ELSEVIER

Agricultural Economics 27 (2002) 317–332

AGRICULTURAL ECONOMICS

www.elsevier.com/locate/agecon

Temporal and spatial modelling of tropical deforestation: a survival analysis linking satellite and household survey data

Colin Vance [a,*], Jacqueline Geoghegan [b]

[a] *US EPA National Centre for Environmental Economics, Washington, DC 20460, USA*
[b] *Department of Economics, Clark University, Worcester, MA 01610, USA*

Abstract

We estimate a spatially explicit model of the forest clearance process among smallholder farmers in an agricultural frontier of southern Mexico. Our analysis takes as its point of departure a simple utility-maximising model that suggests many possible determinants of deforestation in an economic environment characterised by missing or thin markets. Hypotheses from the model are tested on a data set that combines a time series of satellite imagery with data collected from a survey of farm households whose agricultural plots were geo-referenced using a global positioning system (GPS). We implement a survival analysis to identify the effect of household level explanatory variables on the probability of deforestation. This approach allows us to introduce a measure of the time until clearance as a covariate, thereby affording a control for the effect of potentially important explanatory variables that vary through time but are not directly observable. In addition to identifying several variables relevant for policy analysis, including household demographics, proximity to roads, and government provision of agricultural support, model results suggest that the deforestation process is characterised by non-linear duration dependence, with the probability of forest clearance first decreasing and then increasing with the passage of time.
Published by Elsevier Science B.V.

JEL classification: Q15; R14; C41

Keywords: Deforestation; Survival analysis; Agricultural households

1. Introduction

Tropical deforestation plays a central role in many of the most acute environmental threats of our time, including global climate change, habitat degradation, and unprecedented species extinction. Scientific and public concerns about these and other potentially massive ecological disruptions have incited a growing number of studies that aim to quantify the social and biophysical determinants of deforestation processes, as well as their interactions over time and space. An emerging methodological approach to these issues combines high-resolution satellite imagery, geographic information systems (GIS), and socio-economic and geophysical data to model the human–environment interactions that drive land-use change (e.g. Liverman et al., 1998). In the economics literature, the primary focus of this research has been on identifying the socio-economic forces that explain the spatial patterns of landscape development, but less attention has been given to capturing the temporal dynamics from which these patterns emerge. To the extent that both the location and timing of forest clearance matter for assessing environmental outcomes, this de-coupling of spatial and temporal dimensions compromises the implementation of appropriate

* Corresponding author.

0169-5150/02/$ – see front matter. Published by Elsevier Science B.V.
PII: S0169-5150(02)00074-9

policy responses to deforestation. Accordingly, the purpose of the present paper is to advance an empirical methodology that supports analysis of how, over time and space, individual land managers respond to changing economic and ecological conditions.

Our study focuses on land-use change in an agricultural frontier spanning the southern Mexican states of Campeche and Quintana Roo, a region that contains one of the largest and oldest expanses of tropical forests in the America outside of Amazonia. Over the past 30 years, these forests have been under sustained pressure following the construction of a highway in 1972 that opened the frontier to settlement. The road was part of a larger development effort to promote agricultural colonisation and has contributed to a prolonged period of land transformation that has been captured by Thematic Mapper (TM) satellite imagery. We model these landscape dynamics by assembling a spatial database that links the pixels from three TM images spanning the years 1986–1997 with a random sample of farm households whose agricultural plots were geo-referenced using a global positioning system (GPS).[1]

Following a brief overview of the study region, our analysis takes as its point of departure a simple utility-maximising model that suggests many possible determinants of forest clearance in an economic environment characterised by missing or thin markets, as typifies frontier regions in the nascent stages of economic development. We subsequently test the significance of these determinants using survival analysis, also known as duration analysis or hazard modelling, a statistical technique that estimates the instantaneous probability of a transition between two states—in this case land-use states—conditional on the time elapsed until the occurrence of the transition. The final two sections of the paper discuss policy implications and suggest extensions to the current analysis for future research.

2. The region

The southern Yucatán peninsular region occupies roughly 22,000 km^2 of southwestern Quintana Roo and southeastern Campeche, north of the Mexican–Guatemala border (Fig. 1). A rolling karstic terrain of semi-deciduous tropical forests covers the landscape, with elevations in the centre reaching a peak of about 250–300 m. The zone corresponds to what was once a portion of the Maya lowlands, and was nearly completely deforested 1000 years ago during the Classic Period of Lowland Maya domination (100–900 A.D.; Turner, 1983). Following the collapse of Maya civilisation in 800–1000 A.D., the region experienced a period largely free of settlement that, continuing past the birth of the Mexican nation state in 1821, allowed the return of the forests. By the first half of the 20th century, human intervention here re-emerged but was primarily limited to the selective logging of tropical woods, particularly mahogany (*Swietenia macrophyla*) and cedar (*Cedrela odorata*), as well as the extraction of chicle, a tree resin (from *Manilkara zapota*) used in the production of chewing gum.

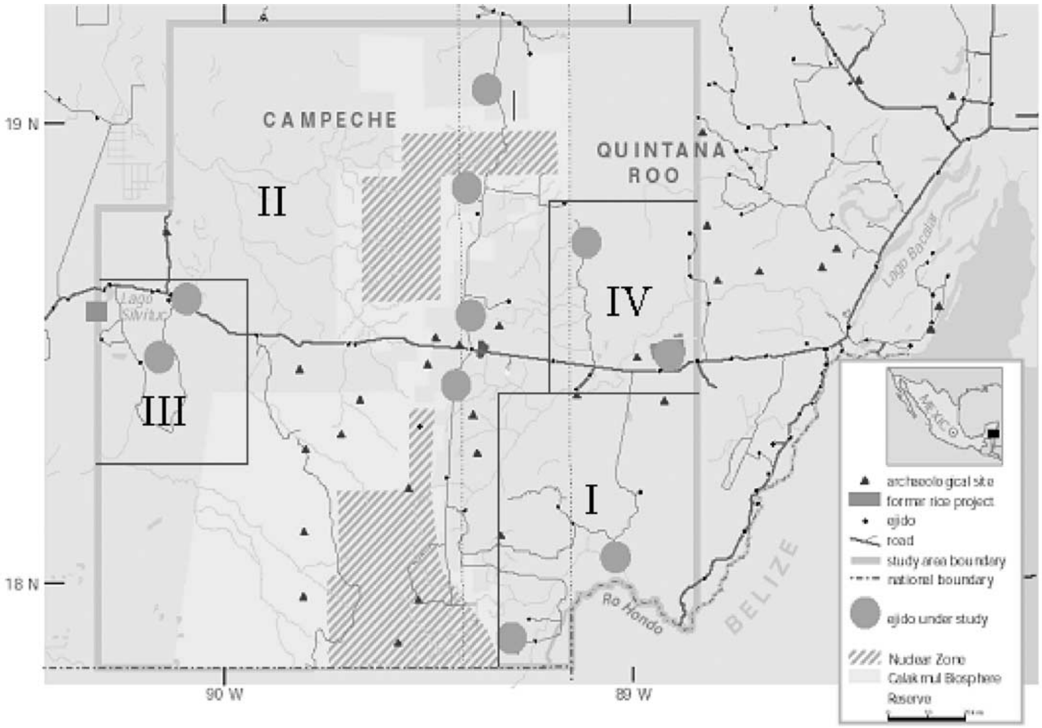
More extensive deforestation followed with the construction of a two-lane highway across the centre of the region in 1972, which opened the frontier to agricultural colonisation primarily via the extension of *ejido* land grants from the federal government. The *ejido* sector was created following the Mexican Revolution (1910–1917), a political and social upheaval with roots in inequitable land distribution. Within *ejido* communities, land is communally regulated by an elected committee, but in this area of southern Mexico, *ejido* members (*ejidatarios*) typically enjoy usufruct access to a single parcel that is permanently allocated to their use.[2] The size of these parcels varies considerably, ranging from 10 to 350 ha, with an average size in the sample of roughly 110 ha.[3] Our study focuses specifically on the *ejido* sector as it historically has been the predominate form of land tenure in the study region.

