The Dynamics of Deforestation:

evidence from Costa Rica

Suzi Kerr, Motu Economic and Public Policy Research

Alexander S.P. Pfaff, Columbia University

Arturo Sanchez, University of Alberta

September, 2002

Abstract

We estimate a deforestation equation for Costa Rica during the 20th century, using an econometric approach following from a dynamic microeconomic model. While the theoretical model is similar to Stavins and Jaffe (1990), we use forest/non-forest transitions during discrete time periods, rather than forest shares at points in time. This permits us to empirically capture dynamics which reasonable theories posit may exist: exogenous development; endogenous development following early clearing; and short-run adjustment costs. Our results confirm the importance of agricultural productivity and distances as found in earlier research. Further they suggest that, at least within a development setting, agricultural/forest land use will be out of equilibrium as defined with respect to observable variables so that a dynamic approach provides additional insight and important controls. The results help to understand agricultural extensification and clearing patterns, and could provide a basis for carbon baseline projections within global regulation.

Key words: land use, deforestation, climate change, transitions, development, Costa Rica

We acknowledge financial support from the National Science Foundation Grant No. 9980252, The Tinker Foundation, the Harvard Institute for International Development, the National Center for Environmental Analysis and Synthesis at UC-Santa Barbara, and CERC and CHSS at Columbia University. Many thanks to Antonio Bento, Nancy Bockstael, Richard Eckhaus, Ken Richards, Rob Stavins and others for useful comments at conferences. Thanks to William Power, Juan Andres Robalino, Jason Timmins, Joanna Hendy, David Kennedy and Steve Cournane for research assistance. All opinions are our own, and we are responsible for all errors and omissions.

1 Introduction

Deforestation leads to habitat loss, release of carbon to the atmosphere, soil degradation and flooding. However, it allows agricultural development and the expansion of human settlement. While we have considerable understanding of land-use patterns in developed countries, and even of land use at points in time within some developing countries, we have little understanding of the dynamics of deforestation along the development path. They can lead to rapid massive losses of forest (including in species-dense areas), or to movements toward the protection of remaining forest. Thus, we find it difficult to predict future deforestation patterns in the developing world.

One reason to predict forest change arises because of global climate change regulation and the resulting focus on net carbon emissions. If we want to reward developing countries for protecting forests and the carbon that they sequester, we must have some way of predicting what countries would have done with their forests in the absence of rewards. Otherwise we may simply reward countries for things they would have done anyway or, out of fear of doing that, not provide them with incentives to protect forest though these incentives might have been effective.

In this paper we estimate a deforestation equation for Costa Rica, across five time periods and 1128 forest observations over space, using an econometric method that follows directly from a dynamic microeconomic model. Costa Rica is an excellent country to use for this type of analysis because it has passed through many different development and deforestation stages, it is small enough that it faces clear exogenous shocks from international markets, and its relevant data sets are perhaps unsurpassed. Our work builds on the rapidly growing literature which uses revealed preference within past land-use behaviors to learn about the drivers of land-use change (Stavins and Jaffe 1990 is an early example, Bell and Bockstael 2000 a more recent one). In the developing country context Chomitz and Gray 1996, Nelson and Hellerstein 1997, Cropper et al. 1997, Pfaff 1999, and Geoghegan et al. 2001 offer a range of analyses, while Kaimowitz and Angelson 1998 offer a comprehensive survey of deforestation modeling papers of different types and scales. Several quite well-known analytical studies of deforestation have looked at Costa Rica specifically, notably including regression analysis by Rosero-Bixby and Palloni 1996. Most, though, have taken approaches somewhat different from our model-derived regressions (see, for instance, Sader and Joyce (1988) on rates of deforestation, Harrison 1991's focus on the effects of population, and the varied, rule-based extrapolations in Pontius, Cornell and Hall 2001).

Our study differs from previous work in several ways. Because of our emphasis on dynamics, we estimate our model from forest to non-forest transitions rather than forest shares. We derive a deforestation regression equation from a dynamic model, so the coefficients are linked to the theory.[1] For a developing country context we have data of extraordinary richness and quality, which allows us to explore a relatively wide range of spatial and dynamic determinants of deforestation. We include time effects to capture some of the dynamic changes in unobservables that arise from the development process and permit fixed factors to have ongoing effects on land use change. We also make efforts to capture or control empirically for other changes that dynamic theories suggest occur over time: endogenous increases in local returns following early production activity and clearing; and short-run adjustment costs that lead to partial adjustment. We use different measures of past clearing behavior in the regressions to try to capture these effects.

The rest of the paper proceeds as follows. Section Two presents a model of individual land-use choices and discusses how aggregate land use and the conditions affecting land use

[1] Many papers (e.g. Cropper and Griffiths 1994) include deforestation regressions without this type of theory link.

have evolved in Costa Rica. The third section derives our dynamic econometric specification, drawing on other economic literature that analyses transitions. Section Four discusses the land-use data and explanatory variables we use, while Section Five presents and discusses our results.

2 A Dynamic Model of Deforestation in Costa Rica

2.1 A DYNAMIC MODEL

In our use of a dynamic theoretical model we follow Stavins and Jaffe (1990), Parks and Hardie (1995) and Ehui and Hertel (1989). The manager of each hectare *i* faces a dynamic optimization problem. Risk neutral by assumption, the land manager selects *T*, the time when land is cleared, in order to maximize the expected present discounted value of returns from the use of hectare *i*:

$$\mathrm{Max}_T \int_0^T S_{it}\, e^{-rt}\, dt \; + \int_T^\infty R_{it}\, e^{-rt}\, dt - C_T\, e^{-rt} \tag{1}$$

where:

S_{it} = expected return to forest uses of the land

R_{it} = expected return to non-forest land uses

C_T = cost of clearing net of obtainable timber value and including lost option value[2]

r = the interest rate

Two conditions are necessary for clearing to occur at time *T*. First, clearing must be profitable:

$$\int_T^\infty (R_{it} - S_{it})\, e^{-rt}\, dt \; - C_T > 0 \tag{2}$$

[2] The correct model includes uncertainty, risk aversion and forward-looking knowledge of the ability to shift back and forth optimally between cleared and uncleared states. Uncertainty alone, when combined with the irreversibility suggested by our one-way model of clearing (or alternatively, simply by a sunk cost of clearing) and the ability to learn over time, imples an option value to waiting to clear.

