanatory variables need to be collected and incorporated into the model. Crop rent will be estimated with crop budget data. Some policy variables, such as government payments, need to be added to the model, in order to make policy simulation analysis. For example, by considering the different magnitudes of marginal effect of government payment on the transition probability, and the different magnitudes of transition probabilities themselves, we can compare the efficiency of government payments to different kinds of operations. For example, if the land use transition probability from corn and soybeans to developed land is P_{cd} , and the marginal effect of government payment on P_{cd} is M_{cd} . We expect M_{cd} is negative, so that increasing government payment to farmers may reduce the probability that agricultural land being converted to developed land. Similarly, suppose the land use transition probability from wheat to developed land is P_w , and the marginal effect of government payment on P_w is M_w . If the absolute value of M_{cd} is bigger than that of M_w , that means the transition probability P_{cd} is more sensitive to government payment than P_w . Combined with the magnitude of transition probability itself and the land use shares, the model may help guide policy makers in allocation funds. For example, given limited resources, we may

be able infer which kinds of farmers should be funded first in order to preserve agricultural land.

Some restrictions can be imposed on the parameters to make the model fit reality. For example, we can restrict the transitional probabilities from developed land to other land use categories as zero to reflect the land irreversibility. This will help to make any policy predictions more useful.

Table 1. Changes in total acreage of the Ten Land Use Categories in Midwestern U.S. (100 acres)

Variables	1982	1987	1992	1997	Percent change (97-82)
High Value Crop	10071	11967	12460	19514	94%
Row Crops	1243022	1199857	1244607	1296549	4%
Close Grown Crops	579670	503561	464510	434007	-25%
Hayland	261791	277196	263594	274492	5%
Other Cropland	160091	244528	150928	111684	-30%
Grazeland	1281334	1235040	1209399	1179322	-8%
Forestland	669252	678240	683826	697702	4%
Other Rural Land	150608	149014	148838	147343	-2%
CRP land	0	53138	167502	158814	--
Developed Land	220083	230965	243921	265183	20%

Sources: 1997 NRI

Table 2. Land Use Change Frequencies

Land Use Categories	Lag Land Use Categories									
	High Vaue Crops	Row Crops	Close Grown Crops	Hay Land	Other Cropland	Grazing Land	Forest	Other Rural Land	CRP	Developed Land
High Vaue Crops	59.36	0.35	0.43	0.20	0.18	0.01	0.01	0.06	0.02	0.00
Row Crops	26.23	79.85	26.53	24.30	19.25	1.56	0.11	1.21	4.87	0.12
Close Grown Crops	3.33	8.56	48.72	8.39	41.19	0.69	0.02	0.35	0.66	0.02
Hay land	4.82	5.18	4.96	56.79	3.92	1.56	0.06	0.47	1.37	0.02
Other Cropland	1.09	2.71	14.41	1.97	27.62	0.31	0.01	0.14	0.17	0.01
Grazing Land	1.95	1.03	1.16	4.98	1.73	93.85	0.56	1.29	1.96	0.10
Forest	0.64	0.03	0.01	0.30	0.26	1.20	98.44	2.27	0.03	0.07
Other Rural Land	1.09	0.18	0.15	0.32	0.41	0.21	0.15	93.70	0.03	0.00
CRP	0.41	1.72	3.47	2.23	5.22	0.35	0.01	0.11	90.88	0.00
Developed Land	1.09	0.39	0.16	0.51	0.23	0.27	0.62	0.39	0.02	99.65
Total	100	100	100	100	100	100	100	100	100	100

*All numbers in the table are in percentage (%).

Table 3. Average Share of Each Land Use

Year	Agriculture	Forest	Developed	Other Uses
1982	71.2%	15.5%	5.5%	7.8%
1987	69.5%	15.7%	5.8%	8.9%
1992	66.8%	15.9%	6.2%	11.1%
1997	66.1%	16.2%	6.7%	11.0%

Table 4. Descriptive Statistics of Variables Used In The Analysis

Variable	Mean	Std. Dev.	Description
Y1	0.675	0.244	Share of land in agricultural use
Y2	0.159	0.192	Share of land in forest use
Y3	0.062	0.091	Share of land in urban use
Y4	0.103	0.083	Share of land in other uses
LAGY1	0.692	0.245	Lag value of agricultural land use share
LAGY2	0.157	0.191	Lag value of forest land use share
LAGY3	0.058	0.085	Lag value of developed land use share
LAGY4	0.093	0.083	Lag value of other land use share
ALCC	3.322	0.936	Average land capability class
LCC_12	0.441	0.253	Share of land in LCC I and II (%)
OWN_PER	0.572	0.147	Share of owned farmland (%)
POP_DEN	0.162	0.511	Population density (persons/acre)
POP_GROWTH	-0.001	0.061	Population growth rate
BLDGAC	0.973	0.639	Average market value of land and buildings per acre (\$1000)
MACHAC	54.837	22.513	Average market value of all machinery/equipment per farm (\$1000)
ROWCROP_R	169.953	61.632	Land rent for row crops
CLOSECROP_R	77.508	33.107	Land rent for close grown crops
CATTLE_N	37.507	33.295	Cattle inventory (1000)

Table 5. Predicted One-period Transition Probability Matrix*

Land Use Categories	Lag Land Use Categories			
	Agriculture	Forest	Developed	Other Uses
Agriculture	0.9806	0.0003	0.0036	0.0070
Forest	0.0032	0.9994	0.0000	0.0001
Developed	0.0036	0.0001	0.9964	0.0048
Other Uses	0.0126	0.0002	0.0000	0.9881

*Evaluated at the means of explanatory variables.