Most agriculture within *ejidos* is based on a traditional slash-and-burn or swidden system of temporary cultivation and continuous rotation through forest

[1] A pixel refers to the unit of spatial resolution, or area on the ground, in a remotely sensed image. For TM data, pixel size is a square with length of 28.5 m.

[2] Roughly 7% of households had access to multiple, non-contiguous plots. In the majority of these cases, cultivation occurred on only one of these plots for the year of questioning.

[3] By the time the majority of ejidos were established in the 1970s, commercial logging interests had abandoned most tracts following the removal of valuable tree species. The resulting relative abundance of land made the region particularly attractive to colonists who migrated from regions of the country where there was land scarcity, most notably Veracruz, Tabasco, and Chiapas.



Adapted from: Evaluation of the Conservation of Selva Maya. Wildlife Conservation Society, 1995

Fig. 1. The Southern Yucatán Peninsular region study area. Note: The thick gray line represents the boundaries of the general region of study, comprising about 22,000 km^2. Roman numerals indicate the approximate locations of the TM satellite zones listed in Table 2. Reprinted from Land Use Policy, Vol. 18, P. Klepeis and B.L. Turner II, Integrated Land History and Global Change Science: The Example of the Southern Yucatán Peninsular Region Project, pp. 27–39, Copyright 2001, by permission of Elsevier Science.

fallow (Ewell and Merrill-Sands, 1987). The system, referred to locally by the Maya word *milpa*, is dominated by maize but is often inter-cropped with squash and beans. In recent years, a growing number of farmers have additionally introduced commercial crops, most notably chili, and pasture into their land-use portfolios. While the incorporation of such uses has increased the complexity and variation of swidden practices, earlier research in the region by Turner (1983) suggests that farmers employ a plot-fallow rotation converting to a roughly 1:3–4 ratio, cycles usually predicated on 3 years of cultivation and 9–12 years of fallow (Klepeis et al., 2003). Although the field survey conducted for the present study was unable to confirm whether such a rotation prevails today, analysis of the satellite imagery suggests that a substantial portion of clearance between the mid 1980s and late 1990s was of forest not yet incorporated into the fallow cycle.[4] As the figures in Table 1 indicate, farmers in the sample cleared an average of 1.42 ha of forest older than 15 years for agricultural production per year by the mid 1990s, in addition to clearing an average of 0.97 ha per year of successional growth 7–15 years of age.

In the region as a whole, deforestation continued unabated through the decades of the 1980s and 1990s, with available satellite imagery revealing a total of 617 km^2 of mature forest cut between 1987 and 1997. In 1989, the Calakmul Biosphere Reserve

4 Respondents were asked about crop rotation, with most reporting that they could approximate a 1:1 crop-fallow cycle. Discussions with key informants, however, suggested that these numbers be deemed suspect, especially given that most respondents relied only marginally on chemical inputs.

Table 1
Average annual conversion of mature and secondary forest to agriculture among sampled households

	1984–1994	1993–1997
Hectares of mature forest older than 15 years cleared for agriculture	0.47	1.42
Hectares of secondary forest between 7 and 15 years cleared for agriculture	0.32	0.97

(723,185 ha) was established in the middle of the region, partly in response to extensive deforestation along the highway and associated international pressures to impede further clearance. Various land uses surround the reserve, predominately: *ejidos*, on which slash-and-burn subsistence production has prevailed but increasingly is giving way to commodity production; *ejidos* coupled with NGO-sponsored agricultural and forest projects; and a small number of private lands largely devoted to livestock production. Cumulatively, these activities have earned the region a designation as a "hot-spot" of forest and bio-diversity loss by various sources (Achard et al., 1998). Whether this prognosis is supported by future trends will depend in large part on current plans surrounding the *Mundo Maya*, an international development scheme to create an ecotourism–archaeological tourist economy stretching across portions of southern Mexico, Belize, Guatemala, El Salvador and Honduras. This plan has led to recent hotel and road expansion in the region.

Tourist sector development notwithstanding, the region retains many of the features of an agricultural frontier. Population density is low, infrastructure is poorly developed, ties to outside markets remain few, and the social relations of production are, by and large, organised around the family farm. These features motivate the theoretical and empirical approach taken in this paper, which emphasise the importance of household-specific characteristics, particularly demographic composition and farm capital, in explaining land-use decision making.

3. The theoretical model

As a predominately agrarian economy, land is an input in virtually all of the economic activities within the *ejido* sector of the southern Yucatán peninsular region. Some of these activities, such as bee keeping, agro-forestry, and hunting, rely on land under forest, while others, such as agricultural production and animal husbandry, require that the land be cleared of trees for use. Whether or not a farmer decides to clear a given tract of land depends on a complex multiplicity of factors, including the market value of output from the land in alternative uses, the availability of labour, the household's consumption requirement, and the farmer's perception of the potential future benefits derivable from the land.

The existing, albeit small, spatially explicit economic literature on land-use change has bifurcated into two separate modelling approaches. The current literature that has focused exclusively on the location of land-use change has analysed deforestation in tropical countries: Belize (Chomitz and Gray, 1996); northern Mexico (Nelson and Hellerstein, 1997); Brazil (Pfaff, 1999); Thailand (Cropper et al., 1999; Cropper et al., 2001); and Panama (Nelson et al., 2001). Models that include the temporal dimension as well as location of land-use change have focused on urban fringe development in the United States (Geoghegan and Bockstael, 2000; Irwin and Bockstael, 2002). While this latter approach has been applied to an essentially irreversible land conversion type, that of impervious surface expansion, it can be modified to analyse tropical deforestation given the high costs of forest clearance and the elimination of alternative forestry based land-use options that may accompany habitat degradation.[5]

Adapting the model found in Irwin and Bockstael (2002), we model the decision of the farmer to clear land for agricultural use. Let $A(i, t)$ be the net benefits to agricultural use for each time period after each pixel of land i, is cleared in time period T. Let $F(i, t)$ be the net benefits to the farmer for leaving pixel i in forestry use in each time period, and let $C(i, T)$ be the one-time clearing costs associated with clearing the

[5] The assumption of irreversibility seems reasonable in the case of our data given that only 1.4% of the pixels reverted to forest older than 15 year over the period under study; 7.1% of pixels reverted to secondary forest.

pixel in time period T and δ is the discount rate. Then the benefits to the farmer of clearing pixel i in period T are:

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

For T to be the optimal time period for clearing, the following two conditions must hold:

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

$$\sum_{t=0}^{\infty} A(i, T+t)\delta^{T+t} - \sum_{t=0}^{\infty} F(i, T+t)\delta^{T+t} - C(i, T) > \sum_{t=1}^{\infty} A(i, T+1)\delta^{T+t} - \sum_{t=1}^{\infty} F(i, T+1)\delta^{T+t} - \delta C(i, T+1) \tag{2b}$$

The first condition is that net benefits to clearing are positive. The second condition considers that although clearing may yield net positive benefits at time T, there may still be benefits to waiting because of the potential for even higher benefits at some future date. Such a circumstance could arise, for example, in anticipation of improved technologies that reduce clearing costs. This very simple model ignores fallow-cycle dynamics, clearly a limitation of the current theoretical framework.