For clearing to occur, the present discounted rents from non-forest uses will have to more than compensate the manager for the lost returns from forestry uses and the net cost of land clearing. However, even if clearing is profitable at time *t*, it may be even more profitable to wait and clear at *t+1*; for example, clearing costs may fall.[3] Thus the following 'arbitrage' condition must hold.

$$R_{it} - \mathrm{S}_{\mathrm{it}} - \mathrm{r}_{\mathrm{t}}\,\mathrm{C}_{\mathrm{t}} + \frac{dC_T}{dt} > 0 \tag{3}$$

If the relative returns to non-forest uses are increasing over time while the rate of fall in costs is decreasing, such that the second-order condition (4) holds, then if either of these necessary conditions holds it also will be sufficient for clearing to be preferred. In the case of land-use change within a developing country, population growth combined with economic growth and improved infrastructure is likely to make this the case.[4]

$$\frac{dR_{it}}{dt} - \frac{dS_{it}}{dt} + \frac{d^2}{dt^2}C_t > 0 \tag{4}$$

Whether or not these conditions hold (or have held previously) will determine the forest status on each land parcel at each point in time. Different land parcels will have different outcomes, because they have different land quality and hence agricultural productivity, as well as different access to markets and hence different input and output prices; thus they have different agricultural returns. Returns and costs will also change systematically over time for both exogenous and endogenous reasons. If land-use change is rapid, costs of clearing may rise temporarily during the adjustment phase. The distribution of land quality and market access, combined with systematic changes in returns and costs, feed through individual decisions to determine the aggregate patterns of regional and national deforestation.

---

[3] In addition, future returns on cleared land may depend on the time of clearing; land may degrade with use, e.g..

2.2 THE COSTA RICAN SETTING

Costa Rica has an area of approximately five million hectares, i.e. it is about 1% as large as the Brazilian Amazon. It borders Nicaragua and Panama but has largely been free of the political unrest in these countries. Deforestation in Costa Rica arises primarily through transformation of forests into agricultural land rather than through logging pressure. The most valuable timber has been extracted without complete deforestation. Thus we focus on agricultural extensification, rather than changes in the returns to logging.

The major export crops produced in Costa Rica are coffee, bananas, sugar cane and beef. Within Costa Rica the predominant crops vary greatly across space and time. The spatial variation mostly represents the spatial dispersion of "lifezones", or bands of precipitation and temperature, as certain crops grow better in certain lifezones. The choice of crops has changed a great deal over the nearly century and a half since the time when the only clearing was found in the Central Plateau (site of the capital, San José, and a good location for growing coffee) and around one of the western ports. The economy has opened up a great deal since 1960 and export prices have changed dramatically over time. These have altered both returns and the crop mix. Additional variation in returns has been introduced by changes in yields and the technologies of production. Both income per capita and population grew dramatically from 1900 through 2000.

Starting as early as 1850, two areas favorable for economic activities saw limited forest clearing: the Central Plateau, favored for agricultural production and the location of San José the capital, and the area around Puntarenas, a Pacific port. Clearing was slow over the next century, but the rate increased considerably from 1950 until the early 1980s. This was fueled at least in

---

[4] At a certain stage of development, however, this trend may be reversed. For instance, as development proceeds, environmental protection may become more stringent, returns to ecotourism may well rise, and agriculture can be

part by significant population growth, including significant immigration from Nicaragua. Cattle production expanded greatly during this time, as did coffee exports. Both coffee and beef prices rose sharply, peaked in the late 1970s, and then fell somewhat in the mid 1980s. Coffee prices fell heavily in the late 1980s. The productive infrastructure appears to have developed slowly. The effect of these changes on forests varied greatly over space, probably because of the spatial distribution of the more productive lifezones and of accessibility.

The late 1980s and 1990s brought a different story as deforestation dropped dramatically. Parts of this change may relate to: the fall in coffee prices; creation and enforcement of national parks and other policies; and the transformation of the Costa Rican economy from an agricultural base toward a manufacturing and services base. Alternatively the most valuable agricultural lands may have already been cleared by that time.

3 Econometric Specification

3.1 Deriving a Regression Equation

In deriving an econometric specification, we are willing to assert that the second-order condition (4) holds over our sample period. We present a derivation of our regression equation based on the arbitrage equation (3). Our reduced form, however, is also interpretable in light of the profit condition. We are not attempting to determine whether (2) or (3) is the dominant condition driving behavior, and thus the basis of the right or best model. Rather, we are testing whether measured factors have theoretically predicted effects on transition choices.

Our econometric approach differs from much of the earlier econometric work because it considers transitions from forest to non-forest, rather than forest shares. The use of changes

more capital intensive and require less land. Agricultural returns could fall relative to forest returns on some land.

permits us to distinguish deforestation from reforestation, and to treat them differently. This is not possible with levels equations, which determine how much forest will be present as a function of determinants, regardless of whether the area in question was all forest or all cleared previously. The theoretical reason to empirically separate these transitions is our belief that reforestation transitions are not simply automatic outcomes of a fall in deforestation pressure. Deforestation is not instantly or costlessly reversible because trees take time to grow, especially on land used for agriculture. In addition, once the costs of clearing and developing a parcel have been incurred, marginal returns on that land have been raised. After an export boom that induces clearing, even if export prices fall back to a level that would have been insufficient to induce clearing, the fall may not be sufficient to induce abandonment and reforestation of the recently cleared land.

The theoretical model predicts time of clearing, i.e. the timing of a transition. Therefore we draw on econometric literature on the analysis of duration (e.g. Kiefer 1988, Lancaster 1990). Our intuitions are similar to those in technology adoption, where the variable of interest is the probability that a firm adopts, and to studies of the probability that an individual ends a spell of unemployment by finding a job. In this fashion, our model predicts clearing behavior on individual parcels. When aggregated, its predictions apply to observed deforestation rates on larger areas of land. The deforestation rate in an area at time T is the number of parcels that satisfy (3) at T but did not before, and so are cleared at T, divided by the number that had not satisfied (3) before T, i.e. were not previously cleared.