Reference:

- Adelman, Irma G. "A Stochastic Analysis of the Size Distribution of Firms." *Journal of the American Statistical Association*. 284 (Dec., 1958): 893-904.
- Bhattarai, Gandhi, Upton Hatch, and Daowei Zhang. "Socioeconomic Influences on Land Use Distribution at Watershed Level: A Multinomial Logit Analysis of Land Use Distribution in West Georgia." AAEA Annual Meeting, 2004, Denver, Colorado
- Choi, Suk-won, Brent Sohngen, and Ralph Alig. "Land-Use Change and Carbon Sequestration in the Forests of Ohio, Indiana, and Illinois: Sensitivity to Population and Model Choice." AAEA Annual Meeting, 2001, Chicago, IL
- Disney, W. Terry, Patricia A. Duffy, and William E. Hardy, Jr. "A Markov Chain Analysis of Pork Farm Size Distributions in the South." *Southern Journal of Agricultural Economics* (December 1988): 57-64
- Elliott, Robert J., William C. Hunter and Barbara M. Jamieson. "Financial Signal Processing: A Self-Calibrating Model." *International Journal of Theoretical and Applied Finance* 4(4): 567-584.
- Ethridge, Don E., Sujit K. Roy, and David W. Myers. "A Markov Chain Analysis of Structural Changes in the Texas High Plain Cotton Ginning Industry." *Southern Journal of Agricultural Economics* (December 1985): 11-20
- Greene, William H. *Econometric Analysis*. Fourth Edition. 2000 Prentice Hall International, Inc.
- Haab, T.C. "Temporal correlation in recreation demand models with limited data." *Journal of Environmental Economics and Management* 45 (2003): 195-212
- Hardie, Ian W., and Peter J. Parks. "Land Use with Heterogeneous Land Quality: An Application of an Area Base Model." *Amer. J. Agr. Econ* 79 (May 1997):299-310.
- Heckman, James J., and Robert J. Willis. "A Beta-logistic Model for the Analysis of Sequential Labor Force Participation by Married Women." *The Journal of Political Economy* 85 (1) (Feb., 1977): 27-58
- Hite, Diane, Brent Sohngen, and Josh Templeton. "Zoning, Development Timing, and Agricultural Land Use at the Suburban Fringe: A Competing Risks Approach." *Amer. J. Agr. Econ* 76 (April 2003): 145-157.
- Hsieh, Wen-hua, Elena G. Irwin, and Lynn Forester. "Spatial Dependence among County-Level Land Use Changes." AAEA Annual Meeting, 2000, Tampa, FL

Irwin, E. G., and N. E. Bockstael. (2002) "Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Patterns." *Journal of Economic Geography*, 2(1): 31-54.

MacRae, Elizabeth Chase. "Estimation of Time-Varying Markov Processes with Aggregate Data." *Econometrica*, 45(Jan., 1977): 183-198.

McMillen, Daniel P., and John F. McDonald. "A Markov Chain Model of Zoning Change." *Journal of Urban Economics* 30 (1991): 257-270

Mellor., C.J. "An Application and Extension of the Markov Chain Model to Cereal Production." *Journal of Agricultural Economics* 35(2) (May 1984): 203-217

Orazem, Peter F., and John A. Miranowski. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *Amer. J. Agr. Econ* 76 (Aug 1994): 385-395.

Plantinga, Andrew J., Thomas Mauldin, and Douglas J. Miller. "An Economic Analysis of The Costs of Sequestering Carbon in Forests." *Amer. J. Agr. Econ* 81 (Nov 1999): 812-824.

Qihao Weng. "Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling." *Journal of Environmental Management* 64 (2002): 273-284

Skaggs, Rhonda, and Soumen Ghosh. "Assessing Changes in Soil Erosion Rates: A Markov Chain Analysis." *Journal of Agricultural and Applied Economics* 31 (December 1999): 611-622

Sohngen, Brent, Diane Hite, and Josh Templeton. "Land Use Change and Property Taxes: An Empirical Study of the Effect of Property Taxes on the Timing of Land Conservation from Agriculture to Development." AAEA Annual Meeting, 2001, Chicago, IL

Tanaka, Katsuya, and Junjie Wu. "An Evaluation of Conservation Policies for Reducing Nitrogen Loads to the Mississippi River and Gulf of Mexico." AAEA Annual Meeting, 2004, Denver, Colorado

Telser, Lester G. "The Demand for Branded Goods as Estimated from Consumer Panel Data." *Review of Economics and Statistics* 44(3) (Aug., 1962): 300-324

Templeton, Scott R. "Demographic, Economic, and Political Determinants of Land Development in the U.S." AAEA Annual Meeting, 2004, Denver, Colorado

USDA -- ERS: US Department of Agriculture -- Economic Research Service, Commodity costs and returns data. <http://www.ers.usda.gov/data/costsandreturns/testpick.htm>

USDA – NASS: US Department of Agriculture – National Agricultural Statistical Service Census of Agriculture, 1982-1997. <http://www.nass.usda.gov:81/ipedb/>

USDA-NRI: US Department of Agriculture. National Resource Inventory Data (NRI), 1982-1997.

USDC – BEA: US Department of Commerce – Bureau of Economic Analysis
<http://www.bea.gov/beahome.html>