Let the characteristics of pixel i be $X(i)$. The optimal time for clearing this pixel then is the first time period in which the following holds:

$$A(X(i), T) - F(X(i), T) - \delta C(X(i), T+1) \geq 0 \tag{3}$$

Given this theoretical framework, our empirical model aims to explain why certain pixels, in certain locations, and under certain land managers, become deforested.

4. The empirical model

We add an error term to Eq. (3) to account for unobservable characteristics:

$$A(X(i), T) - F(X(i), T) - \delta C(X(i), T+1) - \varepsilon(i) \geq 0 \tag{4}$$

The hazard rate—or probability that pixel i will be deforested in period T—can then be expressed as:

$$h(i, T) = \frac{G[W(i, T+1)] - G[W(i, T)]}{1 - G[W(i, T)]} \tag{5}$$

where G is the cumulative distribution function for the error term, and

$$W(i, T) = A(X(i), T) - F(X(i), T) - \delta C(X(i), T+1) \tag{6a}$$

$$W(i, T+1) = A(X(i), T+1) - F(X(i), T+1) - \delta C(X(i), T+2) \tag{6b}$$

We use a survival model to test hypotheses concerning the effect of explanatory variables, $X(i)$, on the hazard, $h(i,T)$, of deforestation. Survival models are a class of statistical methods that focus on the timing of an event (Allison, 1995), which in this paper is designated as the change from forest cover to agricultural land use. These models estimate the conditional probability of exiting a state given that the state has been occupied for some length t. The dependent variable, the duration, is the length of time that elapses from the beginning of the state until its end or until measurement is taken and therefore truncates the observation.

Survival models have a long history of application within the engineering and bio-medical sciences. They have been used to model such processes as the length of time until component failure or the survival times of patients diagnosed with certain diseases. Over the past two decades, these models have been recognised by social scientists to be a powerful tool for investigating a wide range of social phenomena, including the length of unemployment spells, the spacing of births, and the duration of strikes. By and large, however, they have not received application to the issue of landscape dynamics. Two recent exceptions are a study by Coomes et al. (2000) and one by Irwin and Bockstael (2002), cited earlier, both of which specify variants of Cox's proportional hazards model to investigate fallow-cycle length among Amazonian farmers in Peru and urban expansion in Maryland, respectfully.

For this paper, the duration of interest is the length of time that an individual pixel remains in forest before being converted to cropland or pasture. We estimate the effect of static and time interval-varying covariates on this duration by specifying a fully parametric model, the complementary log–log model. This specification assumes that the underlying process that generates the data is continuous, but that the data are grouped into discrete time intervals, an assumption that is well suited for the particular case of our data. Deforestation is a continuous process over time and space, but we only observe the event at select discrete times as dictated by the limited availability of cloud-free TM data for the tropics. Therefore, we do not know the exact timing of the deforestation event, only that it occurred during some time interval defined by the dates of the imagery. The complementary log–log model accommodates this feature of the data and additionally allows for different specifications of the role of time as a covariate, making it possible to test alternative functional forms of the hazard function.[6] For example, by including the logarithm of time as a covariate, the model corresponds to the Weibull (see Allison, 1995 for further discussion).

For the n pixels that are observed to be deforested, the likelihood function is:

$$\prod_{i=1}^{n} P_{i,t_t}(1-P_{i,t_i-1})(1-P_{i,t_i-2})\cdots \tag{7}$$

where $P_{i,t}$ is the probability that deforestation occurs to pixel i in interval t, given that the pixel was not deforested in any earlier periods, and t_i the time period in which pixel i is deforested. The complementary log–log model specification resulting from this likelihood function is:

$$\log[-\log(1-P_{it})] = \beta' X(i,T) \tag{8}$$

where

$$X(i,t) = A(X(i),T) - F(X(i),T) - \delta C(X(i),T+1) \tag{9}$$

That is, X are the exogenous variables and is a β vector of parameters to be estimated using maximum likelihood methods, with the assumption that the underlying survival models is distributed as type I extreme value (Irwin, 1998; Hosmer and Lemeshow, 1999).

A final estimation issue arises from the possible existence of unobserved heterogeneity that would result from an incomplete specification of the empirical model. Two different, but related issues concerning unobserved heterogeneity may result from the dual nature (i.e. temporal and spatial dimensions) of the data. In the context of the temporal dimension, if this incomplete specification results in unaccounted for systemic individual differences in duration distributions, the consequences include a downward biased estimate of the effect of time—or duration dependence in the hazard terminology—and to misleading inferences about the effects of included explanatory variables (Kiefer, 1988). In the context of the spatial dimension, the possibility exists that the features of the landscape affecting the probability that a pixel will be deforested are spatially correlated. If any of these variables are not included in the estimation and are spatially correlated with included explanatory variables, then the estimated coefficients will be biased on these variables (Irwin and Bockstael, 2001).

To control for these two types of unobserved heterogeneity, a fixed-effects specification that includes dummy variables for the spatial unit of each *ejido* is used (Hite et al., 2001). An *ejido* dummy was chosen as the unit for the fixed effect because of potential unobserved institutional differences between *ejidos*, such as their establishment dates, that could affect the baseline hazard rate. This approach allows baseline hazard rates to vary among individual *ejidos* but constrains these rates to be the same within *ejidos*, thereby controlling for differences in duration distributions at the *ejido* level. In addition, to the extent that there are unobservable characteristics that are correlated across space within the *ejido* boundaries, this specification controls for the some of the potential effects of spatial autocorrelation.

5. The data

The econometric model presented in this paper is estimated using Landsat TM satellite data on land cover as the dependent variable and household survey data

[6] The complementary log–log model is similar to the logit and probit models in ensuring that predicted probabilities lie in the [0, 1] interval. Unlike these models, however, whose functions are symmetric around zero, the complementary log–log function is skewed to the right. One practical implication is that the logit model is appropriate for estimating a hazard rate when events can only occur at discrete points in time (e.g. elections, job promotions), while the complementary log–log model should be used when events occur on the continuous time-scale.

Table 2
Dates of Thematic Mapper Satellite Imagery

Zone I	Zones II and III	Zone IV
11 November 1984	1 April 1987	14 January 1985
21 February 1993	29 October 1994	7 November 1994
31 January 1997	5 February 1996	31 January 1997

and other biophysical spatial data for the independent variables. The data sources for each variable are discussed briefly. The unit of observation for the model is the TM pixel, an admittedly arbitrary unit of analysis but one that nevertheless is likely an increment of the area over which land is cleared for farming as suggested by the figures on forest clearance presented in Table 1.[7] The satellite images were obtained across four contiguous zones spanning the study region, the dates for each of which are given in Table 2. The process of imagery classification included the normal preparatory steps of geo-referencing, haze removal, adding NDVI information, and principal component analysis. These steps were followed by texture analysis, which lead to the creation of a seven-band image for signature development and classification. Signature development was facilitated by a combination of ground truth data derived from GPS-assisted field visits and topographic, vegetation and land-use maps. Maximum likelihood supervised classification methods produced six land cover classes. Excluding clouds, shadows and water, these include: mature lowland and upland forest older than 15 years of age; one stage of upland successional growth—predominantly secondary forest—between 7 and 15 years of age; agriculture (including pasture); an invasive fern; and inundated savannas.[8] For further detail on these methods see Geoghegan et al. (2001) and Turner et al. (2001).