In the theoretical model the predicted hazard rate is deterministic, since benefits and costs of clearing are exhaustively described by the variables in (3), i.e. returns and changes in clearing costs. However, we cannot perfectly observe those variables currently, let alone their expected future values. Thus, actual returns to land uses and changes in costs will vary even after we have controlled for all observable factors. Assuming that explanatory factors enter in a linear form, the model can be expressed in terms of a latent variable y^*, the value of clearing this period:

$$y_{it}^* = R_{it} - S_{it} - r_t C_t + \frac{d}{dt} C_{it} = \beta' X_{it} - \varepsilon_{it}$$

$$y_{it} = 1 \text{ if } y^*_{it} > 0 \ ; \quad y_{it} = 0 \text{ otherwise}$$

$$\text{Prob}(y_{it}=1) = \text{Prob}(\varepsilon_{it} < \beta' X_{it}) \tag{5}$$

The measured factors X_{it} in (5) for which the coefficients β will be estimated are not direct measures of the variables in (3), so this is not a structural equation. Other than a direct measure of returns (averaged over each time period), the X variables are proxies for returns to cleared land uses relative to returns to forest during a period, for expectations beyond that period, and for costs of clearing.

Land parcels with a lower ε_{it} (i.e. lower unobserved relative returns to forested land uses) are more likely to be cleared. If we define $\varepsilon_{it}*$ as the level of unobservable returns that exactly equals $\beta' X_{it}$ so that the parcel is just worth clearing, the probability that parcel i is cleared in period t conditional on not having been cleared before t, i.e. the hazard rate for parcel i at period t, is:

$$h_{it} = \frac{f(e^*(X_{it}, t))}{1 - F(e^*(X_{it}, t))} \tag{6}$$

where $F(.)$ is the cumulative distribution function for unobservable returns ε. The behavior of the hazard over time depends on $F(.)$, on the way that X_{it} changes over time, and on the rate at which the mean of the $F(.)$ distribution moves over time because of common unobserved factors. We are dealing with multi-year periods, rather than instantaneous hazard rates, so the estimated hazard rate during a period will increase if the rate at which we move through the distribution increases.

The shape of *F(.)* depends on the distributions of unobservable variables that affect cleared and forested land-use returns and clearing costs. These could include unobservable characteristics of land quality, timber quality, unobservable differences in accessibility, availability of family labor, the quality of local distribution services, access to credit, and slope and aspect of the plot. For a given vector of measured factors X_{it}, the hazard rate will depend on how far through the distribution we have moved, i.e. what has previously been cleared and what is left to be considered for clearing. Thus, current deforestation depends on past deforestation.

If the cumulative distribution of ε_{it} is logistic, we have a logit model:

$$F(\beta' X_{it}) = \frac{1}{1+\exp(\beta' X_{it})} \tag{7}$$

Given that we have grouped, not plot-level data, we estimate this model using the minimum logit chi-square method also known as 'grouped logit'.[5] The $\hat{h}_{it}$ is the deforestation rate for a given observation. It is a simple estimate of the parcel-level hazard rate averaged across our multi-parcel units of observation. Thus, we estimate the equation:

$$\log \frac{\hat{h}_{it}}{1-\hat{h}_{it}} = \beta' X_{it} + \mu_{it} \tag{8}$$

where the variance of μ_{it} can be estimated by $\frac{1}{n\hat{h}_i(1-\hat{h}_i)}$, and the estimator is consistent and asymptotically normal.[6] The equation is estimated by weighted least squares.

[5] Berkson (1953) cited in Maddala (1996).
[6] Maddala (1996) p. 30

3.2 SPATIAL ERROR CORRECTION

As written, the errors in (8) could be interpreted as being independent over space and time. This may not be the case. We know the spatial relationships among the 1128 observations in space that are used in our regressions. We use this information to both test and correct for spatial autocorrelation of the errors at each point in time. A natural basis for the hypothesis that such correlations exist is that we do not observe everything about land parcels' natural productivity, and the unobserved elements are likely to be correlated over space. There may be contiguous patches of land with unobserved good qualities and contiguous patches which are unobservably poor. Unobserved elements of returns may also be socioeconomic, such as demand and supply conditions that affect crop prices unobservably across a locality. Even if we find statistical evidence of correlation, we can not assert a single interpretation. However, because it is sufficiently likely spatial correlation that exists, we always correct for it.[7]

Specifically, we implement the maximum likelihood estimation of the linear regression with spatially dependent errors described in Anselin (1988), which requires a judgement about which observations may be related to which others. In the regressions presented below, we have assumed that the observations that may be linked are no more than five kilometers apart. We also test robustness using ten, thirty and fifty kilometers, and find essentially no change in the basic patterns of the results.

3.3 PROXIES FOR UNOBSERVED DYNAMICS

Observed changes in deforestation over time could be explained by the fact that observed factors in returns (X_{it}) change over time, and that land users are working their way through a distribution

[7] See Anselin (1988) for discussion of testing and correction for spatial linkages of different types among errors.

of unobservable elements of returns. An implication of a logistic distribution is that an increase in returns from very low initial levels will yield little clearing; only the small amount of highest quality land will be cleared. However, an equal increase in these returns could induce significant clearing farther along the development path, when a large amount of land is marginal. Later, even a large increase in returns will yield less clearing, as little usable land remains in the tail of the land-quality distribution. If the rate of increase in returns over time were consistent, this distribution alone would imply an 'inverse-U-shaped' pattern of deforestation rates over time.

If these were the only reasons for shifts over time in clearing patterns, however, fixed explanatory factors such as climate and soil quality, though affecting which pieces of land were marginal, could not on their own explain forest transitions. Transition would occur because of changes in X_{it}. We posit here, in contrast, that shifts over time in clearing patterns can also be explained by unobserved factors changing systematically over time.

Unobservable shifts in relative returns can come from a number of sources, including: changes exogenous to individual land managers and localities, such as those due to national-level development processes; changes exogenous to individuals but endogenous to localities, such as those resulting from local development processes; and changes endogenous to the individual land manager's clearing choices, such as convex adjustment costs that can induce partial adjustment.