From the above classes, we generate two binary variables to register the forest conversion process. The first assumes a value of one if upland, lowland, or secondary growth is converted to agricultural land over an observed interval and zero otherwise. The second applies a stricter definition of deforestation, assuming a value of one only with the conversion of upland or lowland forest to agriculture and zero otherwise. Assuming that the length of fallow in the region approximately ranges between 9 and 12 years, we are unable to fully ensure that the former of these deforestation variables excludes land clearance that occurs as part of a fallow cycle. Indeed, it is highly probable that those pixels under secondary growth are in fact in fallow as most logging and related forest extraction activities in the region had terminated by 1960. Consequently, we estimate the model on both the deforestation variables defined above, recognising that the stricter definition of deforestation is likely to be largely purged of fallow-cycle dynamics given its exclusion of vegetation <15 years old.

Data for the explanatory variables are from a household survey that was carried out in the region during two separate field seasons during 1997–1998 and were linked to the satellite data, as will be further explained below. Selection of households in the sample proceeded according to a stratified, two-stage cluster sample (Warwick and Luinger, 1975; Deaton, 1997), with *ejidos* as the first stage unit and *ejidatarios* as the second stage unit. This resulted in the random selection of 11 *ejidos* followed by the random selection of 188 *ejidatario* households. A standardised questionnaire, administered to the household head, was used to elicit the socio-economic and land-use data. The questionnaire was organised into two sections. The first section covered migration history, farm production and inputs, ethnicity, educational attainment, access to credit and the demographic composition of the household. By collecting information on the births, deaths, and permanent out migration of children of the head, it was possible to reconstruct the biological household's age composition through time. In addition, dichotomous data was collected on ownership of farm capital (e.g. vehicle, chain saw) for the years 1986, 1990, 1993, 1996, and 1997. Using the figures for each of these years, data for the interim years were interpolated. In this way, the percent of time for which the relevant dichotomous variable was in effect could be approximated for any given interval corresponding to the dates of the satellite imagery.

Completion of the second section involved a guided tour of the agricultural plot of the respondent. Using a GPS, the interviewer created a geo-referenced sketch map detailing the configuration of land uses (Fig. 2).

[7] Note that a pixel is roughly 0.076 the size of a hectare.

[8] The primary difference across lowland and upland forest is not in species composition but in soil quality. Bajo soils, which support lowland forests, are deep vertisols of thick clay found in low-lying depressions. Upland forests are supported by *rendzina* soils, which are shallow but agriculturally fertile mollisols.

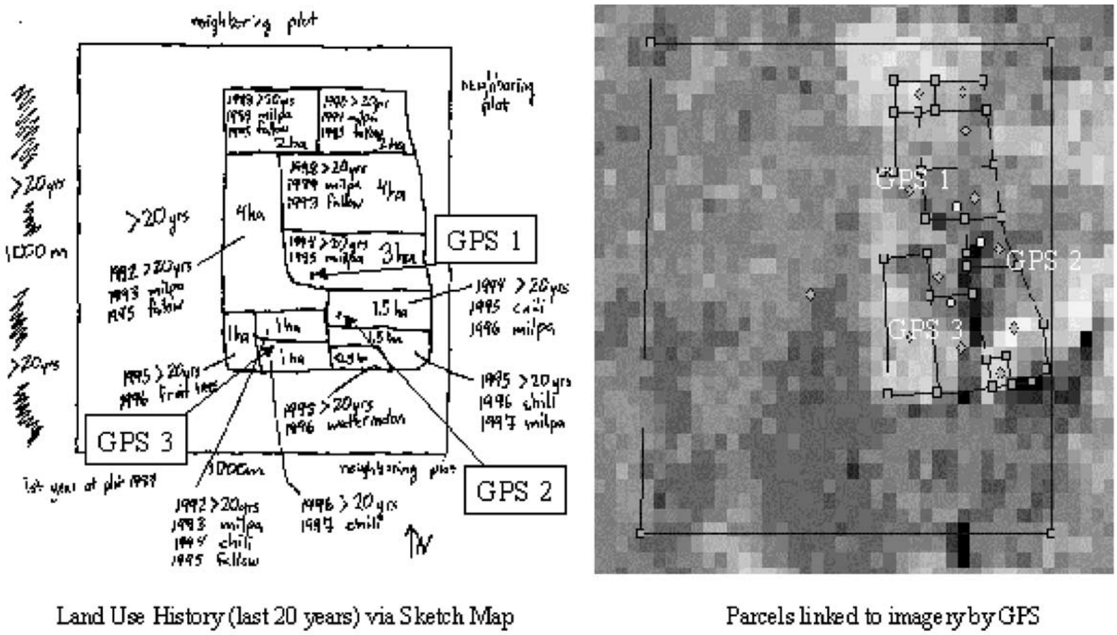


Fig. 2. Sketch map and corresponding backdrop from TM imagery. Note: the sketch map in the left panel depicts a parcel of 100 ha. The image backdrop in the right panel is roughly 120 ha, with pixels of 28.5 m^2. Discrepancies between the land use patterns depicted in the sketch map and those of the image may be partly attributed to the fact that the image is from February 1996 while the sketch map was created in June 1998. From Fig. 6 in chapter 9 of Integrated Land-Change Science and Tropical Deforestation in the Southern Yucatán: Final Frontiers, edited by B.L. Turner II, J. Geoghegan and D.R. Foster (Oxford University Press, forthcoming). Reprinted by permission of Oxford University Press.

Having several GPS points recorded within the borders of the plot made it possible to digitise the borders on a backdrop of the most recent date of satellite imagery.[9] The digitised borders were then extracted and superimposed on available images from previous years, thereby yielding a longitudinal database of land-use change. Thus, only those pixels associated with households from which socio-economic data were elicited are included in the sample. By overlaying this database with other GIS layers containing features such as: the road network, digitised from a 1:50,000 map published by the Instituto Nacional de Estadistica Geografica e Informatica (INEGI, 1985); soil quality measures, digitised from a 1:250,000 INEGI map (1987); slope and elevation from a digital elevation model; and rainfall extrapolated from rain gauge data; it was possible to create spatial explanatory variables to augment the data collected during the interview.

6. Variables used in the analysis

As suggested in the theoretical model presented above, the land clearance decision is based on a comparison of discounted utilities from forest and non-forest land uses. There are several testable socio-economic and environmental factors that could influence this comparison, which we conceptually group into four categories derived from the data elicited by the questionnaire: household demographic composition; physiographic characteristics of the

[9] The emphasis of the dialogue during the sketch mapping exercise was placed on accurately geo-referencing the plot boundaries, seen as the digitized outer rectangle in the image backdrop of Fig. 2, rather than the land use patterns contained therein. While the present studies relies on satellite imagery interpretation to describe those patterns, other work (see Klepeis and Vance, 2000) draws data from the narratives given by farmers to describe the spatial-temporal configurations of their plots, as depicted in the left panel of Fig. 2.

Table 3
Descriptive statistics of the variables used in the model

Explanatory variable	Mean	Standard deviation
Household members >11 (members)	3.048	1.699
Household children <12 (members)	1.618	1.796
Mature forest (1, 0)	0.799	0.401
Upland soil (1, 0)	0.754	0.431
Elevation (m)	164.040	69.100
Slope (°)	1.252	2.733
Precipitation (mm)	56.281	4.186
Plot size (numbers of pixels)	1357.718	1053.586
Percent of interval owning chain saw (%)	0.256	0.372
Percent of interval owning vehicle (%)	0.133	0.306
Education of household head (years)	3.377	3.905
Number of household members *w*/>8 years education (members)	1.039	1.537
Native Spanish speaker (1, 0)	0.848	0.359
PROCAMPO (1000 s pesos)	2.472	3.196
Distance from household to plot (km)	9.156	7.735
Distance from *ejido* to nearest market (km)	21.273	19.143
Duration of occupancy (years)	20.447	12.729

Units are given in parenthesis.

plot; farm capital (human and physical); and the political-economic environment. Descriptive statistics for these variables are presented in Table 3.

The influence of demographic composition is captured by two variables that partition the household according to age: family members over age 11 years, and children under 12 years. As both indices are measured as the average number of members in the respective age categories over the corresponding time interval of the imagery, they vary across households and time intervals. Following results identified elsewhere in the literature, it is expected that the demographic variables exert a positive effect on the hazard of deforestation through both consumption and labour supply effects.

Six time-invariant variables control for the effects of physiographic characteristics: a soil dummy which serves to distinguish between higher quality upland soils and lowland soils; the elevation and slope of the pixel; the 30 year average of rainfall; the size of the plot to which the pixel belongs; and a dummy indicating whether the pixel was categorised as mature forest (forest older than 15 years of age) at the start of the interval. We expect that slope has a negative effect on the hazard of deforestation due to the greater difficulty of cultivation on hillsides, while superior soils are expected to have a positive effect. Rainfall and elevation are also expected to have positive effects given the absence of irrigation in the region and the presence of seasonally inundated lowlands where farmers generally avoid cultivation. The land endowment is expected to have a negative effect as, all else equal, a larger land endowment reduces the hazard that any given pixel is cleared. Finally, the dummy indicating mature forest is expected to have a positive effect on the hazard as such vegetation is generally supported by more fertile soils relative to secondary vegetation.

Chain saw ownership, vehicle ownership, the education of the household head, the number of members in the household having completed a high school education as of 1997, and a dummy indicating whether the head is a native Spanish speaker control for the effects of physical and human capital. While the first two measures vary across households and time intervals, the latter three are time invariant. Ownership of a chain saw and vehicle are both expected to increase the hazard of deforestation given their roles in lowering the labour costs of forest clearance and in accessing the plot. To the extent that more education implies a higher opportunity cost of on-farm labour due to increased wage-earning potential, the two variables measuring educational attainment are expected to have negative effects. No a prior expectation is attached to the Spanish language dummy.

The influence of the political-economic environment is captured by four variables: the *ejido* dummies cited above to control for fixed effects, an interval-varying measure of the amount of agricultural payments received by the household from a government program referred to as Programa de Apoyo Directo al Campo/The Direct Rural Support Program (PROCAMPO), a measure of the distance separating the household from the plot, and a measure of on-road distance from the *ejido* centre to the nearest market. We expect the latter two variables to negatively affect the hazard of deforestation given that they increase access costs to both the plot and the market and effectively reduce the farm-gate price received by the farmer for the plot's output. Consistent with the objectives of the program, the effect of PROCAMPO is expected to also decrease the hazard of deforestation. Initiated in 1994, the program extends agricultural support via payments for the continued cultivation of a fixed area of land over 15 years, a period corresponding to the time in which agrarian price supports are to be phased out under the North American Free Trade Agreement (NAFTA). The payments are extended on a per hectare basis and in 1996 were set in real terms at 484 N pesos (US$ 64) for the remaining life of the program. While the farmer has considerable flexibility in selecting the crops planted, the number of hectares eligible is based on the area that was cultivated in any of nine basic crops in 1994, these being maize, beans, wheat, soybeans, sorghum, rice, cotton, safflower, and barley. As the area and location covered is fixed over the life of the program, one of its primary intents, as stated by the agency administering the subsidy Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación/Secretariat of Agricultural, Cattle Ranching, Rural Development, Fishing, and Food, 2002 (SAGARPA), is to intensify production and thereby decrease pressure on remaining forest (Klepeis and Vance, 2000).

To capture the effect of time as a determinant of forest clearance, we include an interval-varying variable that measures the duration of the household's occupancy as of the end of the time interval. The square of this variable is also included to allow for non-linearities in duration dependence. The inclusion of time as a covariate is a distinguishing feature of the model, as it serves to control for a range of inter-temporal factors that are important to the land-use decision but that are generally not observable (Boscolo et al., 1998). The effect of learning is one such factor. For example, we would expect that continued acclimation to the ecological and economic conditions of the frontier environment would partly determine the path of land allocation shifts over time. Given that we have no direct measures of learning effects, we rely on the variable measuring the household's duration of occupancy as a proxy. In addition to the inclusion of this variable and its square, the specification also includes dummy variables for each interval to control for the fact that they are of differing lengths (Allison, 1995).

7. Results

Tables 4 and 5 present measures of predictive power and the coefficient estimates, respectively, from two complementary log–log models that estimate the effect of the above determinants on the hazard of forest clearance. Model I in the two tables is estimated on the entire sample of observations while Model II is limited to just those pixels classified as lowland or upland forest at the start of an interval. In both models, the standard errors of the coefficient estimates are corrected for heteroskedasticity using Huber/White estimates of variance. As the coefficients from the complementary log–log hazard model are difficult to interpret directly, we derive a more intuitive interpretation through calculation of the "risk ratio". Let β_i be the coefficient associated with explanatory variable, X_i. Then the risk ratio associated with β_i is $\exp(\beta_i)$. For dummy variables the risk ratio is the ratio of the hazard rate for a pixel with the dummy variable equal to one to the hazard rate for a pixel with the dummy variable equal to zero, again holding other variables constant. For the continuous variables, we subtract one from the risk ratio and multiply by 100 (Allison, 1995), which gives the percent change in the hazard rate with a one unit change in X_i holding the other variables constant.

To examine the predictive ability of the model we employ two approaches, the results of which are presented in Table 4. The first, standard in the literature,

Table 4
Indicators of predictive performance

	Model I: forest >6 years			Model II: forest >15 years		
	0	1	Total	0	1	Total
P (clearance) <0.2	74,549	5,859	80,408	66,965	4,280	71,245
P (clearance) ≥0.2	18,707	11,002	29,709	10,918	5,782	16,700
Total	93,256	16,861	110,117	77,883	10,062	87,945
Goodman and Kruskal's gamma				0.611		