Regarding exogenous development changes, as a country develops, its infrastructure and institutions (legal and economic) improve. Improved legal systems can make tenure more secure and induce productive investment, while better credit and insurance systems can reduce risk and lower discount rates. Improved distribution networks can lower input costs, and reduce the gap between farm-gate and export prices, while better technologies can lower costs and improve the

---

[8] 'Grouped logit' can be run as OLS, with weighting to correct for the heteroskedasticity by forest area.

quality of output. These factors raise agricultural returns in ways poorly reflected in observed export prices, crop yields and production costs. Thus we may observe deforestation without changes in the observed X_{it} variables. However, the effects of changes in infrastructure may vary with X_{it}. For example, land which is more productive will benefit more from higher farm-gate prices because it has higher yields. Thus land in good lifezones or with good soils might have higher deforestation than poor land in all periods even though these factors are fixed. More generally, fixed factors could have ongoing effects in transition regressions.

Such development probably accelerated from 1950 through the late 1970s, as Costa Rica grew rapidly and opened up to international markets. In early years, when production was labor intensive, new markets and improved transport leading to higher farmgate prices probably raised profitability without changing production methods, yielding extensification and more clearing. In later years, in contrast, the focus may have shifted toward more capital and technology intensive, higher yield, higher quality production implying agricultural intensification. It may be worth investing heavily only on high quality land. Although agricultural returns may have risen on average, the returns on marginal land may have risen much less or not at all. From the mid 1980s, when the environmental movement in Costa Rica became much stronger, increasing governmental and private pressure to protect forests will have raised the cost of clearing, and increased the return to forest, lowering relative agricultural returns. Thus we must allow for a non-linear relationship between time and shifts in clearing patterns. To empirically permit such effects, we include a quadratic in time in the deforestation-transition regressions.

Some aspects of economic and institutional development affect the country as a whole; others may be locally endogenous. One possible process that is exogenous to individuals but endogenous to a small region is that as the forest is cleared and economic activity increases, this itself stimulates further investment in infrastructure such as in credit agencies and roads and in services such as agricultural input supply. These changes would raise returns in ways we cannot observe. Thus as more land in an area is cleared, the hazard rate may rise. This suggests one measurable proxy for endogenous local development, past clearing in the area.

The positive prior for such a proxy is clouded by an alternative dynamic story. If local land users know more than do we about local land quality, the first land cleared will be the best along unobservable dimensions of quality. The extent of past clearing could be a proxy for the lower quality of the land that remains after this 'selection' effect. This suggests that past clearing could have a negative effect on the hazard rate. The net effect is an empirical question. If this selection effect were to take effect only when most of the productive land is cleared, we might expect the negative effect to be more important when the percentage cleared is high. Thus we include our cleared percentage measure in quadratic form.

Finally, whenever agricultural returns shift, adjustment to the new equilibrium level of cleared land may not be instantaneous, in particular if rapid land adjustment is costly. Costs may rise with the amount of adjustment in a given period of time (convex adjustment costs) due to limited local labor and labor mobility. Local wages may rise with an increase in demand for labor for clearing and, because of lack of credit access, capital costs may rise once local sources are exhausted. The partial adjustment to new equilibria that can result would imply that past changes in returns have an ongoing effect on hazard rates, suggesting persistence in forest clearing rates. We include lagged hazards as a proxy to capture such effects. Unfortunately for its interpretation, we cannot distinguish partial adjustment from local temporal autocorrelation in shocks.

4 Data

Table 1 gives the means and standard deviations of the variables used, weighted by the forested area per observation. The following sections describe the data sources and variable definitions.

4.1 Forest Cover & Dependent Variable

4.1.1 Forest Cover Data

We observe forest cover at five points in time: 1963, 1979, 1986, 1997 and 2000. We also assume no deforestation before 1900, for the purpose of calculating 1900-1963 transitions.[9] We observe the data at the level of homogeneous ecological zones, or lifezones, within each of 436 districts. The Holdridge Life Zone System (Holdridge, 1967) divides Costa Rica into 12 ecological zones reflecting levels of precipitation and temperature. In total we have 1229 basic units of observation each period, or 6145 in total. The data used to create the dependent variable come from several different sources.[10]

The 1963 data come from aerial photos (translated into maps) that were digitized by the University of Alberta EOSL to distinguish forest and non-forest.[11] The 1979 data were produced from Landsat satellite images by the National Meteorological Institute of Costa Rica (IMN, 1994), with support from the Ministry of Environment and Energy (MINAE) and the Agriculture Ministry. Final products were printed at a 1:200,000 scale.[12] They extracted data for several

---

[9] Clearly this is a strong assumption. In fact, some areas had been cleared to some extent by that time. However population was only around 200,000. We find little difference using 1850 instead as the "earliest clearing date".

[10] The data for the years up to and including 1997 are described in more detail in Kerr et al (2001). The reason that there are not 12 x 436 (= 5232) observations per period is that not every lifezone is present in every district.

[11] In fact, the maps appear to represent a range of years centered around 1963, another source of measurement error.

[12] They used remote sensing from two different sensors: Landsat 4 (Multispectral Scanner, 80x80m of spatial resolution and 4 spectral bands) and Landsat 5 (Thematic mapper, 25x25 m of spatial resolution and 7 spectral

land cover types which we aggregate at the district level as 'potential forest' (not water, mangroves or exposed rock). We assume that all these areas were forested in 1900. Within potential forest areas, for 1979 they distinguish ‘forest’ and a series of other land-uses which we group as ‘non-forest’. These data were improved, and gaps were completed, by University of Alberta.

The 1986 and 1997 data are from a study done by the Centro Cientifico Tropical and the Research Center for Sustainable Development of the University of Costa Rica for the National Fund for Forestry Financing (FONAFIFO, 1998). These data were derived from Landsat images. The final product differentiates between forest, non-forest, mangroves and secondary forest (land that was not classified as forest in 1986 but was so in 1996). The final map was produced at 1:250,000 scale. The 2000 data also come from Landsat images and were processed by the University of Alberta EOSL to be consistent with the 1986 and 1997 maps.

The 1986 and 1997 data sometimes mis-classify forest in areas where there is deciduous tropical forest, primarily in the Guanacaste region. Satellite images are generally collected during the dry season, when there are few clouds but no leaves in the deciduous forest. Forest cover can be mistaken for either bare soil or pasture land. This introduces measurement error in the dependent variable, though these errors do not appear to be systematic.[13]

4.1.2 Creation of Dependent Variable

From the land-use data we generate measures of the area deforested, in hectares, in each lifezone in each district during each time period. For periods before 1986 we cannot distinguish net from gross changes in forest, and assume that deforestation is equal to the net deforestation rate if this

---

bands). In general, they visually interpreted data from black and white photographic products, and manually extracted fractal boundaries between classes (no image processing).

is positive.[14] In the few cases where it is negative, we set deforestation to zero. Between 1986 and 2000 we can directly identify gross deforestation for each observation. The 1986, 1997 and 2000 maps all include some cloud cover. We have calculated the area deforested from the visible portions of each observation, using pairs of images with consistent cloud masks.

The deforestation rate is the area deforested divided by the total forest area at risk. We assume that once National Parks and Biological Reserves were created their areas were not at risk of deforestation, given the no-usage mandate for such protected areas. They did in fact remain forested.[15] We have GIS maps of these areas in each period, which permit us to exclude their areas from the forest at risk.[16] This reduces our observations to a maximum of 1128 per period. The analysis also drops observations with zero forest at the start of a period, as there is then no risk of deforestation. This yields 4370 total observations for the pooled regressions.[17]

Our time intervals are of varying lengths, so to make the deforestation rates comparable we convert each to an annualized deforestation or hazard rate. If λ is the deforestation rate over the observed interval and n is the number of years in the interval we calculate the hazard rate as:

$$\hat{h}_{it} = 1-(1-\lambda_{it})^{\frac{1}{n}} \tag{9}$$

We thus implicitly assume that this annualized hazard rate was constant within each time period.

---

[13] We reran our regressions for the subset that does not include this area. The results do not change significantly.

[14] Anecdotal evidence suggests reforestation was not widespread before 1986, so this is probably not a major problem.

[15] For discussion of the parks and their forest outcomes see Sanchez et al. (2002 in press).

[16] These areas were delineated by the government (see Castro-Salazar and Arias-Murillo 1998), and later digitized.

[17] Our 'grouped logit' estimation approach would also drop all observations where no clearing is observed. To avoid dropping zero-clearing observations we add one hectare of clearing to all observations. This addition is not material because it is smaller than the minimum mapping unit, so well within the error range.

4.2 EXPLANATORY VARIABLES

4.2.1 Direct Measure of Returns

The annual return to a hectare of land *i* used to grow crop *j* at time *t* is the crop price per kilo, p_{jt}, times the annual yield per hectare, y_{ijt}, minus the costs of production per hectare, $cost_{ijt}$, minus the transport cost per hectare, t_{ijt}:

$$r_{ijt} = p_{jt}\, y_{ijt} - cost_{ijt} - t_{ijt} \tag{10}$$

We estimate the return to non-forest uses of land for each of the four major export crops: coffee, bananas, sugar and beef. Returns are then averaged across the years in each period. This variable then represents the average return to that crop on one hectare of cleared land during the period in question. It is measured in 1997 US$. The data sources and the techniques used to estimate these crop-level return variables are described in detail in Kerr et al (2001). We have data only from 1950 onward and the quality of the data improves significantly in the later years.

Any deforested plot will only be used for one crop at a time. We define s_{ij} as the probability of each crop being chosen as the use of newly cleared land. These shares are used as weights in creating a measure of expected annual return, by district and year:

$$R_{it} = E(r_{it}) = \sum_j s_{ij} r_{ijt} \tag{11}$$

We calculate the weights for implementation of (11) using data on actual production patterns across areas in the 1970s and 1980s as well as information on the suitability of different lifezones for different crops. Thus, for example, in a humid, lower-montane area we represent the land-manager's choices by assuming that cleared land will be used for coffee or for something with a similar return. The weighted sum that results is our summary measure of returns, RETURN.

Historical returns data is extremely difficult to find, however, particularly when we want that data to measure returns consistently across different crops. As a result, our variable inevitably contains a high level of measurement error. Our return variable is most accurately thought of as an indicator of longer-term returns, including future expectations, rather than the returns from one specific year. Higher returns should lead to higher clearing.

4.2.2 Returns Proxies

Given the difficulty of directly measuring returns, we also consider characteristics of the land that could affect the returns in different land uses. Some of our proxies affect the cost of clearing and preparing for agriculture also. For example, areas that are difficult to access, and hence face lower farmgate prices, might also be harder to clear because it will be harder to access machinery and labor.

Lacking a direct dollar measure of transport costs, we use the minimum linear distance in kilometers, DISTANCE, from the center of each observation to the closest of three major cities and ports, San José, Puntarenas and Limon. We also include this distance interacted with time, DISTANCE*TIME. We expect that greater distance lowers deforestation but that this effect diminishes with time because of improved roads and vehicles. In the regressions for the most recent years (1986 – 2000), we also use the district's road density, ROADSDEN, i.e. the total length of roads in the mid 1980s divided by the district area, as an additional proxy for access to markets and thus returns.

To control for local market size (larger markets mean more local demand for output and lower input prices), we include district level population density, POPDEN, and its square POPDEN^2. The population variable comes from census data measured at the district level in 1950 and 1984.

It is simply divided by the area of the district. Because population is potentially endogenous to other factors that lead to economic activity and deforestation, we use lagged population density.

We use ecological variables to control for variation in agricultural productivity. We expect that more productive land will have higher hazard rates. We create dummy variables for three groups of lifezones: good, medium and bad. GOODLZ includes all humid (medium precipitation) areas, which also have moderate temperatures. MEDLZ includes very humid areas (higher precipitation) in moderate to mountain elevations (hence moderate temperature). BADLZ includes very humid areas with high temperatures (tropical), very dry hot areas, and rainy lifezones, all of which are less productive.

We also have data on seven different soil types.[18] Different soil types have different fertility and drainage and hence different agricultural productivity. We have data on soils only outside of national parks. We create one measure, BADSOIL, which measures the proportion of the non-park area in the observation with low productivity entisol soils. We expect soils to be empirically less important than lifezones because of poor data quality.

Finally, as motivated in section 3.3.2, to take account of unobservable changes in net returns that result from exogenous improvements over time in infrastructure and the general development process, we use the number of years from 1900 to the middle of the period, TIME, as well as its square, TIME^2. The creation and protection of parks could be a significant reason why the deforestation rate fell so dramatically in the 1990s. We would like to include this factor. Parks are, however, very likely endogenous, and we do not have adequate instruments, so we exclude them to avoid biased coefficients. To allow for endogenous local development effects we include the percentage of forest that has been previously cleared, or cumulative deforestation,

---

[18] This comes from the Ministry of Agriculture of Costa Rica, from a joint project of theirs with the UN FAO.