Table 5
Complementary log–log models of forest clearance

Explanatory variable	Model I: forest >6 years		Model II: forest >15 years	
	Estimated coefficient	Risk ratio	Estimated coefficient	Risk ratio
Household members >11	0.026[a] (4.428)	2.668	0.074[a] (8.846)	7.646
Household members <12	0.023[a] (4.394)	2.313	−0.006 (−0.876)	−0.598
Upland soil	0.270[a] (10.821)	1.310	0.416[a] (12.387)	1.515
Elevation	−0.011[a] (−23.786)	−1.055	−0.010[a] (−19.071)	−1.027
Slope	−0.023[a] (−6.396)	−2.229	−0.024[a] (−6.330)	−2.393
Precipitation	0.159[a] (21.990)	17.268	0.106[a] (11.375)	11.191
Plot size	−0.0002[a] (−16.124)	−0.023	−0.0003[a] (−16.210)	−0.029
Percent of interval owning chain saw	−0.001 (−0.039)	0.999	−0.148[a] (−3.560)	0.862
Percent of interval owning vehicle	0.407[a] (11.856)	1.502	0.720[a] (16.599)	2.054
Education of household head	0.022[a] (8.176)	2.233	0.023[a] (5.906)	2.292
Number of members w/>8 years education	0.029[a] (4.366)	2.944	0.008 (0.883)	0.784
Native Spanish speaker	−0.045 (−1.702)	0.956	−0.244[a] (−6.300)	0.784
PROCAMPO subsidy	0.023[a] (16.492)	2.339	0.030[a] (16.979)	3.053
Distance from household to plot	−0.045[a] (−28.266)	−4.381	−0.046[a] (−21.873)	−4.495
Distance from *ejido* to nearest market	−0.048[a] (−24.882)	−4.715	−0.040[a] (−17.174)	−3.959
Duration of occupancy	−0.041[a] (−11.751)	−4.009	−0.052[a] (−11.817)	−5.028
Duration of occupancy squared	0.0005[a] (8.042)	0.052	0.001[a] (11.358)	0.091
Mature forest	−0.645[a] (−33.609)	0.525		
Constant	−6.697[a] (−14.855)		−4.266[a] (−7.741)	
χ^2-statistic for interval dummies	2,240.61[a]		1,595.68[a]	
χ^2-statistic for *ejido* dummies	1,328.88[a]		880.30[a]	
Number of observations	110,117		8,7945	
Wald χ^2	15,809[a]		9,410[a]	
Log likelihood	−38022		−25610	

[a] Denotes significance at 1% level.

references a matrix showing the number of predicted versus actual changes. Following Chomitz and Gray (1996), who argue that even low predicted probabilities convey information, we select 0.2 as the threshold probability for designating observations as predicted changes. Using this criterion, Model I accurately predicts 65% of deforested pixels and 80% of forested pixels, with corresponding figures of 57% and 86% for Model II. As Chomitz and Gray and others have noted, however, the choice of a threshold probability value is entirely arbitrary, and results can change dramatically simply by selecting a different value. An alternative, though rarely used, measure of model fit for binary choice models is given by Goodman and Kruskal's gamma (Goodman and Kruskal, 1954, 1959, 1963). This symmetric, non-parametric measure varies from −1 to +1 and is based on the difference between concordant (C) and discordant (D) pairs of predicted and actual values as a percentage of all pairs ignoring ties, computed as $(C-D)/(C+D)$.[10] The value of gamma for Model I is 0.61 and is only slightly higher for Model II, indicating that the model reduces our error in predicting the outcome of the dependent variable by just over 60%.

Turning to Table 5, it is seen that the estimated coefficients of Model I are generally of the expected sign. Both of the demographic indices are positive and statistically significant determinants of the hazard of

[10] Concordant pairs are those for which the values are higher (or lower) on all elements in one set compared with another. Discordant pairs are those for which the value of one element is higher in one set while the value of some other element is higher in the second set. As an illustration, consider the following hypothetical list of predicted and actual values: a. (0.9, 0); b. (0.3, 0); c. (0.8, 1); d. (0.2, 1). From this list, *ac*, *ad*, and *bd* are discordant pairs while *bc* is a concordant pair. Pairs *ab* and *cd* are ties and are therefore ignored in the calculation of gamma. Gamma in this example equals −0.5.

deforestation, an unsurprising result given that the majority of households in the region are semi-subsistent producers, for whom which family members simultaneously represent a source of labour as well as an overhead cost. Specifically, each additional member over 11 increases the hazard of deforestation by 2.67%, with a similar magnitude for children.

All of the physiographic measures are statistically significant, and, in some cases, confirm findings identified elsewhere in the literature. For example, the estimated coefficients for slope and elevation are both negative, similar to findings of Chomitz and Gray (1996) and Nelson and Hellerstein (1997) in their studies of deforestation in Belize and Mexico, respectively. As expected, rainfall and superior soils both have a positive effect on the hazard of deforestation, while a larger plot size has a negative effect. Finally, the negative coefficient on the mature dummy suggests that farmers prefer to clear secondary growth, an interesting finding given that secondary growth is generally supported by less fertile soils. Specifically, the estimate suggests that the hazard of clearance for mature forest in any period is about 53% of the hazard for secondary vegetation. One possible explanation is a desire to avoid higher clearance costs associated with mature forest even given higher weeding costs associated with inferior soils, a trade-off analysed at length in the fallow-cycle literature (e.g. Dvorak, 1992).

Vehicle ownership increases the hazard of deforestation, as expected, while the effect of chain saw ownership is statistically insignificant. A somewhat surprising result is the positive effect of the two education variables on deforestation. This may reflect the effect of higher labour productivity from increased managerial talent (Tao Yang, 1997), which could result in a relatively lower marginal productivity of land and hence greater clearance. Being a Spanish speaker decreases the hazard of deforestation, though the coefficient estimate is just out of the range of significance at the 5% level.

With regard to the measures of the political-economic environment, the estimated negative coefficients on the two distance variables are consistent with the intuition that higher travel costs decrease the returns from agricultural land use through a reduced farm-gate price of output. Perhaps the most interesting finding, from a policy perspective, is the positive and statistically significant coefficient estimate on PROCAMPO, a result directly at odds with the program's intent to decrease deforestation. Specifically, the estimate indicates that each 1000 pesos extended by the program increases the hazard of deforestation by 2.34%.

Different specifications of the variable measuring duration of occupancy were explored by means of a nested likelihood ratio test, whereby the quadratic functional form was determined to be optimal in terms of fit and parsimony.[11] The estimates indicate non-linear duration dependence of the deforestation process; the conditional probability of forest conversion decreases with the passage of time at a decreasing rate, with some evidence of a reversal after 41 years. This result may reflect a confluence of factors, including the rotation period of the fallow cycle, family life cycle dynamics, and, as suggested above, learning effects from adaptation to local market opportunities and ecological constraints.

Turning to the results of Model II, estimated on only those pixels with vegetation older than 15 years, some significant discrepancies emerge. The most noteworthy of these is the counterintuitive sign reversal on the coefficient estimate for ownership of a chain saw. A plausible explanation for this finding is not immediately forthcoming, other than to speculate that it reflects a shortcoming of the model to adequately control for wealth effects. In this regard, anecdotal evidence from the field survey suggests that those farmers who owned chain saws were more diversified in their income generating activities, relying less on the natural resource base for their livelihoods. Other differences across the two models are seen with respect to the demographic variables and human capital indicators. The number of children under age 12 years, which is positive and significant in the model estimated on the entire sample of pixels, is now insignificant, while the magnitude of the estimate on members older than 11 increases over three-fold. This finding may reflect the fact that the clearance of older vegetation is more strongly determined by the availability

[11] The nested likelihood ratio test compares the log likelihoods of two nested models using the χ^2-statistic. Specifically, if one model is nested within another, the fit of the nested (or constrained) model can be evaluated by taking twice the absolute difference in the log likelihoods for the two models. This difference gives the likelihood ratio χ^2-statistic, which, if significant, indicates that the nested model should be rejected.

of labour rather than by demand-side factors. Finally, the variable measuring the number of members in the household having completed a high school education as of 1997 and the Spanish language dummy, respectively lose and gain significance levels in Model II. With regard to the latter, the negative sign of the coefficient estimate may be indicative of a greater reliance of indigenous farmers on the resource base as opposed to off-farm wage-earning opportunities.