%CLEARED. We also include its square, %CLEARED^2, to allow the 'selection' effect to dominate at high clearing levels. To allow for adjustment costs, we include a one-period lagged hazard (PREVHAZARD). Initially this is identical to %CLEARED but in later years they can have independent effects. For that reason, we include lagged hazard only in the last two periods.

5 Results and Discussion

We see in Figure 1 that the unconditional hazard rate begins relatively low, rises to a peak in the period 1979-1986 and then falls to almost zero by 2000. Some of this will be driven by changes in the explanatory variables, some by the shape of the distribution of land quality and some by the speed of the underlying, unobserved development processes that increase pressure on land by raising all economic returns. The regressions help to separate these effects.

5.1 RETURNS MEASURE & STANDARD PROXIES

Table 2 gives results from the grouped logit with all five time periods pooled. In the pooled regression, the direct measure of returns has a negative and insignificant effect. This suggests not that returns do not matter, but that our direct measure of returns contributes little when we also control for returns through the proxy measures of productivity, time and clearing. This is upheld in the cross section regressions, Table 3, where returns are positive and significant in the later years when the data is of significantly better quality.

Transport costs and access to markets appear to play an important role in deforestation. In the pooled regression, the further land is from a market center the less likely it is to be cleared. This effect falls over time, as hypothesized, which could result from development in transport infrastructure and vehicles over time. In the cross sections in Table 3, increased distance initially reduces the probability of deforestation significantly. Between 1963 and 1986, though, this effect is reversed and more isolated places appear to face more deforestation pressure. This may reflect the 'frontier' nature of deforestation during this period where new agricultural areas are opening up rapidly partly as a result of government policy. In the later periods 1986 – 2000, however, distance is once again a slight deterrent to clearing.

We calculate marginal effects in 1986 for an observation with medium lifezone, good soil and otherwise average characteristics (weighted by forest area). Because of the pooled regression's restrictive functional form, the positive interaction term dominates in the later years, even though the effect of distance returns to a more expected negative effect in the cross section. The marginal effect of a one standard deviation (38km) increase in distance to major markets is a 21% increase in the hazard. This suggests some limitations of the simple interactive term in the pool.

The density of roads in a district could not be included in the pooled regression because it is measured only once, in the mid 1980s. In the cross section regressions, it is significant only in the period 1997 – 2000, where it is positive as we would expect. More roads reduce average travel time. Our final measure of market access is local population density. We include this as a quadratic. Before 1986, higher population density leads to more deforestation, though the negative quadratic term suggests that it is the initial populating of an area (rather than increasing the density, e.g. from a rural area with small towns to a city) that is most significant. In recent years, the effect becomes insignificant and the signs reverse. Possibly cities are protecting their few remaining forests, or densely populated areas have exhausted the available productive land.

---

[19] We do not use fixed effects because most of the variation in our explanatory variables is cross sectional.

Our indirect measures of productivity and thus also returns are stronger explanatory factors than is the direct measure. Good lifezones are significantly more likely to be cleared than medium lifezones (the omitted category). This result can also be seen in most of the cross section regressions. Areas with bad soil and in poor lifezones are significantly less likely to be cleared in both the pooled and cross section regressions. For 1986, moving from a medium lifezone to a good lifezone increases the deforestation rate by 21% while moving to a bad lifezone lowers it by 33%. Moving from good to bad soil lowers the deforestation rate by 12%.

5.2 PROXIES FOR UNOBSERVED DYNAMICS

We find that, as expected, time has an initially positive but concave and hence ultimately negative effect on the hazard rate. A one-decade increase in time in 1986 would lead the deforestation rate to fall by 52%. Using time dummies for a pooled run, in contrast to the shape of the unconditional hazard seen in Figure 1 we estimate a relatively constant baseline hazard until 1986 followed by a rapid decline in deforestation in the 1990s. This dramatic regime shift may reflect political and regulatory change as well as an economic shift away from agriculture.

The percentage of forest cleared has a significant positive effect, supporting arguments about endogenous growth, although this result also could indicate that parcels are positively spatially correlated in unobserved ways within our observational units. It is also consistent with partial adjustment or temporal autocorrelation in the early years. The marginal effect of a one standard deviation increase in percentage cleared is a 42% increase in the deforestation rate. The effect of clearing is remarkably consistent across periods. The quadratic cleared percentage term is negative in all but the most recent period. This suggests that the negative selection effect plays a role as clearing rises. In sum, clearing may induce or simply predict more clearing, but when very little forest is left clearing may slow, as remaining land is of low value.

In the cross sections, we can see the effect of the lagged hazard, which we included in an attempt to empirically capture partial adjustment or 'persistence'. A higher level of recent clearing within the observation significantly increases the probability of future clearing. This suggests that equilibrium may not necessarily be reached within each period.

Finally, the spatial correlation coefficient lambda is positive and significant in the pooled regression, and in three of the five cross sections as well. Correcting for spatial correlation does not alter most of the variables significantly. The largest effect is on cleared percentage and its square. These are intended to capture endogenous spatial correlation within observations, so this result across observations is not too surprising. Cleared percentage still has a positive significant effect but the coefficient falls relative to an uncorrected regression. The quadratic term loses its significance in the pooled run, and in the most recent cross section it changes sign.

Overall, our results conform well with theory and we can explain a significant percentage of the variation in deforestation. Consistent with previous literature, factors that vary only across space have strong explanatory power. The long time period covered by our data, however, allows us to go beyond previous work. We find that some of our 'dynamic' variables are significant. The baseline deforestation rate changes significantly over time, with a drop in the 1990s after a long stable period. We find evidence suggestive of partial adjustment, and suggesting that current clearing may endogenously induce future clearing nearby. We also find that the time period does matter for some effects. For instance, while theory and the literature suggest that distance will always be correlated with lower deforestation, we find this to be true only in some periods of development. In sum, in developing countries it appears that deforestation is a dynamic process, within which the effects of observable variables can change through time.

References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Studies in Operational Regional Science, Kluwer Academic Publishers, Dordrecht, 284p.

Bell, K.P. and N.E. Bockstael (2000). "Applying the generalized moments estimation approach to spatial problems involving micro-level data". *Review of Economics & Statistics* 82:72-82.

Castro-Salazar, R. and G. Arias-Murillo (1998). *Costa Rica: toward the sustainability of its forest resources*. Technical Report, FONAFIFO (Fondo Nacional de Financiamento Forestal). San José, Costa Rica

Chomitz, K.M. and D.A. Gray (1996). "Roads, Land Use and Deforestation: A Spatial Model Applied to Belize". *World Bank Economic Review* 10(3):487-512.

Cropper, M. and C. Griffiths (1994). "The Interaction of Population Growth and Environmental Quality". *American Economics Review: Papers and Proceedings* 84(2):250-254.

Cropper, M, Griffiths, C and Mani, M. (1997) "Roads, population pressures, and deforestation in Thailand, 1976-89". Policy Research Working Paper 1726. World Bank, Washington, DC.

FONAFIFO (1998). *Mapa de Cobertura Forestal de Costa Rica*. San José, Costa Rica.

Geoghegan, J., S.C. Villar, P. Klepeis, P.M. Mendoza, Y. Ogneva-Himmelberger, R.R. Chowdhury, B.L. Turner II and C. Vance (2001). "Modeling tropical deforestation in the southern Yucatan peninsular region: comparing survey and satellite data". *Agriculture Ecosystems & Environment* 85:25-46.

Harrison, Susan (1991). "Population Growth, Land Use and Deforestation in Costa Rica, 1950-1984." *Interciencia* 16(2):83-93.

Holdridge, L. 1967. *Life zone ecology*. Tropical Science Center, San José, Costa Rica.

IMN (Instituto Meteorologico Nacional) (1994). *Mapa de Uso de la Tierra de Costa Rica*. San José, Costa Rica.

Kaimowitz D. and A. Angelsen (1998). *Economic Models of Tropical Deforestation: A Review*. CIFOR, Indonesia.

Kerr, S., A.S.P. Pfaff, and A. Sanchez (2001) "The Dynamics of Deforestation and the Supply of Carbon Sequestration: Illustrative Results from Costa Rica" in Theodore Panayotou ed. *Environment for Growth: Environmental Management for Sustainability and Competitiveness in Central America*. Harvard Studies in International Development, Harvard University Press pp. 409 - 431

Kiefer, N.M. (1988) "Economic Duration Data and Hazard Functions" *Journal of Economic Literature* Vol. XXVI, June, 646-679

Lancaster, T. (1990) *The Econometric Analysis of Transition Data* Econometric Society Monograph No. 17, (Cambridge University Press)

Maddala, G., (1983) *Limited-Dependent and Qualitative Variables in Econometrics* Cambridge: Cambridge University Press

Nelson, G.C. and D. Hellerstein (1997). "Do Roads Cause Deforestation? Using Satellite Images in Econometric Analysis of Land Use" *American J. of Agricultural Economics* 79: 80-88.

Pfaff, A.S.P (1999). "What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data". *J. of Environmental Economics and Mgmt* 37(1):26.

Pontius, R.G. Jr., J.D. Cornell and C.A.S. Hall (2001). "Modeling the spatial pattern of land-use change with GEOMOD2: application and validation for Costa Rica". *Agriculture*

Ecosystems & Environment 85:191-203.

Rosero-Bixby L. and A. Palloni (1996). Population and Deforestation in Costa Rica. Paper presented at the Annual Meeting of the Population Association of America in New Orleans.

Sader, S.A. and Joyce, A. T. (1988) "Deforestation rates and trends in Costa Rica" *Biotropica* 20: 11-19

Saloner, G. and A. Shepard. 1995. Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines. *RAND Journal of Economics* 26(3):479–501.

Sanchez-Azofeifa, G.A., G.C. Daily, A.S.P. Pfaff, and C. Busch (in press 2002). "Integrity and Isolation of Costa Rica's national parks and biological reserves: examining the dynamics of land-cover change". *Biological Conservation*.

Stavins, R.N., A. Jaffe (1990). "Unintended Impacts of Public Investments on Private Decisions: The Depletion of Forested Wetlands". *American Economic Review*, 80(3):337-352.