While the remaining coefficient estimates for Models I and II are for the most part similar with respect to sign and statistical significance, those differences that do emerge suggest that the age of vegetation cleared may have an important mitigating effect on the determinants of land-use decisions. The extent to which these differences are related to Model II's exclusion of land under fallow is not possible to fully explore given the limited number of dates and the length of the temporal window captured by available satellite imagery. One intuitively appealing interpretation to the results of Models I and II is to regard the fallow cycle as the baseline rate of change, whereby the effect of the explanatory variables is to accelerate or slow down this underlying rate depending on whether a positive or negative coefficient is observed. Some caution must be exercised, however, in extending this interpretation too far, particularly if the dependent variable is capturing the clearance of mature forest. In this case, a positive effect of an explanatory variable could actually correspond to a longer cycle due to the incorporation of more land.

8. Discussion

Several policy implications emerge from the results of the model. Perhaps most important from the perspective of agricultural commodity price interventions is the statistical significance of the demographic measures, suggesting potential non-independence of production and consumption decisions that would result from only partial integration in output and/or factor markets. A large and still growing body of literature addresses the implications of incomplete markets for agricultural supply response. This literature establishes that farmers will not, in general, exhibit behaviour consistent with profit maximisation given prohibitively high transaction costs in market participation (e.g. Singh et al., 1986; Benjamin, 1992; Saha, 1994; Omamo, 1998; Key et al., 2000; Vance and Geoghegan, 2002). Based on this observation, a broad consensus has emerged that market integration is an important prerequisite to the implementation of structural adjustment policies as it ensures that farmers are within the reach of standard policy instruments and, moreover, that the response to such instruments will be predictable. This perspective gives rise to recommendations for policies that accelerate market articulation so as to reduce the incidence of sluggish or perverse price responses that may characterise farm systems in the intermediate stages of development (Anderson Medellin et al., 1994).

Whether policies that encourage integration into the market are consistent with both economic and environmental objectives is, however, far from clear. Proximity to roads is one policy related variable that could be used to increase participation with the market, but one which is also identified in this study and others to increase the probability of forest clearance (e.g. Chomitz and Gray, 1996; Nelson and Hellerstein, 1997). Von Thuenen's work on locational rent functions provides a theoretical model for explaining this result. In this model, the value of land for agriculturists and those pursuing other forest resources is largely determined by its location with respect to roads or ports, whereby forest clearance occurs until the point where harvest and access costs equals the additional value garnered from land-based outputs (Katzman, 1975; Hyde et al., 1996). This model is largely applicable to the southern Yucatán, where the paving of the highway in 1972 subsidised transport costs between the frontier and regional urban centres, instigating the unprecedented period of land-use change modelled in this study. Nevertheless, some caution is warranted in extrapolating the negative effect of roads to predict the effects of future road building projects. To the extent that roads reduce the costs of accessing off-farm wage-earning opportunities, they can potentially reduce pressure on forests by drawing labour off the farm. Given that this possibility is predicated on the availability of such opportunities, road construction associated with the Mundo Maya tourism scheme referred to above should proceed with an eye toward simultaneous generation of employment in the ecotourism sector.

A final policy-relevant finding to be drawn from the model is the positive coefficient of PROCAMPO in increasing the hazard of deforestation, which again may be related to farmers' partial integration into markets. Rather than using the support to purchase land-intensive chemical inputs, markets for which in the region are shallow or non-existent, farmers may be using the support to expand the extent of cultivation through the purchase of more readily available inputs such as labour. An additional explanation for the positive coefficient on PROCAMPO may be due to the specific terms of the program, which stipulate that the area and location supported must have been cultivated under some staple for one of the 3 years prior to 1994. Provided that that same tract is maintained under a "productive" use not limited to staples following 1994, the participating farmer is eligible to receive PROCAMPO payments. The basic structure of the support is thus at odds with an agricultural system based on forest/fallow. By requiring that the land be maintained under production, the program precludes the use of fallow as a means of supporting repeat cultivation on the area enrolled, evidently increasing pressures on remaining hectares (Klepeis and Vance, 2000). This suggests the importance of designing support programs not only in consideration of the prevailing market structure within which farmers operate, but also in consideration of the managerial strategies applied to agricultural production.

9. Conclusion and research extensions

This paper has presented two principal advancements to the spatial and temporal modelling of tropical deforestation in agrarian regions. Methodologically, the paper develops the use of GPS-assisted field surveys as a means of linking farmer plots to available satellite imagery. The creation of geo-referenced sketch maps makes it possible to not only spatially associate individual farmers with their plots, but also to track over time the forest conversion process within these plots. Empirically, the paper presents an application of a survival model as a means of analysing continuous time processes using discrete time data. By specifying the optimal timing of forest clearance as the choice variable, the empirical model identified several potentially important ecological and political-economic determinants of deforestation. After estimating the complementary log–log specification of the survival model, it was determined that the data is characterised by non-linear duration dependence, with the hazard of deforestation first decreasing and then increasing with time as seen by the coefficient estimates on the year and year-squared covariates.

There are several potential avenues for future research using this style of model to address deforestation processes in agrarian contexts. The first would entail more explicit treatment of fallow-cycle dynamics. In order to do this, a more sophisticated theoretical model of decision making is required. Empirically, this could be pursued through alternative specifications of the effect of time as well as through tests of parameter consistency according to whether mature or secondary forest is cleared. A second extension would involve further exploration of the extent to which limited exchange opportunities from missing or shallow markets affects the decision to clear land. Such an analysis would be complicated by the fact that the decision to enter the market is, itself, an endogenous choice, requiring accommodation of potential sample selection biases. The final issue meriting further investigation regards the empirical implications of spatial autocorrelation. While the present analysis links the satellite imagery and independent variables at the level of the individual land manger, as noted above the pixel is nonetheless an arbitrary unit of analysis, and the possibility that unobserved farmer-specific attributes introduce biases must be acknowledged. In this regard, one possibility of controlling for spatial autocorrelation given the discrete nature of the present data would be to develop a spatial lag variable as one of the covariates.

Acknowledgements

This work was undertaken through the auspices of the Southern Yucatán Peninsular Region (SYPR) project with core sponsorship from NASAs LCLUC (Land-Cover and Land-Use Change) program (NAG 56406) and the Centre for Integrated Studies on Global Change, Carnegie Mellon University (CIS-CMU; NSF-SBR 95-21914). Additional funding from NASAs New Investigator Program (NAG5-8559) also supported the specific research in this article. SYPR is a collaborative project of El Colegio de la Frontera

Sur (ECOSUR), Harvard Forest—Harvard University, the George Perkins Marsh Institute-Clark University and CIS-CMU. The views expressed in this article are those of the authors and do not necessarily represent those of the US Environmental Protection Agency. We thank Laura Schneider for her assistance with data preparation for this paper. We also thank Julie Hewitt, two anonymous referees, and the editor for providing valuable comments.

References

Achard, F., Eva, H., Glinni, A., Mayaux, P., Richards, T., Stibig, H.J. (Eds.), 1998. Identification of Deforestation Hot Spot Areas in the Humid Tropics. Trees Publication Series B, Research Report No. 4, Space Application Institute, Global Vegetation Monitoring Unit, Joint Research Centre, European Commission, Brussels.

Allison, P.D., 1995. Survival Analysis Using the SAS System: A Practical Guide. SAS Institute Inc., Carey, North Carolina.

Anderson Medellin, M., Apedaile, L.P., Pachico, D., 1994. Commercialisation and price response of a Bean-growing farming System in Columbia. Econ. Dev. Cultural Change 42, 795–816.