<u>Figure 1</u> -- Unconditional Hazard Rates Over Time

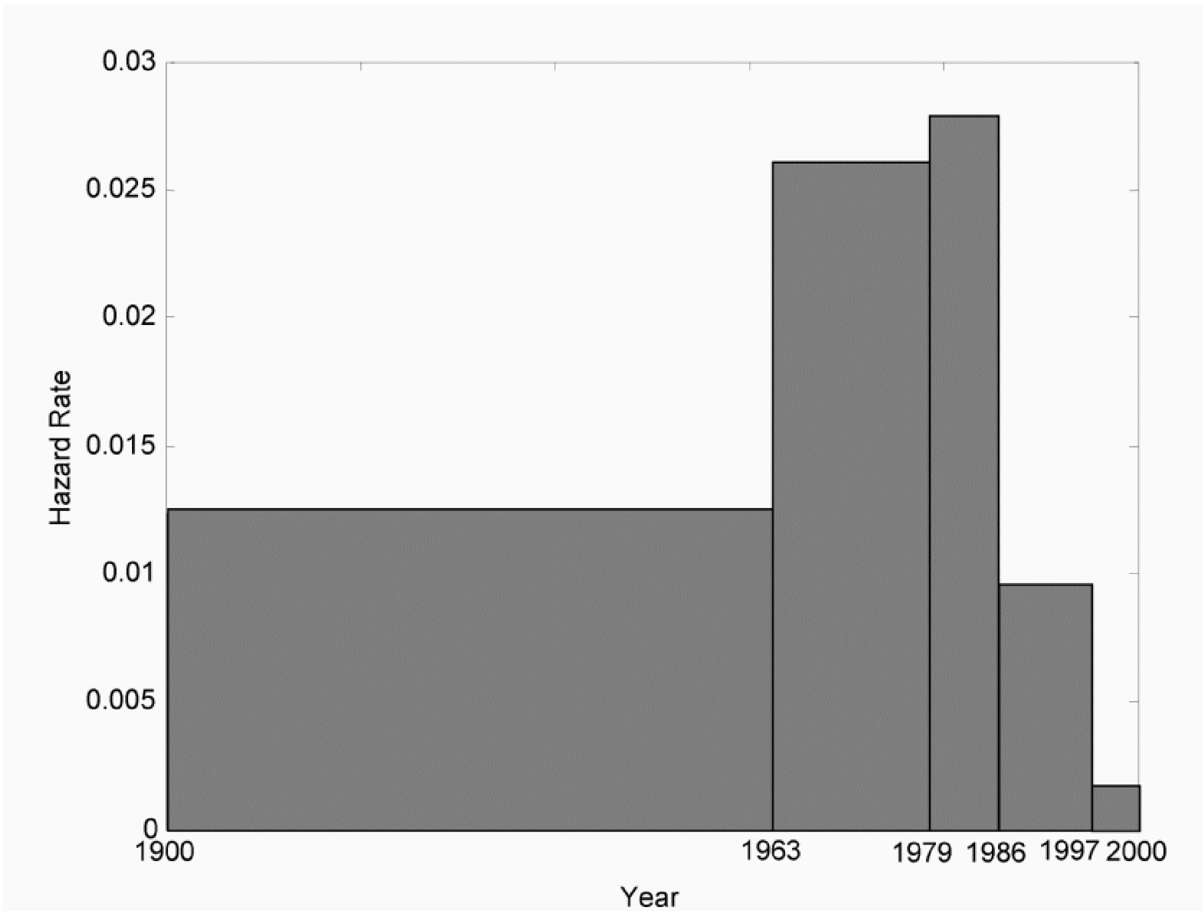

Table 1 -- Variable Definitions and Descriptive Statistics (weighted by forest area)

Variable	Name	Mean	Standard Deviation	Minimum	Maximum
Deforestation rate	Dep.Var.	0.016	0.048	0.00001	1
Returns					
Returns/ha US$ 1997	RETURN	659	1154	0	5047
Distance to major markets (km)	DISTANCE	73	38	0	186
Roads density	ROADSDEN	0.0026	0.0022	0	0.049
Population density (#/ha)	POPDEN	0.089	0.38	0	107
Ecology					
Dummy for humid lifezones	GOOD LZ	0.24	0.43	0	1
Dummy for very humid (pre montane, lower montane) and montane lifezones	MEDIUM LZ	0.23	0.42	0	1
Dummy for very humid (tropical), dry (tropical), and rainy lifezones	BAD LZ	0.54	0.50	0	1
Proportion of known soil types that are entisol	BADSOIL	0.11	0.23	0	1
Dynamics					
Time measured from midpoint of period	TIME	66	26	33	100
Proportion of forest cleared	% CLEARED	0.21	0.26	0	0.99996
Lagged Hazard (1979 onward)	PREVHAZARD	0.0080	0.017	0	0.46

Table 2 -- Pooled Regression Results

	Grouped Logit (spatial error correction, 5km)			
Years	1900 – 2000, pooled transitions			
Dep.Variable	annualized deforestation probability			
Explanatory Variables	**Coefficients** (t statistics)	**Defaults** &	**Marginal Δs**	**1986 Marginal Effects** (1986 default hazard = 0.048)
RETURN	-1.4 E-05 (-0.79)	1232	1685	-0.0011 (2% of default)
DISTANCE	-0.018 (-14)	71km	38 km	0.0100 (21%)
DISTANCE*TIME	2.6 E-04 (14)	(implied	by the above)	(joint with above)
GOOD LZ	0.21 (6.1)	0	1	0.0100 (21%)
BAD LZ	-0.42 (-11)	0	1	-0.0160 (33%)
BADSOIL	-0.13 (-1.8)	0	1	-0.0056 (12%)
TIME	0.13 (21)	86	10	-0.0250 (52%)
TIME^2	-0.0012 (-24)	(implied	by the above)	(joint with above)
% CLEARED	1.9 (8.1)	37%	28%	0.0200 (42%)
% CLEARED^2	-0.78 (-3.0)	(implied	by the above)	(joint with above)
_CONS	-6.2 (-34)			
LAMBDA	0.31 (19)			
R^2	0.36			
N	4370			

Table 3 -- Cross-section Regression Results

	Grouped Logit (spatial error correction. 5km)				
Years	1900-1963	1963-1979	1979-1986	1986-1997	1997-2000
Dependent Variable	annualized def. prob.	annualized def. prob.	annualized def. prob.	annualized def. prob.	annualized def. prob.
Explanatory Variables			**Coefficients** (t statistics)		
RETURN		2.5 E-04 (2.2)	-1.3E-04 (-4.5)	2.7E-04 (13)	2.4E-04 (5.2)
DISTANCE	-0.011 (-16)	0.0035 (3.7)	0.0012 (0.93)	-0.0042 (-4.5)	-0.0017 (-1.2)
POPDEN		0.44 (2.6)	0.94 (3.3)	-0.074 (-0.46)	-0.14 (-2)
POPDEN^2		-0.029 (-2)	-0.06 (-2.6)	0.007 (0.39)	0.0046 (2.3)
ROADSDEN				21 (0.93)	87 (3.7)
GOOD LZ	0.20 (3.8)	0.078 (0.93)	0.23 (2.6)	0.15 (1.7)	0.082 (0.57)
BAD LZ	-0.63 (-8.2)	-0.40 (-4.4)	-0.59 (-6.0)	-0.59 (-8)	-0.60 (-4.4)
BADSOIL	0.0061 (0.056)	-0.40 (-2.4)	-0.14 (-0.82)	-0.89 (-5.3)	-0.38 (-1.4)
% CLEARED		2.2 (4.7)	1.9 (3.6)	2.9 (7.2)	0.33 (0.47)
$\%\ \text{CLEARED}^2$		-0.34 (-0.64)	-1.7 (-2.8)	-0.97 (-2.4)	1.29 (1.9)
PREVHAZARD				2.3 (2.6)	14 (7.4)
_CONS	-3.1 (-56)	-3.7 (28)	-2.8 (-15)	-5.1 (-34)	-5.9 (-25)
LAMBDA	0.21 (5.7)	0.21 (3.7)	0.21 (3.4)	-0.044 (-0.58)	0.22 (3.3)
R^2	0.31	0.29	0.29	0.51	0.39
N	1128	800	644	650	788