Benjamin, D., 1992. Household composition, labour markets, and labour demand: testing for separation in agricultural household models. Econometrica 60, 287–322.

Boscolo, M., Kerr, S., Pfaff, A., Sanchez, A., 1998. What Role for Tropical Forests in Climate Change Mitigation? The Case of Costa Rica, Paper Prepared for the HIID/INCAE/BCIE Central America Project.

Chomitz, K., Gray, D., 1996. Roads, land use and deforestation: a spatial model applied to Belize. World Bank Econ. Rev. 10, 487–512.

Coomes, O.T., Grimard, F., Burt, G.J., 2000. Tropical forests and shifting cultivation: secondary forest fallow dynamics among traditional farmers of the Peruvian Amazon. Ecol. Econ. 32, 109–124.

Cropper, M., Griffiths, C., Mani, M., 1999. Roads, population pressures, and deforestation in Thailand. Land Econ. 75, 58–73.

Cropper, M., Puri, J., Griffiths, C., 2001. Predicting the location of deforestation. Land Econ. 77, 172–186.

Deaton, A., 1997. The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Johns Hopkins University Press, Baltimore.

Dvorak, K., 1992. Resource management by west African farmers and the economics of shifting cultivation. Am. J. Agric. Econ. 74, 809–815.

Ewell, P., Merrill-Sands, D., 1987. Milpa in Yucatán: a long-fallow maize system and its alternatives in the Maya peasant economy. In: Turner II, B.L., Brush, S.B. (Eds.), Comparative Farming Systems, Guilford Press, New York, pp. 95–129.

Geoghegan, J., Bockstael, N., 2000. Smart Growth and the Supply of Sprawl. Paper Presented at the Association of Environmental and Resource Economists Workshop on the Effectiveness of Resource and Environmental Regulation, La Jolla, CA.

Geoghegan, J., Cortina Villar, S., Klepeis, P., Macario Mendoza, P., Ogneva-Himmelberger, Y., Roy Chowdhury, R., Turner II., B.L., Vance, C., 2001. Modelling tropical deforestation in the Southern Yucatán peninsular region: comparing survey and satellite data. Agric. Ecosys. Environ. 85 (1–3), 25–46.

Goodman, L., Kruskal, W.H., 1954. Measures for association for cross-classification. Part I. J. Am. Stat. Assoc. 49, 732–764.

Goodman, L., Kruskal, W.H., 1959. Measures for association for cross-classification. Part II. J. Am. Stat. Assoc. 54, 123–163.

Goodman, L., Kruskal, W.H., 1963. Measures for association for cross-classification. Part III. J. Am. Stat. Assoc. 58, 310–364.

Hite, D., Sohngen, B., Templeton, J., 2001. Property Tax Impacts on the Timing of Land Use Conversion. Paper Presented at the Annual Meeting of the American Agricultural Economics Association, Chicago, August, 2001.

Hosmer, D., Lemeshow, S., 1999. Applied Survival Analysis: Regression Modelling of Time to Event Data. Wiley, New York, NY.

Hyde, W., Amacher, G., Magrath, W., 1996. Deforestation and forest land use: theory, evidence, and policy implications. World Bank Res. Observer 11, 223–248.

INEGI, 1985–1987. Cartas topogrficas. 1:50,000. Instituto Nacional de Estadistica Geografica e Informatica, Aguascalientes.

Irwin, E.G., 1998. Do Spatial Externalities Matter? The Role of Endogenous Interactions in the Evolution of Land Use Pattern. Department of Agricultural and Resource Economics, University of Maryland, College Park, Unpublished Ph.D. Dissertation.

Irwin, E., Bockstael, N., 2001. The problem of identifying land use spillovers: measuring the effects of open space on residential property values. Am. J. Agric. Econ. 83, 699–705.

Irwin, E., Bockstael, N., 2002. Interacting agents, spatial externalities, and the endogenous evolution of residential land use pattern. J. Econ. Geogr. 2, 31–54.

Katzman, M., 1975. The Brazilian frontier in comparative perspective. Comparative Stud. Soc. Hist. 17, 266–285.

Key, N., Sadoulet, E., de Janvry, A., 2000. Transaction costs and agricultural household response. Am. J. Agric. Econ. 82, 245–259.

Kiefer, N., 1988. Economic duration data and hazard functions. J. Econ. Lit. 26, 646–649.

Klepeis, P., Vance, C., 2000. Neoliberal Policy and Deforestation in Southeastern Mexico: Land Reform, PROCAMPO, and Land-Use/Land-Cover Change. Paper Presented at the Conference of Latin American Geographers, Austin, TX.

Klepeis, P., Vance, C., Keys, E., Macario Mendoza, P., Turner II, B.L., 2003. Subsistence Sustained: Swidden or Milpa Cultivation. In: Turner II, B.L., Geoghegan, J., Foster, D. (Eds.), Integrated Land-Change Science and Tropical Deforestation in the Southern Yucatán: Final Frontiers, Oxford: Clarendon Press of Oxford University Press (in press).

Liverman, D., Moran, E.F., Rinfuss, R.R., Stern, P.C. (Eds.), 1998. People and Pixels: Linking Remote Sensing and Social Science. National Academy Press, Washington, DC.

Nelson, G., Hellerstein, D., 1997. Do roads cause deforestation? Using satellite images in econometric analysis of land use. Am. J. Agric. Econ. 79, 80–88.

Nelson, G., Harris, V., Stone, S., 2001. Deforestation, land use and property rights. Land Econ. 77, 187–205.

Omamo, S., 1998. Transport costs and smallholder cropping choices: an application to Siaya district, Kenya. Am. J. Agric. Econ. 80, 116–123.

Pfaff, A., 1999. What drives deforestation in the Brazilian Amazon? J. Environ. Econ. Manage. 37, 26–43.

SAGARPA, 2002. http://www.procampo.gob.mx/. Secretaría de Agricultura, Ganadería, Desarrollo Rural. Pesca y Alimentación/Secretariat of Agricultural, Cattle Ranching, Rural Development, Fishing, and Food.

Saha, A., 1994. A two-season agricultural household model of output and price uncertainty. J. Dev. Econ. 45, 245–269.

Singh, I., Squire, L., Strauss, J. (Eds.), 1986. Agricultural Household Models: Extensions, Applications, and Policy. Johns Hopkins University Press, Baltimore.

Tao Yang, D., 1997. Education and off-farm work. Econ. Dev. Cultural Change 45, 613–632.

Turner II, B.L., 1983. Once Beneath the Forest: Prehistoric Terracing in the Rio Bec Region of the Maya Lowlands. Dellplain Latin American Studies No. 13, Westview Press, Boulder.

Turner II, B.L., Cortina Vilar, S., Foster, D., Geoghegan, J., Keys, E., Klepeis, P., Lawrence, D., Macario Mendoza, P., Manson, S., Ogneva-Himmelberger, Y., Pérez Salicrup, D., Roy Chowdhury, R., Savistsky, B., Schneider, L., Schmook, B., Vance, C., 2001. Deforestation and agricultural change in the southern Yucatán peninsular region: an integrative approach. Forestry Ecosyst. Manage. 154, 353–370.

Vance, C., Geoghegan, J., 2002. Semi-Subsistent and Commercial Land-Use Determinants in an Agricultural Frontier of Southern Mexico: A Switching Regression Approach. International Regional Science Review, Forthcoming.

Warwick, D.P., Luinger, C.A., 1975. The Sample Survey: Theory and Practice. McGraw Hill, New York.